Bell's Theorem

April 18, 2013

Bell's theorem states: *No local hidden variables theory can ever reproduce the predictions of quantum mechanics.*

By a "local hidden variables" theory, what we mean is that there is come classical property of the system, for example physical electron spin, that evolves in a classical way but is, as yet, unknown. For example, if the spin of the electron definitely pointed in some direction at the initial time, and evolved continuously but randomly, Bell's theorem shows that it could not reproduce the probability distribution expected within quantum mechanics.

First, let's look at a general probability argument showing that there is an essential difference between classical and quantum probabilities. Then we turn to a discussion of the Einstein-Podolsky-Rosen problem, followed by an example of Bell's theorem where we can clearly see the the theorem at work.

1 Basic probability argument

Consider the conditional probability, $P_{A \to B}$, of event B given that A has occured, and the conditional probability, $P_{B \to C}$, that C occurs, given B. Classically, the probability, $P_{A \to C}$, is then the product of these summed over all possible intermetiate states B:

$$P_{A \to C} = \sum_B P_{A \to B} P_{B \to C}$$

In quantum physics, this is computed differently, using probability amplitudes. Now, let $\psi_{A \to B}$ be the probability amplitude for event B to occur given that A has occured, and let $\psi_{B \to C}$ be the probability amplitude C to occur, given B. Then the probabilitiy amplitude for A leading to C is again the sum over all intermediate states of the product,

$$\psi_{A \to C} = \sum_B \psi_{A \to B} \psi_{B \to C}$$

This means that the *quantum probability* for C given A is the absolute *square*,

$$
\begin{aligned}
P_{A \to C}^{quant} &= |\psi_{A \to C}|^2 \\
&= \left| \sum_B \psi_{A \to B} \psi_{B \to C} \right|^2 \\
&= \sum_B \psi_{A \to B} \psi_{B \to C} \sum_{B'} \psi_{A \to B'}^* \psi_{B' \to C}^* \\
&= \sum_B \sum_{B'} \psi_{A \to B} \psi_{A \to B'}^* \psi_{B \to C} \psi_{B' \to C}^* \\
&= \sum_B \psi_{A \to B} \psi_{A \to B}^* \psi_{B \to C} \psi_{B \to C}^* + \sum_B \sum_{B' \neq B} \psi_{A \to B} \psi_{A \to B'}^* \psi_{B \to C} \psi_{B' \to C}^*
\end{aligned}
$$

$$= \sum_B P_{A \to B} P_{B \to C} + \sum_B \sum_{B' \neq B} \psi_{A \to B} \psi^*_{A \to B'} \psi_{B \to C} \psi^*_{B' \to C}$$

$$= P^{cl}_{A \to C} + \sum_B \sum_{B' \neq B} \psi_{A \to B} \psi^*_{A \to B'} \psi_{B \to C} \psi^*_{B' \to C}$$

where the first sum on the right consists of the diagonal terms and the second is the sum over all terms with B not equal to B'.

Both of the probabilities here lie between 0 and 1 for any normalized wave function. However, the conditional probability for the quantum transition from A to C is different from the classical probability. This means that in some physical systems, there may be a greater chance of certain events than would be predicted classically. We now explore a simple, yet disturbing, example.

2 Einstein-Podolsky-Rosen (EPR)

Einstein suggested that there must be an underlying reality to our experience, in the sense that, if a particle travels from a to b, and it has a property p at a, and if no process alters that property during its travel time, then it will be measured to have property p at b. Einstein, Podolsky and Rosen explored this idea and subsequent authors have honed the arguments. In the simplest version, we consider a particle pair, created with equal but opposite spin, and use the conservation of angular momentum to consider what happens. If Einstein's view is right, the particles have some property (angular momentum)

Consider the decay of a single, charge $q = -1$, scalar particle (i.e., spin 0, angular momentum state $|0,0\rangle$) into a pair of electrons. If the pair is emitted with no orbital angular momentum, then the final angular momentum is described by the combination of the two spins of the emitted electrons, and from our considerations of addition of angular momentum, we must have

$$|0,0\rangle = \frac{1}{\sqrt{2}} \left(\left|\frac{1}{2}, -\frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, \frac{1}{2}\right\rangle_2 - \left|\frac{1}{2}, \frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_2 \right)$$

Let the emitted particles fly apart for a time, after which observer A measures electron 1, then moments later observer B measures electron 2. Suppose A measures the component of electron 1 spin in the z-direction. Then A has a $50-50$ chance of getting spin up or spin down. However, if B subsequently measures the spin of the second electron, B necessarily gets the opposite result.

Mathematically, if the first measurement by A shows that electron 1 has spin up, then the resulting normalized state is

$$- \left|\frac{1}{2}, \frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_2$$

and the probability for the B to also measure spin up is zero,

$$
\begin{aligned}
P_{1,up \to 2,up} &= \left| \left(\left\langle\frac{1}{2}, +\frac{1}{2}\right|_1 \left\langle\frac{1}{2}, +\frac{1}{2}\right|_2 \right) \left(-\left|\frac{1}{2}, \frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_2 \right) \right|^2 \\
&= \left| \left(-\left\langle\frac{1}{2}, +\frac{1}{2}\right|_1 \left|\frac{1}{2}, \frac{1}{2}\right\rangle_1 \left\langle\frac{1}{2}, +\frac{1}{2}\right|_2 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_2 \right) \right|^2 \\
&= \frac{1}{\sqrt{2}} \left| (-1 \cdot 0) \right|^2 \\
&= 0
\end{aligned}
$$

The probability that B measures the the second electron to have spin down is one,

$$P_{1,up \to 2,down} = \left| \left(\left\langle\frac{1}{2}, +\frac{1}{2}\right|_1 \left\langle\frac{1}{2}, -\frac{1}{2}\right|_2 \right) \left(-\left|\frac{1}{2}, \frac{1}{2}\right\rangle_1 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle_2 \right) \right|^2$$

$$= \left| -\left\langle \frac{1}{2}, +\frac{1}{2} \right|_1 \left| \frac{1}{2}, \frac{1}{2} \right\rangle_1 \left\langle \frac{1}{2}, -\frac{1}{2} \right|_2 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 \right|^2$$

$$= |(-1 \cdot 1)|^2$$

$$= 1$$

So far, there is no conceptual problem. However, suppose just before making the measurement, A decides to measure the x-component of spin instead and finds it to be up. $\left| x, +\frac{1}{2} \right\rangle_1$. Then the normalized state of the system after A's measurement is

$$
\begin{aligned}
|After\ A\rangle &= \frac{1}{\sqrt{2}} \left(\left| x, +\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, \frac{1}{2} \right\rangle_2 - \left| x, +\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 \right) \\
&= \frac{1}{2} \left(\left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle_1 + \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_1 \right) \left| \frac{1}{2}, \frac{1}{2} \right\rangle_2 - \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle_1 + \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_1 \right) \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 \right) \\
&= \frac{1}{2} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, \frac{1}{2} \right\rangle_2 + \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, \frac{1}{2} \right\rangle_2 - \left| \frac{1}{2}, \frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 - \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 \right)
\end{aligned}
$$

This time, if B measures the z-component of spin, the probabilities are

$$
\begin{aligned}
P_{1,x\ up \to 2,z\ up} &= \left| \left\langle x, +\frac{1}{2} \right|_1 \left\langle \frac{1}{2}, +\frac{1}{2} \right|_2 |After\ A\rangle \right|^2 \\
&= \left| \left\langle x, +\frac{1}{2} \right|_1 \left\langle \frac{1}{2}, +\frac{1}{2} \right|_2 \frac{1}{\sqrt{2}} \left(\left| x, +\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, \frac{1}{2} \right\rangle_2 - \left| x, +\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 \right) \right|^2 \\
&= \left| \frac{1}{\sqrt{2}} \left(\left\langle \frac{1}{2}, +\frac{1}{2} \right|_2 \left| \frac{1}{2}, \frac{1}{2} \right\rangle_2 - \left\langle \frac{1}{2}, +\frac{1}{2} \right|_2 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 \right) \right|^2 \\
&= \frac{1}{2}
\end{aligned}
$$

and

$$
\begin{aligned}
P_{1,x\ up \to 2,z\ down} &= \left| \left| \left\langle x, +\frac{1}{2} \right|_1 \left\langle \frac{1}{2}, -\frac{1}{2} \right|_2 |After\ A\rangle \right|^2 \right|^2 \\
&= \left| \left\langle x, +\frac{1}{2} \right|_1 \left\langle \frac{1}{2}, -\frac{1}{2} \right|_2 \frac{1}{\sqrt{2}} \left(\left| x, +\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, \frac{1}{2} \right\rangle_2 - \left| x, +\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 \right) \right|^2 \\
&= \frac{1}{2} \left| \left\langle x, +\frac{1}{2} \right|_1 \left\langle \frac{1}{2}, -\frac{1}{2} \right|_2 \left(-\left| x, +\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 \right) \right|^2 \\
&= \frac{1}{2}
\end{aligned}
$$

This is truly strange because the probability for B to measure a given spin depends on which component A decides to measure. Since this measurement can take place an instant before the electron arrives at A's location, there is no time for a signal to propagate from A to B to change the second electron's state. This is not something that can happen in a classical system with finite propagation speed.

Einstein, Podolsky and Rosen argue that since no message could have been conveyed from A to B in the time between A deciding what to measure and B making a measurement, that the property A measures must have existed throughout the experiment. They went on to argue that the result means that quantum mechanics is incomplete and that there must be some additional "hidden variables" which characterize that property.

To make this last statement clearer, suppose there is some classical, but unknown, physical property of the electron that determines the outcomes of both A's and B's measurements – a hidden variable. We

may even suppose that it is statistical in origin, so that (1) the hidden variable takes its various allowed values with certain probabilities, and (2) the probability of the various possible outcomes of A's and B's measurements depends on which value the hidden variable takes on. Bell's theorem proves an inequality that applies to the distribution of outcomes associated with any such hidden variables picture. We now examine this inequality, and some experiments that test it.

Einstein wrote to Bohr (quoted in Mermin, Physics Today, April 1985),

> That which really exists in B should . . . not depend on what kind of measurement is carried out in part of space A; it should also be independent of whether or not any measurement at all is carried out in space A. If one adheres to this program, one can hardly consider the quantum-theoretical description as a complete representation of the physically real. If one tries to do so in spite of this, one has to assume that the physically real in B suffers a sudden change as a result of a measurement in A. My instinct for physics bristles at this.

3 Bell's inequality

After substantial googling, one learns a lesson that is worth remembering: the clearest presentation of an idea is often found in the original work. In the present discussion, we turn to Bell's original 1964 article, originally published in Phycics, 1, 195-200 (1964), and follow the argument there closely.

Consider the EPR experiment with a Stern-Gerlach device at each of A and B, measuring the spin in directions $\vec{a}$ and $\vec{b}$, respectively. Furthermore, suppose there are one or more hidden variables. Bell writes:

> Let this more complete specification be effected by means of parameters λ. It is a matter of indifference in the following whether λ denotes a single variable or a set, or even a set of functions, and wehther the variables are discrete or continuous. However, we write as if λ were a single continuous parameter.

Denote by $A(\vec{a}, \lambda) = \pm 1$ the result spin up or spin down when the direction $\vec{a}$ is measured at A and by $B(\vec{b}, \lambda) = \pm 1$ the result spin up or spin down when the direction $\vec{b}$ is measured at B. As described by Bell,

> The vital assumption . . . is that the result B for particle 2 does not depend on the setting $\vec{a}$, of the magnet for particle 1, nor A on $\vec{b}$.

Now let $\rho(\lambda)$ be the probability distribution for λ, and write the expectation value of the product of the outcomes,

$$P\left(\vec{a}, \vec{b}\right) = \int d\lambda \rho(\lambda) A(\vec{a}, \lambda) B\left(\vec{b}, \lambda\right)$$

We know that $\rho(\lambda)$ is normalized,

$$\int d\lambda \rho(\lambda) = 1$$

and since $A(\vec{a}, \lambda)$ and $B\left(\vec{b}, \lambda\right)$ take only values± 1, the probability cannot be less than -1. The only way the value -1 can be achieved when A and B both measure the same direciton $\vec{a}$ is if

$$A(\vec{a}, \lambda) = -B(\vec{a}, \lambda)$$

Therefore, if perfect anticorrelation is possible at all, the result of measurement is independent of whether we are at A or at B. Continuing with this equality, we rewrite the probability as

$$
\begin{aligned}
P\left(\vec{a}, \vec{b}\right) &= \int d\lambda \rho(\lambda) A(\vec{a}, \lambda) B\left(\vec{b}, \lambda\right) \\
&= -\int d\lambda \rho(\lambda) A(\vec{a}, \lambda) A\left(\vec{b}, \lambda\right)
\end{aligned}
$$

Let $\vec{c}$ be a third orientation of the Stern-Gerlach device. Then

$$
\begin{aligned}
P\left(\vec{a},\vec{b}\right) - P\left(\vec{a},\vec{c}\right) &= -\int d\lambda\rho\left(\lambda\right) A\left(\vec{a},\lambda\right) A\left(\vec{b},\lambda\right) + \int d\lambda\rho\left(\lambda\right) A\left(\vec{a},\lambda\right) A\left(\vec{c},\lambda\right) \\
&= \int d\lambda\rho\left(\lambda\right) A\left(\vec{a},\lambda\right) A^2\left(\vec{b},\lambda\right) A\left(\vec{c},\lambda\right) - \int d\lambda\rho\left(\lambda\right) A\left(\vec{a},\lambda\right) A\left(\vec{b},\lambda\right) \\
&= \int d\lambda\rho\left(\lambda\right) A\left(\vec{a},\lambda\right) A\left(\vec{b},\lambda\right) \left(A\left(\vec{b},\lambda\right) A\left(\vec{c},\lambda\right) - 1 \right)
\end{aligned}
$$

Since

$$
A\left(\vec{a},\lambda\right) A\left(\vec{b},\lambda\right) \leq 1
$$

the integrand on the right is non-positive. Dropping this product under the integral, then reversing the sign on the right and taking the absolute value on the left,

$$
\begin{aligned}
\left| P\left(\vec{a},\vec{b}\right) - P\left(\vec{a},\vec{c}\right) \right| &\leq \int d\lambda\rho\left(\lambda\right) \left(1 - A\left(\vec{b},\lambda\right) A\left(\vec{c},\lambda\right) \right) \\
&= \int d\lambda\rho\left(\lambda\right) - \int d\lambda\rho\left(\lambda\right) A\left(\vec{b},\lambda\right) A\left(\vec{c},\lambda\right) \\
&= 1 + P\left(\vec{b},\vec{c}\right)
\end{aligned}
$$

and we arrive at Bell's inequality,

$$
1 + P\left(\vec{b},\vec{c}\right) \geq \left| P\left(\vec{a},\vec{b}\right) - P\left(\vec{a},\vec{c}\right) \right|
$$

4 Violation of Bell's inequality

Now consider the predictions of quantum mechanics. We know that the normalized quantum expectation value for the two measurements is $\langle\alpha|\left(\boldsymbol{\sigma}_B\cdot\mathbf{b}\right)\left(\boldsymbol{\sigma}_A\cdot\mathbf{a}\right)|\alpha\rangle$ and since the particles are in the singlet state,

$$
\begin{aligned}
|\alpha\rangle &= \frac{1}{\sqrt{2}}\left(\left|\frac{1}{2},-\frac{1}{2}\right\rangle_A \left|\frac{1}{2},\frac{1}{2}\right\rangle_B - \left|\frac{1}{2},\frac{1}{2}\right\rangle_A \left|\frac{1}{2},-\frac{1}{2}\right\rangle_B \right) \\
&= \frac{1}{\sqrt{2}}\left(|-\rangle_A |+\rangle_B - |+\rangle_A |-\rangle_B \right) \\
&= \frac{1}{\sqrt{2}}\left(\chi_{A,-}\chi_{B,+} - \chi_{A,+}\chi_{B,-} \right) \\
\langle\alpha| &= \frac{1}{\sqrt{2}}\left(\chi_{A,-}^\dagger\chi_{B,+}^\dagger - \chi_{A,+}^\dagger\chi_{B,-}^\dagger \right)
\end{aligned}
$$

and with

$$
\begin{aligned}
\mathbf{a}\cdot\boldsymbol{\sigma}_A &= a_z |+\rangle_A \langle+|_A - a_z |-\rangle_A \langle-|_A \\
&\quad + \left(a_x - ia_y\right)|+\rangle_A \langle-|_A + \left(a_x + ia_y\right)|-\rangle_A \langle+|_A \\
&= \begin{pmatrix} a_z & a_x - ia_y \\ a_x + ia_y & -a_z \end{pmatrix}
\end{aligned}
$$

and similarly for $\mathbf{b}\cdot\boldsymbol{\sigma}_B$, we have

$$
\begin{aligned}
\langle\alpha|\left(\mathbf{b}\cdot\boldsymbol{\sigma}_B\right)\left(\mathbf{a}\cdot\boldsymbol{\sigma}_A\right)|\alpha\rangle &= \frac{1}{\sqrt{2}}\left(\chi_{A,-}^\dagger\chi_{B,+}^\dagger - \chi_{A,+}^\dagger\chi_{B,-}^\dagger \right)\left(\mathbf{b}\cdot\boldsymbol{\sigma}_B\right)\left(\mathbf{a}\cdot\boldsymbol{\sigma}_A\right)\frac{1}{\sqrt{2}}\left(\chi_{A,-}\chi_{B,+} - \chi_{A,+}\chi_{B,-} \right) \\
&= \frac{1}{2}\left(\chi_{B,+}^\dagger\left(\mathbf{b}\cdot\boldsymbol{\sigma}_B\right)\chi_{B,+}\chi_{A,-}^\dagger\left(\mathbf{a}\cdot\boldsymbol{\sigma}_A\right)\chi_{A,-} - \chi_{B,-}^\dagger\left(\mathbf{b}\cdot\boldsymbol{\sigma}_B\right)\chi_{B,+}\chi_{A,+}^\dagger\left(\mathbf{a}\cdot\boldsymbol{\sigma}_A\right)\chi_{A,-} \right)
\end{aligned}
$$

$$-\frac{1}{2}\left(\chi_{B,+}^{\dagger}\left(\mathbf{b}\cdot\boldsymbol{\sigma}_B\right)\chi_{B,-}\chi_{A,-}^{\dagger}\left(\mathbf{a}\cdot\boldsymbol{\sigma}_A\right)\chi_{A,+}-\chi_{B,-}^{\dagger}\left(\mathbf{b}\cdot\boldsymbol{\sigma}_B\right)\chi_{B,-}\chi_{A,+}^{\dagger}\left(\mathbf{a}\cdot\boldsymbol{\sigma}_A\right)\chi_{A,+}\right)$$

$$=\frac{1}{2}\left(\left(b_z\left(-a_z\right)-\left(b_x+ib_y\right)\left(a_x-ia_y\right)\right)-\left(\left(b_x-ib_y\right)\left(a_x+ia_y\right)-\left(-b_z\right)a_z\right)\right)$$

$$=\frac{1}{2}\left(-a_zb_z-a_xb_x+ia_yb_x-ib_ya_x-b_ya_y-a_xb_x-ia_yb_x+ib_ya_x-b_ya_y-b_za_z\right)$$

$$=\frac{1}{2}\left(-a_zb_z-a_xb_x-b_ya_y-a_xb_x-b_ya_y-b_za_z\right)$$

$$=-\mathbf{a}\cdot\mathbf{b}$$

Now, again following Bell, let the directions $\mathbf{a}$,$\mathbf{b}$ and $\mathbf{c}$ be at angles $0,\theta$ and 2θ, respectively, with $\theta\ll 1$. Then the correlation probabilities are

$$\begin{aligned}P\left(\mathbf{a},\mathbf{b}\right)&=-\cos\theta\\&\approx-1+\frac{1}{2}\theta^2\\P\left(\mathbf{b},\mathbf{c}\right)&=-\cos\theta\\&\approx-1+\frac{1}{2}\theta^2\\P\left(\mathbf{a},\mathbf{c}\right)&=-\cos 2\theta\\&\approx-1+2\theta^2\end{aligned}$$

and therefore, substituting,

$$\begin{aligned}1+P\left(\vec{b},\vec{c}\right)&\geq\left|P\left(\vec{a},\vec{b}\right)-P\left(\vec{a},\vec{c}\right)\right|\\1-1+\frac{1}{2}\theta^2&\geq\left|-1+\frac{1}{2}\theta^2-\left(-1+2\theta^2\right)\right|\\\frac{1}{2}\theta^2&\geq\frac{3}{2}\theta^2\end{aligned}$$

giving a clear contradiction. Bell goes on to show that the quantum violation of the Bell inequality is not always infinitesimal. It is sufficient, however to consider the angles $0,\frac{\pi}{6}$ and $\frac{\pi}{3}$, for then

$$\begin{aligned}P\left(\mathbf{a},\mathbf{b}\right)&=-\cos\frac{\pi}{6}\\&=-\frac{\sqrt{3}}{2}\\P\left(\mathbf{b},\mathbf{c}\right)&=-\frac{\sqrt{3}}{2}\\P\left(\mathbf{a},\mathbf{c}\right)&=-\cos\frac{\pi}{3}\\&-\frac{1}{2}\end{aligned}$$

from which we find the finite violation of the inequality

$$\begin{aligned}1+P\left(\vec{b},\vec{c}\right)&\geq^?\left|P\left(\vec{a},\vec{b}\right)-P\left(\vec{a},\vec{c}\right)\right|\\1-\frac{\sqrt{3}}{2}&\geq^?\left|-\frac{1}{2}+\frac{\sqrt{3}}{2}\right|\\\frac{2-\sqrt{3}}{2}&\ngeq\frac{\sqrt{3}-1}{2}\\.134&\ngeq.366\end{aligned}$$

5 An example

In his Physics Today article, D. Mermin presents a non-quantum example of the violation of the Bell inequality. The set-up is simple, and corresponds the bare essentials of what is seen quantum mechanically. There are two remote observers, A and B, and each can choose one of three things to measure, $1, 2, 3$. Whatever they choose, there are two possible outcomes, R, G. Therefore, a complete run consists of a set such as $12RG$, where the first number, 1 in this example, is the setting observer A chooses, the second number, 2, is what B chooses, then A gets the result R and B gets the result G. The experiment is repeated many times, each run yielding a sequence of two numbers and two letters, $32RR, 33GG, 13GR, \ldots$.

Mermin makes it clear that we should picture some physical thing traveling from the central apparatus, out to A and B, arguing that their devices respond only when that apparatus is triggered, and they respond after a time proportional to their distance from the center. No signal occurs if the line-of-sight is blocked.

There are two features of the results that lead to the conflict with a simple picture of what is happening. Both of the following occur:

1. If both A and B choose the same setting, $11, 22$ or 33, they get the same result every time, either RR or GG.

2. If we ignore the settings of the devices and just look at the outcomes, they are completely random.

The first means that one of the following always occurs if the settings agree:

$$11GG \quad 11RR \quad 22GG$$
$$22RR \quad 33GG \quad 33RR$$

while the second condition means that RR, RG, GR, GG are equally probable.

Now, the hidden variables hypothesis says that because of the perfect correlation of point 1, the physical emanation from the center must carry a corresponding property that determines which outcome occurs for each setting, for example, $1R2R3G$, meaning if A or B sets 1 or 2 the outcome will be R, if one, the other or both set 3, they will get G. If this were not the case there is no guarantee (without "spukhafte Fernwirkungen") that if A and B both choose 2 they will both get R.

This "property" of the "particle" has to be specified for all three settings and every run, because A and B can choose their settings anytime before the particle reaches them. The production of the particle occurs before the choice is made.

Now compare probabilities. There are eight different outcome properties the particle may carry:

$$1G2G3G$$
$$1R2G3G$$
$$1G2R3G$$
$$1G2G3R$$
$$1R2R3G$$
$$1R2G3R$$
$$1G2R3R$$
$$1R2R3R$$

These insure that if the same property is measured by A and B, the same result occurs. However, suppose the property is, for example, $1R2G3G$. Then for the various settings A and B choose, we get the outcomes

$$11RR$$
$$12RG$$
$$13RG$$

7

$$22GG$$
$$23GG$$
$$24GG$$

so that for this property, RR and GG are more probable than RG and GR. The same is true for every set of properties, so the results cannot be random.

Let's do the full counting. There are six possible settings of the measurement devices, $11, 12, 13, 22, 23, 33$, and eight possible sets of properties. For either $1G2G3G$ or $1R2R3R$ we always get either GG or RR, for all six setttings, giving 12 measurement settings where we always get GG or RR. For any of the remaining six property sets, there are four ways to get RR or GG and only two to get RG or GR, so the tally is:

$$
\begin{aligned}
(RR \text{ or } GG) &= 6 + 6*4 + 6 = 36 \\
(RG \text{ or } GR) &= 0 + 6*2 + 0 = 12
\end{aligned}
$$

so that different colors occur only $1/4$ of the time. The results are inconsistent with the particles carrying property sets.

6   The difference between classical and quantum descriptions

The use of hidden variables is the principal difference between the classical and quantum descriptions of nature. In the classical description of a particle trajectory, for example, we assume that the particle follows a unique trajectory from its initial to its final points. This trajectory – the one given by Newton's second law – is often not actually measured, although in many cases the (approximate!) path may be directly visible to us. However, especially when we no longer deal with the macroscopic realm, we only *assume* such a unique path. This is an example of using a *hidden variable* – we assume the particle has definite physical properties (certain values of position, momentum, energy, angular momentum, etc.) throughout its trajectory, even though we do not *know* these, where by "knowing" we mean measuring. In Mermin's example, we assume the particle or emanation, has one of the sets of properties $1G2G3G - 1R2R3R$ listed above, and this conflicts with EPR results. In the same way, we are tempted by our classical thinking to assume that an electron has a definite spin vector, even though we have measured only, say, the z-component. This also turns out to be wrong.

In the quantum mechanical picture, the state provides a list that perfectly characterizes what we do and do not know about the system. Each item in the list is constrained by an measurements we have previously performed, or conservation laws which we know hold. At the same time, there are multiple items in the list. These multiple items reflect the fact that we do *not* know which of various things will be found. Thus, in the state for the electron-positron pair in an EPR experiment, where we concern ourselves only with spin variables, the state is the singlet,

$$
\begin{aligned}
|0,0\rangle &= \frac{1}{\sqrt{2}} \left(\chi_{e^-}^- \chi_{e^+}^+ - \chi_{e^-}^+ \chi_{e^+}^- \right) \\
&= \frac{1}{\sqrt{2}} \left(\left| \frac{1}{2}, -\frac{1}{2} \right\rangle_{e^-} \left| \frac{1}{2}, \frac{1}{2} \right\rangle_{e^+} - \left| \frac{1}{2}, \frac{1}{2} \right\rangle_{e^-} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_{e^+} \right)
\end{aligned}
$$

where $\chi_{e^-}^+ = \left| \frac{1}{2}, \frac{1}{2} \right\rangle_{e^-}$ is the state of a spin-up electron, and so on. Each term in this sum is constrained by conservation of momentum, $m_{e^-} + m_{e^+} = 0$, since we know this must hold. We then have a sum of all possible terms satisfying this one known constraint, $\chi_{e^-}^- \chi_{e^+}^+$ and $\chi_{e^-}^+ \chi_{e^+}^-$, but not the terms $\chi_{e^-}^+ \chi_{e^+}^+$ or $\chi_{e^-}^- \chi_{e^+}^-$. The *relative* phase, $e^{i\pi} = -1$, insures that the total angular momentum (j) is zero, while we know that the *overall* phase is arbitrary, i.e., we could equally well write

$$
|0,0\rangle = \frac{1}{\sqrt{2}} e^{i\varphi} \left(\chi_{e^-}^- \chi_{e^+}^+ - \chi_{e^-}^+ \chi_{e^+}^- \right)
$$

for any angle φ. The rules for making predictions from quantum mechanical states insure that this factor will not contribute to predictions.

The path integral approach to quantum mechanics or quantum field theory makes this clear as well. To compute the probability amplitude for a system to evolve from a state A to a state B, this formalism tells us to sum over all possible internal configurations, inserting appropriate phases for each possibility. This is written as

$$\int \mathcal{D}\left[x\left(t\right)\right] \exp\left(\frac{i}{\hbar} S\left[x\left(t\right)\right]\right)$$

where $S\left[x\left(t\right)\right]$ is the action functional for the system and $\int \mathcal{D}\left[x\left(t\right)\right]$ means that we sum or integrate over *all* paths $x\left(t\right)$ that satisfy the initial and final conditions. Again, we see that we enforce the measured conditions at the endpoints while summing over all of the unmeasured paths that we do not measure. Recall that the classical paths are those which make the variation of the action vanish, $\delta S = 0$, and it is this which makes the classical picture emerge from the quantum one. The phase in the exponential, having $\hbar$ in the denominator, will oscillate rapidly if S is changing. Near the stable points of the action, however, the phases add coherently, so the principal contribution to a path integral comes from paths very close to the classical one. Symmetries of the action, which lead to conservation laws by Noether's theorem, contribute to the result strongly as well, since S does not change when we perform a symmetry transformation.