Amplitudes vs Mixtures

Consider the state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, which corresponds to following computational basis (Z basis) measurement distribution:

$$\mu_+(0) = |\langle 0|+\rangle|^2 = \frac{1}{2} \quad , \quad \mu_+(1) = |\langle 1|+\rangle|^2 = \frac{1}{2}$$

For any state $|R(\phi)\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\phi}|1\rangle\right)$, the relative phase $e^{i\phi}$ between 0 and 1 does not affect the computational basis measurement distribution:

$$\mu_{R(\phi)}(0) = |\langle 0|R(\phi)\rangle|^2 = \frac{1}{2} \quad , \quad \mu_{R(\phi)}(1) = |\langle 1|R(\phi)\rangle|^2 = \left|\frac{e^{i\phi}}{\sqrt{2}}\right|^2 = \frac{1}{2}$$

Amplitudes vs Mixtures

In particular, both the states $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$

have the same computational basis measurement distribution:

$$\mu_{\pm}(0) = \frac{1}{2} \quad , \quad \mu_{\pm}(1) = \frac{1}{2}$$

Both states have a uniform distribution of measurement outcomes in the computational basis. Any **mixture** of two uniform distributions with be uniform.

Consider a quantum state that combines $|+\rangle, |-\rangle$ in equal parts,

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

Is the computational basis probability distribution of this state uniform?

Amplitudes vs Mixtures

As we've seen, $|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) = |0\rangle$, so the measurement distribution is

$$\mu_\psi(0) = 1 \quad , \quad \mu_\psi(1) = 0$$

This example is the simplest possible illustration of **quantum interference**. With regard to the distribution of measurements, the whole is not a weighted sum of the parts.

A normalized complex linear combination of quantum states is called a **superposition**

$$z_1|\psi_1\rangle + z_2|\psi_2 \ , \ z_1, z_2 \in \mathbb{C} \ , \ |z_1|^2 + |z_2|^2 = 1$$

We will frequently contrasts the terms **superposition** and **mixture** because (as we've seen) interference is one of the main differences between QM and classical probability theory.

Incompatible Measurements

In our discussion of axiomatic QM, we said it was natural to update the state of a probability theory after measurement reveals additional information about the state, and QM simply retains this principle.

This is technically true, but the reason measurement is considered “weird” in QM is because a measurement of a particular event can unexpectedly change the distribution of incompatible events.

Return to the state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and perform a computational basis measurement. It has an equal probability of returning 0 or 1, let’s say it returns 0. So the state of our qubit is now: $|0\rangle$

So far this is identical to what happens in a probability theory.

Incompatible Measurements

In a probability theory, learning outcomes of additional events would never cause us to update our record about past outcomes. This is because all the events in a probability theory are compatible.

Suppose now we measure our qubit in a different basis, the +/- basis, which corresponds to learning about outcomes of additional events. Our qubit was originally in the $|+\rangle$ state, but then we measured it in the computational basis and updated our state to $|0\rangle$. To find the probabilities for the +/- measurement we compute:

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

So we see that the +/- measurement yields each outcome with probability ½. Suppose we the measurement yields "+", then we update the state to be

$$|+\rangle$$

We've gone in a circle!

Incompatible Measurements

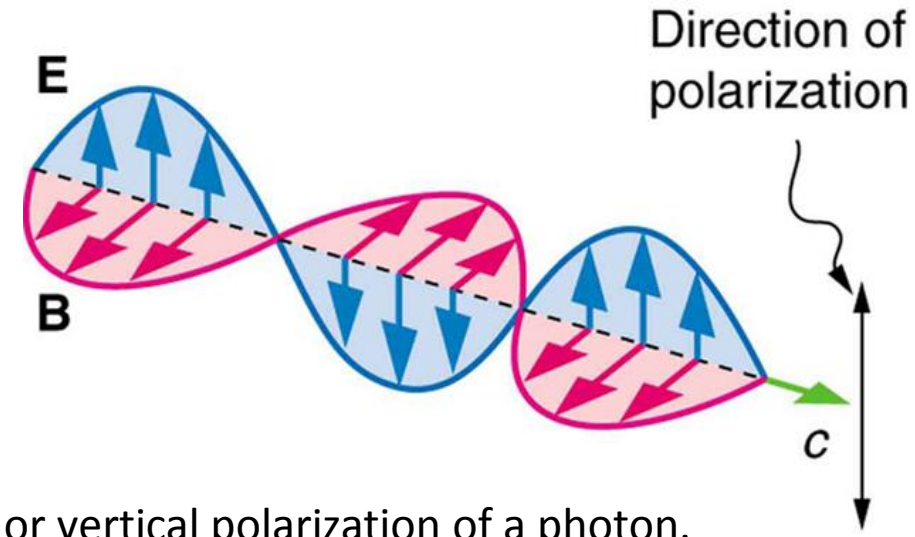
To appreciate why this example is strange, suppose we measured the initial state in the computational basis and got “0”. Then if we did the same measurement again, we would never see a “1”.

However, by first measuring “0” and then measuring the incompatible event “+”, both of which happened with probability ½, then we’ve gone full circle.

Pairs of measurements like X and Z which are as “incompatible as possible” are called **complementary**.

Incompatible Measurements

Polarization states of photons provide a physical implementation (and perhaps a more familiar context) for this example.



Let the states $|0\rangle, |1\rangle$ correspond to the horizontal or vertical polarization of a photon.

Then the state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ is polarized along the axis that is 45 degrees between H and V .

If light passes through a horizontal polarization detector, and then a vertical detector, no light is observed to pass through the vertical detector. But if we place a 45 degree detector between the other two, then we still observe ¼ of the total light getting through!

Composite Systems: Two Qubits

"Stepping up from one qubit to two is a bigger leap than you might expect. Much that is weird and wonderful about quantum mechanics can be appreciated by considering the properties of the quantum states of two qubits." --- J. Preskill

The state of a two qubit system is a vector in $\mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^4$. The computational basis is:

$$|00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad , \quad |01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad |10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad , \quad |11\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Composite Systems: Two Qubits

Observables: a complete basis for 4 x 4 Hermitian matrices can be obtained by taking all tensor products of Paulis

$$\mathcal{P}^2 = \mathcal{P} \otimes \mathcal{P} = \{I, X, Y, Z\} \otimes \{I, X, Y, Z\}$$

Sometimes we label observables with subscripts to indicate which system they act on $X_1 \otimes Y_2$, Other times that is clear from context.

Subscripts are also frequently used to suppress identity factors in the tensor product e.g.

$$Z \otimes I = Z_1 \qquad\qquad I \otimes X = X_2$$

The same notation is also used for tensor product unitaries (and the Pauli operators are unitary).

Composite Systems: Two Qubits

In our review of probability theories, we saw that the joint state of a composite system is a tensor product of the marginal state on the component subsystems iff the subsystems are independent.

$$\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \otimes \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

Therefore any correlated probability distribution yields a vector that is not a tensor product:

$$\begin{bmatrix} 1/2 \\ 0 \\ 0 \\ 1/2 \end{bmatrix} = \frac{1}{2}\mathbf{00} + \frac{1}{2}\mathbf{11} \neq \mathbf{a} \otimes \mathbf{b}$$

Composite Systems: Two Qubits

If a vector in a tensor product vector space cannot be decomposed into a tensor product of vectors, then non-zero scalar multiples of the vector cannot be decomposed as tensor products either.

Therefore the vector describing a correlated probability distribution also describes a quantum state (which happens to have all positive amplitudes) that can't be decomposed as a tensor product.

$$\begin{bmatrix} 1/2 \\ 0 \\ 0 \\ 1/2 \end{bmatrix} \equiv_{ray} \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \neq |\psi_1\rangle \otimes |\psi_2\rangle$$

Composite Systems: Two Qubits

Definition: A quantum state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is **entangled** (across the bipartition A/B) if it cannot be written as $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ for any $|\psi_A\rangle \in \mathcal{H}_A$, $|\psi_B\rangle \in \mathcal{H}_B$.

Using the previous example imported from probability theory, we have our first entangled state:

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

This state is so important that it has the standardized notation $|\Phi^+\rangle$. It is also sometimes called after various people (“EPR pair” , “Bell pair”).

Composite Systems: Two Qubits

We've built an entangled state of two qubits from a simple correlated probability distribution. If entanglement is nothing but correlation, then why give it a new name?

Computational basis measurements of the two qubits in the state $|\Phi^+\rangle$ have perfectly correlated outcomes: if one qubit is measured to be 0, the other will also be 0, and similarly for the outcome 1.

Suppose instead we measure in the +/- basis. After a bit of algebra one finds:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}\left(|++\rangle + |--\rangle\right)$$

Therefore the measurement outcomes in the +/- basis are also perfectly correlated!

Composite Systems: Two Qubits

Beginning from a product state $|00\rangle$, how do we arrive at an entangled state like $|\Phi^+\rangle$?

We could consider tensor products of our single-qubit unitary transformations, but it turns out that these cannot generate any entangled states:

$$U_1 \otimes U_2|00\rangle = (U_1|0\rangle) \otimes (U_2|0\rangle) = |\psi_1\rangle \otimes |\psi_2\rangle$$

Evidently we will need a unitary U that cannot be written as a tensor product,

$$U \neq U_1 \otimes U_2 \text{ for all } U_1, U_2$$

Composite Systems: Two Qubits

An important example of a unitary which is capable of creating entanglement is called CNOT.

$$\mathrm{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

The name is short for "controlled" not. A semi-standard notation is $\Lambda(X)$, where Λ is a general piece of notation that means "controlled" , and Pauli X is a NOT gate (bit flip).

$$\Lambda(X)|00\rangle = |00\rangle \ , \ \Lambda(X)|01\rangle = |01\rangle \ , \ \Lambda(X)|10\rangle = |11\rangle \ , \ \Lambda(X)|11\rangle = |10\rangle$$

"If the first bit is a 1, then apply an X (NOT)". In general $\Lambda(U)$ is the first bit is 1, apply the unitary U. This is called a controlled unitary.

Combining CNOT and Hadamard, $\Lambda(X)\,(H \otimes I)\,|00\rangle = |\Phi^+\rangle$ and so CNOT is indeed capable of transforming a product state into an entangled state.

Composite Systems: Two Qubits

If U is a two-qubit unitary that is not written as a tensor product, is it always the case that U is entangling?

Hint: what is the “most classical” thing you can do to two qubits, that you can’t do to one alone?

Composite Systems: Two Qubits

If U is a two-qubit unitary that is not written as a tensor product, is it always the case that U is entangling?

Hint: what is the "most classical" thing you can do to two qubits, that you can't do to one alone?

Answer: Have them trade places! Apply the swap gate:

$$\mathrm{SWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \frac{1}{2}\left(I + \vec{S_1} \cdot \vec{S_2}\right)$$

$$\mathrm{SWAP}|00\rangle = |00\rangle \ , \ \mathrm{SWAP}|01\rangle = |10\rangle \ , \ \mathrm{SWAP}|10\rangle = |01\rangle \ , \ \mathrm{SWAP}|11\rangle = |11\rangle$$

Composite Systems: Two Qubits

In addition to classifying which two-qubit gates are entangling, we also may ask about "universal" sets of two-qubit quantum gates, which suffice to generate any entangled state of two qubits.

Theorem: all two-qubit gates of the form $(V_1 \otimes V_2)\, SWAP\, (W_1 \otimes W_2)$, where V_1, V_2, W_1, W_2 , are unentangling. All two qubit gates not expressed in this form.

Here we say a two-qubit gate is universal if it can be used, in combination with arbitrary single-qubit gates, to generate any two-qubit unitary.

Theorem: every two-qubit entangling gate is universal. [Phys. Rev. Lett. 89, 247902 (2002)]

Proofs of universality tend to use the algebraic structure of the unitary group, for example there is a canonical decomposition of any two-qubit gate

$$U = (V_1 \otimes V_2)\, e^{i(\alpha_x XX + \alpha_y YY + \alpha_z ZZ)}\, (W_1 \otimes W_2)$$

Composite Systems: Two Qubits

The fact that any entangling gate is universal means that the CNOT gate is universal.

The fact the CNOT is entangling, and even universal, is surprising given the way it acts "classically" on the computational basis:

$$\Lambda(X)|00\rangle = |00\rangle \ , \ \Lambda(X)|01\rangle = |01\rangle \ , \ \Lambda(X)|10\rangle = |11\rangle \ , \ \Lambda(X)|11\rangle = |10\rangle$$

CNOT maps each computational basis states to another computational basis states (as opposed to superpositions thereof). The unitary matrices with this property are permutation matrices, and they are all gates that can be used in the context of classical reversible (deterministic) computation.

Even though CNOT can be regarded as a classical gate, it gets its power from the way that interacts over superpositions of basis states $\Lambda(X)\,(H \otimes I)\,|00\rangle = |\Phi^+\rangle$.

An analogous classical statement holds: if we apply CNOT to an uncorrelated probability distribution on two pbits, then it can generate correlations.

Composite Systems: Two Qubits

In addition to being a permutation gate that is also entangling and universal, the other reason to give CNOT so much attention is that it is a Clifford gate.

Recall that the single-qubit Clifford group is the set of unitaries which map Paulis to Paulis:

$$\sigma \in \mathcal{P} \implies U\sigma U^\dagger \in \mathcal{P}$$

Our examples of Clifford unitaries were the Paulis themselves, Hadamard which maps X to Z, and the phase gate P which maps X to Y. We saw the single-qubit Clifford group is generated by H and P.

The two-qubit Clifford group is the set of unitaries which map two-qubit Paulis to two-qubit Paulis:

$$\sigma \in \mathcal{P}^2 \implies U\sigma U^\dagger \in \mathcal{P}^2$$

Exercise: prove CNOT is a Clifford gate. Find a minimal set of generators for all two-qubit Cliffords.

Component Subystems

So far we have discussed two-qubit states, observables, and unitary evolutions.

But suppose two qubits A and B are originally prepared in some state, which may be entangled, and A and B are sent to two keepers in different locations: Alice and Bob.

Each of the qubit keepers is allowed to probe their qubit as they see fit, but right now they are not allowed to communicate or share notes about their results.

Each keeper has a component subsystem of the same joint quantum state. If one of the keepers measures their subsystem, they will see outcomes. Our best description of the reduced state of the subsystem should account for all the observed probabilities of these outcomes.

In a probability theory, this reduced state of the subsystem would be a marginal distribution. But in quantum theory we have a distribution for each choice of basis, so it seems much more difficult to marginalize a quantum state in such a way to obtain the correct marginals on every subsystem.

Component Subystems

Take our example of an entangled state, $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|++\rangle + |--\rangle)$

Now suppose Alice holds the first qubit, and measures it in the 0/1. We can compute the probability for her to see 0 and to see 1 by finding the full measurement distribution in the 0/1 basis and computing its marginal. We find an identical result in the +/- basis:

$$\mu^A_{\Phi^+}(0) = \frac{1}{2}\ ,\ \mu^A_{\Phi^+}(1) = \frac{1}{2}\ ,\ \mu^B_{\Phi^+}(0) = \frac{1}{2}\ ,\ \mu^B_{\Phi^+}(0) = \frac{1}{2}$$

What quantum state $|\psi\rangle = \begin{bmatrix}\alpha\\ \beta\end{bmatrix}$ on Alice's subsystem is consistent with these measurements?

Component Subystems

There is no quantum state, at least of the kind we have discussed so far, which has these outcomes!

The reason this happens is that in losing or ignoring Bob's subsystem, Alice now has incomplete information about the full system. For all she knows, Bob could have measured or interacted with his qubit in any way whatsoever. Therefore the object she holds is not a quantum state, it is a **probability distribution of possible quantum states**.

How do we describe a prob distribution over quantum states? Well we could try combining $|\psi_1\rangle, |\psi_2\rangle$ with probabilities p_1, p_2 like we did for mixtures:

$$\sqrt{p_1}|\psi_1\rangle + \sqrt{p_2}|\psi_2\rangle$$

But this only works to create a mixture distribution in at most one choice of basis. In every other choice of basis we get unwanted interference instead of a simple mixture of outcomes.

Component Subystems

Suppose A is an observable, and we have a quantum system in the state $|\psi_1\rangle$ with probability p_1 , And the state $|\psi_2\rangle$ with probability p_2 . Then the expectation of A should be

$$\langle A\rangle = p_1\langle\psi_1|A|\psi_1\rangle + p_2\langle\psi_2|A|\psi_2\rangle$$

In contrast, any superposition of $|\psi_1\rangle$ and $|\psi_2\rangle$ will give rise to unwanted cross terms like $\langle\psi_1|A|\psi_2\rangle$

As long as $|\psi_1\rangle, |\psi_2\rangle$ are orthogonal states, we can express the expectation above as

$$\langle A\rangle = \mathrm{tr}\,(A\rho) \quad , \quad \rho = p_1|\psi_1\rangle\langle\psi_1| + p_2|\psi_2\rangle\langle\psi_2|$$

This new object ρ, which represents a probability distribution over quantum states, is a Hermitian matrix with nonnegative eigenvalues that sum to 1. It is called a **density matrix** and it will suffice to describe the **reduced state of component subsystems**.

Component Subystems

To see that density matrices are the right choice for describing component subsystems, we can consider an observable A which acts only on the first qubit of an entangled state:

$$A \otimes I \quad , \quad |\psi\rangle = a|00\rangle + b|11\rangle$$

The expectation value is $\langle A\rangle = |a|^2\langle 0|A|0\rangle + |b|^2\langle 1|B|1\rangle$

Which can equivalently be expressed as

$$\langle A\rangle = \mathrm{tr}\,(A\rho) \quad , \quad \rho = |a|^2|0\rangle\langle 0| + |b|^2|1\rangle\langle 1|$$

Component Subystems

To compute the marginal of a probability distribution, we sum over all possible events on the subsystem being ignored. The density matrix ρ_A of a subsystem A , for the joint state $|\psi\rangle$ on A/B is:

$$\rho_A = \sum_{a_i, a_j} \rho_{ij}^A |a_i\rangle\langle a_j| \quad , \quad \rho_{ij}^A = \sum_{b_k} \langle a_i b_k|\psi\rangle\langle\psi|a_j b_k\rangle$$

where $\{a_i\}$ is a basis for $\mathcal{H}_A$ and $\{b_i\}$ is a basis for $\mathcal{H}_B$.

This operation which maps $|\psi_{AB}\rangle$ to ρ_A is called the **partial trace** (over the subsystem B, in this case). It is as fundamental in quantum theory as the notion of a marginal distribution in probability.

We have already seen that a joint distribution with intricate correlations can have marginal distributions that are flat and boring. A similar thing happens for entangled states! (e.g. $|\phi^+\rangle$).

Component Subsystems

A key reason why we cannot ignore density matrices (and why would you want to?) is that closed systems are an idealization. Every quantum subsystem in the real world is a subsystem (of the universe).

The quantum states we have described so far are called **pure states**. A pure state corresponds to a density matrix whose distribution is fully concentrated on that one state:

$$\Psi = |\psi\rangle\langle\psi|$$

Whether we can describe a physical system by a pure state depends on whether it shares correlations with any other systems. If it does not (at least to a good approximation), then

$$|\psi_{\text{universe}}\rangle \approx |\psi_{\text{system}}\rangle \otimes |\psi_{\text{environment}}\rangle$$

So that the reduced state of the subsystem is relatively pure, $\Psi_{\text{system}} = |\psi_{\text{system}}\rangle\langle\psi_{\text{system}}|$.
Engineering this near total lack of correlations is a major challenge, and so in general we must regard our quantum systems as open (not closed) and describe them by density operators.

Component Subystems

A key reason why we cannot ignore density matrices (and why would you want to?) is that closed systems are an idealization. Every quantum subsystem in the real world is a subsystem (of the universe).

The quantum states we have described so far are called **pure states**. A pure state corresponds to a density matrix whose distribution is fully concentrated on that one state:

$$\Psi = |\psi\rangle\langle\psi|$$

Whether we can describe a physical system by a pure state depends on whether it shares correlations with any other systems. If it does not (at least to a good approximation), then

$$|\psi_{\text{universe}}\rangle \approx |\psi_{\text{system}}\rangle \otimes |\psi_{\text{environment}}\rangle$$

So that the reduced state of the subsystem is relatively pure, $\Psi_{\text{system}} = |\psi_{\text{system}}\rangle\langle\psi_{\text{system}}|$. Engineering this near total lack of correlations is a major challenge, and so in general we must regard our quantum systems as open (not closed) and describe them by density operators.

For an open quantum system, states are not rays in Hilbert space, but instead are described density matrices. It will turn out that the evolution of density matrices need not be unitary, but instead belongs to a more general class of physical evolutions called **Quantum Channels**. Axioms next time!