[Physics FAQ] - [Copyright]

Updated May 1996 by PEG (thanks to Colin Naturman).
Updated August 1993 by SIC.
Original by John Blanton.

Does Bell's Inequality Principle rule out local theories of quantum mechanics?

In 1935 Albert Einstein and two colleagues, Boris Podolsky and Nathan Rosen (EPR) developed a thought experiment to demonstrate what they felt was a lack of completeness in quantum mechanics. This so-called "EPR paradox" has led to much subsequent, and still on-going, research. This article is an introduction to EPR, Bell's inequality, and the real experiments that have attempted to address the interesting issues raised by this discussion.

One of the principal features of quantum mechanics is that not all the classical physical observables of a system can be simultaneously known with unlimited precision, even in principle. Instead, there may be several sets of observables which give qualitatively different, but nonetheless complete (maximal possible) descriptions of a quantum mechanical system. These sets are sets of "good quantum numbers," and are also known as "maximal sets of commuting observables." Observables from different sets are "noncommuting observables".

A well known example is position and momentum. You can put a subatomic particle into a state of well-defined momentum, but then you cannot know where it is. It's not just a matter of your inability to measure, but rather, an intrinsic property of the particle. Conversely, you can put a particle in a definite position, but then its momentum is completely ill-defined. You can also create states of intermediate knowledge of both observables: if you confine the particle to some arbitrarily large region of space, you can define the momentum more and more precisely. But you can never know both, exactly, at the same time.

(Actually, some of the above statements are not quite correct. For one thing, observables that don't commute can still have mutual eigenstates. Such subtleties are very important to those who examine the derivation of Bell's inequality in great detail in order to find hidden assumptions. For the purposes of this short article we'll overlook these finer points.)

Position and momentum are continuous observables. But the same situation can arise for discrete observables such as spin. The quantum mechanical spin of a particle along each of the three space axes is a set of mutually noncommuting observables. You can only know the spin along one axis at a time. A proton with spin "up" along the x-axis has undefined spin along the y and z axes. You cannot simultaneously measure the x and y spin projections of a proton. EPR sought to demonstrate that this phenomenon could be exploited to construct an experiment that would demonstrate a paradox which they believed was inherent in the quantum-mechanical description of the world.

They imagined two physical systems that are allowed to interact initially so that they will subsequently be defined by a single Schrodinger wave equation. (For simplicity, imagine a simple physical realization of this idea - a neutral pion at rest in your lab, which decays into a pair of back-to-back photons. The pair of photons is described by a single two-particle wave function.) Once separated, the two systems (read: photons) are still described by the same wave equation, and a measurement of one observable of the first system will determine the measurement of the corresponding observable of the second system. (Example: the neutral pion is a scalar particle - it has zero angular momentum. So the two photons must speed off in opposite directions with opposite spin. If photon 1 is found to have spin up along the x-axis, then photon 2 *must* have spin down along the x-axis, since the total angular momentum of the final-state, two-photon, system must be the same as the angular momentum of the initial state, a single neutral pion. You know the spin of photon 2 even without measuring it.) Likewise, the measurement of another observable of the first system will determine the measurement of the corresponding observable of the second system, even though the systems are no longer physically linked in the traditional sense of local coupling.

However, QM prohibits the simultaneous knowledge of more than one mutually noncommuting observable of either system. The paradox of EPR is the following contradiction: for our coupled systems, we can measure observable A of system I (for example, photon 1 has spin up along the x-axis; photon 2 must therefore have x-spin down), and observable B of system II (for example, photon 2 has spin down along the y-axis; therefore the y-spin of photon 1 must be up), thereby revealing both observables for both systems, contrary to QM.

QM dictates that this should be impossible, creating the paradoxical implication that measuring one system should "poison" any measurement of the other system, no matter what the distance between them. (In one commonly studied interpretation, the mechanism by which this proceeds is 'instantaneous collapse of the wave function'. But the rules of QM do not require this interpretation, and several other perfectly valid interpretations exist.) The second system would instantaneously be put into a state of well-defined observable A, and, consequently, ill-defined observable B, spoiling the measurement. Yet, one could imagine the two measurements were so far apart in space that special relativity would prohibit any influence of one measurement over the other. For example, after the neutral-pion decay, we can wait until the two photons are a light year apart, and then "simultaneously" measure the x-spin of photon 1 and the y-spin of photon 2. QM suggests that if say the measurement of the photon 1 x-spin happens first, then this measurement must instantaneously force photon 2 into a state of ill-defined y-spin, even though it is light years away from photon 1.

How do we reconcile the fact that photon 2 "knows" that the x-spin of photon 1 has been measured, even though they are separated by light years of space and far too little time has passed for information to have travelled to it according to the rules of special relativity? There are basically two choices. You can accept the postulates of QM as a fact of life, in spite of its seemingly uncomfortable coexistence with special relativity, or you can postulate that QM is not complete, that there *was* more information available for the description of the two-particle system at the time it was created, carried away by both photons, and that you just didn't know it because QM does not properly account for it.

So, EPR postulated that the existence of hidden variables, some so-far unknown properties, of the systems should account for the discrepancy. Their claim was that QM theory is incomplete: it does not completely describe the physical reality. System II knows all about System I long before the scientist measures any of the observables, thereby supposedly consigning the other noncommuting observables to obscurity. Furthermore, they claimed that the hidden variables would be *local*. No instantaneous action-at-a-distance is necessary in this picture, which postulates that each System has more variables than are accounted by QM. Niels Bohr, one of the founders of QM, held the opposite view and defended a strict interpretation, the Copenhagen Interpretation, of QM.

In 1964 John Bell proposed a mechanism to test for the existence of these hidden variables, and he developed his inequality principle as the basis for such a test. He showed that if his inequality was ever not satisfied, then it was not possible to have a local theory that accounted for the spin experiment.

Using the example of two photons configured in the singlet state, consider this: after separation, each photon will have spin values for each of the three axes of space, and each spin can have one of two values; call them up and down. Call the axes x, y and z and call the spin in the x axis x+ if it is up in that axis, otherwise call it x-. Use similar definitions for the other two axes.

Now perform the experiment. Measure the spin in one axis of one particle and the spin in another axis of the other photon. If EPR were correct, each photon will simultaneously have properties for spin in each of axes x, y and z.

Next, look at the statistics. Perform the measurements with a number of sets of photons. Use the symbol N (x+, y-) to designate the words "the number of photons with x+ and y-". Similarly for N(x+, y+), N(y-, z+), etc. Also use the designation N(x+, y-, z+) to mean "the number of photons with x+, y- and z+", and so on. It's easy to demonstrate that for a set of photons

```
(1)   N(x+, y-) = N(x+, y-, z+) + N(x+, y-, z-)
```

because all of the (x+, y-, z+) and all of the (x+, y-, z-) photons are included in the designation (x+, y-), and nothing else is included in N(x+, y-). You can make this claim if these measurements are connected to some real properties of the photons.

Let n[x+, y+] be the designation for "the number of measurements of pairs of photons in which the first photon measured x+, and the second photon measured y+." Use a similar designation for the other possible results. This is necessary because this is all that it is possible to measure. You can't measure both x and y for the same photon. Bell demonstrated that in an actual experiment, if (1) is true (indicating real properties), then the following must be true:

```
(2)    n[x+, y+] <= n[x+, z+] + n[y-, z-].
```

Additional inequality relations can be written by just making the appropriate permutations of the letters x, y and z and the two signs. This is Bell's Inequality Principle, and it is proved to be true if there are real (perhaps hidden) variables to account for the measurements.

At the time Bell's result first became known, the experimental record was reviewed to see if any known results provided evidence against locality. None did. Thus an effort began to develop tests of Bell's Inequality. A series of experiments was conducted by Aspect ending with one in which polarizer angles were changed while the photons were `in flight'. This was widely regarded at the time as being a reasonably conclusive experiment confirming the predictions of QM.

Three years later Franson published a paper showing that the timing constraints in this experiment were not adequate to confirm that locality was violated. Aspect measured the time delays between detections of photon pairs. The critical time delay is that between when a polarizer angle is changed and when this affects the statistics of detecting photon pairs. Aspect estimated this time based on the speed of a photon and the distance between the polarizers and the detectors. Quantum mechanics does not allow making assumptions about *where* a particle is between detections. We cannot know *when* a particle traverses a polarizer unless we detect the particle *at* the polarizer.

Experimental tests of Bell's Inequality are ongoing but none has yet fully addressed the issue raised by Franson. In addition there is an issue of detector efficiency. By postulating new laws of physics one can get the expected correlations without any nonlocal effects unless the detectors are close to 90% efficient. The importance of these issues is a matter of judgment.

The subject is alive theoretically as well. Eberhard and later Fine uncovered further subtleties in Bell's argument. Some physicists argue that there are assumptions in derivations of Bell's Inequality and that it may be possible to construct a local theory that does not respect those assumptions. The subject is not yet closed, and may yet provide more interesting insights into the subtleties of quantum mechanics.

References

1. A. Einstein, B. Podolsky, N. Rosen: "Can quantum-mechanical description of physical reality be considered complete?" Physical Review 41, 777 (15 May 1935). (The original EPR paper)
2. D. Bohm: Quantum Theory, Dover, New York (1957). (Bohm discusses some of his ideas concerning hidden variables.)
3. N. Herbert: Quantum Reality, Doubleday. (A very good popular treatment of EPR and related issues)
4. M. Gardner: Science - Good, Bad and Bogus, Prometheus Books. (Martin Gardner gives a skeptics view of the fringe science associated with EPR.)
5. J. Gribbin: In Search of Schrodinger's Cat, Bantam Books. (A popular treatment of EPR and the paradox of "Schrodinger's cat" that results from the Copenhagen interpretation)
6. N. Bohr: "Can quantum-mechanical description of physical reality be considered complete?" Physical Review 48, 696 (15 Oct 1935). (Niels Bohr's response to EPR)
7. J. Bell: "On the Einstein Podolsky Rosen paradox" Physics 1 #3, 195 (1964).
8. J. Bell: "On the problem of hidden variables in quantum mechanics" Reviews of Modern Physics 38 #3, 447 (July 1966).
9. D. Bohm, J. Bub: "A proposed solution of the measurement problem in quantum mechanics by a hidden variable theory" Reviews of Modern Physics 38 #3, 453 (July 1966).
10. B. DeWitt: "Quantum mechanics and reality" Physics Today p. 30 (Sept 1970).
11. J. Clauser, A. Shimony: "Bell's theorem: experimental tests and implications" Rep. Prog. Phys. 41, 1881 (1978).
12. A. Aspect, Dalibard, Roger: "Experimental test of Bell's inequalities using time- varying analyzers" Physical Review Letters 49 #25, 1804 (20 Dec 1982).
13. A. Aspect, P. Grangier, G. Roger: "Experimental realization of Einstein-Podolsky-Rosen-Bohm gedanken experiment; a new violation of Bell's inequalities" Physical Review Letters 49 #2, 91 (12 July 1982).
14. A. Robinson: "Loophole closed in quantum mechanics test" Science 219, 40 (7 Jan 1983).
15. B. d'Espagnat: "The quantum theory and reality" Scientific American 241 #5 (November 1979).
16. "Bell's theorem and delayed determinism", Franson, Physical Review D, pages 2529-2532, Vol. 31, No. 10, May 1985.
17. "Bell's theorem without hidden variables", P. H. Eberhard, Il Nuovo Cimento, 38 B 1, pages 75-80, (1977).
18. "Bell's theorem and the different concepts of locality", P. H. Eberhard, Il Nuovo Cimento 46 B, pages 392-419, (1978).