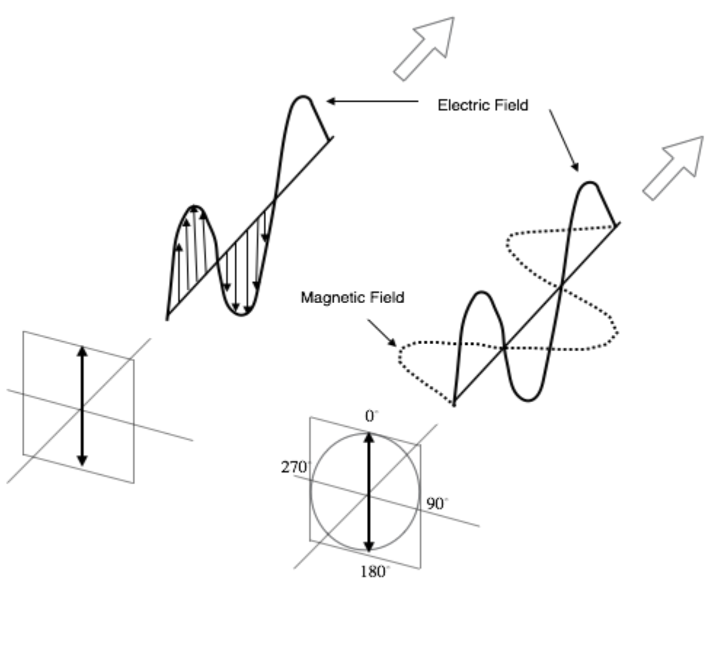
Bell's Inequality 1

Quantum physics is an incredibly successful theory. Its predictions about the universe have yet to be disproven by experiment. However, there is much debate about why it works so well. This is largely due to the nonintuitive nature of its assertions, examples of which will be discussed below. Indeed, because of this nonintuitiveness, scientists over the years, including Albert Einstein, have questioned the soundness of the theory, suggesting that the theory is incomplete or should be abandoned all together. This was especially true in the early stages of the theory's development. A debate over the theory's validity raged for years but remained philosophical because no one could figure out how to do an experiment to put the theory to the test—until John Bell. In 1966, Bell published a paper (to be discussed in the second installment on this subject) that proposed a way to test the tenets of quantum physics[1]. Central to this paper is a mathematical equation which has become known as Bell's Inequality. It is this inequality and the concepts laid out in that paper that are the subject of this article.

Now Bell's Inequality is intimately involved with a phenomenon called quantum entanglement. Therefore, in order to understand Bell's Inequality, we need to examine quantum entanglement first.

Indeed, entanglement is one of the most fascinating and enigmatic phenomena in quantum mechanics and in all of science. What is quantum entanglement? Well, to quote Wikipedia,

> quantum entanglement is a physical phenomenon which occurs when pairs or groups of particles are generated, interact, or share spatial proximity in ways such that the quantum state of each particle cannot be described independently of the state of the other(s), even when the particles are separated by a large distance—instead, a quantum state must be described for the system as a whole.[2]



Consider a photon, a particle of light. It's electromagnetic energy and consists of an electrical field and a magnetic field. The strength (or amplitude) of the electric field wavers back and forth regularly (that is, oscillates) in the x-z plane. The amplitude of the magnetic field oscillates in the y-z plane and the photon moves in the z direction. The electric and magnetic fields always oscillate in directions perpendicular to each other and the direction of motion of the photon is always perpendicular to the direction of oscillation of the electric and magnetic fields. According to the conventions shown in the diagram, if the electric field oscillates in the x-z plane, we say that the photon's plane of polarization is at 0 degrees. Or, another way of saying it is that the photon is polarized in the zero-degree direction. Now suppose we rotate the plane of polarization clockwise such that the new plane of polarization makes a 45-degree angle with the x-z plane. The angle of polarization of the light is now said to be 45 degrees. Rotate the polarization plane 90 degrees and the angle of polarization is 90 degrees; Rotate it 123 degrees and the angle of polarization is 123 degrees, and so forth.

There are devices called polarization filters that function as follows: they will let a photon through 100% of the time if it is polarized at the angle at which the device is set; it will block the photon 100% of the time if it is set at an angle 90 degrees different from the angle at which the photon is polarized; and it will let the photon through some but not all of the time, if the angle of polarization differs from the filter's setting by some angle other than 90 degrees, the probability of it getting through being a function of the angle of difference. Individual photons will either get through or not get through but if you send in enough photons, then the percentage that get through will be the same as the probability of an individual photon getting through (or close to it). And one more thing: once a photon passes through a filter, it assumes the polarization angle at which the filter was set. That is, a photon that's polarized at 45° before it reaches a 90° filter will emerge from the filter polarized at 90°, if it passes through. In the vernacular of standard interpretation of quantum mechanics, the interaction with the filter (which essentially constitutes a measurement, if we care to look) causes collapse of the photon's wave function to 90°.

Now photons can be split into what are called entangled pairs. The details of how this is done are not important. What's important is how the entangled pairs behave. Say you create the pair in a lab in Chicago and send each member of the pair off in opposite directions, one to LA and the other to New York. You put a filter in the path of each photon in both the LA and New York laboratories, and just behind each filter, you put a detector. If the photon gets through the filter, it will register on the detector. Set the filters in both laboratories to 0°. Check to see if the photon got through in each lab. Do this over and over again for many photons. If all of the photons get through and register on the detector in one lab, then all will get through and register on the detector in the other lab. If no photons get through in the lab on the east coast, then they won't get through in the lab on the west coast either. Now set the filters to another angle but make sure it's the same in both labs and check the detectors. The same thing happens; if all of the photon gets through in one lab, they'll also all get through in the other, and vice versa.

This may suggest to you that, when the photons become entangled, they communicate to each other some sort of program about how they're going to behave when measured in a certain way. If you are thinking that, then you are in good company. Albert Einstein was probably the greatest proponent of this viewpoint.

This viewpoint has come to be known as **local realism—realism** meaning that particles have definite values for all of their properties at all times (in this case polarization); and **local** meaning that particles can only influence each other by direct communication (i.e., transmission of a force). Thus, to explain entanglement, Einstein believed that either entangled particles 1) have definite states programmed into them, during their interaction, through direct (local) communication of forces, by unknown means (i.e., through so-called **local hidden variables**) or 2) they communicate with each other by some sort of signal at the time of measurement. Now velocity = change in distance / change in time, and the speed of light equals 3×10^{10} cm/second. That means that if the distance in centimeters separating the particles at the time of measurement is greater than or equal to 3×10^{10} times the time interval between measurements in seconds, then, for one particle to send a signal to its entangled partner (e.g. to tell it what polarization angle to assume), that signal would have to travel faster than the speed of light. Since nothing has ever been demonstrated to travel faster than the speed of light, Einstein felt that the only explanation for the entanglement phenomenon was alternative 1—local realism orchestrated by local hidden variables.

Opposing Einstein's point of view, and instead, supporting quantum theory was Neils Bohr. In a famous paper commonly referred to as EPR[3] (after its authors Einstein, Podolsky and Rosen), Einstein's group defined reality as follows:

> If, without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.

Bohr[4] contested this definition. While this definition was consistent with common sense (and classical physics), he felt that it was arbitrary; that there was nothing to say that quantum mechanics, with its probabilities and indeterminacy, didn't actually represent a new reality that should replace classical physics—just as relativity had previously replaced the classical view of space and time. To Bohr, the correlated behavior seen in entanglement is just an intrinsic property of reality. That's just the way it is. No need to look for underlying mechanisms (or hidden variables).

As mentioned above, the dispute between the two schools of thought went on for decades without making any headway until Bell devised a way to settle the issue. But Bell's original paper is not easy to digest, especially for the non-professional audience at whom this article is directed. Therefore, it will be better to ease into the subject. The best introduction to this subject that I have seen can be found at a website by David R. Schneider, http://drchinese.com/Bells_Theorem.htm

The link involving easy math is the one to follow for this discussion. It can be found here:

http://drchinese.com/David/Bell_Theorem_Easy_Math.htm

The discussion below is based on the information on this website. It refers to the following table and is reproduced with permission.

Case	A=0°	B=120°	C=240°	AB	BC	AC	Sum AB+BC+AC	Avg (AB+BC+AC)/3
1	A+	B+	C+	1(++)	1(++)	1(++)	3	1
2	A+	B-	C+	0	0	1(++)	1	.333
3	A+	B+	C-	1 (++)	0	0	1	.333
4	A+	B-	C-	0	1(++)	0	1	.333
5	A-	B+	C+	0	1(++)	0	1	.333
6	A-	B-	C+	1(++)	0	0	1	.333
7	A-	B+	C-	0	0	1(++)	1	.333
8	A-	B-	C-	1 (- -)	1 (- -)	1 (- -)	3	1

We'll work with entangled photons and make the assumption that there is some program or some set of instructions that tells the photons how they should be polarized when measured at any given angle. Furthermore, we'll take the stance that this state is a definite state that exists even in the absence of us measuring it. For simplicity, let's choose just three angles: angle A = 0°, B = 120° and C = 240°. You could, of course, examine each degree of a circle, or every half degree or for every possible measurement. The argument we're going to make would still be valid.

We want to set the filter in the lab in New York to one of the three angles and set the filter in the lab in LA to another different one of the three. We shoot one of a pair of entangled photons to each lab and see if the photon gets through the filter and reaches the detector in each lab. If there's some program, some hidden set of properties that a photon possesses, then it definitely—with 100% certainly—will or will not pass through a polarizing filter at a given setting. For three angles, a photon can posses only one of eight possible programs, as listed in the table. For example, A+B+C+ means the photon will pass through filters set at any of our three angles; A-B+C+ means it will not pass through a filter set at 'A' but will pass through filters set at 'B' and 'C', and so on. We can only measure one angle at a time and we randomly choose at which of the three angles we're going to set the filter in each lab. There are three possibilities: 1) set the filter in one lab to A = 0° while setting the filter in the other lab to B = 120° (listed as AB in the table) 2) set the filter in one lab to B = 120° and the filter in the other lab to C = 240° (listed as BC in the table) or 3) set the filter in one lab to A = 0° and the filter in the other lab to C = 240° (listed as AC in the table).

Now we send a few million pairs of entangled photons to each lab with the above three combinations of filter settings, collect data then use a computer program to analyze it. The program works as follows: it checks each experiment and sets up a table with two columns and a row for each experiment. In the left hand column, it puts which two filters were used. It doesn't matter which lab uses which filter (e.g. it doesn't matter whether you use the A filter in New York and B filter in LA or the A filter in LA and the B filter in New York; it puts AB in the left-hand column). If the results in the labs in both New York and LA match (i.e., photons in both labs either both reach the detector or both don't reach the detector), then it puts a 1 in the right-hand column. If, on the other hand, the results differ in the labs (e.g. the photon reaches the detector in New York but not in LA), it puts a 0 in the right-hand column. Then it sorts the data by detector combination, and for each detector combination, figures out what percentage of the experiments result in a match for each detector combo. If there is some program which specifies definitive, hidden patterns of polarization properties for the photons, then the probability of a match should be 1/3 or 33.3%.

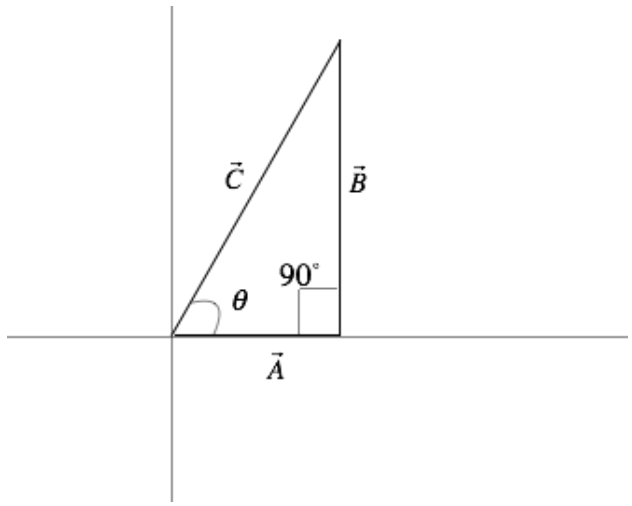
You can see this from the table. For example, suppose the entangled photons are working under program 6, or case 6 in the table, A-B-C+. That means that it will not pass through and register on the detector if filters A or B are used but will if filter C is used. Thus, if the AB filter combination is used (e.g., filter A is used in New York and B is used in LA), neither the photon in New York or the photon in LA will get through. Thus, the results will match and the computer will register a 1. However, if the BC or AC combos are used, there will be no match and the computer will register a 0. Now we said that the detector combos (AB, BC or AC) are chosen randomly. That means that if you repeat the experiment enough times, each filter combination will be chosen an equal number of times. Since every time the AB filter combination is chosen there will be a match, and every time the BC or AC filter combinations are used, there will not be a match, on average, 1 out of every 3 experiments will result in a match. And if you look at the table, this is true for cases 2 through 7, as well. For case 1 (A+B+C+) and case 8 (A-B-C-), the photons are either polarized at all three angles or not polarized at any of the three angles. So there'll be 100% agreement no matter at what angle you measure them.

It's clear from all this, that for any combination of filters used, there'll be at least 33.3% agreement, if local realism is true.

Next, we need to see what quantum mechanics predicts will be the percentage of agreement for the various combinations of filters.

According to quantum mechanics, the probability of agreement in measurement of entangled photons in LA and New York depends on the difference between the angle of measurement in the two places. Specifically, it varies according to square of a function called a cosine. To understand why this is so, we'll have to discuss some trigonometry—sine and cosine—then talk about vectors, then see how these things can be used to calculate probabilities in quantum mechanics.

First up is some trigonometry.



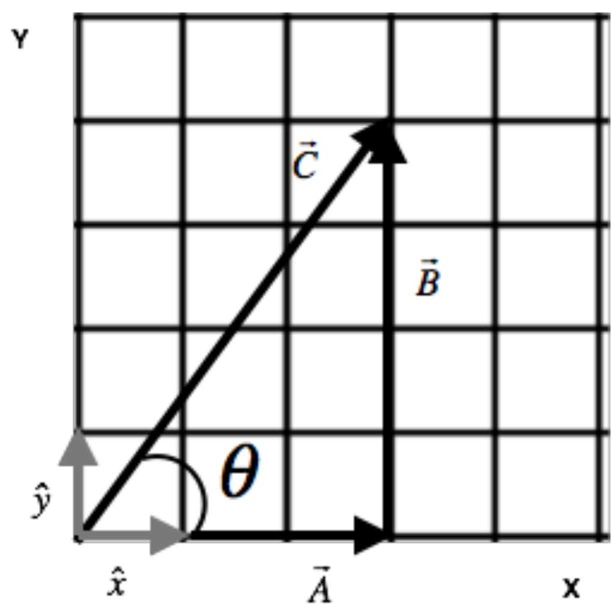
Let's start with a right triangle, like the one in the diagram. A right triangle is defined as one that has a 90° angle.

If you have such a right triangle, with one of its angles other than the 90° angle being θ, we can define two useful functions—actually, several—but there are really only two that we need to talk about now: sine and cosine. Cosine equals the length of the side of the triangle that touches (or is, as they say, adjacent to) the angle θ divided by the length of the hypotenuse (which is the side opposite the right angle). In this case: $\cos\theta = \frac{\text{Adjacent}}{\text{Hypotenuse}} = \frac{\|\vec{A}\|}{\|\vec{C}\|}$ where the little bar over A and C indicates that they are vectors (see below) and $\|\vec{A}\|$ and $\|\vec{C}\|$ are the lengths of vectors A and C, respectively. Similarly, sine equals the length of the side of the triangle opposite the angle divided by the length of the hypotenuse. In this case: $\sin\theta = \frac{\text{Opposite}}{\text{Hypotenuse}} = \frac{\|\vec{B}\|}{\|\vec{C}\|}$.

Another important relationship associated with right triangles is called the Pythagorean theorem. The Pythagorean theorem states that that the square of the length of the hypotenuse equals the sum of the square of the lengths of sides adjacent to the right angle. In the diagram, $\|\vec{A}\|^2 + \|\vec{B}\|^2 = \|\vec{C}\|^2$.

Next, a few words about vectors. Vectors, in a general sense, are just an ordered collection of elements. Usually, the elements give the length along an axis that specifies a specific dimension. The easiest way to visualize this is to consider vectors in our everyday physical space. In this space, the dimensions would be something like forward, sideways and upward. We could use the x-axis to represent position in the sideways direction, the y-axis to represent the forward direction and the z-axis to represent the upward direction. In our familiar three dimensional physical world, the following facts are pretty much true: 1) the axes are straight lines 2) the axes are perpendicular to each other 3) the distance between each unit on each axis is the same and 4) the standard Euclidian geometry that's taught in high school applies. A coordinate system such as this is referred to as a Cartesian coordinate system. To make things simple, we'll just consider the x and y axes.

"In this simple setup, we can consider a vector to be a mathematical entity that has size (or magnitude) and direction. Here is an example:



As suggested by the diagram, you can add vectors graphically. To add vectors A and B by this method, you lay out vector A and place vector B with it's origin at vector A's end. You then connect the origin of A with the end of B. The line that forms this connection is the vector sum of A and B.

Vectors can also be broken down into components. Each vector can be thought of as a linear combination of unit basis vectors. A unit basis vector is a vector 1 unit long in the direction of one of the axes and is usually written with a hat over it. In this case, the basis vector in the x-direction is $\hat{x}$ and the basis vector in the y-direction is $\hat{y}$. Referring to the diagram, we can express $\vec{A}, \vec{B},$ and $\vec{C}$ as follows:

$$\vec{A} = a_x\hat{x} + a_y\hat{y}$$
$$\vec{B} = b_x\hat{x} + b_y\hat{y}$$
$$\vec{C} = c_x\hat{x} + c_y\hat{y}$$

Where a_x, a_y, b_x, b_y, c_x, and c_y are called coefficients, numbers by which you multiply the basis vectors to tell how long the component vectors are that make up the vector.

In the diagram, $a_x = 3$, $a_y = 0$; $b_x = 0$, $b_y = 4$; $c_x = 3$, $c_y = 4$. Frequently, vectors are represented by putting their components within parentheses or brackets. For example, in the diagram, $\vec{C}$ is represented as (3 4). Furthermore, vectors can be expressed as column vectors or row vectors. The technical differences aren't important for our purposes here. For our purposes, we'll just see what they look like. $\vec{C}$ in row form is $\begin{pmatrix} 3 & 4 \end{pmatrix}$. In column form, it's $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

What we've been talking about so far are vectors in space, vectors where the units on the coordinate systems used are things like centimeters, meters or miles. However, we could represent anything on those axes. In quantum mechanics, the thing represented on the axes is called probability amplitude of a certain property of a quantum particle.

In the conventional interpretation of quantum mechanics, a particle or quantum is not in a definite state until it's measured. For an electron, for example, there may be a 30% chance that it's at position **a**, 20% chance that it's at position **b**, 5% chance that it's at position **c**, and so on. The only constraint is that the probabilities have to add up to 100%. And you get those probabilities from squaring entities called probability amplitudes that make up a thing called the wave function. Position is a continuous variable. Therefore, to get the probability function, you have to square the value of the wave function (i.e., the probability amplitude) at every position.

The case of polarization of a photon is simpler. We choose two angles of polarization that are orthogonal to (i.e., at right angles to) each other (called a basis). These are the angles that you're going to measure. The probability amplitudes for those angles define the state of polarization of a photon. This can be represented by a vector called a state vector. If you square a probability amplitude, you wind up with the probability that the photon will pass through a filter oriented at the angle associated with that probability amplitude.

For example, take a photon polarized at 45°. We're going to measure it in the 0°-90° basis. Since we're measuring in the 0°-90° basis, the state vector must contain probability amplitudes that will yield correct predictions about how often the photon will pass through a filter set at 0° and how often it will pass through a filter set to 90°. The state vector that does this is $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$. So the probability of the photon passing through the 0° filter is $\left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$ and the probability of the photon passing through the 90° filter is $\left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$. As expected, the total probability is 1:

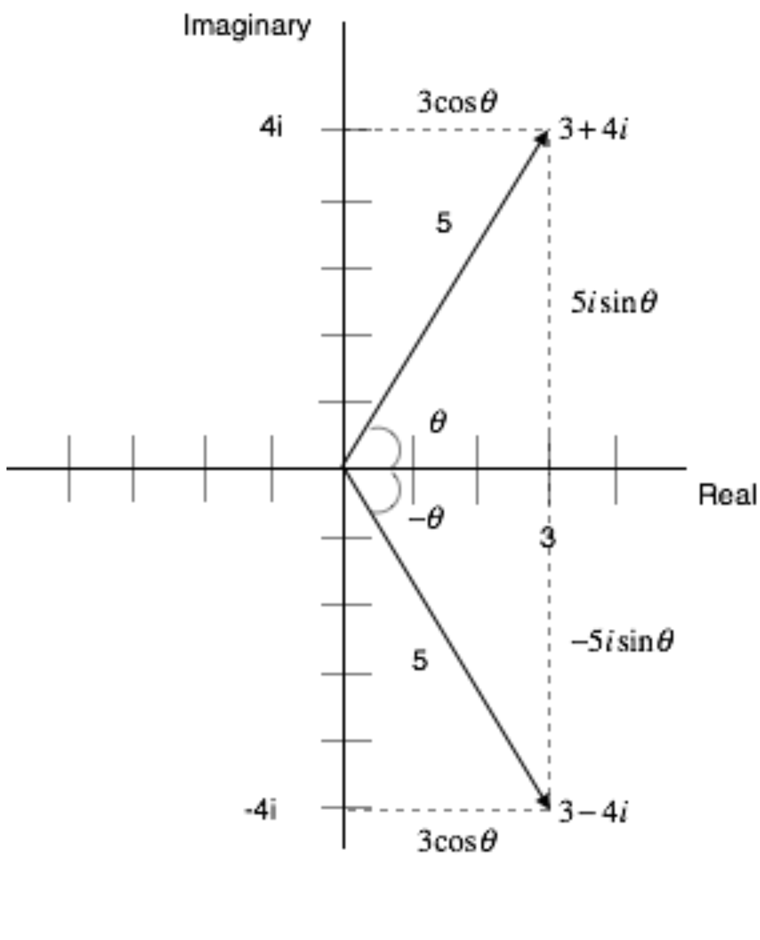
$$\frac{1}{2} + \frac{1}{2} = 1$$

or, in terms of percentages:

$$\frac{1}{2}\cdot 100\% + \frac{1}{2}\cdot 100\% = \left(\frac{1}{2}+\frac{1}{2}\right)\cdot 100\% = 1\cdot 100\% = 100\%$$
$$50\% + 50\% = 100\%$$

As an aside, in quantum mechanics, the coefficients of the vector states are actually complex numbers. That is, they consist of a real and an imaginary component. The imaginary component is just a real number multiplied by $\sqrt{-1}$ which is represented by the letter 'i'.

Complex numbers can be graphed in what's referred to as the complex plane where we plot the imaginary component on the y-axis and the real component on the x-axis.



We can think of each point in the complex plane as being represented by a vector with real and imaginary components. Each complex number has what's called a complex conjugate which consists of the same real component but the opposite complex component. For example, the complex conjugate of the complex number $a+bi$ is $a-bi$. By definition, squaring a complex number means multiplying it by its complex conjugate. Notice that if you do this, you get a real number:

$$\begin{aligned}(3+4i)(3-4i) &= 9-12i+12i-16i^2 \\ &= 9+0-16\left(\sqrt{-1}\right)^2 \\ &= 9-16(-1) \\ &= 9+16=25\end{aligned}$$

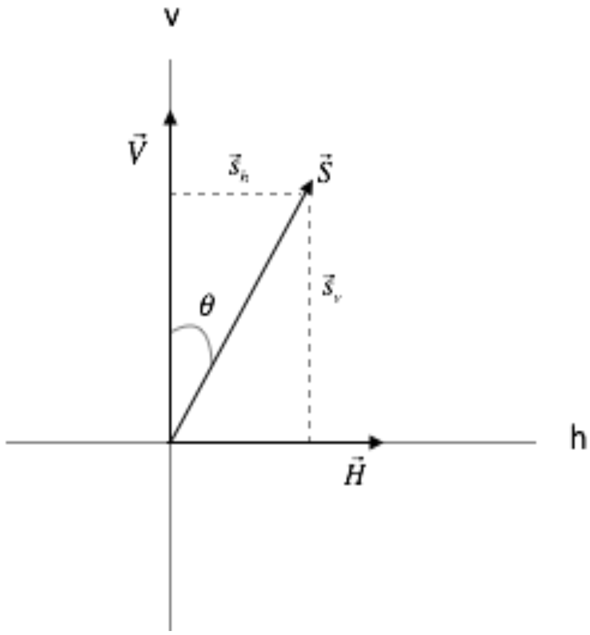
or more generally,

$$(a+bi)(a-bi)=a^2+b^2$$

Obtaining a real number under these circumstances is fortunate because, as mentioned, probabilities in quantum mechanics are the square of probability amplitudes which, in turn, are complex numbers. It would not be a good thing if the number we got for a probability were imaginary because no one knows what an imaginary probability means. But here's the good news: in the case we're considering, the imaginary parts of the coefficients are zero so all we have to deal with is real numbers.

Now moving back toward our goal of calculating quantum probabilities. In quantum mechanics, a quantum state can be written as a linear combination of basis vectors—which are unit vectors pointing in the direction of each of the axes (properties) we're considering. Each basis vector is multiplied by a probability amplitude which reflects the chances of that property being present. To simplify the math, basis vectors are usually chosen so that they are each one unit long and are all orthogonal to each other (a so-called orthonormal basis).

In the case of polarization, consider a photon polarized at some angle, θ, about to be measured in the 0°-90° basis. Its state can be described by a vector, $\vec{S}$, in terms of 0° (vertical) and 90° (horizontal) orthonormal basis vectors $\vec{V}$ and $\vec{H}$, respectively.



The equation for this vector is

$\vec{S} = \vec{s}_v + \vec{s}_h = \alpha_v \vec{V} + \alpha_h \vec{H}$ where

α_v and α_h are probability amplitudes (which are complex numbers)

$\alpha_v^* \alpha_v =$ the probability of $\vec{S}$ being polarized in the vertical direction

$\alpha_h^* \alpha_h =$ the probability of $\vec{S}$ being polarized in the horizontal direction

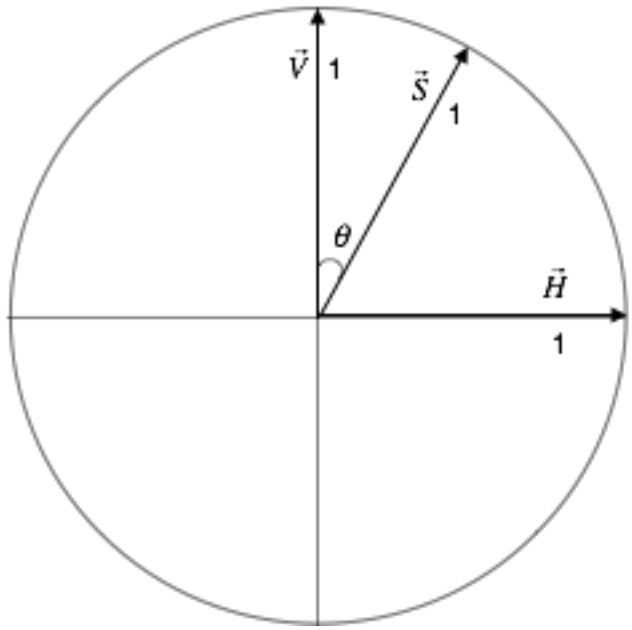
$\alpha_v^* \alpha_v + \alpha_h^* \alpha_h = 1$ (i.e. the probabilities, in decimal form, have to add to 1)

The asterisk (*) indicates the complex conjugate of the entity that it's placed after.

$\alpha_v = \left\|s_v\right\| = \left\|\vec{S}\right\| \cos\theta$ which means that

$\alpha_v * \alpha_v = \left\|\vec{S}\right\|^2 \cos^2\theta$

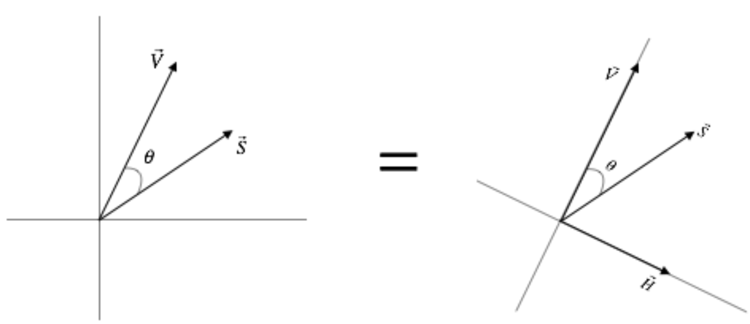
Say the endpoints of the vectors V, H and S all lie on a unit circle (that is, a circle with a radius 1 unit in length).



That means that vectors V, H and S are all 1 unit long. Therefore,

$$\alpha_v * \alpha_v = \left|\left|\vec{S}\right|\right|^2 \cos^2\theta = 1^2\cos^2\theta = \cos^2\theta$$

That means that the probability of the photon passing through a vertical filter is given by $\cos^2\theta$. That's a general result. The thing that determines the probability of the photon passing is the angle between the polarization of the photon and the filter setting. And that filter can be set at any angle. Think about it. If we rotated both vectors 26° clockwise maintaining a separation between the vectors of θ°, we'll get the same result. That's because if you also rotate the axis system clockwise, we'll be figuring out the same problem.



Now let's get back to the original table and original problem with which we started. Recall that in our experimental setup, we're dealing with entangled photons. Frequently, that means that the two photons have polarizations that are orthogonal (that is, their polarization angles differ by 90°). However, because of the manner in which we've prepared them, in our case, they'll both have the *same* polarization. Remember also that once a photon is measured at a particular angle, it becomes polarized at that angle. So if one of a pair of entangled photons passes through a filter in LA set at A = 0°, then it becomes vertically polarized.

Next let's consider the probability that its entangled counterpart will pass through a filter set at B=120° in New York. Because it is entangled with the photon in LA, it is polarized at the same angle, 0°. The difference, θ, between the polarization angle of the photon (0°) and the angle at which it is being measured (120°) is 120°. Therefore, the probability that it will pass through the filter is given by $\cos^2\theta = \cos^2 120^\circ$. You can look up $\cos 120^\circ$. It's -0.5 or -1/2. Thus, $\cos^2 120^\circ = (-0.5)^2$ = 0.25. Expressed as a percentage, that's 25%, or as a fraction, 1/4. So the probability that both photons will pass through filters in LA and New York (and thus, that their measurements will agree) is 0.25.

Now what if the photon in LA does not pass through the A = 0° filter. That means that it's polarized at 90°. The next question, then is what is the probability that its entangled partner will not pass through the B = 120° filter in New York. Well, it's 1 minus (the probability that it will pass through the filter.) To figure that out, we've got to determine the probability that the photon will pass through the 120° filter. Fortunately, we have a formula for that. The difference between the angle of photon polarization and the angle at which it's being measured is 90° - 120° = -30°. The cosine of -30° is $\frac{\sqrt{3}}{2} \approx 0.866$ and $\cos^2(-30)^\circ \approx (0.866)^2 = 0.75$. So the probability that it will pass through the 120° filter is 0.75. Therefore, the probability that it won't is 1 - 0.75 = 0.25. Ergo, the probability that both of a pair of entangled photons in LA and New York will not pass through A = 0° and B = 120° filters (i.e., that measurements will agree), again, is 0.25.

Note that the difference in angle between BC and AC is also 120°. Thus, the probability that both photons of an entangled pair will either pass through or not pass through both filters when the filter configurations are BC or AC (i.e., the probability that the measurements will agree) is also 0.25 (or 25%).

The bottom line, then, is that—according to quantum mechanics—the percentage of events in which both photons from an entangled pair will pass through filters, widely separated in space, whose angle settings differ by 120°, is 25%. This is in contrast to 33 and 1/3%, the percentage predicted by a so-called local realism/local hidden variables model in which photons are conceived as being in a definite state prior to measurement.

So *which is it?* you might be asking. *What percentage of entangled photon polarization measurements agree with each other if polarization filters used for measurement differ by 120° or 240°?* And by extension, *which theory—Einstein's local realism or Bohr's quantum mechanics—does experimental evidence agree with?*

Well, it turns out that, to my knowledge, the experiment described in the above example has never actually been done. However, experiments have been done that test different formulations of Bell's inequality. *What are the results?* These are described in the third part of this article.

Before we get there, however, the second portion of this article will describe Bell's original paper on the subject.

References

1. http://www.drchinese.com/David/Bell_Compact.pdf

2. https://en.wikipedia.org/wiki/Quantum_entanglement

3. https://journals.aps.org/pr/pdf/10.1103/PhysRev.47.777

4. http://cds.cern.ch/record/1060284/files/PhysRev.48.696.pdf