QI Lecture 11

Bell inequalities

It is worryingly easy to believe ideas that are simultaneously both obviously true and objectively false. For example, it's obvious that the Earth is flat, but evidently it's not.[1] It's obviously true that everything experiences time at the same rate, but in fact they don't.[2]

One of the absolutely-so-obvious-it-shouldn't-even-get-a-name assumptions of physics until the 20th century is "**realism**" or "counter-factual definiteness". Realism is the belief that things have a specific state, independent of your knowledge of that state. If you flip a coin and cover it, the result is already there, you just don't happen to know what it is. Applied to particles, realism says that at any given time every particle is in some definite state at some definite place and that the strange behavior we witness must be due to some properties that perhaps we can't take into account or simply don't know about. These unknown or unknowable properties are "**hidden variables**". Models premised on hidden variables and the impossibility of instantaneous interactions over distance are called "**local hidden variable theories**".

Bell Inequalities

For a single measurement it turns out that there's no way to determine if the phenomena you're looking at is displaying classical randomness due to hidden variables or fundamental randomness from something being in a superposition of states. It may be that there are things you're not taking into account that's dictating the results, and if so, then a diagonally polarized photon going through a vertical polarizer is random in the same way that a hidden coin is random.

Definite but unknowable states are described using probability distributions. For example, a coin (with sides ±1) is described by $p(x)$ where $p(1) = p(-1) = \frac{1}{2}$. More complex probability distributions are described by probabilities with multiple inputs, $p(a, b, \ldots)$. For example, for two dice $p(2, 4) = \frac{1}{36}$.

You can figure out the expectation values of the arguments or even functions of the arguments, $f(a, b, \ldots)$, by summing over the probability distribution:

[1] See Eratosthenes, 240 BC.
[2] See Einstein, 1905 AD.

$$E[f] = \sum_{a,b,\ldots} f(a,b,\ldots)\, p(a,b,\ldots)$$

A Bell inequality is a statement about regular probabilities that holds regardless of what the underlying probability distribution is. If something is in a definite but unknown state, then we can describe it using a probability distribution and therefore its statistics must obey all (applicable) Bell inequalities.

However, measurements of quantum phenomena involving entanglement can violate Bell inequalities![3] This means that these phenomena cannot be described using probabilities alone, regardless of whether there are hidden variables.

We take this with a wink and a nod, because we already know that quantum mechanics is rooted in superpositions and probability amplitudes as opposed to classical definite states ad probabilities.

Figure 1: When things are in a definite (but unknown) state we can use probabilities. So if quantum phenomena are always in a definite state, we should be able to describe it using probabilities, even if we don't know what the underlying "hidden variables" are.

Suppose that Alice and Bob have a steady supply of $|\Phi_+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$ states and they each make a measurement at some angle. Given the state $|\Phi_+\rangle$, the probability that Alice will see $|\theta\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle$ and Bob will see $|\phi\rangle = \cos(\phi)|0\rangle + \sin(\phi)|1\rangle$ is:[4]

[3] I say "can" because there's presently no known method for determining if a given Bell inequality will be violated by some quantum phenomena.

[4] Here we use the trig identities $\cos(x) = \cos(-x)$, $\sin(x) = -\sin(-x)$, and $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$.

$$
\begin{aligned}
p &= \left|\langle\theta|_a\langle\phi|_b\left(\frac{|0\rangle_a|0\rangle_b+|1\rangle_a|1\rangle_b}{\sqrt{2}}\right)\right|^2 \\
&= \frac{1}{2}\left|\langle\theta|0\rangle\langle\phi|0\rangle+\langle\theta|1\rangle\langle\phi|1\rangle\right|^2 \\
&= \frac{1}{2}\left|\cos(\theta)\cos(\phi)+\sin(\theta)\sin(\phi)\right|^2 \\
&= \frac{1}{2}\left|\cos(\theta)\cos(-\phi)-\sin(\theta)\sin(-\phi)\right|^2 \\
&= \frac{1}{2}\left|\cos(\theta-\phi)\right|^2 \\
&= \frac{1}{2}\cos^2(\theta-\phi)
\end{aligned}
$$

Now assume that Alice and Bob use the observables

$$
A = |\theta\rangle\langle\theta| - |\theta\perp\rangle\langle\theta\perp| \qquad\qquad B = |\phi\rangle\langle\phi| - |\phi\perp\rangle\langle\phi\perp|
$$

which means that $|\theta\rangle_a$ and $|\phi\rangle_b$ correspond to "1" while $|\theta\perp\rangle_a = |\theta+\frac{\pi}{2}\rangle_a$ and $|\phi\perp\rangle_b = |\phi+\frac{\pi}{2}\rangle_b$ correspond to "-1".

The joint observable is

$$
\begin{aligned}
A\otimes B &= \left[|\theta\rangle_a\langle\theta|_a - |\theta\perp\rangle_a\langle\theta\perp|_a\right]\left[|\phi\rangle_b\langle\phi|_b - |\phi\perp\rangle_b\langle\phi\perp|_b\right] \\
&= \left[|\theta\rangle_a|\phi\rangle_b\langle\theta|_a\langle\phi|_b + |\theta\perp\rangle_a|\phi\perp\rangle_b\langle\theta\perp|_a\langle\phi\perp|_b\right] - \left[|\theta\rangle_a|\phi\perp\rangle_b\langle\theta|_a\langle\phi\perp|_b + |\theta\perp\rangle_a|\phi\rangle_b\langle\theta\perp|_a\langle\phi|_b\right]
\end{aligned}
$$

which means that for the joint observable "1" corresponds to Alice and Bob getting the same result ("1,1" or "-1,-1") and "-1" corresponds to them getting opposite results. So the probability of Alice and Bob getting the same result is $p(same) = \frac{\cos^2(\theta-\phi)}{2} + \frac{\cos^2\left(\theta+\frac{\pi}{2}-\phi-\frac{\pi}{2}\right)}{2}$ and therefore

$$
\boxed{p(same) = \cos^2(\theta-\phi) \qquad\qquad p(different) = \sin^2(\theta-\phi)}
$$

This is a whole lot of math and notation[5] to say "If Alice and Bob are measuring photons in the state $|\Phi_+\rangle$ and they set up their polarizers at angles θ and ϕ, the probability that they'll get the same result is $\cos^2(\theta-\phi)$."

This tidy formula turns out to be incompatible with any probability distribution. We'll see why twice, first with a thought experiment and then an actual experiment.

Polarizers and Time Machines

Assume that, like a covered coin, the result of a measurement is set in stone whether you make the measurement or not. This is the realism assumption. Clearly, if you have access to a time machine, then you can loop back and do every measurement you like; if you do

[5] And good practice!

the same measurement on the same object, then you'll always get the same result, but you have the added benefit of being able to do every other possible measurement.

First we'll derive Bell inequality that holds true for any probability distribution. Assume that a, b, and c can only take one of two values. It follows that

$$\begin{aligned}
& P(a=c) \\
= \; & P(a=b \cap b=c) + P(a \neq b \cap b \neq c) \\
\geq \; & P(a=b \cap b=c) \\
= \; & P(a=b) + P(b=c) - P(a=b \cup b=c) \\
\geq \; & P(a=b) + P(b=c) - 1
\end{aligned}$$

So for a, b, and c only taking one of two values, regardless of the probability distribution, we can always say that:

$$P(a=c) \geq P(a=b) + P(b=c) - 1$$

If realism holds, then whether or not we do a measurement, the result of our polarizer experiments "exist" (like a hidden coin). We can only measure a photon with a polarizer once before changing its state, but with a time machine we can go back and do as many measurements as we like; every time the initial conditions are the same, so the result of any given measurement should be the same.

Define $P(\theta, \phi)$ as the probability of doing a measurement at angles θ and ϕ and getting the same result (both photons go through or are stopped).

Applying the inequality several times:

$$\begin{aligned}
& P\left(0, \tfrac{\pi}{2}\right) \\
\geq \; & P\left(0, \tfrac{3\pi}{8}\right) + P\left(\tfrac{3\pi}{8}, \tfrac{\pi}{2}\right) - 1 \\
\geq \; & \left[P\left(0, \tfrac{\pi}{4}\right) + P\left(\tfrac{\pi}{4}, \tfrac{3\pi}{8}\right) - 1\right] + P\left(\tfrac{3\pi}{8}, \tfrac{\pi}{2}\right) - 1 \\
\geq \; & \left[\left[P\left(0, \tfrac{\pi}{8}\right) + P\left(\tfrac{\pi}{8}, \tfrac{\pi}{4}\right) - 1\right] + P\left(\tfrac{\pi}{4}, \tfrac{3\pi}{8}\right) - 1\right] + P\left(\tfrac{3\pi}{8}, \tfrac{\pi}{2}\right) - 1 \\
= \; & P\left(0, \tfrac{\pi}{8}\right) + P\left(\tfrac{\pi}{8}, \tfrac{\pi}{4}\right) + P\left(\tfrac{\pi}{4}, \tfrac{3\pi}{8}\right) + P\left(\tfrac{3\pi}{8}, \tfrac{\pi}{2}\right) - 3 \\
= \; & 4\cos^2\left(\tfrac{\pi}{8}\right) - 3 \\
= \; & \sqrt{2} - 1 \\
\approx \; & 0.414
\end{aligned}$$

By imagining that we can do many measurements by leaping back in time and adjusting one polarizer at a time, we find that we reach a contradiction. Looking at the correlations, and assuming that it makes sense to measure at each polarization from 0 to $\frac{\pi}{2}$ spaced $\frac{\pi}{8}$

apart, we find that $P\left(0, \frac{\pi}{2}\right) \geq 0.414$. But at the same time, it is an empirical fact that $P\left(0, \frac{\pi}{2}\right) = \cos^2\left(\frac{\pi}{2}\right) = 0$.

Clearly this is a contradiction. Evidently it doesn't make sense to even imagine that the photons in $|\Phi_+\rangle$ have a particular, pre-determined result for every given angle.

CHSH

Lacking time machines, we need an experiment that can actually be done. In order to use the CHSH inequality Alice and Bob can each chose between one of two observables

$$A = |0\rangle\langle 0| - |\frac{\pi}{2}\rangle\langle\frac{\pi}{2}| \qquad A' = |\frac{\pi}{4}\rangle\langle\frac{\pi}{4}| - |\frac{3\pi}{4}\rangle\langle\frac{3\pi}{4}|$$

$$B = |\frac{\pi}{8}\rangle\langle\frac{\pi}{8}| - |\frac{5\pi}{8}\rangle\langle\frac{5\pi}{8}| \qquad B' = |\frac{3\pi}{8}\rangle\langle\frac{3\pi}{8}| - |\frac{7\pi}{8}\rangle\langle\frac{7\pi}{8}|$$

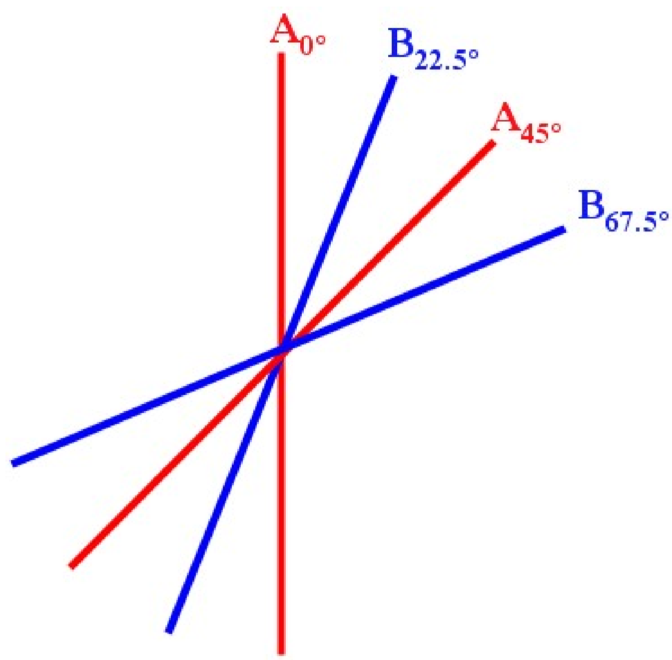

Figure 2: The "1" orientations for each of Alice and Bob's measurements.

where the results of each of these observables are $a, a', b, b' = \pm 1$. In practice Alice and Bob can each only make one measurement, but if realism holds, then all four variables have a value even if they aren't measured.[6]

[6]Like a coin that isn't so much hidden as lost.

Now consider this function:

$$\boxed{S = ab + a'b + a'b' - ab' = (a' + a)\,b + (a' - a)\,b'}$$

Notice that one of $(a' + a)$ and $(a' - a)$ is always zero and the other is always ± 2, and since $b, b' = \pm 1$ we have that

$$|S| = \left|ab + a'b + a'b' - ab'\right| = 2$$

and therefore, regardless of the probability distribution over all of the observables $p\,(a, a', b, b')$,

$$E[S] = \sum_{a,a',b,b'=-1}^{1} \left[ab + a'b + a'b' - ab'\right] p\left(a, a', b, b'\right) \leq \sum_{a,a',b,b'=-1}^{1} 2p\left(a, a', b, b'\right) = 2$$

This is the CHSH[7] inequality:

$$\boxed{E\,[ab + a'b + a'b' - ab'] \leq 2}$$

This is a Bell inequality, which means that it holds true for absolutely any probability distribution. Any two pairs of two-result experiments where the probabilities and uncertainty come solely from a lack of knowledge must adhere to this equation.

Although we can't actually measure every one of these variables for every experiment, we can do the experiment in each of four different ways and, by measuring the expectation value for each alignment of the polarizers, we can build up the overall expectation value:

$$E\left[ab + a'b + a'b' - ab'\right] = E\,[ab] + E\left[a'b\right] + E\left[a'b'\right] - E\left[ab'\right]$$

$ab = 1$ when the results are the same and $ab = -1$ when they're different, so these expectation values are the probability of the results being the same minus their probability of being different:

$$p(same) - p(different) = \cos^2(\theta - \phi) - \sin^2(\theta - \phi) = \cos\left(2(\theta - \phi)\right)$$

Recalling the angular difference between each pair of measurements:

$$\begin{array}{rcccc}
E\,[ab] & = & \cos\left(2\left(0 - \frac{\pi}{8}\right)\right) & = & \frac{1}{\sqrt{2}} \\
E\,[a'b] & = & \cos\left(2\left(\frac{\pi}{4} - \frac{\pi}{8}\right)\right) & = & \frac{1}{\sqrt{2}} \\
E\,[a'b'] & = & \cos\left(2\left(\frac{\pi}{4} - \frac{3\pi}{8}\right)\right) & = & \frac{1}{\sqrt{2}} \\
E\,[ab'] & = & \cos\left(2\left(0 - \frac{3\pi}{8}\right)\right) & = & -\frac{1}{\sqrt{2}}
\end{array}$$

[7] John Clauser, Michael Horne, Abner Shimony, and Richard Holt

Each of these expectation values are objective measurements that can be made, in a lab, using many copies of the $|\Phi_+\rangle$ state. Combining them we find that

$$E\left[ab + a'b + a'b' - ab'\right] = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - \left(-\frac{1}{\sqrt{2}}\right) = 2\sqrt{2}$$

So, if entangled particles are in a specific state, no matter how complicated, and the results of every experiment exists even if they aren't measured, then they can be described by a probability distribution and therefore $E[S] \leq 2$. However, we find experimentally that $E[S] = 2\sqrt{2} \approx 2.828$.

"**Bell Tests**" like this are a standard method for demonstrating the "quantumness" (typically taken to be superposition) of a given system.

Loopholes

Realism is a very comforting, intuitive idea and nobody really wants to give it up. So there have been a lot of attempts to poke holes in Bell tests.

Communication Loophole

If you and a friend had to create the results of a Bell test by hand, there's nothing stopping you from doing it. You'd say to each other "Hey, my polarizer is at angle θ, how about you?" and then you could compare notes to make sure you faked the right correlations.[8]

The communication loophole says that maybe, somehow, the results of the experiment are due to parts of the experiment communicating with each other and colluding to create the illusion of superposition, when in fact the system is always in a single, definite state. So when one photon passes through a polarizer it quickly lets the other one know how to adjust its probability distribution.

No intervening materials have any impact on Bell tests, so to ensure that no signal passes from one measurement to the other, we ensure that they're randomized after the entangled particles are generated and so that the randomization and measurement at one location is outside of the light cone of the randomization and measurement at the other. In other words, we ensure that even at the speed of light, nothing about one measurement can make it to the other. We use the speed of light as a barrier to prevent collusion.

However, it may still be possible that somehow some kind of information is sneaking around before the experiment is even set up and that this affects the correlations. That is, the randomized polarizer orientations aren't really random, they're dictated (in a remarkably specific way) by some hidden variable that's already present. After all, the whole

[8] Sort of the "walk the skee ball up the ramp and drop it in" of quantum physics.

experiment is in the forward light cone of everything on Earth from a little under a tenth of a second ago.

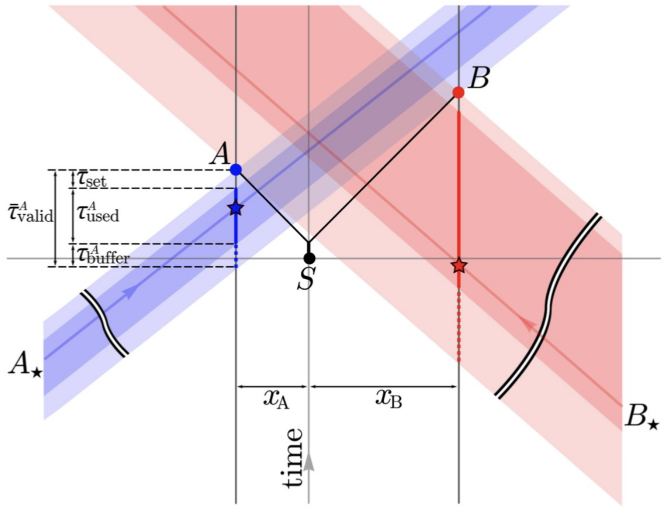

Figure 3: The spacetime diagram for a Bell test that uses starlight to randomize the polarizers. S generates a pair of entangled photons and sends them to A and B where they're measured. In this set up, the light from each star arrives in time to be used by one polarizer, but too late for the other to know about it.

In order to make sure that there isn't some kind of "synchronizing signal" that's somehow affecting every method we use to randomize the polarizers, we need sources of randomness that can't be affected by a common cause in the past. A good way to do this is to point telescopes at specific stars a couple hundred light years apart, in opposite directions, and use their starlight as a random number generator. That way, in order for something to correlate the randomness of both experiments, it would have needed to start planning it nefarious and extremely precise scheme hundreds of years ago.

Fair Sampling Assumption

When we physically carry out a Bell test, we don't always detect both of the entangled particles, and that means that we can't look at the correlations between them. In fact, for photons a detection efficiency of 30% or less is normal.[9]

The fair sampling assumption is that whether we detect or miss a particle has no impact on how it behaves, or would have behaved, when measured. In other words, the particles

[9] Since photons are cheap, fast, and aren't too bothered by air, glass, or the vacuum of space, it would be nice if we could catch them consistently. Research into single-photon detectors is intense and ongoing. That 30% is a hard-won victory that someone is proud of.

we detect are a representative sample of all of the particles. Under the fair sampling assumption, the CHSH inequality stays the same:

$$E\left[ab + a'b + a'b' - ab'\right] \leq 2$$

But if we abandon the fair sampling assumption, we have to worry about how non-detections might be included in the probability distribution. In particular, we have to worry about the possibility that (for whatever reason) unusually correlated results are easier to detect. If the efficiency (the probability of detection) is η, then the "revised" CHSH inequality is:

$$E\left[ab + a'b + a'b' - ab'\right] \leq \frac{4}{\eta} - 2$$

In other words, as η decreases, the requirement for the Bell test becomes more strict and for $\eta < 0.83$, $2\sqrt{2}$ is no longer high enough. Fortunately, tests on several other systems such as trapped ions, nitrogen-vacancy flaws in diamonds, and super conducting qubits[10] have all surpassed this efficiency threshold and passed the Bell test.

[10] All possible quantum computer architectures!

Exercises

#1) Best of Many Worlds.

Alice and Bob are being sent an endless stream of photons in the state $|\Phi_+\rangle = \frac{|0\rangle_a|0\rangle_b+|1\rangle_a|1\rangle_b}{\sqrt{2}}$. Alice and Bob each make two sets of measurements on their photons.

Measurement:	Basis:	Labeling of results:	
A	$\{\|0\rangle, \|1\rangle\}$	$\|0\rangle \to a = 1$	$\|1\rangle \to a = -1$
A'	$\{\|\theta\rangle, \|\theta + \frac{\pi}{2}\rangle\}$	$\|\theta\rangle \to a' = 1$	$\|\theta + \frac{\pi}{2}\rangle \to a' = -1$
B	$\{\|\phi\rangle, \|\phi + \frac{\pi}{2}\rangle\}$	$\|\phi\rangle \to b = 1$	$\|\phi + \frac{\pi}{2}\rangle \to b = -1$
B'	$\{\|\mu\rangle, \|\mu + \frac{\pi}{2}\rangle\}$	$\|\mu\rangle \to b' = 1$	$\|\mu + \frac{\pi}{2}\rangle \to b' = -1$

Find the set of angles $\theta, \phi, \mu \in [0, \pi)$ that maximizes the experimental violation of the CHSH inequality, $E[ab + a'b + a'b' - ab'] \leq 2$.

#2) Wigner's Inequality

Wigner's inequality applies to a situation with three pairs of experiments with one of two results, $a_1, b_1, a_2, b_2, a_3, b_3 \in \{-1, 1\}$, in which Alice and Bob always have perfectly anti-correlated results whenever they make the same measurement. So if $a_j = 1$, then $b_j = -1$. However, different measurements can take any value.

Define $p(a_1, b_1, a_2, b_2, a_3, b_3)$ to be the probability distribution over all possible results and (for example) define

$$p(a_1, b_2) = \sum_{a_2, b_1, a_3, b_3 = -1}^{1} p(a_1, b_1, a_2, b_2, a_3, b_3)$$

as the marginal distribution over a_1 and b_2. Wigner's inequality states that

$$p(a_1 = 1, b_2 = 1) + p(a_2 = 1, b_3 = 1) \geq p(a_1 = 1, b_3 = 1)$$

a) First show that $p(a_j, b_k)$ for $j \neq k$ can be written as the sum of only two non-zero probabilities from the full distribution $p(a_1, b_1, a_2, b_2, a_3, b_3)$.

b) Use the result from part a to prove Wigner's inequality.

c) Assume that Alice and Bob each have one qubit from the entangled state

$$|\Psi_-\rangle = \frac{|0\rangle_a|1\rangle_b - |1\rangle_a|0\rangle_b}{\sqrt{2}}$$

Alice's measurement operators and their eigenvalues and eigenstates are

$$
\begin{array}{ll}
A_1 = \frac{1}{2} Z_a + \frac{\sqrt{3}}{2} X_a & \begin{cases} a_1 = 1 & \frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle \\ a_1 = -1 & \frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle \end{cases} \\
A_2 = Z_a & \begin{cases} a_2 = 1 & |0\rangle \\ a_2 = -1 & |1\rangle \end{cases} \\
A_3 = \frac{1}{2} Z_a - \frac{\sqrt{3}}{2} X_a & \begin{cases} a_3 = 1 & \frac{\sqrt{3}}{2}|0\rangle - \frac{1}{2}|1\rangle \\ a_3 = -1 & \frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle \end{cases}
\end{array}
$$

and Bob's are the same (just with "B's" instead of "A's").

Show that the probabilities of the results of these measurements applied to $|\Psi_-\rangle$ do not obey Wigner's inequality.