EPR Paradox and Bell's Theorem

By Bill Daniel

Introduction

In their paper, EPR explored the subtle consequences of entanglement between a pair of particles that are allowed to interact and then move a considerable distance apart. EPR pointed out that measurements on one particle can instantly provide information about the other, unmeasured, particle. Since special relativity forbids superluminal infirmation transfer, they concluded that there must be some, as yet unknown, aspect of quantum mechanics that connected the two measurements - a so called "hidden variable." This process of extracting information about an unmeasured part of a system via measurements on another part has come to be called "counter-factual definiteness." Most physicists accept that the results of such implied "measurements" are just as real as actual measurements, but some do not.

EPR raised their objections in 1935, but it wasn't until 1964 that Bell was able to shed some light on the paradox. Bell's theorem says that *any physical theory that assumes local realism* (as EPR did) *cannot also predict all of the results of quantum mechanics*. Local hidden variable theories and quantum mechanics are fundamentally incompatible.

EPR envisioned an experiment involving entangled photons whose polarizations must be perpendicular. This is the example worked out in *Basic Concepts in Physics* §6.17.3-6. I find that treatment to be very confusing due to complexities involved in polarization that are unnecessary to the argument. In this case I find, as I often do, that the clearest explanation of an idea is in the original paper[1]. I propose we consider Bell's own, much simpler, version. I'll outline his argument, using his notation.

Bell's inequality

Bell considers "a pair of spin one-half particles formed somehow in the singlet spin state and moving freely in opposite directions." Simply for the sake of concreteness, I'll suppose they are an electron/positron pair produced by the decay of a π^0 meson. In the rest frame of the meson, the e^- and e^+ travel in opposite directions (by conservation of momentum). Since the π^0 has no spin, the spin of the e^- must be opposite that of the e^+ (because they are in the singlet state). We can't predict the outcome of an individual spin measurement, but, if the detectors used to measure the spins of the two particles are oriented in the same direction, we can predict that we will always measure opposite spins: if the e^- is spin up (↑), the e^+ will be spin down (↓) and vice versa.

The measurements of the e^- and e^+ are made at locations **A** and **B**, a considerable distance apart, using spin detectors oriented at angles **a** and **b** to some arbitrary reference direction. The **a** = **b** configuration (detectors aligned) is the only one considered by EPR and we know that, while the spin orientation of any specific measurement at **A** is unpredictable (either ↑ or ↓), every ↑ measurement at **A** will necessarily entail a ↓ measurement at **B**. The measurements at **A** and **B** are said to be perfectly "anti-correlated" or, said another way, they exhibit a -1 correlation.

Bell's breakthrough came from the idea of analyzing the situation when the spin detectors are *differently* oriented. If the detectors are *not* oriented in the same way (**a** ≠ **b**), the measurements at **A** and **B** will still be correlated, but no longer perfectly so. Averaging over many measurements, we will find a correlation between measurements at **A** and **B** somewhere between +1 and -1. A perfect, +1 correlation would mean that a measurement of ↑ at **A** would always be paired with an ↑ measurement at **B**, while a -1 correlation is the EPR case when an ↑ spin at **A** always implies a ↓ measurement at **B**.

Assume that there exists a local theory with hidden variable(s), λ. This λ could be anything - real or complex number(s), vector(s), spinor(s), tensor(s), ..., and it could be discrete or continuous. It varies in some unknown way from one meson decay to the next. Next, define a discrete function $A(\boldsymbol{a},\ \lambda) = \pm 1$ that represents the spin measurement at **A** $(\uparrow \Rightarrow A(\boldsymbol{a},\ \lambda) = +1,\ \downarrow \Rightarrow A(\boldsymbol{a},\ \lambda) = -1)$. A similar function $B(\boldsymbol{b},\ \lambda) = \pm 1$ expresses the spin measure ment at **B**. As we have seen, EPR pointed out that, for **a** = **b**:

$$A(\boldsymbol{a},\ \lambda) = -B(\boldsymbol{a},\ \lambda). \qquad (1)$$

or equivalently,

$$A(\boldsymbol{a},\ \lambda)\, B(\boldsymbol{a},\ \lambda) = -1. \qquad (2)$$

With *arbitrary* detector orientations, the product of these measurement functions will be ±1, depending on whether the spin measurement results are parallel or anti-parallel:

$$A(\boldsymbol{a},\ \lambda)\, B(\boldsymbol{b},\ \lambda) = \pm 1. \qquad (3)$$

The average value of this product over many π-meson decays is given by:

$$P(\boldsymbol{a},\ \boldsymbol{b}) = \int d\lambda\, \rho(\lambda)\, A(\boldsymbol{a},\ \lambda)\, B(\boldsymbol{b},\ \lambda), \qquad (4)$$

where $\rho(\lambda)$ is the unknown probability density of the hidden variable, λ. Even though we don't know the details of how ρ varies with λ, the simple fact that it is a probability density requires that,

$$0 \le \rho(\lambda) \le 1 \text{ and } \int d\lambda\, \rho(\lambda) = 1. \qquad (5)$$

Note that $P(\boldsymbol{a},\ \boldsymbol{b})$ is a continuous function (unlike the discrete $A(\boldsymbol{a},\ \lambda)$ and $B(\boldsymbol{b},\ \lambda)$) with:

$$-1 \le P(\boldsymbol{a},\ \boldsymbol{b}) \le +1. \qquad (6)$$

In fact, $P(\boldsymbol{a},\ \boldsymbol{b})$ is the "expectation value" of the product of the two measurements. (In quantum mechanics this poorly named quantity is *not* the "most likely outcome," as its name would seem to imply, but is instead the measured value of the quantity averaged over many trials.)

By equation (1), substituting **b** for **a**, we can eliminate function **B**:

$$P(\boldsymbol{a},\ \boldsymbol{b}) = -\int d\lambda\, \rho(\lambda)\, A(\boldsymbol{a},\ \lambda)\, A(\boldsymbol{b},\ \lambda). \qquad (7)$$

Now, suppose detector **B** is rotated to a new angle, **c**, and we cauculate the difference between the expectation values $P(\boldsymbol{a},\ \boldsymbol{b})$ and $P(\boldsymbol{a},\ \boldsymbol{c})$ at the two different **B** detector settings:

$$P(\boldsymbol{a},\ \boldsymbol{b}) - P(\boldsymbol{a},\ \boldsymbol{c}) = -\int d\lambda\, \rho(\lambda)[A(\boldsymbol{a},\ \lambda)\, A(\boldsymbol{b},\ \lambda) - A(\boldsymbol{a},\ \lambda)\, A(\boldsymbol{c},\ \lambda)]. \qquad (8)$$

Since, $(A(\boldsymbol{b},\ \lambda))^2 = 1$, we can rewrite equation (8) as:

$$P(\boldsymbol{a}, \boldsymbol{b}) - P(\boldsymbol{a}, \boldsymbol{c}) = -\int d\lambda\, \rho(\lambda)[1 - A(\boldsymbol{b}, \lambda)\, A(\boldsymbol{c}, \lambda)]\, A(\boldsymbol{a}, \lambda)\, A(\boldsymbol{b}, \lambda) \qquad (9)$$

or,

$$P(\boldsymbol{a}, \boldsymbol{b}) - P(\boldsymbol{a}, \boldsymbol{c}) = -\int d\lambda\, \rho(\lambda)[1 - A(\boldsymbol{b}, \lambda)\, A(\boldsymbol{c}, \lambda)] \times \int d\lambda\, \rho(\lambda)\, A(\boldsymbol{a}, \lambda)\, A(\boldsymbol{b}, \lambda). \qquad (10)$$

By (7) and (6), the absolute value of the second integral must be ≤ 1, so:

$$|P(\boldsymbol{a}, \boldsymbol{b}) - P(\boldsymbol{a}, \boldsymbol{c})| \leq \int d\lambda\, \rho(\lambda)[1 - A(\boldsymbol{b}, \lambda)\, A(\boldsymbol{c}, \lambda)]\,, \qquad (11)$$

and, by (5) and (7), substituting **b** for **a** and **c** for **b**,

$$|P(\boldsymbol{a}, \boldsymbol{b}) - P(\boldsymbol{a}, \boldsymbol{c})| \leq 1 + P(\boldsymbol{b}, \boldsymbol{c})\,. \qquad (12)$$

This is the famous Bell inequality. The only assumption that went into its derivation was locality: that measurements at **A** cannot instantaneously influence measurements at **B**. No assumptions were made about the nature of the hidden variable λ or its probability distribution $\rho(\lambda)$.

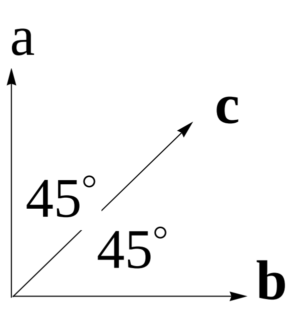
Relation to quantum mechanics

For arbitrary orientations, **a** and **b**, of the detectors at **A** and **B**, quantum mechanics predicts,

$$P(\boldsymbol{a}, \boldsymbol{b}) = -\boldsymbol{a} \cdot \boldsymbol{b} = -\cos\theta, \qquad (13)$$

where θ is the angle between **a** and **b**. (This derivation is beyond the scope of our discussion.)

Consider the case below.



$$P(\boldsymbol{a}, \boldsymbol{b}) = -\cos 90^\circ = 0 \qquad P(\boldsymbol{a}, \boldsymbol{c}) = P(\boldsymbol{b}, \boldsymbol{c}) = -\cos 45^\circ = -0.707.$$

Inserting these expectation values into Bell's inequality gives,

$$|P(\boldsymbol{a}, \boldsymbol{b}) - P(\boldsymbol{a}, \boldsymbol{c})| = |0 + 0.707| = 0.707 \;\nleq\; 1 + P(\boldsymbol{b}, \boldsymbol{c}) = 1 - 0.707 = 0.293\,.$$

Having made no assumptions about λ and having shown that Bell's inequality fails for one set of orientations is enough to prove that *no local hidden variable theory can reproduce this quantum mechanical prediction.*

(Actually, the inequality fails for almost all θ, but this one instance is sufficient for the proof.)

Implications

Bell’s Theorem showed that EPR’s paradox is far more fundamental that its originators had understood. There are only two ways out of the grip of Bell’s inequality. Either,

1) quantum mechanics is not only *incomplete*, as EPR claimed, it’s flat out *wrong*, or
2) the universe is fundamentally *nonlocal* and no hidden variable theory can save the day.

Subsequent experiments by Clauser, Aspect and others have shown to the satisfaction of all but a few physicists that conclusion 2) is the correct choice.

Much has been made in the popular literature about the nonlocality of nature, and many people have drawn wildly incorrect implications from it. Let’s consider what Bell’s theorem actually tells us about nonlocality.

Does the measurement at **A** *influence* the measurement at **B**? Surely it must, or else how can we account for the correlation between them? But, does one measurement *cause* the result observed at the other location? Certainly not. The fundamental principle of Bell’s theory is that *the observer at **A** cannot use his measurement to send a signal to **B***, since the result of his measurement is *a priori* unknown. Observer **A** cannot force his measurement to come out in a particular spin direction. If he could, he could send a message to **B**. But, the only thing observer **A** can control is whether or not he makes a measurement at all. This exercise of “free will” has no effect on measurements made at **B**, though, since the lists of spin measurement results at both locations will be purely random in either case. It’s only when the lists are *compared* (which can only take place at light speed or less) that the correlations are seen. Nonlocality, then, only applies to influences of a very subtle nature - those that do not involve the exchange of energy or information, and for which the evidence only becomes apparent when the results of distant experiments are compared.

Bell’s theorem completely ruled out *local* hidden variable theories, but it does not eliminate the possibility of *nonlocal* ones. A 2007 experiment[2] excluded a large class of such nonlocal theories, but there remains room for certain nonlocal hidden variable theories to be possible models of nature. In fact, in the 1950s, David Bohm resurrected one such model originally proposed by de Broglie that is exactly such an allowed theory. The transactional interpretation of quantum mechanics, first introduced in 1986 by John Cramer, is another candidate.

Still other physicists believe that Bell’s inequality can be circumvented by rejecting counterfactual definiteness or by embracing the many worlds model of quantum mechanics. All of this is very speculative and probably not worth our attention.

1. 1964: Bell, J. S., "On the Einstein Podolski Rosen Paradox," *Physics*, 1 (3): pp.195-200, Reprinted in J. S. Bell, *Speakable and unspeakable in quantum mechanics*, 1993, Cambridge University Press, pp.14-21.

2. Gröblacher, Simon, et.al. (2007). "An experimental test of non-local realism". *Nature*. 446 (7138): 871–5. arXiv:0704.2529.