Chapter 8
Bell's Theorem

The result which has come to be known as "Bell's Theorem" – but which Bell himself instead referred to as the "locality inequality theorem" [1] – first appeared in Bell's 1964 paper, "On the Einstein–Podolsky–Rosen paradox" [2]. Following Bell's own presentation, we begin here by recalling (from Chap. 4) the EPR argument, in the updated form introduced by Bohm in 1951.

8.1 EPRB Revisited

In Bohm's re-formulation, we consider a pair of spatially separated spin 1/2 particles in the spin "singlet" state

$$\psi_s = \frac{1}{\sqrt{2}} \left[\psi_{+z}^1 \psi_{-z}^2 - \psi_{-z}^1 \psi_{+z}^2 \right]. \tag{8.1}$$

With the state written in this form, it is apparent that, from a measurement of the z-component of particle 1's spin, we can immediately infer the z-component of particle 2's spin. The two particles' z-spins are perfectly anti-correlated: a $+1$ outcome on one side implies a -1 outcome on the other side, and vice versa. But according to the locality assumption, measuring the z-spin of particle 1 (say, nearby) should not disturb the physical state of the (say, distant) particle 2. And so, according to the reasoning introduced by EPR in 1935, the distant particle must *already possess* a definite z-spin value (which is then simply revealed when its z-spin is measured). The only alternative is that the distant particle's z-spin somehow comes into existence (crystallizing out of a prior fog, so to speak) as a result of our measurement on the nearby particle; but that would constitute a violation of local causality. The EPR claim is that the only way to avoid non-locality is to attribute a pre-determined z-spin value to the distant particle.

The singlet state ψ_s can also be re-written in terms of the single particle eigenstates for the x-component of the spin:

$$\psi_{\pm x}^1 = \frac{1}{\sqrt{2}} \left[\psi_{+z}^1 \pm \psi_{-z}^1 \right] \tag{8.2}$$

and identically for particle 2. Solving for $\psi_{\pm z}$ in terms of $\psi_{\pm x}$ and plugging in gives

$$\psi_s = \frac{1}{\sqrt{2}} \left[\psi_{+x}^1 \psi_{-x}^2 - \psi_{-x}^1 \psi_{+x}^2 \right]. \tag{8.3}$$

It is a special property of the singlet state that it takes exactly the same form, written in terms of the states of definite x-spin, as it does written in terms of the states of definite z-spin. This allows the same EPR reasoning to be run again, this time about x-spin values: the distant particle must *also* possess a definite pre-measurement value for its x-spin since we could determine this value, with certainty, and without disturbing the physical state of the distant particle, by measuring the x-spin of the nearby particle.

And note – contra Bohr's reply to EPR – that the fact that we could only measure *either* the nearby particle's x-spin *or* its z-spin, does not block the inference to the pre-determinateness of *both* properties of the distant particle. It is true that we could only *learn* about one or the other of the two properties on a given particle pair. But the assumption is that we have a genuinely free choice about which property (if either!) to measure on the nearby particle. *If* we choose to measure the z-spin of the nearby particle, we would learn the value of the z-spin of the distant particle (without disturbing its physical state in any way) and hence be in a position to infer its existence. Whether we in fact *do* so choose is immaterial: for the existence of the distant particle's z-spin to depend upon whether or not we choose to measure the nearby particle's z-spin, would be a violation of locality. So, again, the claim is that to avoid non-locality ("spooky action-at-a-distance") we must attribute definite, pre-measurement values to both the x- and z-spin of the distant particle.

And, finally, the argument is symmetric with respect to the two particles: if the *distant* particle must already possess definite values of both x-spin and z-spin (because we, here, could determine those values indirectly, without in any way influencing the physical state of the particle), so must the *nearby* particle (because someone over there could determine the values indirectly, without in any way influencing the physical state of the nearby particle).

So the upshot of the EPR-Bohm argument is that when a pair of particles is prepared in the spin state ψ_s, each particle in the pair must already possess a definite value for both x- and z-spin, and these values (for the two particles) must apparently be correlated to guarantee that any of the possible subsequent measurements agree with the quantum mechanical statistics. Locality thus requires a "hidden variable" theory of the sort summarized in the following table:

Pair type	Particle 1	Particle 2	Frequency
1	$\{+1, +1\}$	$\{-1, -1\}$	25%
2	$\{+1, -1\}$	$\{-1, +1\}$	25%
3	$\{-1, +1\}$	$\{+1, -1\}$	25%
4	$\{-1, -1\}$	$\{+1, +1\}$	25%

According to this model, the description of the particle pair in terms of the quantum state ψ_s is decidedly *incomplete*. There are facts about the x- and z-spins of both particles that are not contained in ψ_s! In particular, each pair of particles (prepared in the quantum state ψ_s) is in fact of one of the four sub-types described in the four rows of the table. Which of the four types a given pair of particles ends up is somehow just random, with equal 25% probabilities for each of the four types.

The exact nature of the four types is described in the "Particle 1" and "Particle 2" columns. For example, $\{+1, -1\}$ means that the particle in question is spin-up ("+1") along the x-direction and spin-down ("−1") along the z-direction. Notice that both the x-spin and z-spin values are perfectly anti-correlated within each pair type. For example, if Particle 1 is spin-up along z, then Particle 2 is spin-down along z. This ensures that, if the same property (x-spin or z-spin) is measured on both particles, the results will always be opposite (as predicted by QM). (This also explains why there are exactly four allowed "pair types". The other logical possibilities, for example $\{+1, +1\}$ for particle 1 and $\{+1, -1\}$ for particle 2, would violate the perfect correlation property for at least one possible set of measurements – here, if the x-spin is measured on both particles. Such pair types, if included, would need to be assigned frequencies of exactly zero in order for the model to reproduce the quantum predictions, so we might as well just exclude them entirely from the discussion.)

And notice also that the equal 25% frequencies for all four pair types are required to match the rest of the quantum predictions. For example, what happens if the x-spin of particle 1 is measured and the z-spin of particle 2 is measured? The quantum statistics can be read off from ψ_s re-written in this form:

$$\psi_s = \frac{1}{2} \left[\psi_{+x}^1 \psi_{-z}^2 + \psi_{-x}^1 \psi_{-z}^2 - \psi_{+x}^1 \psi_{+z}^2 + \psi_{-x}^1 \psi_{+z}^2 \right]. \tag{8.4}$$

The four possible joint outcomes (up-up, down-up, up-down, and down-down) thus have equal, 25%, probabilities. We can thus reproduce the complete slate of quantum mechanical statistical predictions (for any set of measurements along the x- and z-axes) by letting each of the four pair types in our hidden variable model occur with 25% frequency.

It is clear that this sort of "hidden variable" model, in which particles carry pre-determined values for possible spin measurements along the x- and z-directions, can reproduce the quantum predictions but without the non-locality associated with ordinary quantum theory's collapse postulate (combined with the claim that the wave function provides a complete state description). The matter effectively stood there for several decades, with EPR having shown that such a model is needed

to account for the quantum correlations in a local way, but with most physicists believing that Bohr had somehow refuted the EPR argument and therefore ignoring the issue entirely. As we will see, though, Bell moved the issue forward in the 1960s by asking: could this same local hidden variable model continue to reproduce the quantum mechanical predictions in a more general setting, where spin measurement along more and different axes are also allowed?

8.2   A Preliminary Bell Inequality

As we just showed, it is rather straightforward to reproduce the quantum mechanical predictions, for all possible spin measurements along the x- and z-directions, on a pair of entangled spin 1/2 particles, with a hidden variable model in which each particle's x-spin and z-spin are pre-determined. But let us broaden the discussion. Suppose that instead of restricting ourselves to measuring the spins of the particles along the x- and z-directions, we allow spin measurements in arbitrary (not necessarily orthogonal!) directions; and suppose that instead of considering only two possible directions (along which to measure the particles' spins) we allow the experimenter on each side to choose from among *three* possible axes. Let's call the three axes $\hat{a}$, $\hat{b}$, and $\hat{c}$. Notice that, since the singlet state ψ_s can be written

$$\psi_s = \frac{1}{\sqrt{2}} \left[\psi^1_{+a} \psi^2_{-a} - \psi^1_{-a} \psi^2_{+a} \right] \tag{8.5}$$

– and identically for $\hat{b}$ and $\hat{c}$ – it is clear that, according to quantum mechanics, the outcomes should be perfectly anti-correlated (either "up-down" or "down-up") whenever the two experimenters happen to measure their particles' spins along the same axis. Agreement with this aspect of the quantum mechanical predictions – and continuing to insist on locality – again requires a "hidden variable" theory of the sort we considered before, and requires in particular that the values of (for example) a-spin (and then identically for b-spin and c-spin) should be perfectly anti-correlated. A little contemplation reveals that there are now *eight* types of particle pairs which might be produced (with nonzero frequency) when particle pairs are prepared in the quantum state ψ_s. The types are described in the following table:

Pair type	Particle 1	Particle 2	Frequency
1	$\{+1, +1, +1\}$	$\{-1, -1, -1\}$	F_1
2	$\{+1, +1, -1\}$	$\{-1, -1, +1\}$	F_2
3	$\{+1, -1, +1\}$	$\{-1, +1, -1\}$	F_3
4	$\{-1, +1, +1\}$	$\{+1, -1, -1\}$	F_4
5	$\{+1, -1, -1\}$	$\{-1, +1, +1\}$	F_5
6	$\{-1, +1, -1\}$	$\{+1, -1, +1\}$	F_6
7	$\{-1, -1, +1\}$	$\{+1, +1, -1\}$	F_7
8	$\{-1, -1, -1\}$	$\{+1, +1, +1\}$	F_8

As before, the lists in the "Particle 1" and "Particle 2" columns tell us how a particle which is a member of the indicated pair type will respond to all three possible questions that might be put to it. So, for example, "$\{+1, -1, +1\}$" means that the particle will be found spin-up along $\hat{a}$ (if so measured!), spin-down along $\hat{b}$ (if so measured!), and spin-up along $\hat{c}$ (if so measured!).

Notice that the frequencies F_i of the 8 different pair types are left unspecified. The hope, of course, will be to pick values (as we were able to do in the previous section) so that the statistical predictions of the local hidden variable theory will agree with those of quantum mechanics.

But, as it turns out, this is impossible. The proof that it is impossible is, of course, "Bell's theorem", which involves showing that the predictions of the local hidden variable theory we are considering are constrained by a certain inequality ("Bell's inequality") that the quantum mechanical predictions do not respect. In short, there will be situations where – no matter exactly how the frequencies F_i are selected – the local hidden variable theory cannot reproduce the quantum mechanical statistics. We develop the proof in the remainder of this section in a way that is a little simpler than what Bell did in his original 1964 paper; in subsequent sections we return to consider Bell's own way of presenting things.

Notice first that we can express probabilities for specific possible outcomes in terms of the frequencies F_i that appear in the table. For example, suppose that particle 1 is measured along the $\hat{a}$ direction and particle 2 is measured along the $\hat{b}$ direction. What, for example, is the probability $P_{ab}(++)$ that both measurements have outcome "spin-up"? To answer, we can simply scan down the table on the previous page and look for the pair types for which this will occur. In particular, here, we need a "$+1$" as the first entry in the Particle 1 column (indicating that Particle 1 will be measured "spin-up" in the $\hat{a}$ direction) and a "$+1$" as the second entry in the Particle 2 column (indicating that Particle 2 will be measured "spin-up" in the $\hat{b}$ direction).

I find the appropriate entries in row 3 and row 5. This means that pairs of type 3 and type 5 will yield the outcomes "particle 1 is spin-up along $\hat{a}$" and "particle 2 is spin-up along $\hat{b}$". (Pairs of all the other types will yield at least one different outcome if the particles' spins are measured along $\hat{a}$ and $\hat{b}$ respectively.) And so the probability of seeing that particular outcome ("$++$") is just the probability that a given particle pair is of type 3 or type 5. That is:

$$P_{ab}(++) = F_3 + F_5. \tag{8.6}$$

Let's practice with a couple of other examples. What is the probability $P_{bc}(++)$ of seeing both particles "spin-up" given that particle 1 is measured along the $\hat{b}$ direction and particle 2 is measured along the $\hat{c}$ direction? I find:

$$P_{bc}(++) = F_2 + F_6. \tag{8.7}$$

And similarly

$$P_{ac}(++) = F_2 + F_5. \tag{8.8}$$

Make sure you see where these equations are coming from (and make sure you agree with what I wrote!).

Now, amazingly, we are already in a position to write down a (preliminary example of a) Bell inequality. Since the F_i's represent the frequencies with which pairs of different types are supposed to be produced when we create a particle pair in the singlet state, they must all be positive and they should add to 1. And so it must be the case that

$$F_2 + F_5 \leq F_3 + F_5 + F_2 + F_6 \tag{8.9}$$

since the right hand side is the same as the left hand side plus two additional terms which cannot be smaller than zero! But this means that, for a local hidden variable theory of the sort being considered here, it must be the case that

$$P_{ac}(++) \leq P_{ab}(++) + P_{bc}(++). \tag{8.10}$$

That is, no matter how we pick the frequencies F_i, a theory in which spin measurements simply reveal pre-existing values will have to make statistical predictions that obey Equation (or actually, Inequality) (8.10).

Now the incredible thing is that it is possible to choose directions $\hat{a}$, $\hat{b}$, and $\hat{c}$ such that this same inequality is *violated* by the quantum mechanical predictions. So let us recall in more detail how the relevant quantum predictions come about. In Chap. 2, we saw that the one-particle spin states $\psi_{\pm n}$ (corresponding to particles being definitely spin up or definitely spin down along the $\hat{n}$ axis, an angle θ down from the z-axis in the x-z-plane) were given by

$$\psi_{+n} = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} = \cos(\theta/2)\psi_{+z} + \sin(\theta/2)\psi_{-z} \tag{8.11}$$

and

$$\psi_{-n} = \begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2) \end{pmatrix} = \sin(\theta/2)\psi_{+z} - \cos(\theta/2)\psi_{-z}. \tag{8.12}$$

It is fairly straightforward to invert this relationship (solving for $\psi_{\pm z}$ in terms of $\psi_{\pm n}$). The result is that

$$\psi_{+z} = \cos(\theta/2)\psi_{+n} + \sin(\theta/2)\psi_{-n} \tag{8.13}$$

and

$$\psi_{-z} = \sin(\theta/2)\psi_{+n} - \cos(\theta/2)\psi_{-n}. \tag{8.14}$$

This allows us to rewrite the singlet state as follows:

$$\psi_s = \frac{1}{\sqrt{2}}\left[\sin\left(\frac{\theta}{2}\right)\psi_{+z}^1\psi_{+n}^2 - \cos\left(\frac{\theta}{2}\right)\psi_{+z}^1\psi_{-n}^2 - \cos\left(\frac{\theta}{2}\right)\psi_{-z}^1\psi_{+n}^2 - \sin\left(\frac{\theta}{2}\right)\psi_{-z}^1\psi_{-n}^2\right]. \tag{8.15}$$

From this, we can read off (as the square of the coefficient in front of the $\psi^1_{+z}\psi^2_{+n}$ term) the probability of seeing two "spin-up" outcomes when we measure particle 1 along the z-axis and particle 2 along a direction that is an angle θ away from the z-axis. This is simply:

$$P_{z,\theta}(++) = \frac{1}{2}\sin^2(\theta/2). \tag{8.16}$$

Since (as we have seen) the singlet state is symmetrical – and since what direction we choose to call the z-direction is ultimately arbitrary – this formula turns out to give the quantum mechanical probability for a "$++$" outcome whenever the two measurement directions have an angle θ between them (whether one of them is the "z-axis" or not). So we can use this general formula now to compute the quantum mechanical prediction for all three of the probabilities that appeared in Eq. (8.10), our baby Bell inequality.

Suppose we pick the three directions $\hat{a}$, $\hat{b}$, and $\hat{c}$ as follows: pick $\hat{b} = \hat{z}$, and then pick $\hat{a}$ and $\hat{c}$ to be tilted at angle θ away from the z-axis, in opposite directions. Then we have that, according to QM,

$$P^{QM}_{ab}(++) = \frac{1}{2}\sin^2(\theta/2) \tag{8.17}$$

and similarly

$$P^{QM}_{bc}(++) = \frac{1}{2}\sin^2(\theta/2). \tag{8.18}$$

What about $P^{QM}_{ac}(++)$? The angle between $\hat{a}$ and $\hat{c}$ is 2θ so the general formula gives

$$P^{QM}_{ac}(++) = \frac{1}{2}\sin^2(\theta). \tag{8.19}$$

Now it is a plain and simple mathematical fact that

$$\frac{1}{2}\sin^2(\theta) \leq \frac{1}{2}\sin^2(\theta/2) + \frac{1}{2}\sin^2(\theta/2) \tag{8.20}$$

is *false* for $\theta \leq \pi/2$. The biggest violation of the inequality occurs for $\theta = \pi/3 = 60°$. In that case we have

$$P^{QM}_{ab}(++) = \frac{1}{8}, \tag{8.21}$$

$$P^{QM}_{bc}(++) = \frac{1}{8}, \tag{8.22}$$

and

$$P^{QM}_{ac}(++) = \frac{3}{8}. \tag{8.23}$$

And, as plain as day, it is not the case that

$$\frac{3}{8} \leq \frac{1}{8} + \frac{1}{8}. \tag{8.24}$$

Bell's inequality is violated by the quantum mechanical predictions.

And so there is, in principle, a certain kind of experiment that we could do to test whether quantum mechanics is right, or the type of local hidden variable theory suggested by EPR is right. The two theories make empirically distinguishable predictions. What is the experiment, exactly? Well, we would need to produce a bunch of spin 1/2 particle pairs, have them fly off in opposite directions toward Stern–Gerlach measuring devices which could be oriented along one of the possible three directions ($\hat{a}$, $\hat{b}$, or $\hat{c}$). For reasons that we will discuss further in the next section, it should ideally be the case that the measurement direction on each side be chosen randomly and at the last possible second before the particles arrive.

Then we would simply keep track of how often, when the $\hat{a}$-spin of particle 1 was measured and the $\hat{b}$-spin of particle 2 was measured, both measurements yielded the "spin-up" outcome. That is, we would empirically measure $P_{ab}(++)$. And similarly for $P_{bc}(++)$ and $P_{ac}(++)$. And then at the end of the day we would compare these probabilities. If Bell's inequality

$$P_{ac}(++) \leq P_{ab}(++) + P_{bc}(++) \tag{8.25}$$

was respected by the experimental data, it would constitute a refutation of Quantum Mechanics and a vindication of the local hidden variables theory; whereas if the measured probabilities agreed with the QM predictions it would constitute an experimental refutation of the local hidden variable theory.

8.3 The Real Bell (and the CHSH) Inequality

Although the experiment just sketched would indeed be possible and gives one the flavor of how a Bell inequality can be experimentally tested, the inequality derived in the previous section is really just a kind of preliminary toy example of a Bell-type inequality. In this section we show how the real Bell inequality (that is, the actual inequality first derived by Bell in his 1964 paper), as well as the closely-related "CHSH Inequality," can be developed. These, as it turns out, suggest far superior types of experimental tests which we will discuss in the following section.

To motivate this discussion, perhaps it is worth thinking about why the experimental test proposed at the end of the last section is somehow less than ideal. Part of the answer is simply that it is very inefficient: if the directions (along which to measure the spins of the two particles) are selected randomly for each particle pair, then only about 1/3 of the time will we happen to make one of the three types of measurements (namely, ab, bc, or ac) that are relevant to Eq. (8.10). So 2/3 of the data – 2/3 of the particle pairs produced – are simply wasted.

You are probably thinking that one could eliminate the waste by just fixing the Stern–Gerlach device on the particle 1 side in the $\hat{a}$ direction, and similarly fixing the Stern–Gerlach device on the particle 2 side in the $\hat{b}$ direction, collecting data for (say) a million particle pairs, then switching the detectors to the bc orientations, collecting more data, and then finally switching to the ac orientations and collecting a last set of data. And that is true. You could do that (and some of the early experiments were along these lines). But (as hinted at previously) it is important that the orientations of the measuring devices be set randomly and at the last possible second before the particles arrive – ideally, so late that the measurement on particle 1 cannot be influenced (by any signal propagating at the speed of light or slower) by the orientation of particle 2's measuring device (or vice versa).

To understand this, imagine that both detectors are just fixed in place – say, in the ac orientation – for a run of many particle pairs. Then each particle pair could "know", already when it is created, and without any violation of local causality, that particle 1 will be measured along the $\hat{a}$ direction and particle 2 will be measured along the $\hat{c}$ direction. But one could imagine that, in such a circumstance, the particle source would be free to emit particles not only of the 8 types captured in our earlier table, but also "rogue" types in which, for example, particle 1 has properties $\{+1, -1, -1\}$ and particle 2 has properties $\{+1, -1, +1\}$. After all, if the particles "know" in advance that they will each be measured along particular, pre-determined directions, there is no reason the pre-existing spin components along all three directions should have to be perfectly correlated. Similarly, if the particle pairs "know", in advance of being emitted, the directions along which their spins will later be measured, it might be possible for the source to adjust the frequencies F_i in response: for example, perhaps when the detectors are in the ab configuration, F_1 is big and F_2 is small, whereas when the detectors are in the bc configuration, F_1 is small and F_2 is big.

It should be clear that in either of these scenarios ("rogue" particle types, or measurement-axis-dependent pair frequencies), the straightforward type of hidden variable model we've described is no longer required and our derivation of the inequality would not go through. Turning this around, then, it should make sense that, in order for the straightforward type of hidden variable model we've described to genuinely be required by locality, we must have in mind an experimental setup in which it is impossible for the particles to know in advance along which axes their spins will be measured. So if at the end of the day we want a clean experimental discrimination between Quantum Mechanics and the sort of local theory implied by the EPR-Bohm argument, the experimental test must implement strict "Einstein locality" conditions in which the device setting on each side is only determined while the particles are in flight, and, indeed, determined sufficiently late that information about it is locally inaccessible to the measurement on the other side.

We will discuss this point a bit more in the following sections; for now, suffice it to say that for a variety of technical and practical reasons, it would be nice to develop a Bell type inequality that doesn't focus so narrowly on one specific outcome for each of just three possible measurements, but instead embraces all possible measurements and outcomes in a more democratic way.

To begin to develop such an inequality, let us consider the "correlation coefficient" $C(\hat{n}_1, \hat{n}_2)$, defined as the expected value of the product of the two measurement outcomes when measurements are made along direction $\hat{n}_1$ and $\hat{n}_2$ on the two particles. Since the product of the outcomes is either $+1$ (if the outcomes are "$++$" or "$--$") or -1 (if the outcomes are "$+-$" or "$-+$"), the correlation coefficient is thus the probability for "$++$" plus the probability for "$--$" but then minus the probability for "$+-$" and minus the probability for "$-+$":

$$C(\hat{n}_1, \hat{n}_2) = P_{n_1 n_2}(++) + P_{n_1 n_2}(--) - P_{n_1 n_2}(+-) - P_{n_1 n_2}(-+). \qquad (8.26)$$

Using the same technique we used in the previous section, i.e., reading off the probabilities as the squares of the coefficients of the four terms in Eq. (8.15), it is easy to work out that the Quantum Mechanical prediction for the correlation coefficient is

$$C^{QM}(\hat{n}_1, \hat{n}_2) = \frac{1}{2}\sin^2\left(\frac{\theta}{2}\right) + \frac{1}{2}\sin^2\left(\frac{\theta}{2}\right) - \frac{1}{2}\cos^2\left(\frac{\theta}{2}\right) - \frac{1}{2}\cos^2\left(\frac{\theta}{2}\right) = -\cos(\theta)$$
$$(8.27)$$

where θ is the angle between $\hat{n}_1$ and $\hat{n}_2$. Note that, for $\theta = 0$, $C = -1$, meaning that the two outcomes are perfectly anti-correlated (always opposite). Whereas for $\theta = 90°$, $C = 0$, meaning that there is no correlation at all between the outcomes. This is all consistent with what we have seen before, namely, that if the two particles' spins are measured along the same direction, the individual outcomes are necessarily opposite, whereas if they are measured along orthogonal axes, a "$+1$" outcome on one side is equally likely to be accompanied by a "$+1$" or a "-1" on the other side, and so on.

What about the local hidden variable theory described in the previous section? Writing

$$C(\hat{n}_1, \hat{n}_2) = P_{n_1 n_2}(++) + P_{n_1 n_2}(--) - P_{n_1 n_2}(+-) - P_{n_1 n_2}(-+) \qquad (8.28)$$

we see that each of the four probabilities on the right hand side can be expressed in terms of the frequencies F_i from our table. For given measurement directions $\hat{n}_1$ and $\hat{n}_2$, each frequency will appear, either with a plus sign or a minus sign. For example:

$$C(\hat{a}, \hat{b}) = F_3 + F_5 + F_4 + F_6 - F_1 - F_2 - F_7 - F_8. \qquad (8.29)$$

(Take a minute and make sure you understand exactly how I got this!) Similarly:

$$C(\hat{a}, \hat{c}) = F_2 + F_5 + F_4 + F_7 - F_1 - F_3 - F_6 - F_8 \qquad (8.30)$$

and

$$C(\hat{b}, \hat{c}) = F_2 + F_3 + F_6 + F_7 - F_1 - F_4 - F_5 - F_8. \qquad (8.31)$$

But then, as before, obviously-true inequalities involving the F_i can be seen to be equivalent to inequalities involving the correlation coefficients for different measurement settings. For example, the trivial inequality

$$|F_3 + F_6 - F_7 - F_2| \leq F_2 + F_3 + F_6 + F_7 \tag{8.32}$$

turns out to be equivalent to

$$\left| C(\hat{a}, \hat{b}) - C(\hat{a}, \hat{c}) \right| \leq 1 + C(\hat{b}, \hat{c}). \tag{8.33}$$

This is actually the original "Bell inequality" that Bell derived (using a somewhat different method) in his original 1964 paper, and it is easy to see that it is violated by the Quantum Mechanical predictions. For $\hat{a}$, $\hat{b}$, and $\hat{c}$ as before (with $\hat{b}$ in the middle and then $\hat{a}$ and $\hat{c}$ θ away in opposite directions) the greatest violation is again achieved for $\theta = \pi/3 = 60°$. So then $C^{QM}(\hat{a}, \hat{b}) = -\cos(60°) = -1/2$ and $C^{QM}(\hat{a}, \hat{c}) = -\cos(120°) = +1/2$ and $C^{QM}(\hat{b}, \hat{c}) = -1/2$ so we have

$$\left| -\frac{1}{2} - \frac{1}{2} \right| \leq 1 - \frac{1}{2} \tag{8.34}$$

which reduces to

$$1 \leq \frac{1}{2} \tag{8.35}$$

which is definitely not true! So the QM predictions violate Bell's Inequality, which is a constraint on the correlations (between the outcomes on the two sides) for the kind of theory implied by the EPR-Bohm argument.

One can develop a second, closely-related Bell-type inequality by starting with the trivial inequality

$$|F_5 + F_4 - F_1 - F_8| \leq F_1 + F_4 + F_5 + F_8 \tag{8.36}$$

which turns out to be equivalent to

$$\left| C(\hat{a}', \hat{b}) + C(\hat{a}', \hat{c}) \right| \leq 1 - C(\hat{b}, \hat{c}) \tag{8.37}$$

where $\hat{a}'$ may be (but need not be!) the same direction as the previous $\hat{a}$.

Note, finally, that by adding Equations (actually, inequalities) (8.33) and (8.37) we arrive at

$$\left| C(\hat{a}, \hat{b}) - C(\hat{a}, \hat{c}) \right| + \left| C(\hat{a}', \hat{b}) + C(\hat{a}', \hat{c}) \right| \leq 2 \tag{8.38}$$

which is a particularly important Bell-type inequality called the "CHSH inequality" (after Clauser, Horne, Shimony, and Holt who first derived it in 1969) [3].

The CHSH inequality is particularly well-suited to experimental test, because only two measurement angles appear in each "wing" of the experiment. That is, suppose that for each particle pair that is created, Particle 1 is sent toward Alice, who will (randomly and at the last possible second) choose to measure the spin of Particle 1 along either the direction $\hat{a}$ or the direction $\hat{a}'$. And similarly, Particle 2 is sent toward Bob, who will (randomly and at the last possible second) choose to measure the spin of Particle 2 along either the direction $\hat{b}$ or the direction $\hat{c}$.

Alice and Bob then record their respective outcomes for that particular particle pair, and get ready for the particles from the next pair to arrive. After collecting data for some time, Alice and Bob meet somewhere and compare notes. Importantly, every single outcome of every single measurement on every single particle pair gets used to determine an empirical value for one of the four correlation coefficients appearing in the CHSH inequality. So none of the data is wasted.

What does QM predict for the CHSH parameter? Well, suppose we pick the directions $\hat{a}$, $\hat{a}'$, $\hat{b}$, and $\hat{c}$ as follows: $\hat{a}$ will be the z-axis, and $\hat{a}'$ will be the x-axis. Then $\hat{b}$ will be halfway between the x- and z-axes, i.e., 45° down from the z-axis toward the x-axis. And $\hat{c}$ will be 45° away from the x-axis in the other direction, i.e., 135° down from the z-axis toward the x-axis, i.e., halfway between the x-axis and the negative z-axis. See Fig. 8.1.

These turn out to be the directions that give the greatest possible discrepancy between the local hidden variable and the quantum mechanical predictions. To see exactly what QM predicts, note that the angle between $\hat{a}$ and $\hat{b}$ is 45°, so, according to Eq. (8.27)

$$C^{QM}(\hat{a}, \hat{b}) = -\cos(45°) = -1/\sqrt{2}. \tag{8.39}$$

Similarly, we have that

$$C^{QM}(\hat{a}, \hat{c}) = +1/\sqrt{2}, \tag{8.40}$$

Fig. 8.1 The four directions that give the greatest discrepancy between the local hidden variable and quantum mechanical predictions for the CHSH inequality
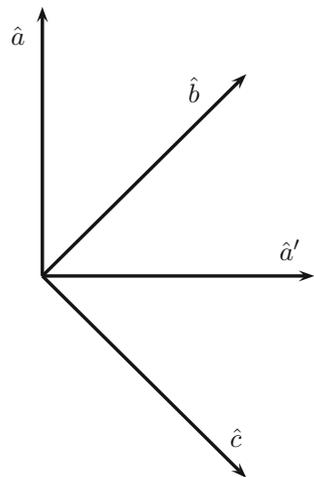
$$C^{QM}(\hat{a}', \hat{b}) = -1/\sqrt{2}, \qquad (8.41)$$

and

$$C^{QM}(\hat{a}', \hat{c}) = -1/\sqrt{2}. \qquad (8.42)$$

Putting these together, we have that

$$\left| C^{QM}(\hat{a}, \hat{b}) - C^{QM}(\hat{a}, \hat{c}) \right| + \left| C^{QM}(\hat{a}', \hat{b}) + C^{QM}(\hat{a}', \hat{c}) \right| = 2\sqrt{2}. \qquad (8.43)$$

This is a factor of $\sqrt{2}$ – i.e., about 40% – bigger than should be allowed if a local hidden variable theory is true.

8.4 Experiments

Let us then review some of the actual experiments of this sort.

The first really systematic test of the CHSH inequality was done by Alain Aspect and collaborators in 1982 [4]. Instead of using pairs of spin 1/2 particles in the spin singlet state, they used pairs of photons (emitted from excited Calcium atoms) whose *polarizations* are entangled in a way that is perfectly analogous to the singlet state we've been discussing. Note, however, that there is a factor-of-2 difference between the angles involved in the spin 1/2 case and the photon polarization case. Whereas, for example, the two possible spin directions of a spin 1/2 particle are *opposite* ("up" and "down", different by 180°) the two possible polarizations of a photon are *orthogonal* (e.g., "horizontal" and "vertical", different by 90°). In addition, whereas spin-1/2 particles in the singlet state display perfect anti-correlation (when their spins are measured along the same axes), the photon pairs instead display perfect (positive) correlation (when their polarizations are measured along the same axes). So the quantum prediction for the polarization correlation coefficient $C^{QM}(\hat{n}_1, \hat{n}_2)$ in the case of photons is $\cos(2\theta)$ rather than the $-\cos(\theta)$ we saw previously for the case of spin-1/2 particles. But otherwise everything is just as we've been discussing.

A schematic diagram of the experiment and a graph of their data are reproduced (from the 1982 paper) in Fig. 8.2. There is essentially perfect agreement between the experimental results and the QM predictions, and the CHSH parameter (that is, the combination of correlation coefficients that can be no greater than 2 for local hidden variable theories) was

$$S_{expt} = 2.697 \pm 0.015, \qquad (8.44)$$

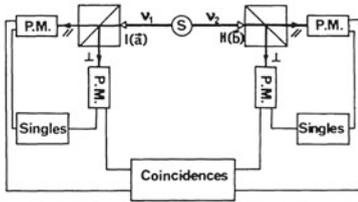i.e., well above the maximum possible value (namely, 2) allowed for the local hidden variable theories.

FIG. 2. Experimental setup. Two polarimeters I and II, in orientations $\vec{a}$ and $\vec{b}$, perform true dichotomic measurements of linear polarization on photons ν_1 and ν_2. Each polarimeter is rotatable around the axis of the incident beam. The counting electronics monitors the singles and the coincidences.
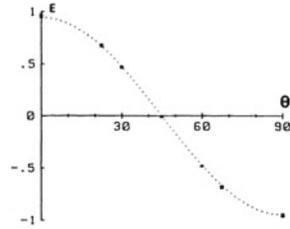
FIG. 3. Correlation of polarizations as a function of the relative angle of the polarimeters. The indicated errors are ± 2 standard deviations. The dotted curve is not a fit to the data, but quantum mechanical predictions for the actual experiment. For ideal polarizers, the curve would reach the values ± 1.

Fig. 8.2 The *left panel* shows a schematic diagram of the setup for the Aspect et al. experiment, Ref. [4]; the *right panel* shows their data. The essentially perfect agreement with $C^{QM} = \cos(2\theta)$ is evident. In this experiment they found the CHSH parameter to be $S_{expt} = 2.697 \pm 0.015$, i.e., well above the maximum value of 2 allowed for local hidden variable theories and in excellent agreement with the QM prediction (which, for detector efficiency and alignment reasons) for their experiment was $S_{QM} = 2.70 \pm 0.05$ (i.e., slightly less than $2\sqrt{2}$). (Reprinted with permission from A. Aspect, P. Grangier, G. Roger, *Physical Review Letters*, 49, 91–94, 1982, by the American Physical Society.)

Here is how Aspect et al. summarize their conclusions:

> ...our experiment yields the strongest violation of Bell's inequalities ever achieved, and excellent agreement with quantum mechanics. Since it is a straightforward transposition of the ideal Einstein–Podolsky-Rosen-Bohm scheme, the experimental procedure is very simple, and needs no auxiliary measurements as in previous experiments with single-channel polarizers. We are thus led to the rejection of realistic [i.e., "hidden variable"] local theories if we accept the assumption that there is no bias in the detected samples: Experiments support this natural assumption.

> Only two loopholes remain open for advocates of realistic theories without action at a distance. The first one, exploiting the low efficiencies of detectors, could be ruled out by a feasible experiment. The second one, exploiting the static character of all previous experiments, could also be ruled out by a 'timing experiment' with variable analyzers now in progress [4].

What is this "assumption that there is no bias in the detected samples"? The idea here is that the committed advocate of a local hidden variable theory could claim that the CHSH inequality wasn't *really* shown to be violated because, actually, only a small fraction of all emitted photon pairs were successfully detected. (This has to do with the fact that the photon source in this experiment emits photon pairs isotropically, so only the occasional pairs which just happen to be aimed right at the detectors, will actually be detected.) "If only" (says the conspiracy theorist here) "*all* of the pairs had been detected, we might have found a CHSH parameter less than 2." And, indeed, that is in principle possible, although it seems exceedingly unlikely. In order for the experimental results to be biased in this way, there would need to be some reason why the polarization correlations between pairs which happen to be aimed at the detectors, are significantly different than those between pairs going in other directions. But neither quantum mechanics nor any other available
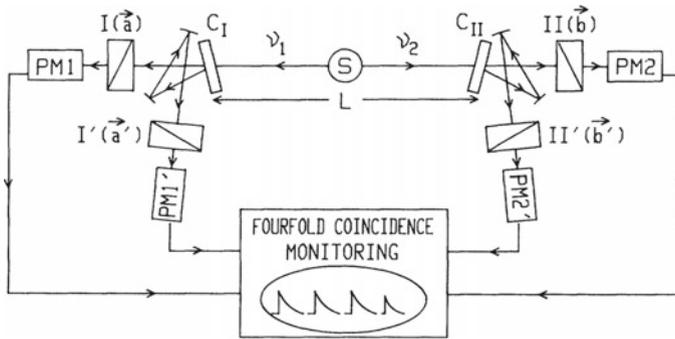
Fig. 8.3 Aspect et al.'s schematic diagram of their "Experimental test of Bell's inequalities using time-varying analyzers" from Ref. [5]. They write: "timing experiment with optical switches. Each switching device (C_I, C_{II}) is followed by two polarizers in two different orientations. Each combination is equivalent to a polarizer switched fast between two orientations." (Reprinted with permission from A. Aspect, J. Dalibard, G. Roger, *Physical Review Letters*, 49, 1804–1807, 1982, by the American Physical Society.)

idea provides any basis for suspecting such a thing. So this way of trying to elude the apparent implications of Aspect's first experiment seems like grasping at straws. Nevertheless, increasing the fraction of particle pairs detected (i.e., working to "close the detector efficiency loophole") is something that experimenters worked on in subsequent experiments.

The second "loophole" that Aspect mentions, however – having to do with the "static character" of this and previous experiments – is a bit of a more serious issue. Indeed, this is precisely the issue I already raised in the previous section: in order to really discriminate quantum mechanics from local hidden variable theories of the sort implied by the EPR-Bohm argument, the measurement devices should not remain static; instead, the settings (that is, for Alice, the choice between measuring along $\hat{a}$ and $\hat{a}'$, and, for Bob, the choice between measuring along $\hat{b}$ and $\hat{c}$) should be made randomly and at the last possible second.

In a follow-up experiment (also published in 1982!) Aspect and his collaborators devised an ingenious mechanism for switching between different measurement settings. (See Fig. 8.3) Instead of physically rotating a polarization measuring device (which would simply not be feasible on the required time scales), each photon was shunted to one or the other of two static measuring devices by an "acousto-optical switch" consisting of a cell of water in which a (roughly) 50 MHz acoustic standing wave is set up. During the part of the cycle in which the amplitude of the standing wave is small, incident photons simply pass straight through. But during the part of the cycle in which the amplitude is large, the water (with now a spatially-varying density) acts like a diffraction grating and photons passing through are deflected at some angle. The 50 MHz standing wave frequency corresponds to about a 10 nanosecond period during which any incident photons are shunted in the one direction, before switching to the other direction. This timescale can be compared to the

time – $L/c \approx 40$ nanoseconds – it would take a signal, propagating at the speed of light, to get from Alice's measuring device to Bob's (or vice versa). If the fast back-and-forth switching between the two measurement directions on either side is considered "effectively random" (and note here that the frequencies were deliberately made incommensurate on the two sides) one can thus say that the choice of measurement settings on the two sides are spacelike separated. As Aspect et al. write:

> The new feature of this experiment is that we change the settings of the polarizers, at a rate greater than c/L. The ideal scheme has not been completed since the change is not truly random, but rather quasiperiodic. Nevertheless, the two switches on the two sides are driven by different generators at different frequencies. It is then very natural to assume that they function in an uncorrelated way.
>
> A more ideal experiment with random and complete switching would be necessary for a fully conclusive argument against the whole class of supplementary-parameter [i.e., hidden variable] theories obeying Einstein's [local] causality. However, our observed violation of Bell's inequalities indicates that the experimental accuracy was good enough for pointing out a hypothetical discrepancy with the predictions of quantum mechanics. No such effect was observed [5].

Thus, again in this improved version of the experiment, results consistent with QM – and inconsistent with a Bell-type inequality – were observed.

Some of the possible improvements described by Aspect et al. were implemented by a new experiment performed in 1998 by Gregor Weihs, Anton Zeilinger, and collaborators [6]. They used a new (more controllable and more efficient) source of polarization-entangled photon pairs (called "type-II parametric down-conversion"). The photons were carried from the central source to the polarization measuring stations (at opposite ends of the campus of the University of Innsbruck in Austria) along fibre optic cables. And, crucially, the experiment used "high speed physical random number generators and fast electro-optic modulators" to arrange for the choice of measurement axis (along which each particle's polarization was measured) to be made, for the first time, genuinely randomly and at unambiguously space-like separation from the similar choice occuring in the other wing of the experiment.[1] Weihs et al. report typical measured CHSH parameter values of $S = 2.73 \pm 0.02$ – in excellent agreement with the quantum mechanical predictions (as applicable to their particular experimental setup) and in clear violation of the CHSH inequality.

The experiment by Weihs et al. remained essentially the state-of-the-art until quite recently, when *three* different groups published results showing that the QM predictions remain correct even when the various "loopholes" described originally by Aspect are simultaneously closed [7–9].

---

[1]Technical detail: actually, instead of using the output of the random number generator to physically rotate the polarization measuring device (which could never be done quickly enough), the output was fed into an electro-optic modulator which rotated (by one of two possible amounts) the polarization of the incoming photon.

8.5   **What Does It Mean?**

I have been describing Bell's inequalities as constraints on the predictions of local hidden variable theories of the sort implied by the EPR-Bohm argument we reviewed at the beginning of the chapter. From this point of view, the experiments we surveyed in the last section prove (to me, at least, quite convincingly) that the predictions of QM are *correct* and the predictions of the local hidden variable theories are simply *wrong*. And leaving aside the conspiracy theorists, everybody agrees about this.

There is, however, a surprising amount of controversy about what, exactly, should be inferred from the empirical violations of Bell inequalities.

In particular, many people have taken Bell's theorem as a proof that hidden variables theories are not viable.[2] Eugene Wigner, for example, wrote (about the possibility of a hidden variables "completion" of ordinary QM) that the proof of von Neumann "uses assumptions which, in my opinion, can quite reasonably be questioned." (Here Wigner was in total agreement with Bell, who recall showed in his 1966 paper that von Neumann's assumptions were in fact totally arbitrary and unwarranted.) But Wigner goes on to state:

> In my opinion, the most convincing argument against the theory of hidden variables was presented by J.S. Bell [10].

A similar remark has been made by the eminent theoretician Rudolf Peierls:

> If people are obstinate in opposing the accepted view they can think of many new possibilities, but there is no sensible view of hidden variables which doesn't conflict with these experimental results [i.e., Aspect's experiments]. That was proved by John Bell, who has great merit in establishing this. Prior to that there was a proof due to the mathematician von Neumann, but he made an assumption which is not really necessary [11].

More recently, Stephn Hawking summarized the situation as follows:

> Einstein's view was what would now be called a hidden variable theory. Hidden variable theories might seem to be the most obvious way to incorporate the Uncertainty Principle into physics. They form the basis of the mental picture of the universe, held by many scientists, and almost all philosophers of science. But these hidden variable theories are wrong. The British physicist, John Bell ... devised an experimental test that would distinguish hidden variable theories. When the experiment was carried out carefully, the results were inconsistent with hidden variables. Thus it seems that even God is bound by the Uncertainty Principle.... God does play dice with the universe [12].

It is easy to multiply examples. In a review article on "One Hundred Years of Quantum Physics", Daniel Kleppner and Roman Jackiw of MIT wrote, about the experiments reviewed in the last section, that "[t]heir collective data came down decisively against the possibility of hidden variables. For most scientists this resolved any doubt about the validity of quantum mechanics" [13]. And in a similar article celebrating the 100 year anniversary of QM, Max Tegmark and John Wheeler wrote the following:

---

[2]Most of the quotes in the following paragraph were collected by Jean Bricmont in Sect. 7.5 of *Making Sense of Quantum Mechanics*.

Could the apparent quantum randomness be replaced by some kind of unknown quantity carried out inside particles, so-called 'hidden variables'? CERN theorist John Bell showed that in this case, quantities that could be measured in certain difficult experiments would inevitably disagree with standard quantum predictions. After many years, technology allowed researchers to conduct these experiments and eliminate hidden variables as a possibility [14].

In their preface to the published proceedings of a conference honoring Bell 10 years after his death, Reinhold Bertlmann (a long-time colleague, collaborator, and friend of Bell's... recall his mis-matched socks from Chap. 4!) and Anton Zeilinger (one of the authors of the Innsbruck experiment paper discussed in the last section and one of the most prominent experimental quantum physicists) explained how, although Bell had seemingly opened the door to hidden variables by refuting von Neumann's supposed impossibility proof,

he immediately dealt them [i.e., hidden variables] a major blow. In 1964 ... he showed that any hidden variables theory, which obeys Einstein's requirement of locality, i.e., no influence travelling faster than the speed of light, would automatically be in conflict with quantum mechanics. [...] While a very tiny [experimental] loophole in principle remains for local realism, it is a very safe position to assume that quantum mechanics has definitely been shown to be the right theory. Thus, a very deep philosophical question, namely, whether or not events observed in the quantum world can be described by an underlying deterministic theory, has been answered by experiment, thanks to the momentous achievement of John Bell [15].

What is going on here? How can all these people claim that the experimental violation of Bell's inequality somehow refutes the possibility of an underlying deterministic or hidden variable completion of quantum mechanics, when such a theory (namely, the de Broglie - Bohm pilot-wave theory) already actually exists and is demonstrably consistent with these experiments?

Part of the answer, to be sure, is that most physicists are simply not as aware as they should be about the existence of the pilot-wave theory. They've heard of it but never looked into it and hence don't actually understand how it works and, as we've seen, they dismiss the broad category of hidden variable theories (of which the pilot-wave theory is just one concrete example) on the grounds that they have been ruled out, experimentally, as shown by Bell. There is a kind of rich and tragic irony here, in citing Bell as having supposedly refuted hidden variables theories (and hence entrenching the unjustified belief that the pilot-wave theory cannot be right and must not be worth looking into) when, as we have seen, Bell's theorem was actually inspired by Bohm's 1952 pilot-wave theory papers, and indeed Bell remained far and away the pilot-wave theory's greatest champion until his death in 1990.

But there is more going on, in the citation of Bell's theorem as refuting the hidden variables program, than mere ignorance of the pilot-wave theory. Some of the people who make this kind of argument do know about the pilot-wave theory, and reject it on the grounds that it is non-local and hence in apparent conflict with relativity. This point of view was perhaps best encapsulated by David Mermin's remark:

To those for whom nonlocality is anathema, Bell's theorem finally spells the death of the hidden-variables program [16].

The idea here, apparently, is that Bell's inequality – which we now know from experiment is *false* – follows from the conjunction of two premises: locality and hidden variables. (Or sometimes, instead, of "hidden variables" people will say "determinism" or "realism" – or something essentially equivalent but a little more cryptic called "counter-factual definiteness".) But if these two premises, together, imply something that is false, at least one of the premises must be wrong. According to this viewpoint, then, we have to *choose* between the following two options:

1. Uphold locality and reject hidden variables, i.e., retain consistency with relativistic causality and admit (as everyone has told us we should have done anyway) that Einstein was wrong and Bohr was right regarding the question of whether the quantum mechanical description of reality can be considered complete.
2. Uphold hidden variables and reject locality, i.e., side with the somewhat senile Einstein in his stubborn, arbitrary, and philosophical demand that "God does not play dice" and insist that, despite being one of the most successful and highly-confirmed theories in the history of science, relativity is somehow wrong.

If those were the two available options, it would indeed be a no-brainer. Obviously we should choose option 1. Selecting option 2 would be crazy.

In this way of looking at the matter, non-locality is the price one has to pay for attempting to restore determinism (and/or "realism") to quantum theory... and the price, obviously, is simply too high. Maintaining consistency with relativistic causality (i.e., locality) is mandatory, and if that means we need to abandon the quest for a more complete underlying model of quantum phenomena, so be it; indeed, most would say, good riddance.

That, I think, captures the viewpoint of the vast majority of physicists today. But, I believe, it is completely and utterly and hopelessly wrong. We do not face anything like the choice between options 1 and 2 above, and indeed, at the end of the day, Bell's theorem tells us absolutely nothing about "hidden variables" or determinism or counter-factual definiteness or "realism" or whether the moon is there or not when you aren't looking [17] or *any* of these sorts of things that people so frequently say it is fundamentally about. Everybody is just simply wrong here, because they have forgotten (or, more commonly, because they never understood in the first place) a crucial part of the broader context of Bell's theorem.

In particular, they have forgotten the EPR argument – which, remember, is supposed to be a proof that deterministic hidden variables are *required, in the first place, precisely in order to avoid non-locality*. The sort of hidden variable theory that Bell's theorem ends up ruling out, that is, is not something that Bell – or for that matter Einstein – just dreamed up. It's not something they just liked or randomly felt like considering. It's something they considered specifically because they recognized it as the only possible hope for maintaining locality in the face of the perfect EPR correlations.

Bell's theorem, then – taken here to mean the proof that local hidden variable theories are wrong – must be understood as the second part of an overall two-part argument, the first part of which is the EPR argument. Schematically, the two-part argument goes like this:

EPR: locality → X
Bell: X → conflict with experiment

Here "X" stands for something like "local deterministic hidden variables", but somehow the logic is easier to grasp by suppressing this. Obviously, if locality → X, and X in turn implies a conflict with experiment, then we cannot maintain X, which means we cannot maintain locality (because locality entails X!).

So according to this view, what should be concluded from the experimentally observed violations of Bell-type inequalities is not that we cannot have hidden variables (we can!), and not even that we must choose between hidden variables and locality. It only appears that we face such a choice if we look *only* at the second, Bell-part of the two-part argument. But if we remember also the first, EPR-part of the argument, we remember that the choice is highly constrained: keeping locality but abandoning hidden variables is not an available option at all. We must, that is, simply conclude that locality – that the prohibition on faster-than-light causation that seems somehow to be implied by relativity theory – is false. Relativistic local causality is wrong, is in conflict with experimental data. Faster-than-light causal influences really exist in Nature!

That is, to be sure, a shocking conclusion and raises all kinds of pressing questions that proliferate in all directions. But we will not be able to pursue them in detail here. I will instead close this section by sharing that this "alternative" view – according to which the upshot of Bell's theorem is that locality is false – is not only my view (and that of some other contemporary physicists and philosophers of science), but was also the view of the person in the best possible position to understand Bell's reasoning: Bell himself.

In his introductory remarks at a 1984 conference, for example, Bell said that "the real problem with quantum theory" is the "essential conflict between any sharp formulation and fundamental relativity" and went on to speak of the "incompatibility, at the deepest level, between the two fundamental pillars of contemporary theory" (meaning quantum theory and relativity theory) [18].

Indeed, Bell even went so far as to suggest, in response to his theorem and the relevant experimental data, the rejection of "fundamental relativity" and the return to a Lorentzian view in which there is a dynamically privileged (though probably empirically undetectable) reference frame:

> "It may well be that a relativistic version of [quantum] theory, while Lorentz invariant and local at the observational level, may be necessarily non-local and with a preferred frame (or aether) at the fundamental level" [19].

And elsewhere:

> "... I would say the cheapest resolution is something like going back to relativity as it was before Einstein, when people like Lorentz and Poincaré thought that there was an aether – a preferred frame of reference – but that our measuring instruments were distorted by motion in such a way that we could not detect motion through the aether. Now, in that way you can imagine that there is a preferred frame of reference, and in this preferred frame of reference things do go faster than light. .... Behind the apparent Lorentz invariance of the phenomena, there is a deeper level which is not Lorentz invariant... [This] pre-Einstein

position of Lorentz and Poincaré, Larmor and Fitzgerald, was perfectly coherent, and is not inconsistent with relativity theory. The idea that there is an aether, and these Fitzgerald contractions and Larmor dilations occur, and that as a result the instruments do not detect motion through the aether – that is a perfectly coherent point of view" [20].

Why did Bell take so seriously these sorts of ideas, which everybody else today regards as completely outmoded and wrong? Because he thinks his theorem (and the associated experimental evidence) proves that nonlocality is a fact of Nature, rather than merely a defect of a type of theory we shouldn't believe in.

And, as I have tried to explain, he thinks that because he sees his theorem as building from where the EPR argument left off. He makes his reasoning particularly clear in his classic 1981 paper, "Bertlmann's Socks and the Nature of Reality", in which he reacts against the confusion described above (namely, forgetting about the EPR argument and hence inferring a completely wrong conclusion from the theorem) by laying out the two-part argument:

"Let me summarize once again the logic that leads to the impasse. The EPRB correlations are such that the result of the experiment on one side immediately foretells that on the other, whenever the analyzers happen to be parallel. If we do not accept the intervention on one side as a causal influence on the other, we seem obliged to admit that the results on both sides are determined in advance anyway, independently of the intervention on the other side, by signals from the source and by the local magnet settings. [That's the first, EPR-part of the argument.] But this has implications for non-parallel settings which conflict with those of quantum mechanics. [That's the second part, what is usually (alone) called "Bell's theorem".] So we *cannot* dismiss intervention on one side as a causal influence on the other" [21].

The last sentence expresses the overall conclusion of the two-part argument, that (something about) the measurement on one side *does* influence, faster than light, the results on the other side.

Earlier in the same paper, Bell rehearses the EPR argument and then underscores its logical structure as follows:

"It is important to note that to the limited degree to which *determinism* plays a role in the EPR argument, it is not assumed but *inferred*. What is held sacred is the principle of 'local causality' - or 'no action at a distance'. Of course, mere *correlation* between distant events does not by itself imply action at a distance, but only correlation between the signals reaching the two places. These signals, in the idealized example of Bohm, must be sufficient to *determine* whether the particles go up or down. For any residual undeterminism could only spoil the perfect correlation.

"It is remarkably difficult to get this point across, that determinism is not a *presupposition* of the analysis. There is a widespread and erroneous conviction that for Einstein[*] determinism was always *the* sacred principle. The quotability of his famous 'God does not play dice' has not helped in this respect" [21].

The footnote referred to after the mention of Einstein reads:

"And his followers. My own first paper on this subject [i.e., Bell's 1964 paper presenting "Bell's theorem"] starts with a summary of the EPR argument *from locality to* deterministic hidden variables. But the commentators have almost universally reported that it begins with deterministic hidden variables" [21].

I personally find this footnote remarkable and extremely revealing. Bell describes himself as a follower of Einstein (meaning, presumably, that for him, like for Einstein, it is 'local causality' rather than determinism which is *the* sacred principle) and then says explicitly that "the commentators have almost universally" misunderstood Bell's theorem (as presented in his original 1964 paper) because they have failed to appreciate the relevance of "the EPR argument *from locality to* deterministic hidden variables."

8.6   (Bell's) Locality Inequality Theorem

As we saw, Bell claimed rather unambiguously in 1981 that his original 1964 paper "starts with a summary of the EPR argument *from locality to* deterministic hidden variables". Whether or not Bell had actually had this two-part argument in mind from the beginning, however, has been the focus of some discussion and debate during the recent 50th anniversary celebration of Bell's paper [22]. Suffice it to say that, on the one hand, it is clear that the introductory sections of Bell's 1964 paper begin by reminding the reader of what Einstein et al. had already established, several decades earlier. But, on the other hand, Bell's summary of the EPR argument is indeed somewhat unfortunately brief and informal.

Happily, though, Bell continued to write and give talks about "Bell's theorem" throughout the period between 1964 and 1990, and in these talks and papers we see a systematic attempt to clarify, sharpen, and make more explicit several aspects of the reasoning, and thereby to pre-empt the sort of misunderstanding discussed above. There are a couple of threads to this development. One is the thing we have just been focusing on: making the EPR argument and its relationship to his new discovery more explicit and clear.

But another thread, the one I want to focus on here, involved eliminating the middle-man, so to speak – that is, constructing a simpler and more direct demonstration of the incompatibility between local causality and experiment (via an empirically testable inequality). That is, whereas in Bell's earlier presentations, he tends to use the "two-part argument" described in the last section (locality implies deterministic local hidden variables, and then deterministic local hidden variables imply a Bell-type inequality), in his later presentations Bell instead lays out much more explicitly what exactly he means by "locality" and then shows *directly* how locality entails (for example) the CHSH inequality. So I thought it would be good to end the chapter by rehearsing this more direct presentation that represents, I think, Bell's mature sense of what he proved and why it's important.

We have already discussed Bell's formulation of "locality" – way back in Chap. 1 and then again in the context of the EPR argument in Chap. 4 (and again in Chap. 5). See Fig. 8.4 for a brief recap of the idea that, in a theory respecting relativistic local causality, certain information (at appropriate space-like separation from a given event) must necessarily be irrelevant for making predictions about the given event, once what happens in the backward light cone of that event is sufficiently specified.
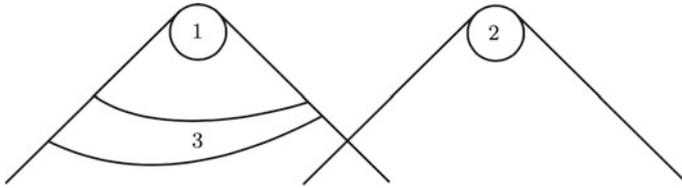
Fig. 8.4 Space-time regions relevant to Bell's formulation of local causality. Bell writes: "Full specification of what happens in 3 makes events in 2 irrelevant for predictions about 1 in a locally causal theory" [23]

Let us then sketch how the empirically-testable CHSH inequality can be derived – *directly* – from Bell's formulation of local causality. Consider the now-familiar sort of experimental setup in which a source creates particle pairs in a state that QM would describe as the singlet state. Of course, we want to be completely general here and allow that perhaps some other (hidden variable type) theory turns out to provide the correct description. So we will say that a given particle pair has state λ. This *might* be just the QM singlet state ψ_s, or it might be the singlet state wave function plus some additional "hidden variables", or it might be something entirely distinct from the QM wave function. We will keep it completely unspecified, completely general.

So then, the particles fly off in opposite directions toward measuring stations womanned, and manned, respectively, by Alice and Bob. Alice uses some kind of random number generator to pick, at the last possible second, an axis $\hat{n}_1$ along which to measure the spin of her particle, and Bob similarly uses an independent random number generator to pick, also at the last possible second, an axis $\hat{n}_2$ along which to measure the spin of his particle. Let's call Alice's outcome A (with, as usual, $A = +1$ meaning "spin up" and $A = -1$ meaning "spin down") and similarly for Bob's outcome, B. All of these things are depicted on the space-time diagram in Fig. 8.5.

Note that we utilize the necessary condition for locality introduced in Chap. 5: the (large!) region 3 in Fig. 8.5 not only shields off the past light cone of region 1 from region 2 (as in the previous figure), but also vice versa. Thus a complete specification of the physical state of things in region 3 should render everything about region 2 (in particular, both Bob's randomly-selected setting $\hat{n}_2$ and his outcome B) irrelevant for predictions about region 1 – and it should also render everything about region 1 (in particular, both Alice's randomly-selected setting $\hat{n}_1$ and her outcome A) irrelevant for predictions about region 2.

This is the key idea behind the *factorization* of the joint probability $P(A, B|\hat{n}_1, \hat{n}_2, \lambda)$ for outcomes A and B (conditioned on all the things these outcomes might depend upon). The definition of conditional probability implies that this joint probability can be written as

$$P(A, B|\hat{n}_1, \hat{n}_2, \lambda) = P(A|\hat{n}_1, \hat{n}_2, B, \lambda) \cdot P(B|\hat{n}_1, \hat{n}_2, \lambda). \qquad (8.45)$$
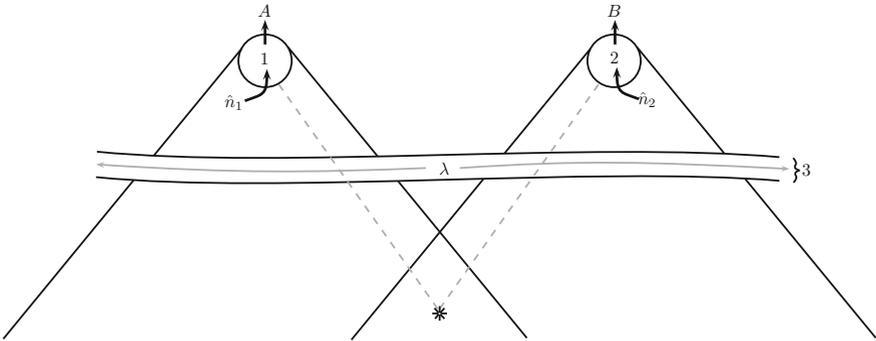
Fig. 8.5 Space-time diagram for the Bell experiment. The particle pair is emitted at the "flash" at the *bottom* of the diagram; world-lines for the two individual particles flying apart in opposite directions are represented by the *gray dashed lines*. The (large!) region 3 encompasses both particles at some intermediate time and shields the two measurement regions, 1 and 2, from their overlapping past light cones in the way that is required in Bell's formulation of locality. (Note that the complete description of the particle pair, λ, in region 3 need not assign specific facts to specific points in space; this allows λ to be something like a two-particle quantum-mechanical wave function which, as discussed in Chap. 5, does not imply a clear ontology in 3D space.) The apparatus settings $\hat{n}_1$ and $\hat{n}_2$ are shown as "inputs" to the measurements occuring in regions 1 and 2, whereas the individual outcomes A and B are shown as "outputs"

But then, in each of the two factors on the right hand side, Bell's definition of local causality implies that the entries relating to the space-like separated region are redundant. In particular, local causality requires that

$$P(A|\hat{n}_1, \hat{n}_2, B, \lambda) = P(A|\hat{n}_1, \lambda) \tag{8.46}$$

and

$$P(B|\hat{n}_1, \hat{n}_2, \lambda) = P(B|\hat{n}_2, \lambda). \tag{8.47}$$

In words: the probability for Alice's experiment to have outcome A should only depend on the state λ of the particle pair (for that run of the experiment) and the setting $\hat{n}_1$ of her apparatus; it should *not* depend on the setting of Bob's apparatus or on the outcome B of his experiment. And similarly, the probability for Bob's experiment to have outcome B should not depend on the setting of Alice's apparatus.

Plugging in we see that, for a locally causal theory, the joint probability factorizes as follows:

$$P(A, B|\hat{n}_1, \hat{n}_2, \lambda) = P(A|\hat{n}_1, \lambda)\, P(B|\hat{n}_2, \lambda). \tag{8.48}$$

From here, it turns out to be a straightforward mathematical exercise to derive the CHSH inequality. Recall first that the correlation coefficient $C(\hat{n}_1, \hat{n}_2)$ is defined as the expected value of the product of the outcomes A and B. This is simply the sum (over all four possible joint outcomes for A and B) of the product of the outcomes, weighted by the probability of that particular joint outcome. Notice also that, when we

speak of the "expected value", we mean: over a run of many particle pairs in which, for all we know, the exact state λ of the particle pair may vary from run to run. So we should also, for each possible joint outcome, average over the possible states λ that might have been produced by the source. (Let's assume, for definiteness but without loss of generality, that a continuously infinite spectrum of different possible λs are possible, with probability density $\rho(\lambda)$.) Our expression for the correlation coefficient then looks like:

$$
\begin{aligned}
C(\hat{n}_1, \hat{n}_2) &= \int \sum_{A,B} A \cdot B \cdot P(A, B|\hat{n}_1, \hat{n}_2, \lambda)\, \rho(\lambda)\, d\lambda \\
&= \int \sum_{A,B} A \cdot B \cdot P(A|\hat{n}_1, \lambda) \cdot P(B|\hat{n}_2, \lambda)\, \rho(\lambda)\, d\lambda \\
&= \int \left[\sum_A A \cdot P(A|\hat{n}_1\, \lambda) \right]\left[\sum_B B \cdot P(B|\hat{n}_2, \lambda) \right] \rho(\lambda)\, d\lambda \\
&= \int \bar{A}(\hat{n}_1, \lambda)\, \bar{B}(\hat{n}_2, \lambda)\, \rho(\lambda)\, d\lambda
\end{aligned}
\tag{8.49}
$$

where, in the first step, we have used the factorized expression for the joint probability (which follows from local causality) and in the last step we have defined

$$
\bar{A}(\hat{n}_1, \lambda) = \sum_A A \cdot P(A|\hat{n}_1, \lambda)
\tag{8.50}
$$

as the average value of A (and similarly for B). Since the only two possible outcomes for A are $+1$ and -1, it is obvious that this average value must be between -1 and $+1$, i.e.,

$$
\left| \bar{A}(\hat{n}_1, \lambda) \right| \leq 1
\tag{8.51}
$$

and similarly

$$
\left| \bar{B}(\hat{n}_2, \lambda) \right| \leq 1.
\tag{8.52}
$$

Now let us consider the combinations of correlation coefficients that appear in the CHSH inequality. To begin with,

$$
C(\hat{a}, \hat{b}) - C(\hat{a}, \hat{c}) = \int \bar{A}(\hat{a}, \lambda) \left[\bar{B}(\hat{b}, \lambda) - \bar{B}(\hat{c}, \lambda) \right] \rho(\lambda)\, d\lambda
\tag{8.53}
$$

so that

$$
\left| C(\hat{a}, \hat{b}) - C(\hat{a}, \hat{c}) \right| \leq \int \left| \bar{B}(\hat{b}, \lambda) - \bar{B}(\hat{c}, \lambda) \right| \rho(\lambda)\, d\lambda
\tag{8.54}
$$

since $\left| \bar{A}(\hat{a}, \lambda) \right| \leq 1$.

In a similar way, we have that

$$\left| C(\hat{a}', \hat{b}) + C(\hat{a}', \hat{c}) \right| \leq \int \left| \bar{B}(\hat{b}, \lambda) + \bar{B}(\hat{c}, \lambda) \right| \rho(\lambda)\, d\lambda. \qquad (8.55)$$

Adding the last two equations, noting that $|x - y| + |x + y|$ is either $2x$, or $-2x$, or $2y$, or $-2y$, and is hence definitely less than or equal to 2 as long as $|x| \leq 1$ and $|y| \leq 1$, and using the fact that $\int \rho(\lambda)d\lambda = 1$, we arrive at the CHSH inequality:

$$\left| C(\hat{a}, \hat{b}) - C(\hat{a}, \hat{c}) \right| + \left| C(\hat{a}', \hat{b}) + C(\hat{a}', \hat{c}) \right| \leq 2. \qquad (8.56)$$

And so, any theory that respects Bell's "local causality" condition must make predictions for the correlations in this kind of experiment which respect the inequality. But since the actual experimental data shows a clear violation of the inequality, it follows that all theories which respect Bell's "local causality" condition are *wrong*. The true theory, whatever that is exactly, must violate "local causality". But that is just a complicated way of saying that Nature itself violates local causality, i.e., the faster-than-light causal influences (which "local causality" prohibits) really exist in the world.

That, as we have already acknowledged, is profound and deeply troubling. And this way of arriving at the conclusion should make much clearer that one cannot escape it simply, for example, by upholding the orthodox completeness doctrine and rejecting hidden variables (or determinism or "realism"). This should help you appreciate why, when Bell referred to his own theorem, he (modestly) called it the "locality inequality theorem" [1].

Projects:

8.1 Create your own (preliminary, toy) Bell inequality like the one discussed in Sect. 8.2.
8.2 Use the table from Sect. 8.2 to show that, indeed, Eqs. (8.32) and (8.33) are equivalent.
8.3 Work through and understand all the detailed steps in the derivation, from Sect. 8.6, of the CHSH inequality (many of which are glossed over hastily in the text).
8.4 In Sect. 8.3, we discussed the need to assume that the numbers F_i (characterizing the fraction of particle pairs that are of each possible type) are independent of the axes along which the particle spins will be measured. (This assumption in the derivation of the inequality is then rendered applicable in an ideal experimental test of the inequality by letting the measurement axes be chosen randomly and only after the particle pairs have been emitted.) A similar assumption is made in the derivation of the CHSH inequality in Sect. 8.6, but the terminology is a little different and the assumption was not highlighted in the text. Explain, in the terminology of Sect. 8.6, what this assumption is, and point out the first equation in the derivation which would be invalid without this assumption.

8.5 Read Bell's 1964 paper, Ref. [2]. Summarize his method of deriving an inequality and comment on whether you think he is presenting a proof that the empirical predictions of quantum mechanics are inconsistent with locality, or inconsistent with the joint assumptions of locality and "realism" (i.e., deterministic hidden variables).

8.6 Read the (first) 1982 paper of Aspect et al., Ref. [4]. Summarize their experimental setup and procedure.

8.7 Read the (other) 1982 paper of Aspect et al., Ref. [5]. Summarize the relationship of this experiment to their earlier one, and comment on any other features you find interesting or surprising.

8.8 Read the 1998 paper of Weihs et al., Ref. [6]. Describe what was novel about their experiment (relative to Aspect's 1982 experiments) and summarize their results.

8.9 Read Bell's "Bertlmann's Socks..." paper, Ref. [21]. Summarize Bell's amusing derivation of a locality inequality (in terms of socks and washing machines).

8.10 Your friend is a sociologist doing her senior thesis on the political opinions of twins. She invites pairs of twins to show up and earn $20 by participating in her study. When a pair arrives, she sends the older twin into the room on the left with her assistant, Alice, and sends the younger twin into the room on the right with her other assistant, Bob. The rooms are extremely well soundproofed and the doors are tightly locked after each subject enters his/her room. After the doors are locked, the assistant (Alice or Bob) rolls a 3-sided die to randomly choose one of three pre-determined yes/no questions to ask the subjects. (What the questions are don't matter here, but you could imagine they are something like: Q1 is "Should we raise the minimum wage?", Q2 is "Should the fed raise interest rates?", and Q3 is "Should Roe-vs-Wade be overturned?") The assistants record the subjects' answers, and the whole process is repeated for several hundred pairs of subjects. Afterwards, your friend collects and analyzes all of the data and notices the following:

- When both twins happen to be asked the *same* question, they always answer the same way (either both "yes" or both "no").
- When the older twin is asked Q1 and the younger twin is asked Q2, the answers are (respectively) "yes" and "no" 20% of the time.
- When the older twin is asked Q2 and the younger twin is asked Q3, the answers are (respectively) "yes" and "no" 15% of the time.
- When the older twin is asked Q1 and the younger twin is asked Q3, the answers are (respectively) "yes" and "no" 40% of the time.

What should you advise your friend to conclude?

8.11 Interview some physicists to find out what they think Bell's Theorem is and proves. If they say that Bell's theorem proves you can't have a deterministic/hidden variables theory, you might consider following up by asking them how they reconcile this with the existence of the de Broglie - Bohm pilot-wave theory. You might also consider asking them if they think that ordinary QM (without any hidden variables) is a local theory and, if so, how they recon-

cile this with the perfect EPR correlations (i.e., ask them exactly how, in their understanding, ordinary QM explains these correlations in a local way).

8.12 Recall the assumption, made in the derivation of Bell inequalities, that, in the notation of Sect. 8.6, the probability distribution $\rho(\lambda)$ is independent of the settings $\hat{n}_1$ and $\hat{n}_2$. We have discussed how an ideal experimental test involves randomly choosing the settings "at the last second" in order to ensure that this condition is satisfied, the idea being that then there is no way for the particle source to have "known" about the settings – they didn't even exist yet when the source emitted the particles! But this is a little too quick. Sometimes people imagine that the settings could actually be determined, directly, by some kind of "free will choices" made at the last second by Alice and Bob. And depending on whether one believes in, and/or how one understands, "free will", that might indeed ensure that $\rho(\lambda)$ is independent of the settings. But nobody has ever performed an experiment like that; the real experiments use various sorts of random-number generators to determine the settings. But (again, depending on what sort of random number generator is used, exactly, and perhaps depending on whether one believes in "hidden variables") the outputs of random number generators are not actually random. In principle, the outputs are determined by something, which was in turn determined by something else, and so on into the past. Sketch a space-time diagram to make it clear how, at least in principle, both the settings $\hat{n}_1$ and $\hat{n}_2$ and the state λ of the particle pair could all be influenced/determined by something in their overlapping past light cones, and could therefore be correlated (such that $\rho(\lambda)$ is different for different values of $\hat{n}_1$ and $\hat{n}_2$) without any funny-business like non-locality or backwards-in-time causation. How seriously do you think this possibility should be taken? It may be helpful to think about the extent to which a similar independence assumption is needed in other scientific experiments that have nothing to do with quantum mechanics, e.g., a controlled drug trial in which patients are randomly assigned to receive either the drug or a placebo.[3]

8.13 Read one (or more!) of the three recent experimental papers reporting improved tests of Bell's inequalities, Refs. [7–9]. There are at least two interesting things to pay attention to in these papers. First, how do their experiments work and how do they represent improvements over the earlier experiments of Aspect (et al.) and Weihs (et al.)? And second, how do the authors talk about what their experiments show? That is, do they regard them as refuting hidden variables, or proving non-locality, or what exactly?

---

[3] Just to give you a sense of the spectrum of views which exist on this issue, the assumption – that $\rho(\lambda)$ is independent of the settings $\hat{n}_1$ and $\hat{n}_2$ – has been called the "no conspiracies" assumption, with the implication that you'd have to be a crazy conspiracy theorist to take it seriously; on the other hand, the Nobel Prize winning particle physicist Gerard 't Hooft, among others, thinks that relativity and quantum theory can and should be reconciled by denying that this assumption applies to the real experiments.

References

1. Bell, Preface to the first edition, in *Speakable and Unspeakable in Quantum Mechanics*, 2nd ed. (Cambridge, 2004)
2. Bell, On the Einstein Podolsky Rosen paradox. Physics **1**, 195–200 (1964). (Reprinted in *Speakable and Unspeakable in Quantum Mechanics*, 2nd ed. (Cambridge, 2004))
3. H. Clauser, Shimony, Holt, Proposed experiment to test local hidden-variable theories. Phys. Rev. Lett. **23**(15), 880–884 (1969)
4. Aspect, Grangier, Roger, Experimental realization of Einstein-Podolsky-Rosen-Bohm *Gedankenexperiment*: a new violation of Bell's inequalities. Phys. Rev. Lett. **49**(2), 91–94 (1982)
5. Aspect, Dalibard, Roger, Experimental test of Bell's inequalities using time-varying analyzers. Phys. Rev. Lett. **49**(25), 1804–1807 (1982)
6. Weihs, Jennewein, Simon, Weinfurter, Zeilinger, Violation of Bell's inequality under strict Einstein locality conditions. Phys. Rev. Lett. **81**, 5039 (1998)
7. Hensen et al., Loophole-free Bell inequality violation using electron spins separated by 1.3 km. Nature **526**, 682–686 (2015)
8. Giustina et al., Significant-loophole-free test of Bell's theorem with entangled photons. Phys. Rev. Lett. **115**, 250401 (2015)
9. Shalm et al., A strong loophole-free test of local realism. Phys. Rev. Lett. **115**, 250402 (2015)
10. Wigner, Interpretation of quantum mechanics, in *Quantum Theory and Measurement*, ed. by Wheeler and Zurek (Princeton University Press, Princeton, 1983)
11. Interview with R. Peierls, in Davies and Brown, in *The Ghost in the Atom* (Cambridge, Cambridge University Press, 1993)
12. Hawking, Does god play dice? (1999), http://www.hawking.org.uk/does-god-play-dice.html
13. Kleppner and Jackiw, One hundred years of quantum physics, in *Science*, August 11, 2000
14. Tegmark and Wheeler, 100 years of the quantum, in *Scientific American*, February 2001
15. Bertlmann and Zeilinger, *Preface to Quantum [Un]speakables: From Bell to QUantum Information* (Springer, Berlin, 2002)
16. Mermin, Hidden variables and the two theorems of John Bell. Rev. Modern Phys. **65**, 803–815 (1993)
17. Mermin, Is the moon there when nobody looks? Reality and the quantum theory. Phys. Today (1985)
18. Bell, Speakable and unspeakable in quantum mechanics, in *Speakable and Unspeakable in Quantum Mechanics*, 2nd ed. (Cambridge, Cambridge University Press, 2004)
19. Bell, Quantum mechanics for cosmologists, in *Speakable and Unspeakable in Quantum Mechanics*, 2nd ed. (Cambridge, Cambridge University Press, 2004)
20. Interview with J.S. Bell, in Davies and Brown, *The Ghost in the Atom* (Cambridge, Cambridge University Press, 1993)
21. Bell, Bertlmann's socks and the nature of reality, in *Speakable and Unspeakable in Quantum Mechanics*, 2nd ed. (Cambridge, Cambridge University Press, 2004)
22. See, for example, H. Wiseman, The two Bell's theorems of John Bell. J. Phys. A Math. Theor. **47**, 424001 (2014), and T. Norsen, Are there really two different Bell's theorems? http://www.ijqf.org/archives/1646
23. Bell, La Nouvelle Cuisine, in *Speakable and Unspeakable in Quantum Mehanics*, 2nd ed. (Cambridge, Cambridge University Press, 2004)