October 2020

# Lecture Notes for 8.225 / STS.042, "Physics in the 20th Century": Bell's Inequality and Quantum Entanglement

David Kaiser

*Center for Theoretical Physics, MIT*

## Introduction

These notes discuss the famous inequality derived by physicist John S. Bell in 1964, regarding correlations of the outcomes of measurements conducted on two or more particles. The first section introduces Bell's inequality and the concept of "local realism." The next section illustrates how systems that are prepared in a quantum-entangled state are predicted to *violate* Bell's inequality, revealing stronger correlations than local realism would allow. The appendix provides a derivation of Bell's inequality.

Reading these notes is *optional*; the notes are meant to fill in some of the gaps in various derivations that we will not cover during our class session.

## Bell's Inequality and "Local Realism"

In 1964, theoretical physicist John S. Bell published a short article in the first volume of an out-of-the-way physics journal, *Physics Physique Fizika*. The journal folded just a few years later, but Bell's article went on to become one of the most influential physics papers of the twentieth century.[1] In his article, Bell scrutinized the famous critque of quantum mechanics by Albert Einstein, Boris Podolsky, and Nathan Rosen from 1935, which is usually referred

[1] John S. Bell, "On the Einstein-Podolsky-Rosen paradox," *Physics Physique Fizika* **1** (1964): 195-200. Many of Bell's related articles and essays on the topic are available in Bell, *Speakable and Unspeakable in Quantum Mechanics* (New York: Cambridge University Press, 1990). On the history of Bell's inequality, see esp. Olival Freire, Jr., "Philosophy enters the optics laboratory: Bell's theorem and its first experimental tests (1965–1982)," *Studies in History and Philosophy of Modern Physics* **37** (2006): 577-616; Louisa Gilder, *The Age of Entanglement: When Quantum Physics was Reborn* (New York: Knopf, 2008); and David Kaiser, *How the Hippies Saved Physics: Science, Counterculture, and the Quantum Revival* (New York: W. W. Norton, 2011). Recent, accessible discussions include the graphic novel by Tanya Bub and Jeffrey Bub, *Totally Random: Why Nobody Understands Quantum Mechanics* (Princeton: Princeton University Press, 2018), and George Greenstein, *Quantum Strangeness: Wrestling with Bell's Theorem and the Ultimate Nature of Reality* (Cambridge: MIT Press, 2019).

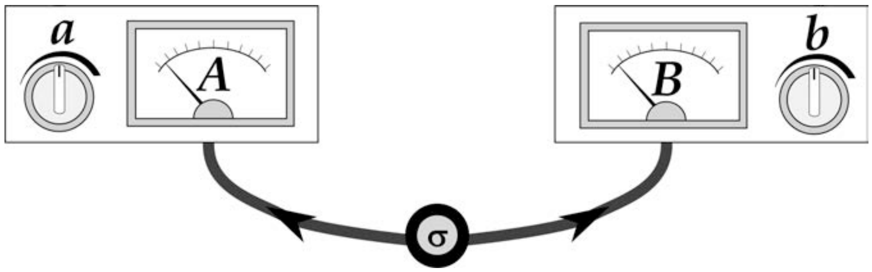


Figure 1: Schematic illustration of the original EPR thought experiment. A source $\sigma$ emits a pair of particles that travel in opposite directions. At each detector, a physicist selects a measurement to be performed by adjusting the detector settings $(a, b)$; each detector then yields a measurement outcome $(A, B)$.

to by the authors' initials: EPR.[2]

In their article, the EPR authors described a thought experiment like the one shown in Fig. 1. A source $\sigma$ emits pairs of particles that travel in opposite directions. The particle traveling to the left encounters a device at which a physicist could select some property of the particle to measure, by adjusting the detector setting $a$. Once the type of measurement to perform has been selected and the particle has arrived at the detector, the detector completes a measurement of the selected property and yields a measurement outcome $A$. The particle traveling to the right encounters an identical device, at which a second physicist could select various types of measurements to perform by adjusting the detector setting $b$. Her device then yields the measurement outcome $B$.

In the context of this thought experiment, the EPR authors articulated two criteria that they argued any reasonable physical theory should meet:

- "Reality criterion":"If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of reality corresponding to this physical quantity."

- "Locality": "Since at the time of measurement the two systems [that is, the pair of particles emitted from source $\sigma$] no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system."[3]

The combination of these criteria is often referred to as "*local realism*": that physical systems, including elementary particles, possess definite values for various properties on their own, prior to and independent of our choice of what to measure, and that no influence can travel across space arbitrarily quickly.

[2] Albert Einstein, Boris Podolsky, and Nathan Rosen, "Can quantum mechanical description of reality be considered complete?," *Physical Review* **47** (1935): 777-780.

[3] Einstein, Podolsky, and Rosen, "Can quantum mechanical description," on pp. 777, 779.

The EPR authors argued that since quantum mechanics is incompatible with local realism, quantum mechanics could not be a complete or final theory of nature. In his 1964 analysis, Bell agreed that quantum mechanics is incompatible with local realism; but he also emphasized that local realism itself was an *assumption*, not (as yet) a demonstrated fact. No matter how reasonable local realism may appear — judged either by common experience or by its easy fit with other well-established physical theories, including Einstein's own relativity — the question remained whether *nature* behaved in accordance with local realism.

Bell proposed a new type of experiment that could test whether the behavior of physical systems at the atomic scale was consistent with local realism. His first step was to suggest that the two detectors in Fig. 1 perform measurements whose outcomes could only ever take one of two values (so-called "dichotomic observables"), such as the *spin* of an electron or the *polarization* of a photon. For example, if one were measuring an electron's spin along $\hat{\mathbf{z}}$, the only possible measurement outcomes would be spin up or spin down along $\hat{\mathbf{z}}$. If one measured the electron's spin in units of $\hbar/2$, then these two outcomes would correspond to $+1$ (spin up) and $-1$ (spin down). Similarly for photon polarization: once one oriented the polarizing filter along some direction, a photon could only ever be measured with polarization parallel ($+1$) or perpendicular ($-1$) to the chosen orientation. For measurements of dichotomic observables, the measurement outcomes can always be represented as simply $A = \pm 1$ and $B = \pm 1$.

Next, Bell continued, the physicists at each detector could adjust the detector settings at their device to select among various measurements to perform. The physicist at the left detector could choose to measure an electron's spin along $\hat{\mathbf{z}}$ or along some direction inclined by an angle $\theta$ from $\hat{\mathbf{z}}$, and similarly for the physicist at the right detector. In other words, the physicist on the left could choose to perform a measurement either with detector setting $a$ or $a'$; and the physicist on the right could choose among detector settings $b$ or $b'$.

Because the outcomes of measurements at each detector could only ever be $A = \pm 1$ and $B = \pm 1$ — no matter what the selection of detector settings $(a, b)$ had been on a given experimental run — Bell suggested that physicists consider the *correlation function*:

$$E(a, b) = \langle A(a)\, B(b) \rangle \,, \tag{1}$$

where the angular brackets denote averaging over many experimental runs in which the particles were subjected to measurements with the pair of detector settings $(a, b)$. Given $A(a) = \pm 1$ and $B(b) = \pm 1$, the product $A(a)\, B(b)$ can only ever be $\pm 1$ on any given experimental run. Upon averaging over many runs in which the detector settings were $(a, b)$, the correlation function would therefore satisfy $-1 \leq E(a, b) \leq 1$. One may then consider a *combination* of these correlation functions, as one varies the detector settings at each side:

$$S \equiv E(a,b) + E(a',b) - E(a,b') + E(a',b') \,. \tag{2}$$

This form of the parameter $S$ is often called the Bell-CHSH parameter: it was derived by physicists John Clauser, Michael Horne, Abner Shimony, and Richard Holt soon after Bell's original derivation.[4]

The correlation functions $E(a,b)$ (and hence the Bell-CHSH parameter $S$) can be measured experimentally by conducting measurements on many pairs of particles, recording the measurement outcomes $A$ and $B$ for a given pair of detector settings $(a,b)$, and repeating the exercise for the other combinations of detector settings, such as $(a,b')$, $(a',b)$, and so on. One can also make *predictions* for the values of $E(a,b)$ (and hence $S$) using various physical theories, and compare those predictions with experimental observations. In general, the theoretical prediction for the value of the correlation function $E(a,b)$ can be written as

$$E(a,b) = \sum_{A,B=\pm 1} A\, B\, p(A,B|a,b) \,, \tag{3}$$

where $p(A,B|a,b)$ is an expression for the probability that the measurement outcomes on a given experimental run will be $(A,B)$ when the detector settings are selected to be $(a,b)$. For example, one may use a particular theoretical model to calculate $p(+1,+1|a,b')$, which is the probability to find $A=+1$ and $B=+1$ when the detector settings are set to $a$ and $b'$.

The question then comes down to how various physical theories make predictions for the correlation functions $E(a,b)$. Bell and the CHSH authors demonstrated that *any* physical theory that satisfies local realism would make predictions for $E(a,b)$, and hence for $S$, in which $S$ was subject to an *inequality*:

$$|S| \le 2 \,. \tag{4}$$

I review Bell's derivation of this inequality in the Appendix. The main take-away message is that quantum mechanics is *not* compatible with local realism, and it predicts that measurements on certain systems could be *more strongly correlated* than local-realist theories would allow. That is, quantum mechanics predicts that physicists should measure *violations* of Bell's inequality, finding $|S| > 2$ under certain conditions, whereas any local-realist theory should obey Eq. (4).

[4]John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt, "Proposed experiment to test local hidden variable theories," *Physical Review Letters* **23** (1969): 880-884. Their derivation followed closely along Bell's original argument in Bell, "On the Einstein-Podolsky-Rosen paradox."

# Quantum Entanglement and Bell's Inequality

## Quantum-Mechanical Description of Polarized Photons

For definiteness, let us consider the case in which the source $\sigma$ emits pairs of photons, and the physicists at each detector choose to measure their photon's linear polarization. Once a physicist fixes the orientation of the polarizing filter at her detector station (along some direction in space described by a unit vector $\hat{\mathbf{n}}$), she can register either horizontally polarized photons (in state $|H\rangle$) or vertically polarized photons (in state $|V\rangle$). A horizontally polarized photon registers as measurement outcome $+1$ (analogous to spin up for an electron), and a vertically polarized photon registers as $-1$ (akin to spin down). These states have no overlap; they are orthogonal to each other:

$$\langle V|H\rangle = 0\,,\ \langle H|V\rangle = 0\,, \tag{5}$$

where $\langle\psi|$ is the "Hermitian conjugate" of the state $|\psi\rangle$.[5] The states are also normalized such that

$$\langle H|H\rangle = 1\,,\ \langle V|V\rangle = 1\,. \tag{6}$$

If a physicist were to *rotate* the polarizing filter at her detector so that it pointed along a new direction $\hat{\mathbf{n}}'$ that made an angle $\theta \equiv \hat{\mathbf{n}}\cdot\hat{\mathbf{n}}'$ with respect to the original orientation, then a *different* set of states $|\tilde{H}\rangle$ and $|\tilde{V}\rangle$ would register as either horizontally or vertically polarized at her detector. These states share some overlap with the original states $|H\rangle$ and $|V\rangle$. In fact, the new states are simply rotations by the angle $\theta$ from the original states:

$$\begin{aligned}\begin{pmatrix}|\tilde{H}\rangle\\|\tilde{V}\rangle\end{pmatrix} &= R(\theta)\begin{pmatrix}|H\rangle\\|V\rangle\end{pmatrix}\\ &= \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}|H\rangle\\|V\rangle\end{pmatrix}\,,\end{aligned} \tag{7}$$

[5]If we represent the state $|\psi\rangle$ as a column vector in its abstract vector space, then $\langle\psi|$ is a row vector in that space, given by the conjugate transpose. For example:

$$|\psi\rangle = \begin{pmatrix}x\\y\end{pmatrix} \longrightarrow \langle\psi| = (x^*, y^*)\,,$$

where the components $x$ and $y$ are numbers (which could be complex). The inner product between two states $|\psi_1\rangle$ and $|\psi_2\rangle$ can then be evaluated:

$$\text{inner product of }\ |\psi_1\rangle = \begin{pmatrix}m\\n\end{pmatrix}\ \text{with}\ |\psi_2\rangle = \begin{pmatrix}s\\t\end{pmatrix}\ \text{is given by}\ \langle\psi_2|\psi_1\rangle = (s^*, t^*)\begin{pmatrix}m\\n\end{pmatrix} = ms^* + nt^*\,.$$

If $|\psi_1\rangle = |H\rangle$ and $|\psi_2\rangle = |V\rangle$, representing photon states with horizontal and vertical polarization (in a particular basis), respectively, then $m = 1$, $n = 0$, $s = 0$, and $t = 1$, and we see that $\langle V|H\rangle = \langle H|V\rangle = 0$. More loosely, we can think of the inner product $\langle\psi_2|\psi_1\rangle$ as the analogy of a dot product among ordinary vectors, keeping in mind that the components of the quantum states $|\psi_i\rangle$ can, in general, be complex numbers.

or

$$
\begin{aligned}
|\tilde{H}\rangle &= |H\rangle \cos\theta + |V\rangle \sin\theta\,, \\
|\tilde{V}\rangle &= -\,|H\rangle \sin\theta + |V\rangle \cos\theta\,.
\end{aligned} \tag{8}
$$

In Eq. (7), $R(\theta)$ is the usual rotation matrix in two spatial dimensions.[6] When measured along the orientation $\hat{\mathbf{n}}'$, these states yield measurement outcomes $+1$ (for state $|\tilde{H}\rangle$) and $-1$ (for state $|\tilde{V}\rangle$). Using Eqs. (5) and (6), it is easy to confirm that the pair of polarization states in the rotated basis are orthogonal to each other and also properly normalized: $\langle\tilde{V}|\tilde{H}\rangle = \langle\tilde{H}|\tilde{V}\rangle = 0$, and $\langle\tilde{H}|\tilde{H}\rangle = \langle\tilde{V}|\tilde{V}\rangle = 1$.

## Maximally Entangled States

We can arrange for the source $\sigma$ to emit pairs of photons in a maximally entangled state, such as

$$
|\Psi\rangle = \frac{1}{\sqrt{2}}\Big\{\,|H\rangle_A\,|V\rangle_B - |V\rangle_A\,|H\rangle_B\,\Big\}\,. \tag{9}
$$

The subscript "$A$" refers to the state of the particle traveling toward the left detector, and "$B$" refers to the particle traveling toward the right detector. Eq. (9) represents an "entangled" state because it cannot be *factorized*: there is no way to describe the behavior of particle $A$ without *also* referring to some aspect of the behavior of particle $B$, and vice versa. The state $|\Psi\rangle$ is "maximally" entangled because the states within the superposition are weighted with coefficients of equal magnitude (in this case, $1/\sqrt{2}$).

When a pair of photons is prepared in the state $|\Psi\rangle$ and each photon is measured by a polarizer oriented in the direction $\hat{\mathbf{n}}$, the measurement outcomes should always be opposite to each other: $A = +1$ at the left detector together with $B = -1$ at the right detector, or $A = -1$ with $B = +1$. These two outcomes for the pair of measurements $(A, B)$ should occur with equal probability, since the coefficients in front of each term on the right side of Eq. (9) are of equal magnitude.

We can unpack those statements a bit more explicitly, which should make it easier to understand how things change once the physicists rotate their polarizers to directions other than $\hat{\mathbf{n}}$. When both polarizers remain aligned along $\hat{\mathbf{n}}$ (that is, $a$ and $b$ are each rotated by $0°$ from the direction $\hat{\mathbf{n}}$), we may calculate $p(A, B|a, b) \to p(A, B|0°, 0°)$ for various outcomes $A = \pm1$ and $B = \pm1$. In general in quantum theory, the probability for a given outcome is given by the absolute square of the corresponding *amplitude*, $\mathcal{A}$. The amplitude, in turn, is calculated from the *overlap* (or inner product) between the incoming quantum state $|\Psi\rangle$ and

[6] We are assuming that the polarizing filters are oriented at some direction within the two-dimensional plane that is perpendicular to the photons' direction of travel; hence we may use this form for $R(\theta)$.

the state corresponding to the relevant measurement outcome. For example, we may consider measurements along $\hat{\mathbf{n}}$ that yield outcomes $A = +1$ and $B = +1$. The state corresponding to this measurement outcome is $|H\rangle_A |H\rangle_B$. Using the expression for the incoming state $|\Psi\rangle$ in Eq. (9), we may calculate the amplitude as

$$\begin{aligned}
\mathcal{A}(+1,+1|0°,0°) &= \Big\{ {}_B\langle H|_A\langle H| \Big\} |\Psi\rangle \\
&= \frac{1}{\sqrt{2}} \Big\{ {}_B\langle H|_A\langle H| \Big\} \Big\{ |H\rangle_A |V\rangle_B - |V\rangle_A |H\rangle_B \Big\} \\
&= \frac{1}{\sqrt{2}} \Big\{ \Big( {}_B\langle H|V\rangle_B \Big) \Big( {}_A\langle H|H\rangle_A \Big) - \Big( {}_B\langle H|H\rangle_B \Big) \Big( {}_A\langle H|V\rangle_A \Big) \Big\} \qquad (10) \\
&= \frac{1}{\sqrt{2}} \Big\{ 0 \times 1 - 1 \times 0 \Big\} \\
&= 0\,.
\end{aligned}$$

In moving from the second line to the third, we have made use of the fact that the states pertaining to particles $A$ and $B$ belong to distinct vector spaces, so we always have ${}_B\langle \psi_2|\psi_1\rangle_A = 0$ and ${}_A\langle \psi_1|\psi_2\rangle_B = 0$ for *any* states $|\psi_1\rangle_A$ and $|\psi_2\rangle_B$. In moving from the third to the fourth line, we have made use of the *orthogonality* of the states $|H\rangle$ and $|V\rangle$, as in Eq. (5): for each particle, ${}_A\langle H|V\rangle_A = {}_A\langle V|H\rangle_A = {}_B\langle H|V\rangle_B = {}_B\langle V|H\rangle_B = 0$. We have also used the fact that these states are *normalized*, as in Eq. (6). Thus we see that for a pair of particles prepared in the state $|\Psi\rangle$ of Eq. (9), the amplitude to measure $A = +1$ and $B = +1$ along the orientation $\hat{\mathbf{n}}$ *vanishes*. Therefore we find the probability

$$p(+1,+1|0°,0°) = \mathcal{A}(+1,+1|0°,0°)^{2} = 0\,. \qquad (11)$$

According to quantum mechanics, there should be *zero* probability of measuring $A = +1$ and $B = +1$ along $\hat{\mathbf{n}}$ for pairs of particles prepared in state $|\Psi\rangle$. A similar calculation yields $p(-1,-1|0°,0°) = 0$ for particles prepared in state $|\Psi\rangle$.

On the other hand, we find a nonzero probability to measure $A = +1$ and $B = -1$:

$$\begin{aligned}
\mathcal{A}(+1,-1|0°,0°) &= \Big\{ {}_B\langle V|_A\langle H| \Big\} |\Psi\rangle \\
&= \frac{1}{\sqrt{2}} \Big\{ {}_B\langle V|_A\langle H| \Big\} \Big\{ |H\rangle_A |V\rangle_B - |V\rangle_A |H\rangle_B \Big\} \\
&= \frac{1}{\sqrt{2}} \Big\{ \Big( {}_B\langle V|V\rangle_B \Big) \Big( {}_A\langle H|H\rangle_A \Big) - \Big( {}_B\langle V|H\rangle_B \Big) \Big( {}_A\langle H|V\rangle_A \Big) \Big\} \qquad (12) \\
&= \frac{1}{\sqrt{2}} \Big\{ 1 \times 1 - 0 \times 0 \Big\} \\
&= \frac{1}{\sqrt{2}}\,.
\end{aligned}$$

So then we find

$$p(+1,-1|0^\circ,0^\circ) = \ \mathcal{A}(+1,-1|0^\circ,0^\circ)\ ^2 = \frac{1}{2}\,. \tag{13}$$

The same series of steps yields $p(-1,+1|0^\circ,0^\circ) = 1/2$. Therefore we see that when both polarizers are oriented along the original direction $\hat{\mathbf{n}}$, quantum mechanics predicts that physicists should find the results $A = +1, B = -1$ half the time (probability $= 1/2$), and $A = -1, B = +1$ half the time. Note that this calculation does *not* enable us to predict which specific result will arise on a *given* experimental run. When using quantum mechanics, we are only able to calculate the *probability* to find a specific outcome on any given run.

## Measurement Correlations with Rotated Polarizers

The real genius of John Bell's work comes from allowing each physicist to rotate the polarizing filter at her detector station by some angle away from the original direction $\hat{\mathbf{n}}$, and then considering how the correlations in measurement outcomes vary. Suppose that the physicist at the left, whose detector station performs measurements on particle $A$, chooses to rotate her polarizing filter by some angle $\theta_a$ away from the original direction $\hat{\mathbf{n}}$, while the physicist at the right, who measures particle $B$, rotates her polarizing filter by some other angle $\theta_b$ away from $\hat{\mathbf{n}}$. The angle between the two polarizing filters is then given by $\theta_{ab} = \theta_a - \theta_b$.

Using Eq. (8), the states that correspond to measurement outcomes $A = \pm 1$ once the physicist at the left detector has rotated her polarizing filter by angle $\theta_a$ may be written

$$\begin{aligned} |\check{H}\rangle_A &= |H\rangle_A \cos\theta_a + |V\rangle_A \sin\theta_a\,, \\ |\check{V}\rangle_A &= -\,|H\rangle_A \sin\theta_a + |V\rangle_A \cos\theta_a\,, \end{aligned} \tag{14}$$

while the states that correspond to measurement outcomes $B = \pm 1$ when the polarizing filter at the right detector has been rotated by angle $\theta_b$ may be written

$$\begin{aligned} |\tilde{H}\rangle_B &= |H\rangle_B \cos\theta_b + |V\rangle_B \sin\theta_b\,, \\ |\tilde{V}\rangle_B &= -\,|H\rangle_B \sin\theta_b + |V\rangle_B \cos\theta_b\,. \end{aligned} \tag{15}$$

Then we may calculate various amplitudes, such as

$$\begin{aligned} \mathcal{A}(+1,+1|\theta_a,\theta_b) &= \Big\{ {}_B\langle\tilde{H}|\,{}_A\langle\check{H}| \Big\}\,|\Psi\rangle \\ &= \Big\{ \Big( {}_B\langle H|\cos\theta_b + {}_B\langle V|\sin\theta_b \Big)\Big( {}_A\langle H|\cos\theta_a + {}_A\langle V|\sin\theta_a \Big) \Big\}\,|\Psi\rangle \\ &= \Big\{ {}_B\langle H|\,{}_A\langle H|\cos\theta_b\cos\theta_a + {}_B\langle H|\,{}_A\langle V|\cos\theta_b\sin\theta_a \\ &\qquad + {}_B\langle V|\,{}_A\langle H|\sin\theta_b\cos\theta_a + {}_B\langle V|\,{}_A\langle V|\sin\theta_b\sin\theta_a \Big\}\,|\Psi\rangle\ . \end{aligned} \tag{16}$$

Given the form of $|\Psi\rangle$ in Eq. (9) and the orhogonality of $|H\rangle$ and $|V\rangle$ states, the only nonvanishing contributions in Eq. (16) come from terms proportional to ${}_B\langle V|_A\langle H|$ and ${}_B\langle H|_A\langle V|$, so we then find

$$\begin{aligned}
\mathcal{A}(+1,+1|\theta_a,\theta_b) &= \frac{1}{\sqrt{2}}\Big\{ -\Big({}_B\langle H|H\rangle_B\Big)\Big({}_A\langle V|V\rangle_A\Big)\cos\theta_b\sin\theta_a \\
&\qquad\qquad + \Big({}_B\langle V|V\rangle_B\Big)\Big({}_A\langle H|H\rangle_A\Big)\sin\theta_b\cos\theta_a\Big\} \\
&= \frac{1}{\sqrt{2}}\Big\{ -\cos\theta_b\sin\theta_a + \sin\theta_b\cos\theta_a\Big\} \\
&= \frac{1}{\sqrt{2}}\sin(\theta_b-\theta_a) \\
&= -\frac{1}{\sqrt{2}}\sin\theta_{ab}\,,
\end{aligned} \tag{17}$$

where the penultimate step makes use of the trigonometric angle-addition formula, and the final step follows upon using $\theta_{ab} = \theta_a - \theta_b$ and $\sin(-\theta_{ab}) = -\sin\theta_{ab}$. Following the same series of steps, we may calculate the other amplitudes:

$$\begin{aligned}
\mathcal{A}(-1,-1|\theta_a,\theta_b) &= \Big\{{}_B\langle\tilde{V}|_A\langle\check{V}|\Big\}\,|\Psi\rangle = -\frac{1}{\sqrt{2}}\sin\theta_{ab}\,, \\
\mathcal{A}(+1,-1|\theta_a,\theta_b) &= \Big\{{}_B\langle\tilde{V}|_A\langle\check{H}|\Big\}\,|\Psi\rangle = \frac{1}{\sqrt{2}}\cos\theta_{ab}\,, \\
\mathcal{A}(-1,+1|\theta_a,\theta_b) &= \Big\{{}_B\langle\tilde{H}|_A\langle\check{V}|\Big\}\,|\Psi\rangle = -\frac{1}{\sqrt{2}}\cos\theta_{ab}\,.
\end{aligned} \tag{18}$$

Taking the absolute squares, we then find the probabilities for these various outcomes:

$$\begin{aligned}
p(+1,+1|\theta_a,\theta_b) = p(-1,-1|\theta_a,\theta_b) &= \frac{1}{2}\sin^2\theta_{ab}\,, \\
p(+1,-1|\theta_a,\theta_b) = p(-1,+1|\theta_a,\theta_b) &= \frac{1}{2}\cos^2\theta_{ab}\,.
\end{aligned} \tag{19}$$

We now have all the pieces we need to calculate the quantum-mechanical prediction for the correlation function $E(a,b)$ as in Eq. (3), applied to a pair of particles prepared in the state $|\Psi\rangle$ of Eq. (9). For an angle $\theta_{ab}$ between the orientations of the polarizers at the left and right stations, the correlation function takes the form

$$\begin{aligned}
E(a,b) &= \sum_{A,B=\pm 1} A\,B\,p(A,B|a,b) \\
&= (+1)(+1)\,p(+1,+1|\theta_a,\theta_b) + (+1)(-1)\,p(+1,-1|\theta_a,\theta_b) \\
&\qquad + (-1)(+1)\,p(-1,+1|\theta_a,\theta_b) + (-1)(-1)\,p(-1,-1|\theta_a,\theta_b) \\
&= \sin^2\theta_{ab} - \cos^2\theta_{ab} \\
&= -\cos(2\theta_{ab})\,,
\end{aligned} \tag{20}$$

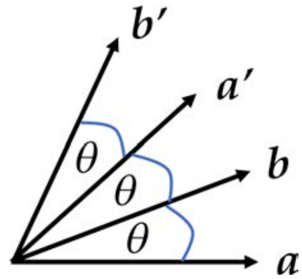


Figure 2: In a test of the Bell-CHSH inequality, physicists can select various angles between their detector settings $(a, a')$ and $(b, b')$. A common choice is to set $\theta_{ab} = \theta_{a'b} = \theta_{a'b'} = \theta$, and then $\theta_{ab'} = 3\theta$.

where the last step follows upon using the usual double-angle formula. As a quick reality check, note that for the case in which the two polarizers are each aligned along $\hat{\mathbf{n}}$, so that $\theta_{ab} = 0$, the expression in Eq. (20) indicates that the two photons should always be measured to have *opposite* polarizations: either $A = +1$ and $B = -1$, or $A = -1$ and $B = +1$, exactly as we would expect from the form of $|\Psi\rangle$ in Eq. (9).

## Predictions for the Bell-CHSH Parameter

Our quantum-mechanical expression for $E(a, b)$ in Eq. (20) holds for *any* value of $\theta_{ab}$. So we may us it to calculate the quantum-mechanical prediction for the Bell-CHSH parameter $S$ defined in Eq. (2):

$$\begin{aligned} S &\equiv E(a,b) + E(a',b) - E(a,b') + E(a',b') \\ &= -\cos(2\theta_{ab}) - \cos(2\theta_{a'b}) + \cos(2\theta_{ab'}) - \cos(2\theta_{a'b'})\,. \end{aligned} \tag{21}$$

We can consider a particular arrangement among the detector settings $(a, a')$ and $(b, b')$, as in Fig. 2. In that case, we set $\theta_{ab} = \theta_{a'b} = \theta_{a'b'} = \theta$, and hence $\theta_{ab'} = 3\theta$. Fig. 3 shows how the value of $-S$ varies with the angle $\theta$.[7] Note that for several choices of the angle $\theta$, the quantum-mechanical prediction for the Bell-CHSH parameter $S$ *exceeds* the Bell-CHSH inequality of Eq. (4): rather than $|S|$ remaining restricted to $|S| \leq 2$, there are scenarios predicted in which measurements on quantum systems should be *more strongly correlated* than the Bell-CHSH inequality would allow, with $|S| > 2$. In particular, the predicted correlations are maximized for $\theta = 22.5°$. For that choice of relative orientations among the detector settings, we have $\cos(2\theta) = \cos(45°) = 1/\sqrt{2}$ and $\cos(6\theta) = \cos(135°) = -1/\sqrt{2}$,

[7]In Fig. 3 I plot $-S$ rather than $S$ because the Bell-CHSH inequality applies to $|S|$, and I find it easier to visualize the main peaks as in Fig. 3. If we had considered correlations among measurements on pairs of particles prepared in some other maximally entangled state, such as $|\Phi\rangle = \{|H\rangle_A |H\rangle_B + |V\rangle_A |V\rangle_B\}/\sqrt{2}$, then our expression for $S$ would have had the opposite overall sign from the expression in Eq. (21), though, obviously, the same behavior of $|S|$.

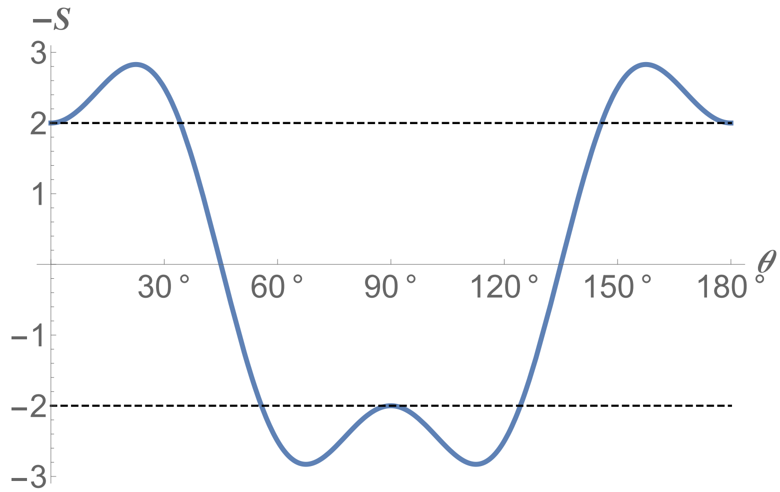


Figure 3: The quantum-mechanical prediction for the Bell-CHSH parameter $-S$ from Eq. (21) when we select the angles between the various detector settings $(a, a')$ and $(b, b')$ as in Fig. 2. Certain choices of the angle $\theta$, such as $\theta = 22.5°$, yield a maximum predicted violation of the Bell-CHSH inequality: $|S| = 2\sqrt{2}$ rather than $|S| \leq 2$.

which yields

$$S_{QM}^{\max} = 2\sqrt{2} = 2.83\,. \tag{22}$$

This is known as the "Tsirelson bound": the maximum value for the Bell-CHSH parameter predicted to occur for quantum systems.[8]

As Bell demonstrated, local-realist theories and quantum mechanics make substantially different predictions for the degree of correlation among measurements on pairs of particles that have been prepared in a certain way. Moreover, the quantitative difference between $|S| \leq 2$ and $|S| = 2\sqrt{2}$ is sufficiently large that Bell hoped even fairly rudimentary experiments might be able to distinguish between these predictions empirically.[9]

The first experimental test of Bell's inequality was conducted by John Clauser and Stuart Freedman at the Lawrence Berkeley Laboratory in California in 1972. They found results consistent with the quantum-mechanical predictions, violating the Bell-CHSH inequality to high statistical significance.[10] Since that time, dozens of additional experimental tests have been conducted, subjecting different types of matter to measurements with different types of detectors. Even more important, recent experiments have closed a series of "loopholes": logical possibilities — however seemingly remote or implausible — by which a local-realist theory could give rise to correlated measurements that mimic the expectations from quantum

[8] B. S. Cirel'son, "Quantum generalizations of Bell's inequality," *Lett. Math. Phys.* **4** (March 1980): 93-100.

[9] Bell, "On the Einstein-Podolsky-Rosen paradox"; Bell, *Speakable and Unspeakable in Quantum Mechanics.*

[10] Stuart J. Freedman and John F. Clauser, "Experimental test of local hidden-variable theories," *Phys. Rev. Lett.* **28**: 938-941.

theory, exceeding Bell's bound. For example, during 2016–2018 my own group conducted a series of "Cosmic Bell" experiments, in which we used real-time astronomical observations of light from distant objects, such as high-redshift quasars, as sources of randomness to determine which detector settings to use at our two detector stations for each experimental run. Our goal was to minimize the possibility that the choices of measurements to be performed on any given experimental run could have had some unanticipated correlation with the properties of the entangled particles emitted from the source.[11] To date, every published test of Bell's inequality, including the recent experiments that address various loopholes, has been consistent with the predictions of quantum mechanics.[12] The experimental evidence is stronger than ever: quantum entanglement seems to be a robust feature of our world, not only a theoretical prediction.

## Faster-Than-Light Communication?

Although Einstein did not live to see either Bell's 1964 paper or the earliest experimental tests of Bell's inequality, he harbored serious concerns about quantum entanglement for years after publishing his 1935 EPR paper. He famously wrote to his friend Max Born in March 1947, for example, that he "cannot seriously believe in [quantum mechanics] because the theory cannot be reconciled with the idea that physics should represent a reality in time and space, free from spooky actions at a distance."[13]

Why did Einstein object to entanglement as "spooky action at a distance"? Much like Bell, Einstein and his EPR co-authors had imagined that each physicist conducting an experiment like that shown in Fig. 1 would have some *choice* as to which specific type of measurement to perform on the particle heading her way from the central source. If the

[11]My colleagues and I proposed this new type of "Cosmic Bell" experimental test in Jason Gallicchio, Andrew S. Friedman, and David I. Kaiser, "Testing Bell's inequality with cosmic photons: Closing the setting-independence loophole," *Phys. Rev. Lett.* **112** (2014): 110405, arXiv:1310.3288. Upon teaming up with experimentalist Anton Zeilinger and his group at the Institute for Quantum Optics and Quantum Information in Vienna, we next conducted a pilot test: Johannes Handsteiner et al. (Cosmic Bell collaboration), "Cosmic Bell test: Measurement settings from Milky Way stars," *Phys. Rev. Lett.* **118** (2017): 060401, arXiv:1611.06985. More recently our group conducted a test using light from extragalactic sources: Dominik Rauch et al. (Cosmic Bell collaboration), "Cosmic Bell test using random measurement settings from high-redshift quasars," *Phys. Rev. Lett.* **121** (2018): 080403, arXiv:1808.05966. For more on our astronomical-randomness technique, see also Calvin Leung, Amy Brown, Hien Nguyen, Andrew S. Friedman, David I. Kaiser, and Jason Gallicchio, "Astronomical random numbers for quantum foundations experiments," *Phys. Rev. A* **97** (2018): 042120, arXiv:1706.02276.

[12]For a review of experimental tests of Bell's inequality, including recent efforts to address various "loopholes," see David Kaiser, "Tackling loopholes in experimental tests of Bell's inequality," in *Oxford Handbook of the History of Interpretations of Quantum Physics*, ed. Olival Freire, Jr. (New York: Oxford University Press, forthcoming).

[13]Albert Einstein to Max Born, 3 March 1947, as translated and published in Max Born, ed., *The Born-Einstein Letters*, trans. Irene Born (New York: Macmillan, 1971), on p. 158.

physicist at the left detector happened to choose to measure her particle's *position*, then — according to the EPR authors' "reality criterion" — there *must* exist an "element of reality" corresponding to the precise position of the other particle, traveling toward the detector on the right side, since such a value could be calculated with certainty based on the outcome at the left detector. If, instead, the physicist at the left detector had chosen to measure the *momentum* of her particle, then there would exist an "element of reality" corresponding to the precise value of momentum of the right-moving particle. Yet according to Heisenberg's uncertainty principle, the right-moving particle could not have simultaneously sharp values of *both* position and momentum. So how would that particle "know" which of its properties should have a definite value by the time it reached its own detector? If the detectors were sufficiently far apart from each other, and the physicist at the left waited until the last possible moment to select the type of measurement her detector would perform, then there would be no time — even for a light signal — to update the right-moving particle as to which property it "should" have. The correlation of the particles' behavior inherent in the quantum-mechanical description, Einstein was convinced, implied some faster-than-light "action at a distance." To this day, entanglement is sometimes referred to as "quantum nonlocality."

Given the strong correlations that persist between entangled particles, even after they have moved an arbitrary distance apart from each other, a natural question to ask — and one that several physicists did indeed ask upon learning about Bell's inequality in the 1960s and 1970s — is whether entanglement can be used to send messages faster than light.[14] According to the quantum-mechanical description of entanglement, the answer, alas, is "no."

Remember that Bell tests involve comparing measurement outcomes $(A, B)$ from distant detector stations, as the detector settings $(a, b)$ are varied. Only when we collect both the lists of detector settings that were actually used on each experimental run, as well as the measurement outcomes for each run, can we identify a distinct signature of quantum entanglement, such as a violation of the Bell-CHSH inequality. If we only have access to information from *one* side of the experiment — say, a list of detector settings $(a)$ and measurement outcomes $(A)$ from the left detector — then we have *no way to know* whether those measurement outcomes showed any particular correlation with whatever might have happened at the right detector. In fact, if all we have is access to the list from one detector, all we can ever measure is a *random* series of $+1$'s and $-1$'s, each occurring with equal frequency but in an unpredictable order. That is, data from only one side of a Bell test is indistinguishable from *random noise*.

In Eq. (19) we calculated the probabilities for various joint outcomes $(A, B)$ given joint

[14] See esp. Kaiser, *How the Hippies Saved Physics*, chap. 9.

detector settings $(a, b)$. If we only have access to information from the left detector, then all we can do is compare the output of the detector with our predictions for various measurement outcomes $A = \pm 1$ when the polarizing filter at that detector is oriented at an angle $\theta_a$. In particular, there should not be any *change* in the probabilities to find various outcomes $A$ at the left detector based on the choice of what detector setting $b$ to use at the distant detector. (If one *could* affect the outcomes $A$ on the left by changing the detector setting $b$ on the right, then one could send signals from right to left arbitrarily quickly.) Our test is therefore whether

$$p(A = +1|a, b) = p(A = +1|a, b') \quad \text{and} \quad p(A = -1|a, b) = p(A = -1|a, b') \,. \tag{23}$$

The probabilty $p(A|a, b)$ is just given by the *sum* over the (unknown) measurement outcomes at the right detector:

$$p(A|a, b) = \sum_{B=\pm 1} p(A, B|a, b) \,. \tag{24}$$

Using Eq. (19), we find

$$\begin{aligned} p(A = +1|\theta_a, \theta_b) &= p(+1, +1|\theta_a, \theta_b) + p(+1, -1|\theta_a, \theta_b) \\ &= \frac{1}{2} \sin^2 \theta_{ab} + \frac{1}{2} \cos^2 \theta_{ab} \\ &= \frac{1}{2} \,, \end{aligned} \tag{25}$$

*independent* of the relative orientation of the polarizing filters at each detector. Hence we find precisely the same answer for when the polarizer at the right detector is rotated to some other angle $\theta_{b'}$ rather than $\theta_b$: $p(A = +1|\theta_a, \theta_{b'}) = 1/2$. Likewise

$$\begin{aligned} p(A = -1|\theta_a, \theta_b) &= p(-1, +1|\theta_a, \theta_b) + p(-1, -1|\theta_a, \theta_b) \\ &= \frac{1}{2} \cos^2 \theta_{ab} + \frac{1}{2} \sin^2 \theta_{ab} \\ &= \frac{1}{2} \,, \end{aligned} \tag{26}$$

again independent of the angle $\theta_b$; so as before we find $p(A = -1|\theta_a, \theta_{b'}) = 1/2$. Not only are the "no-signaling" criteria of Eq. (23) satisfied. We find an even stronger result: the output at the left detector is *nothing other than* a random sequence of $A = +1$ and $A = -1$, each occurring with probability $= 1/2$, the very definition of random noise!

According to quantum mechanics, therefore, no *information* can be conveyed faster than light by exploiting quantum entanglement. The signal that two particles really were entangled can only be ascertained once we use some other means of communication — limited by the speed of light — to collate information about detector settings and measurement outcomes from *both* detectors.

# Appendix: Derivation of Bell's Inequality

The EPR authors had argued that quantum mechanics was incomplete. Bell argued that if a more complete physical description were to remain compatible with local realism, the measurement outcomes $A$ and $B$ would need to be determined by functions of the form

$$A(a, \lambda) = \pm 1\,, \;\; B(b, \lambda) = \pm 1\,. \tag{27}$$

Here $A$ and $B$ are the measurement outcomes at each detector given detector settings $a$ and $b$, and $\lambda$ is shorthand for any other properties of the particles prepared at the source $\sigma$ that could affect the measurement outcomes $A$ and $B$. (The extra properties $\lambda$, not included in a standard quantum-mechanical description, are often called "hidden variables.") As Bell emphasized, the forms of $A(a, \lambda)$ and $B(b, \lambda)$ in Eq. (27) incorporate *locality*: the measurement outcome at the left detector, $A$, could only depend on the detector setting at that detector ($a$) and on the particles' shared properties ($\lambda$); the measurement outcome $A$ could *not* depend on either the choice of detector setting ($b$) or the measurement outcome ($B$) at the right detector, and vice versa.[15]

Bell imagined that for a given detector setting $a$ and a particular value $\lambda_i$, $A(a, \lambda_i) = +1$, while for that same detector setting and a distinct value $\lambda_j$, $A(a, \lambda_j) = -1$, and similarly for $B(b, \lambda)$. The parameters $\lambda$ would vary across experimental runs, governed by some probability distribution $p(\lambda)$; presumably on any given experimental run, both particles would *share* the same value of $\lambda$, since it presumably arose during the process by which that particular pair of particles was generated at the source $\sigma$. On a given experimental run, therefore, the particles would share a particular value $\lambda_i$ and be subjected to measurements with the joint detector settings $(a, b)$. In that case, the product of the measurement outcomes at the two detectors would be $A(a, \lambda_i)\, B(b, \lambda_i)$. The expectation value $E(a, b)$ would then be given by averaging over many trials while holding $(a, b)$ fixed. Across those trials, the quantity that would change would be $\lambda$, so that in a local-realist theory,

$$E(a, b) = \int d\lambda\, p(\lambda)\, A(a, \lambda)\, B(b, \lambda)\,. \tag{28}$$

[15] Bell, "On the Einstein-Podolsky-Rosen paradox," p. 196.

For a local-realist theory, the Bell-CHSH parameter $S$ would then take the form

$$
\begin{aligned}
S &\equiv E(a,b) + E(a',b) - E(a,b') + E(a',b') \\
&= \int d\lambda\, p(\lambda) \Big\{ A(a,\lambda)\, B(b,\lambda) + A(a',\lambda)\, B(b,\lambda) \\
&\qquad\qquad - A(a,\lambda)\, B(b',\lambda) + A(a',\lambda)\, B(b',\lambda) \Big\} \\
&= \int d\lambda\, p(\lambda) \Big\{ A(a,\lambda) \Big[ B(b,\lambda) - B(b',\lambda) \Big] + A(a',\lambda) \Big[ B(b,\lambda) + B(b',\lambda) \Big] \Big\} .
\end{aligned}
\tag{29}
$$

Let us consider the quantity in curly brackets within the integrand in the last line of Eq. (29) for a particular value $\lambda_i$. For any particular value $\lambda_i$, the functions $B(b,\lambda_i)$ and $B(b',\lambda_i)$ can only equal $\pm 1$, as in Eq. (27). Suppose that for this particular value $\lambda_i$, these functions happen to equal $B(b,\lambda_i) = +1$ and $B(b',\lambda_i) = -1$. Then for this particular value of $\lambda$, we have

$$
\Big[ B(b,\lambda_i) - B(b',\lambda_i) \Big] = +2\,, \ \Big[ B(b,\lambda_i) + B(b',\lambda_i) \Big] = 0\,. \tag{30}
$$

Meanwhile, for that same value $\lambda_i$, the functions $A(a,\lambda_i)$ and $A(a',\lambda_i)$ can likewise only equal $\pm 1$. In that case,

$$
\Big\{ A(a,\lambda_i) \Big[ B(b,\lambda_i) - B(b',\lambda_i) \Big] + A(a',\lambda_i) \Big[ B(b,\lambda_i) + B(b',\lambda_i) \Big] \Big\} = 2A(a,\lambda_i) = \pm 2\,. \tag{31}
$$

If, for that same value $\lambda_i$, the functions $B(b,\lambda_i)$ and $B(b',\lambda_i)$ instead happened to equal $B(b,\lambda_i) = -1$ and $B(b',\lambda_i) = -1$, we would again have found that *one* of the two terms in square brackets on the left side of Eq. (31) vanished, while the other term had the value $\pm 2$. Given $|A(a,\lambda_i)| = |A(a',\lambda_i)| = 1$, we would again have found that the term in curly brackets on the left side of Eq. (31) would have the value $\pm 2$. In fact, for *any* assignments $A(a,\lambda_i), A(a',\lambda_i), B(b,\lambda_i)$ and $B(b',\lambda_i)$ as $\pm 1$, we will *always* find that the expression in curly brackets on the left side of Eq. (31) must equal $\pm 2$.

To evaluate the Bell-CHSH parameter $S$ in Eq. (29), we need to average over all the possible values $\lambda$ that could arise on a given experimental run, weighted by the probability distribution $p(\lambda)$. Upon using the fact that the probability distribution $p(\lambda)$ is normalized,

$$
\int d\lambda\, p(\lambda) = 1\,, \tag{32}
$$

we arrive at the Bell-CHSH inequality:

$$
|S| \leq 2 \tag{33}
$$

for *any* local-realist theory.[16]

[16]For generalizations of this derivation, see Andrew S. Friedman, Alan H. Guth, Michael J. W. Hall, David I. Kaiser, and Jason Gallicchio, "Relaxed Bell inequalities with arbitrary measurement dependence for each observer," *Physical Review A* **99** (2019): 012121, arXiv:1809.01307 and references therein.

MIT OpenCourseWare
https://ocw.mit.edu

STS.042J / 8.225J Einstein, Oppenheimer, Feynman: Physics in the 20th Century
Fall 2020

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms