

A few words on mixed and pure states

Santiago Quintero de los Ríos

2018

I see that there is a bit of confusion between mixed and pure states in quantum mechanics. This is because the measurement of arbitrary observables for **pure states** is probabilistic, and this is easily confused with the probabilities associated to a **mixed state**.

So let's begin with the probabilistic nature of measurement of observables of **pure** states.

Pure states

Let \mathcal{H} be a Hilbert space (which for our sanity, we will assume finite-dimensional, of dimension N). Any given vector $|\psi\rangle$ is a **pure state**. For example, in the spin-1/2 case, both $|+\rangle$ and $|-\rangle$ are pure states, but also any complex-linear combination of them is a pure state (since these two vectors span all of \mathcal{H}).

Now let A be an observable with associated hermitian operator \hat{A} . Suppose that \hat{A} has a set of orthonormal eigenstates $\{|a_1\rangle, \dots, |a_n\rangle\}$. This means that $\hat{A}|a_k\rangle = a_k|a_k\rangle$. Assuming that the spectrum does not have any degeneracies (i.e. repeated eigenvalues), then given any arbitrary state $|\psi\rangle$, the **probability** of obtaining the value a_k when measuring the observable A is

$$P(A = a_k|\psi) = |\langle a_k|\psi\rangle|^2.$$

Here we used standard notation from probability, and we explicitly wrote $P(A = a_k|\psi)$ as a conditional probability, since this is the probability to measure the value a_k of the observable A *given the fact* that we **know** that the system is in state $|\psi\rangle$. Note that applying this equation to $|\psi\rangle = |a_j\rangle$ for some j , we obtain

$$P(A = a_k|a_j) = |\langle a_k|a_j\rangle|^2 = \delta_{kj}.$$

This is perfectly consistent with the axioms of quantum mechanics. If we measure the observable A **knowing** that the system is in the eigenstate $|a_j\rangle$, then we are absolutely *certain* that the measurement will return the value a_j . However, if the state is **not** an eigenstate of \hat{A} , then we have *uncertainty* about what the measurement will return. This (un)certainty is what the probability $P(A = a_k|\psi)$ represents.

Then **even if we are absolutely certain that the system is in a given state**, there is uncertainty regarding the outcome of (most) experiments.

But what if we are not even certain about what state the system is in? That's where mixed states come in.

Mixed states

Imagine the following machine: It can emit a particle randomly in *one* of two different states $|\psi_1\rangle, |\psi_2\rangle$, each state $|\psi_j\rangle$ with probability p_j . After being emitted, the particle travels to a detector which measures the observable A . After the machine shoots a particle, we are *uncertain* about what the state of the particle is. It might be in state $|\psi_1\rangle$ or in state $|\psi_2\rangle$, the point is that *we don't know*. In this case we say that the system is in a **mixed state**. Compare this to the case of the previous section, where we were *absolutely* certain that the particle was in a specific state (this might be achieved by *preparing the system* first, e.g. via a measurement).

Now we ask: After the random machine shoots a particle, what is the probability of measuring the value a_k for the observable A ?

We can make use of a bit of probability theory and *marginalize*. We are absolutely certain that the particle will be in **one** of the states $|\psi_j\rangle$, but we don't know which one. The event "the observed value of A is a_k " is then trivially equivalent to the two events

1. The observed value of A is a_k , AND
2. the particle is either in state $|\psi_1\rangle$ or in state $|\psi_2\rangle$.

Of course, statement 2 is rather trivial. But both statements can be written simultaneously as

$$(A = a_k) \wedge (|\psi\rangle = |\psi_1\rangle \vee |\psi\rangle = |\psi_2\rangle).$$

We now use the distributive property of conjunction and disjunction ($p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$) and obtain that our statement is equivalent to

$$(A = a_k \wedge |\psi\rangle = |\psi_1\rangle) \vee (A = a_k \wedge |\psi\rangle = |\psi_2\rangle).$$

Then the probability that we measure a_k is equal to the probability that

1. The observed value of A is a_k AND the state is $|\psi_1\rangle$, OR
2. the observed value of A is a_k AND the state is $|\psi_2\rangle$.

Therefore:

$$P(A = a_k) = P(A = a_k \wedge |\psi\rangle = |\psi_1\rangle) + P(A = a_k \wedge |\psi\rangle = |\psi_2\rangle).$$

But now we use the **definition** of conditional probability: $P(A \wedge B) = P(A|B)P(B)$, that is, the probability both A and B occur is the probability that B occurs times the probability that A occurs given the fact that we know that B occurs. Using this definition, the previous equation becomes (I'm dropping the $A = \dots$ and $|\psi\rangle = \dots$ parts)

$$\begin{aligned} P(A = a_k) &= P(a_k | |\psi_1\rangle)P(|\psi_1\rangle) + P(a_k | |\psi_2\rangle)P(|\psi_2\rangle) \\ &= p_1 P(a_k | |\psi_1\rangle) + p_2 P(a_k | |\psi_2\rangle) \\ &= p_1 |\langle a_k | \psi_1 \rangle|^2 + p_2 |\langle a_k | \psi_2 \rangle|^2 \end{aligned}$$

Here we can identify **two different sources of uncertainty**. The first one is purely quantum-mechanical. This uncertainty cannot be eliminated, even with perfect information about the state of the system. The second one is associated with our imperfect knowledge of the system, and it can be reduced with better information about its state. For example, if we were *certain* that the system is in state $|\psi_1\rangle$, then $p_1 = 1$ and $p_2 = 0$, so that $P(A = a_k) = P(A = a_k | \psi_1) = |\langle a_k | \psi_1 \rangle|^2$, just as in the case of a pure state.

All the previous discussion is generalized to the case of m different possible states, say $|\psi_1\rangle, \dots, |\psi_m\rangle$, with probabilities p_1, \dots, p_m . In this case, the probability of observing the value a_k when measuring A is

$$P(A = a_k) = \sum_{i=1}^m p_i |\langle a_k | \psi_i \rangle|^2.$$

One matrix to rule them all

How do we represent mixed states mathematically? Suppose we have a mixed state with possible outcomes $|\psi_1\rangle, \dots, |\psi_m\rangle$, and probabilities p_1, \dots, p_m . Recall that (basically) the only things that we actually measure and care about in quantum mechanics are the *expectation values* of observables. The expectation value is just the weighed average of the observable, so

$$\begin{aligned}
\langle A \rangle_{\text{mixed}} &= \sum_k a_k P(A = a_k) \\
&= \sum_k a_k \sum_i p_i |\langle a_k | \psi_i \rangle|^2 \\
&= \sum_k a_k \sum_i p_i \langle a_k | \psi_i \rangle \langle \psi_i | a_k \rangle \\
&= \sum_k a_k \langle a_k | \left(\sum_i p_i |\psi_i\rangle \langle \psi_i| \right) | a_k \rangle \\
&= \sum_k \langle a_k | \left(\sum_i p_i |\psi_i\rangle \langle \psi_i| \right) a_k | a_k \rangle \\
&= \sum_k \langle a_k | \left(\sum_i p_i |\psi_i\rangle \langle \psi_i| \right) \hat{A} | a_k \rangle \\
&= \text{Tr}(\hat{\rho} \hat{A}).
\end{aligned}$$

Here we have assumed that the eigenstates of \hat{A} form a complete orthonormal basis, and we have defined **the density matrix** $\hat{\rho}$ as

$$\hat{\rho} := \sum_i p_i |\psi_i\rangle \langle \psi_i|.$$

With this matrix we can calculate the probability of measuring the value a_k of the observable A in the mixed state as

$$P(A = a_k) = \sum_{i=1}^m p_i |\langle a_k | \psi_i \rangle|^2 = \langle a_k | \left(\sum_{i=1}^m p_i |\psi_i\rangle \langle \psi_i| \right) | a_k \rangle = \langle a_k | \hat{\rho} | a_k \rangle,$$

which is the k -th diagonal element of the matrix $\hat{\rho}$ when expressed in the basis of eigenstates of \hat{A} .

Note that $\text{Tr}(\hat{\rho}) = 1$, and $\hat{\rho}$ is an hermitian operator (i.e. $\hat{\rho}^\dagger = \hat{\rho}$, where † denotes the conjugate transpose), therefore it can be diagonalized. Furthermore, since for every vector $|\alpha\rangle$ we have that

$$\langle \alpha | \hat{\rho} | \alpha \rangle = \sum_i p_i |\langle \alpha | \psi_i \rangle|^2 \geq 0,$$

therefore $\hat{\rho}$ is positive semidefinite, so all of its eigenvalues (let's call them ρ_1, \dots, ρ_N) are non-negative. Some of them might be zero, some might be repeated, but all are real and non-negative. We have then, that:

$$\text{Tr}(\hat{\rho}) = \rho_1 + \dots + \rho_N = 1,$$

so it follows that $0 \leq \rho_k \leq 1$ for all $k = 1, \dots, N$. This, in turn, means that

$$\text{Tr}(\hat{\rho}^2) = \rho_1^2 + \dots + \rho_N^2 \leq 1.$$

We have, then, that $\text{Tr}(\hat{\rho}^2) = 1$ if and *only if* $\rho_k = 1$ for some k , and all the other $\rho_j = 0$ for $j \neq k$. What this means is that **we are certain that the system is in a particular state** $|\psi_k\rangle$, so, as in the first section, the system is in a **pure state**. In this case, 1 is an eigenvalue of $\hat{\rho}$, then it must leave its eigenstate $|\psi\rangle$ invariant:

$$\hat{\rho}|\psi\rangle = |\psi\rangle.$$

Now $\hat{\rho}$ cannot have any other (nonzero) eigenvalues, so it can be written in the form

$$\hat{\rho} = |\psi\rangle\langle\psi|.$$

This is the general form of the density matrix for a pure state. Recall that, in this case $\text{Tr}(\hat{\rho}^2) = 1$, but for any other case, the inequality is strict. This means that $\text{Tr}(\hat{\rho}^2)$ is, in some sense, **a measure of how “pure”** the state is.

What we have just seen is that **the density matrix can represent both mixed and pure states**, where pure states are of the form $\hat{\rho}_{\text{pure}} = |\psi\rangle\langle\psi|$, whereas mixed states are a **convex combination** of density matrices of pure states:

$$\hat{\rho}_{\text{mixed}} = \sum_{i=1}^m p_i \hat{\rho}_{i,\text{pure}},$$

with $p_1 + \dots + p_m = 1$.

An example

Consider the case of spin-1/2. Suppose that we are in a mixed state of $|+z\rangle$ with probability p and $|-z\rangle$ with probability $1 - p$. The density matrix for this mixed state is

$$\hat{\rho} = p|+z\rangle\langle+z| + (1 - p)|-z\rangle\langle-z|.$$

In this z -basis, the matrix takes the rather simple form

$$\hat{\rho} = \begin{pmatrix} p & 0 \\ 0 & 1 - p \end{pmatrix},$$

so that the square is simply

$$\hat{\rho}^2 = \begin{pmatrix} p^2 & 0 \\ 0 & 1 - 2p + p^2 \end{pmatrix}.$$

Then we have that

$$\text{Tr}(\hat{\rho}^2) = 2p^2 - 2p + 1.$$

We could say that $I(\hat{\rho}) = 1 - \text{Tr}(\hat{\rho}^2)$ is a measurement of how **impure** the state is: if this quantity is zero, then the state is pure. The maximum of $I(\hat{\rho})$ occurs when $p = 1/2$, which is precisely when there is maximum uncertainty about the state,

$$\hat{\rho} = \frac{1}{2}|+z\rangle\langle+z| + \frac{1}{2}|-z\rangle\langle-z|.$$

The state that this matrix represents is not, I repeat, **not** the same as the “superposed” state

$$|+x\rangle = \frac{1}{\sqrt{2}}(|+z\rangle + |-z\rangle).$$

Let’s analyze the similarities and the differences between the two.

The two states seem quite similar at first: They seem to be “equal parts” $|+z\rangle$ and $|-z\rangle$. Furthermore, and probably most importantly, the probability distribution of the z -spin observable \hat{S}_z is *the same* for both states. For a refresher, (assuming for god’s sake that $\hbar = 1$), the $|\pm z\rangle$ states are precisely the eigenstates of \hat{S}_z with eigenvalues $\pm \frac{1}{2}$:

$$\hat{S}_z|\pm z\rangle = \pm \frac{1}{2}|\pm z\rangle.$$

Therefore in our mixed state represented by $\hat{\rho}$,

$$P\left(S_z = \pm \frac{1}{2}\right)_{\text{mixed}} = \langle \pm z | \hat{\rho} | \pm z \rangle = \frac{1}{2}.$$

Similarly, for the “superposed state” $|+x\rangle$, we have:

$$P\left(S_z = \pm \frac{1}{2} \middle| +x\right) = |\langle \pm z | +x \rangle|^2 = \frac{1}{2}.$$

This means that **if we only were able to measure S_z** , then the two states would indeed be indistinguishable. However, *there are other observables* that distinguish the half-and-half mixed state represented by $\hat{\rho}$ and the “superposed” state $|+x\rangle$. A crystal-clear observable that can distinguish between the two is the x -spin observable, S_x . On one hand, for the mixed state $\hat{\rho}$,

$$P\left(S_x = \frac{1}{2}\right)_{\text{mixed}} = \langle +x | \hat{\rho} | +x \rangle = \frac{1}{2}.$$

However, for the superposed state:

$$P\left(S_x = \frac{1}{2} \middle| +x\right) = |\langle +x | +x \rangle|^2 = 1.$$

Furthermore, if we write the density matrix for the state $|+x\rangle$, i.e. $\hat{\rho}_{+x} = |+x\rangle\langle +x|$, it passes the purity test with flying colors:

$$\text{Tr}(\hat{\rho}_{+x}^2) = 1,$$

whereas, as we had already seen, for our mixed state $\text{Tr}(\hat{\rho}^2) = 1/2$.

The takeaway

The most important things to take away from all this are the following:

1. There is **always** uncertainty in the measurement of observables in quantum mechanics, even if you are **absolutely certain** that the system is in a specific state. This uncertainty is **not** a consequence of systematic or instrumental errors, or lack of information about the system. In this case of absolute certainty, we say that the system is in a **pure state**.
2. When we are not sure of what state the system is in, we represent our lack of knowledge by writing down a probability distribution on the set of probable states. Mathematically, we say that we know with probability p_i that the system is in a state $|\psi_i\rangle$, and say that the system is in a **mixed state**. In this case, there is an *additional* uncertainty in the measurement of observables that comes from our *lack of knowledge* of the precise state that the system is in.
3. Both pure and mixed states can be represented mathematically with a **density matrix**. This matrix has a lot of neat properties that make calculations of expectation values quite easy, even in the case of mixed states.

References

1. Most of this was inspired by a set of lectures that were given by dr. Stefan Vandoren during the Summer School in Theoretical Physics at Utrecht University in 2018.
2. Ballentine, Leslie E. 1990. *Quantum Mechanics*.