

Interpretation of the Wavefunction

In the final lecture of the course, I would like to return to the question: Is the wavefunction real, and does it really collapse? Philosophers have many answers to this question. In this lecture, for the sake of clarity, I will contrast three extreme positions:

1. The wavefunction is real and it never collapses.
2. The wavefunction collapses when a measurement is made. After the measurement, the wavefunction is in an eigenstate of the measured operator. The choice of the value of the measurement and the final state is probabilistic.
3. There is an underlying deterministic process that controls the outcome of all measurements. The rules of quantum mechanics determine the probabilities of the various outcomes, but the actual results follow with certainty from equations at a deeper level.

These options are called (1) the Everett or Many-Worlds interpretation, (2) the Copenhagen interpretation, and (3) the theory of Hidden Variables.

The original interpretation of the quantum-mechanical wavefunction was the Copenhagen interpretation, formulated by Heisenberg, Bohr, and, especially, Max Born. The rule that, for measurements of the operator \mathcal{O} in the state $|\psi\rangle$, the probability of the outcome λ_n is

$$P(n) = |\langle n | \psi \rangle|^2$$

is called the *Born rule*. This rule does an excellent job of describing the reproducing the results of experiments on quantum mechanical systems.

Einstein refused to believe that the laws of nature were fundamentally formulated in terms of probabilities. In a 1926 letter to Born, he wrote

Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but it does not really bring us closer to the secrets of the "old one". I, at any rate, am convinced that He does not throw dice.

In 1935, in a very important paper with Podolsky and Rosen (Phys. Rev. 47, 777 (1935)), Einstein launched a frontal attack on the Copenhagen interpretation. He called attention to the states that are *entangled* in the manner that I described in the previous lecture. Consider again the state

$$\frac{1}{\sqrt{2}} [|\downarrow\rangle_1 |\uparrow\rangle_2 + |\uparrow\rangle_1 |\downarrow\rangle_2]$$

If spin down is measured in the first state, spin up is always measured in the second state, and vice versa. The two measurements seem to affect each other. This is true even though the measurements can be made very far apart, at spacelike separation. In this paper, Einstein, Podolsky, and Rosen gave examples where a measurement on one side seemed to determine a measurement on the other side in violation of the Heisenberg uncertainty principle.

Bohm and Aharonov gave a particularly sharp argument version of the argument, written explicitly in terms of a two-level system (Phys. Rev. 108, 1070 (1957)). Consider a spin 0 particle that decays to two spin $\frac{1}{2}$ particles. Angular momentum is conserved, so we know that the two particles resulting from the decay must be in the state of total spin 0,

$$|0\rangle = \frac{1}{\sqrt{2}} [|\uparrow\rangle_z |\downarrow\rangle_z - |\downarrow\rangle_z |\uparrow\rangle_z]$$

This state is written in the basis of eigenstates of J^z . However, the state is rotationally invariant, so it takes the same form for any choice of axes. You can check this by plugging in the expression for the eigenstates of S^z in terms of eigenstates of S^x

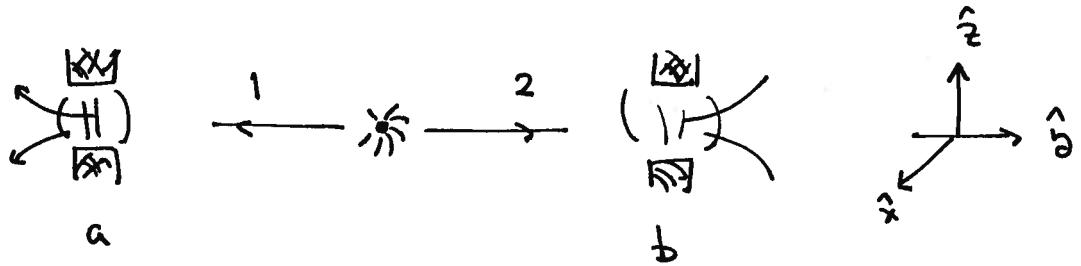
$$|\uparrow\rangle_z = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} [\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix}] = \frac{1}{\sqrt{2}} [|\uparrow\rangle_x + |\downarrow\rangle_x]$$

$$|\downarrow\rangle_z = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} [\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix}] = \frac{1}{\sqrt{2}} [|\uparrow\rangle_x - |\downarrow\rangle_x]$$

and verifying that

$$|0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle_x |\downarrow\rangle_x - |\downarrow\rangle_x |\uparrow\rangle_x)$$

Now consider the following thought experiment: The spin 0 particle decays, and the two spin $\frac{1}{2}$ particles, which I will call 1 and 2, fly off in opposite directions. The spins are measured by two Stern-Gerlach devices, a and b,



The magnets in the Stern-Gerlach devices need not be pointed in the same direction. Call the axes of the two devices \hat{a} and \hat{b} . Then the device a measures $\hat{a} \cdot \vec{S}_1$ and the device b measures $\hat{b} \cdot \vec{S}_2$.

If the two Stern-Gerlach devices are oriented in the same direction, there is perfect correlation between the two measurements. For example, if $\hat{a} = \hat{b} = \hat{z}$, the possible outcomes are

detector a: \uparrow	detector b: \downarrow	prob. 50%
detector a: \downarrow	detector b: \uparrow	prob. 50%

Similarly, if $\hat{a} = \hat{b} = \hat{x}$, the possible outcomes are

detector a: \uparrow	detector b: \downarrow	prob. 50%
detector a: \downarrow	detector b: \uparrow	prob. 50%

However, you can see from the transformation between the z and x eigenstates above that, if we keep $\hat{a} = \hat{z}$ but turn the b device into the direction $\hat{b} = \hat{x}$, there are four possible outcomes

detector a: ↑	detector b: ↓	prob 25%
detector a: ↑	detector b: ↑	prob 25%
detector a: ↓	detector b: ↑	prob 25%
detector a: ↓	detector b: ↓	prob 25%

all of which occur with equal probability.

That is, if we set the b device to $\hat{b} = \hat{z}$ and measure spin up, the outcome of spin up for a is forbidden, while if we set $\hat{b} = \hat{z}$, all outcomes for a occur with equal probability. What is stranger is that we can decide on the orientation of b when this detector is out of causal contact with the device a and even complete the measurement before a light signal has had a chance to communicate the orientation of b to the device a. Indeed, if the wavefunction actually collapses, the collapse of the wavefunction by the detector b must be *immediately* communicated to other detectors, at speeds faster than the speed of light. It would seem that the theory of relativity insists that this should be impossible.

Einstein, Podolsky, and Rosen argued that, if the result of a measurement is determined by previous physical phenomena, that result should correspond to a real, definite physical state. Using this logic, they argued that there must be hidden variables that determine the results of all quantum mechanical measurements.

Let us analyze this position in more detail. I would like to formalize the idea of hidden variables, adding the idea that the hidden variables act locally to determine the result of measurements. Call the quantities measured by the Stern-Gerlach devices

$$A = \frac{\hat{a} \cdot \vec{S}_1}{\hbar/2} \quad B = \frac{\hat{b} \cdot \vec{S}_1}{\hbar/2}$$

We observe that A and B are always measured to be $+1$ or -1 . A hidden variable theory must predict the correct value from the values of hidden variables λ . Then, there must be a function

$$A(\lambda, \hat{a}) = \pm 1$$

that predicts the outcome from the detector a and a similar function

$$B(\lambda, \hat{b}) = \pm 1$$

that predicts the outcome from the detector b. The assumption of locality is introduced by the statement that A cannot depend on \hat{b} but only on the hidden variables and the orientation \hat{a} of the detector a, and, similarly, the measurement outcome at b cannot depend directly on \hat{a} .

We do not know how many hidden variables there are or what values of these variables are produced in each decay of the original spin 0 particle. I would like to represent the spectrum of variables produced in a series of decays by a probability distribution

$$d\lambda \rho(\lambda)$$

which has the properties

$$\rho(\lambda) \geq 0 \qquad \int d\lambda \rho(\lambda) = 1$$

Then the expectation values of the observables A and B and their correlation are given by the expressions

$$\langle A \rangle = \int d\lambda \rho(\lambda) A(\lambda, \hat{a})$$

$$\langle B \rangle = \int d\lambda \rho(\lambda) B(\lambda, \hat{b})$$

$$\langle AB \rangle = \int d\lambda \rho(\lambda) A(\lambda, \hat{a}) B(\lambda, \hat{b})$$

The predictions of quantum mechanics for these expectation values are

$$\langle A \rangle = 0 \quad \langle B \rangle = 0 \quad \langle AB \rangle = -\hat{a} \cdot \hat{b}$$

These expectation values are in good agreement with experiment. (Rather, I should say that more realistic experiments that test these correlations are well-described by the quantum mechanical predictions.) A hidden variable theory should be able to reproduce these results.

The values of the separate expectation values of A and B are obvious, but nevertheless I would like to compute these answers from quantum mechanics in a formal way. The value of the correlation is less trivial, and we can obtain it by the same method.

Begin from

$$\langle A \rangle = \langle 0 | \frac{\hat{a} \cdot \vec{S}_1}{\hbar/2} | 0 \rangle = \langle 0 | \hat{a} \cdot \vec{\sigma}_1 | 0 \rangle$$

Putting in the explicit form of the state $|0\rangle$ using the basis of S^z eigenstates, this is

$$\begin{aligned} \langle A \rangle &= \frac{1}{\sqrt{2}} (\langle \uparrow \downarrow | - \langle \downarrow \uparrow |) \hat{a} \cdot \vec{\sigma}_1 \frac{1}{\sqrt{2}} (| \uparrow \downarrow \rangle - | \downarrow \uparrow \rangle) \\ &= \frac{1}{2} [\langle \uparrow | \hat{a} \cdot \vec{\sigma} | \uparrow \rangle \langle \downarrow | \downarrow \rangle - \langle \uparrow | \hat{a} \cdot \vec{\sigma} | \downarrow \rangle \langle \downarrow | \uparrow \rangle \\ &\quad - \langle \downarrow | \hat{a} \cdot \vec{\sigma} | \uparrow \rangle \langle \uparrow | \downarrow \rangle + \langle \downarrow | \hat{a} \cdot \vec{\sigma} | \downarrow \rangle \langle \uparrow | \uparrow \rangle] \end{aligned}$$

The simple overlaps are

$$\langle \downarrow | \downarrow \rangle = \langle \uparrow | \uparrow \rangle = 1 \quad \langle \downarrow | \uparrow \rangle = \langle \uparrow | \downarrow \rangle = 0$$

Using the values of the σ^i matrices

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

we see that the needed matrix elements of $\hat{a} \cdot \vec{\sigma}$ are

$$\langle \uparrow | \hat{a} \cdot \vec{\sigma} | \uparrow \rangle = + \hat{a}^z \quad \langle \downarrow | \hat{a} \cdot \vec{\sigma} | \downarrow \rangle = - \hat{a}^z$$

Only a^z appears because only σ^z has diagonal elements. Finally,

$$\langle A \rangle = \frac{1}{2} [\hat{a}^z - 0 - 0 + (-\hat{a}^z)] = 0$$

A parallel argument gives

$$\langle B \rangle = 0$$

We can use the same method to evaluate $\langle AB \rangle$,

$$\begin{aligned} \langle AB \rangle &= \frac{1}{2} (\langle \uparrow \downarrow | - \langle \downarrow \uparrow |) \hat{a} \cdot \vec{\sigma}_1 \hat{b} \cdot \vec{\sigma}_2 \frac{1}{2} (\langle \uparrow \downarrow | - \langle \downarrow \uparrow |) \\ &= \frac{1}{2} \langle \uparrow | \hat{a} \cdot \vec{\sigma} | \uparrow \rangle \langle \downarrow | \hat{b} \cdot \vec{\sigma} | \downarrow \rangle - \langle \uparrow | \hat{a} \cdot \vec{\sigma} | \downarrow \rangle \langle \downarrow | \hat{b} \cdot \vec{\sigma} | \uparrow \rangle \\ &\quad - \langle \downarrow | \hat{a} \cdot \vec{\sigma} | \uparrow \rangle \langle \uparrow | \hat{b} \cdot \vec{\sigma} | \downarrow \rangle + \langle \downarrow | \hat{a} \cdot \vec{\sigma} | \downarrow \rangle \langle \uparrow | \hat{b} \cdot \vec{\sigma} | \uparrow \rangle \end{aligned}$$

We need the additional, off-diagonal, matrix elements.

$$\langle \uparrow | \hat{a} \cdot \vec{\sigma} | \downarrow \rangle = \hat{a}^x - i \hat{a}^y \quad \langle \downarrow | \hat{a} \cdot \vec{\sigma} | \uparrow \rangle = \hat{a}^x + i \hat{a}^y$$

$$\begin{aligned} \langle AB \rangle &= \frac{1}{2} [\hat{a}^z (-\hat{b}^z) - (\hat{a}^x - i\hat{a}^y)(\hat{b}^x + i\hat{b}^y) \\ &\quad - (\hat{a}^x + i\hat{a}^y)(\hat{b}^x - i\hat{b}^y) + (-\hat{a}^z) \hat{b}^z] \\ &= -a^z b^z - a^x b^x - a^y b^y = -\hat{a} \cdot \hat{b} \end{aligned}$$

This justifies the result quoted above. The final result should be rotationally invariant, and it is.

Now we have the problem of finding a probability distribution $\rho(\lambda)$ for the hidden variables that reproduces the results predicted by quantum mechanics. In 1964, John Bell proved that this problem cannot be solved. Here is his proof:

First, consider the case $\hat{a} = \hat{b}$. In this case,

$$\langle AB \rangle = -1$$

Then

$$-1 = \int d\lambda \rho(\lambda) A(\lambda, \hat{a}) B(\lambda, \hat{a})$$

But $\rho(\lambda)$ is positive and integrates to 1, and

$$A(\lambda, \hat{a}), B(\lambda, \hat{a})$$

always equal either +1 or -1. For the integral to equal -1, it must always be true that

$$A(\lambda, \hat{a}) B(\lambda, \hat{a}) = -1$$

for all λ and all choices of \hat{a} . Then

$$B(\lambda, \hat{a}) = -A(\lambda, \hat{a})$$

and

$$\langle AB \rangle = - \int d\lambda \rho(\lambda) A(\lambda, \hat{a}) A(\lambda, \hat{b})$$

Let \hat{c} another unit vector, and let C be the quantity measured by a Stern-Gerlach device aligned along the axis \hat{c} . Then

$$\begin{aligned} \langle AB \rangle - \langle AC \rangle &= - \int d\lambda \rho(\lambda) A(\lambda, \hat{a}) [A(\lambda, \hat{b}) - A(\lambda, \hat{c})] \\ &= - \int d\lambda \rho(\lambda) A(\lambda, \hat{a}) A(\lambda, \hat{b}) [1 - A(\lambda, \hat{b}) A(\lambda, \hat{c})] \end{aligned}$$

But

$$A(\lambda, \hat{a}) A(\lambda, \hat{b}) = \pm 1$$

$$[1 - A(\lambda, \hat{b}) A(\lambda, \hat{c})] \geq 0$$

Then

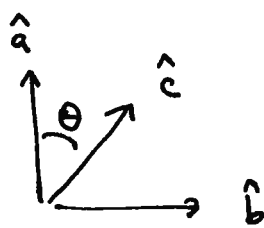
$$\begin{aligned} |\langle AB \rangle - \langle AC \rangle| &\leq \int d\lambda \rho(\lambda) [1 - A(\lambda, \hat{b}) A(\lambda, \hat{c})] \\ &= 1 + \langle BC \rangle \end{aligned}$$

We have proven that, for any local hidden variable theory,

$$1 + \langle BC \rangle \geq | \langle AB \rangle - \langle AC \rangle |$$

This relation is called *Bell's inequality*.

It is instructive to evaluate the two sides of the inequality for the prediction of quantum mechanics given above. For definiteness, consider the arrangement of vectors



$$\hat{a} \cdot \hat{b} = 0$$

$$\hat{a} \cdot \hat{c} = \cos \theta$$

$$\hat{b} \cdot \hat{c} = \sin \theta$$

Then

$$1 + \langle BC \rangle = 1 - \sin \theta$$

$$| \langle AB \rangle - \langle AC \rangle | = \cos \theta$$

For $\theta = \pi/4$, the inequality would read

$$1 - \sin \theta \geq \cos \theta$$

$$\theta = \pi/4 \quad 1 - \frac{1}{\sqrt{2}} \geq \frac{1}{\sqrt{2}} \quad 0.29 \geq 0.71$$

and it is manifestly not satisfied.

I have now presented arguments that would seem to annihilate two of the three positions that I defined at the beginning of the lecture. I have shown, at least, that both the Copenhagen interpretation and the Hidden Variables theory require weird nonlocal interactions to be consistent with the expectations of quantum mechanics. However, we should not be so quick to embrace the third possibility. We have already

seen that, if we take the predictions of the Schrödinger equation absolutely literally, we must make sense of wavefunctions such as the Schrödinger state

$$\frac{1}{\sqrt{2}} \left[| \text{cat alive} \rangle + | \text{cat dead} \rangle \right]$$

How do we make sense of this?

In 1957, Hugh Everett argued that it is sensible to consider states of this type that are linear superpositions of macroscopically distinct outcomes. What we see in nature is that macroscopic outcomes are final. They do not go back into a ghostly mixture with other incompatible outcomes. In Schrödinger's problem, we could build a counterexample to this statement if we could form an interference between the two terms in the wavefunction. A way to do this might be to evolve the two terms to

$$\begin{aligned} | \text{cat alive} \rangle &\rightarrow | \text{cat dead} + 1 \text{ photon} \rangle \\ | \text{cat dead} \rangle &\rightarrow | \text{cat dead} + 1 \text{ photon} \rangle \end{aligned}$$

so that the dead cats are in the same quantum state, and then to arrange for an interference pattern of the photons. Unfortunately, when we refer to a dead cat, we do not refer to a single quantum state but rather to a large ensemble of states which, macroscopically, have almost identical properties. The probability that we end up in the same states for both processes is infinitesimal, of the order of e^{-N_A} , where N_A is Avogadro's number.

Almost certainly, then, the evolution of a Schrödinger state by further integration of the Schrödinger equation simply evolves each half of the wavefunction completely independently of the other half. In the more recent literature, Roland Omnès, Robert Griffiths, and Murray Gell-Mann and James Hartle have attempted to write explicit conditions under which pieces of the quantum mechanical wavefunction can no longer interfere and instead evolve completely independently. These pieces are called *consistent* or *decoherent histories*.

It would be comforting to say that, after a quantum mechanical measurement, the quantum state is in one of these decoherent pieces. However, the literal prediction of quantum mechanics is that the wave function is a linear combination of all of the possibilities. That is why the Everett interpretation is called *Many Worlds*. Imagine what happens to the Schrödinger state when Albert opens the box and observes

the cat, alive or dead. Albert's brain chemistry is affected by the outcome of the observation. We find

$$\frac{1}{\sqrt{2}} \left[| \text{cat alive, Albert happy} \rangle + | \text{cat dead, Albert depressed} \rangle \right]$$

Maybe you become depressed if Albert is depressed, so you quit Stanford and go to live in a cabin in the Yukon. Maybe you are elated when Albert told you that the cat survived, so you sign up for that voyage to Tahiti. Then the further evolution the Schrödinger equation gives

$$\frac{1}{\sqrt{2}} \left[| \text{cat alive, Albert happy, Tahiti} \rangle + | \text{cat dead, Albert depressed, Cabin} \rangle \right]$$

In Everett's interpretation, every time that a quantum process has macroscopic consequences, the wavefunction splits into pieces, with one piece realizing every possible macroscopic outcome. These pieces then develop according to their natural evolution. All around us, quantum processes are coupling to macroscopic systems. So the total wavefunction of the universe is constantly splitting, and different versions of ourselves are being carried down all of the possible pathways that are the outcomes of these interactions.

There is no completely satisfactory conclusion. That is why there is no agreement on the "correct" interpretation of the Schrödinger wavefunction almost a century after its discovery.

Fortunately, all of the interpretations that are based on quantum mechanics, including the simple and mechanical Copenhagen interpretation, give the same predictions for experiment. So, you can adopt any one and continue to study quantum mechanics. We have seen that the observable outcomes of quantum mechanics are themselves surprising and fascinating. You will see more examples of this as the course continues.