Chapter 24

EPR Paradox and Bell Inequalities

24.1 Bohm Version of the EPR Paradox

Einstein, Podolsky, and Rosen (EPR) were concerned with the following issue. Given two spatially separated quantum systems A and B and an appropriate initial entangled state, a measurement of a property on system A can be an indirect measurement of B in the sense that from the outcome of the A measurement one can infer with probability 1 a property of B, because the two systems are correlated. There are cases in which either of two properties of B represented by noncommuting projectors can be measured indirectly in this manner, and EPR argued that this implied that system B could possess two incompatible properties at the same time, contrary to the principles of quantum theory.

In order to understand this argument, it is best to apply it to a specific model system, and we shall do so using Bohm's formulation of the EPR paradox in which the systems A and B are two spin-half particles a and b in two different regions of space, with their spin degrees of freedom initially in a spin singlet state (23.2). As an aid to later discussion, we write the argument in the form of a set of numbered assertions leading to a paradox: a result which seems plausible, but contradicts the basic principles of quantum theory. The assertions E1 to E4 are not intended to be exact counterparts of statements in the original EPR paper, even when the latter are translated into the language of spin-half particles. However, the general idea is very similar, and the basic conundrum is the same.

E1. Suppose S_{az} is measured for particle a. The result allows one to predict S_{bz} for particle b, since $S_{bz} = -S_{az}$.

E2. In the same way, the outcome of a measurement of S_{ax} allows one to predict S_{bx} since $S_{bx} = -S_{ax}$.

E3. Particle b is isolated from particle a, and therefore it cannot be affected by measurements carried out on particle a.

E4. Consequently, particle b must simultaneously possess values for both S_{bz} and S_{bx}, namely the values revealed by the corresponding measurements on particle a, either of which could be carried out in any given experimental run.

E5. But this contradicts the basic principles of quantum theory, since in the two-dimensional spin space one cannot simultaneously assign values of both S_z and S_x to particle b.

Let us explore the paradox by asking how each of these assertions is related to a precise quantum mechanical description of the situation. We begin with E1, and employ the notation in Sec. 23.4, with the particles initially in a spin singlet state $|\psi_0\rangle$, and an apparatus designed to measure S_{az} initially in the state $|Z_a^\circ\rangle$ at time t_0. The interaction of particle a with the apparatus during the time interval from t_1 to t_2 gives rise to the unitary time transformation (23.21). We then need a consistent family which includes the possible outcomes Z_a^+ and Z_a^- of the measurement, corresponding to $S_{az} = +1/2$ and $-1/2$, together with the values of S_{bz}.

It is useful to begin with the family in (23.29), since it comes the closest among all the families in Sec. 23.4 to representing how physicists would have thought about the problem in 1935, when the EPR paper was published. In this family the initial state evolves unitarily until after the measurement has occurred, when there is a split (or "collapse") into the two possibilities $Z_a^+ z_b^-$ and $Z_a^- z_b^+$. Using this family one can deduce $S_{bz} = -1/2$ from the measurement outcome Z_a^+, and $S_{bz} = +1/2$ from Z_a^-; the results can be expressed formally as conditional probabilities, (23.25). This means that E1 is in agreement with the principles of quantum theory.

Even stronger results can be obtained using the family (23.22) in which the stochastic split takes place at an earlier time. In this family it is possible to view the measurement of S_{az} as revealing a pre-existing property of particle a at a time before the measurement took place, a value which was already the opposite of S_{bz}. In addition, the value of S_{bz} was unaffected by the measurement of S_{az}, a fact expressed formally by the conditional probabilities in (23.26). Thus this family both confirms E1 and lends support to E3. Additional support for E3 comes from the family (23.31), which shows that a measurement of S_{az} does not have any effect upon S_{bx}, and of course one could set up an analogous family using any other component of spin of particle b, and reach the same conclusion.

Next we come to E2. It is nothing but E1 with S_z replaced by S_x for both particles, so the preceding discussion of E1 will apply to E2, with obvious modifications. The family (23.28) with its apparatus for measuring S_{ax} must be used in place of (23.22), and from it one can deduce the counterparts of (23.23) to (23.26) with z and Z replaced by x and X. And of course the S_{ax} measurement will not alter any component of the spin of particle b, which confirms E3.

Assertion E4 would seem to be an immediate consequence of those preceding it were it not for the requirement that quantum reasoning employ a *single* framework in order to reach a sound conclusion, Sec. 16.1. Assertions E1 and E2 have been justified on the basis of two distinct consistent families, (23.22) and (23.28). Are these families compatible, i.e., can they be combined in a single framework? One's first thought is that they cannot be combined, because the projectors for the properties associated with S_{az} and S_{bz} at t_1 (the intermediate time) in (23.22) obviously do not commute with those in (23.28), which are associated with S_{ax} and S_{bx}, and the same is true of the projectors at t_2. However, the situation is not so simple. The projectors representing the complete histories in (23.22) are orthogonal to, and hence commute with, the history projectors in (23.28), because the initial states $|Z_a^\circ\rangle$ and $|X_a^\circ\rangle$ for the apparatus will be orthogonal. This follows from the fact that an apparatus designed to measure S_z will differ in a visible (macroscopic) way from one designed to measure S_x; see the discussion following (17.10)

Consequently, (23.22) and (23.28) can be combined in a single consistent family with two distinct initial states: the spin singlet state of the particles combined with either of the measuring apparatuses. However, the resulting framework does *not* support E4. The reason is that the two initial states are mutually exclusive, so that only one or the other will occur in a particular experimental run. Consequently, the conclusion that S_{bz} will have a particular value, at t_1 or t_2, as determined by the measurement outcome, is only correct for a run in which the apparatus is set up to measure S_{az}, and the corresponding conclusion for S_{bx} only holds for runs in which the apparatus is set up to measure S_{ax}. But E4 asserts that particle b *simultaneously* possesses values of S_z and S_x, and this conclusion obviously cannot be reached using the framework under consideration.

To put the matter in a different way, E1 is correct in a situation in which S_{az} is measured, and E2 in a situation in which S_{ax} is measured. But there is no way to measure S_{az} and S_{ax} simultaneously for a single particle, and therefore no situation in which E1 and E2 can be applied to the same particle. Einstein, Podolsky and Rosen were aware of this type of objection, as they mention it towards the end of their paper, and they respond in a fashion which can be translated into the language of spin-half particles in the following way. If one allows that an S_{az} measurement can be used to predict S_{bz} and an S_{ax} measurement to predict S_{bx}, but then asserts that S_{bx} does not exist when S_{az} is measured, and S_{bz} does not exist when S_{ax} is measured, this makes the properties of particle b depend upon which measurement is carried out on particle a, and no reasonable theory could allow this sort of thing.

There is nothing in the analysis presented in Sec. 23.4 to suggest that the properties of particle b depend in any way upon the type of measurement carried out on particle a. However, the type of property considered for particle b, S_{bz} as against S_{bx}, depends upon the choice of framework. There are frameworks, such as (23.22) and (23.29), in which a measurement of S_{az} is combined with values of S_{bz}, and other frameworks, such as (23.30) and (23.31), in which a measurement of S_{az} is combined with values for S_{bx}. Quantum theory does not specify which framework is to be used for a situation in which S_{az} is measured. However, only a framework which includes S_{bz} can be used to correlate the outcome of an S_{az} measurement with some property of the spin of particle b in a way which constitutes an indirect measurement of the latter.

Thus implicit in the analysis given in the EPR paper is the assumption that quantum theory is limited to a single framework in the case of an S_{az} measurement, one corresponding to a wave function collapse picture, (23.29), for this particular measurement. Once one recognizes that there are many possible frameworks, the argument no longer works. One can hardly fault Einstein and his colleagues for making such an assumption, as they were seeking to point out an inadequacy of quantum mechanics as it had been developed up to that time, with measurement and wave function collapse essential features of its physical interpretation. One can see in retrospect that they had, indeed, located a severe shortcoming of the principal interpretation of quantum theory then available, though they themselves did not know how to remedy it.

24.2 Counterfactuals and the EPR Paradox

An alternative way of thinking about assertion E4 in the previous section is to consider a case in which S_{az} is measured (and thus S_{bz} is indirectly measured), and ask what would have been the case, in this particular experimental run, if S_{ax} had been measured instead, e.g., by rotating the direction of the field gradient in the Stern-Gerlach apparatus just before the arrival of particle a. This requires a counterfactual analysis, which can be carried out with the help of a quantum coin toss in the manner indicated in Sec. 19.4. Let the total quantum system be described by an initial state

$$|\Phi_0\rangle = |\psi_0\rangle|Q\rangle, \tag{24.1}$$

where $|\psi_0\rangle$ is the spin singlet state (23.2), and $|Q\rangle$ the initial state of the quantum coin, servomechanism, and the measuring apparatus. (As it is not important for the following discussion, the center of mass wave function $|\omega_t\rangle$, (23.1), has been omitted, just as in Ch. 23.) It will be convenient to assume that the quantum coin toss corresponds to a unitary time development

$$|Q\rangle \mapsto \bigl(|X_a^\circ\rangle + |Z_a^\circ\rangle\bigr)/\sqrt{2}, \tag{24.2}$$

during the interval from t_1 to t_2, and that the measurement of S_{ax} or S_{az} takes place during the time interval from t_2 to t_3, rather than between t_1 and t_2 as in Ch. 23. Here $|X_a^\circ\rangle$ and $|Z_a^\circ\rangle$ are states of the apparatus in which it is ready to measure S_{ax} and S_{az}, respectively, and the servomechanism, etc., is thought of as included in these states. Thus the overall unitary time development from the initial time t_0 to the final time t_3 is given by

$$\begin{gathered}|\Phi_0\rangle \mapsto |\Phi_0\rangle \mapsto |\psi_0\rangle\bigl(|X_a^\circ\rangle + |Z_a^\circ\rangle\bigr)/\sqrt{2} \mapsto \\ \bigl(|x_b^-\rangle|X_a^+\rangle - |x_b^+\rangle|X_a^-\rangle + |z_b^-\rangle|Z_a^+\rangle - |z_b^+\rangle|Z_a^-\rangle\bigr)/2.\end{gathered} \tag{24.3}$$

The final step from t_2 to t_3 is obtained by assuming that (23.27) applies when the apparatus is in the state $|X_a^\circ\rangle$, and (23.21) when it is in the state $|Z_a^\circ\rangle$ at t_2.

A consistent family $\mathcal{F}_1$ which provides one way of analyzing the counterfactual question posed at the beginning of this section has for its support six histories for times $t_0 < t_1 < t_2 < t_3$. It is convenient to arrange them in two groups of three:

$$\Phi_0 \odot z_a^+ z_b^- \odot \begin{cases} Z_a^\circ \odot \quad Z_a^+ z_b^-, \\ X_a^\circ \odot \begin{cases} X_a^+ z_b^-, \\ X_a^- z_b^-, \end{cases}\end{cases} \quad \Phi_0 \odot z_a^- z_b^+ \odot \begin{cases} Z_a^\circ \odot \quad Z_a^- z_b^+, \\ X_a^\circ \odot \begin{cases} X_a^+ z_b^+, \\ X_a^- z_b^+. \end{cases}\end{cases} \tag{24.4}$$

Suppose the coin toss resulted in S_{az} being measured, and the outcome was Z_a^+, implying $S_{bz} = -1/2$. To answer the question of what would have happened if S_{ax} had been measured instead, use the procedure of Sec. 19.4 and trace the outcome $Z_a^+ z_b^-$ in the first set of histories in (24.4) backwards to the pivot $z_a^+ z_b^-$ and then forwards through the X_a° node to the corresponding events at t_3. One concludes that had the quantum coin toss resulted in a measurement of S_{ax}, the outcome would have been X_a^+ or X_a^-, each with probability 1/2, but *in either case* S_{bz} would have had the value $-1/2$, corresponding to z_b^-, that is to say, the same value it had in the actual world in which S_{az}, not S_{ax}, was measured. This conclusion seems very reasonable on physical grounds, for one would not expect a last minute choice to measure S_x rather than S_z for particle a to have any influence on the distant particle b, since the measuring apparatus does not interact in any way with particle b. To put the matter in another way, the conclusion of this counterfactual analysis agrees with the discussion of E3 in Sec.24.1.

On the other hand, (24.4) by itself provides no immediate support for E4, for it supplies no information at all about S_{bx}. Of course, this is only one consistent family, and one might hope to do better using some other framework. One possibility might be the consistent family $\mathcal{F}_2$ with support

$$\Phi_0 \odot \Phi_0 \odot \begin{cases} Z_a^\circ \odot \begin{cases} Z_a^+ z_b^-, \\ Z_a^- z_b^+, \end{cases} \\ X_a^\circ \odot \begin{cases} X_a^+ x_b^-, \\ X_a^- x_b^+, \end{cases} \end{cases} \tag{24.5}$$

which corresponds pretty closely to the notion of wave function collapse. Once again assume that the quantum coin toss leads to an S_{az} measurement, and that the outcome of this measurement is Z_a^+. Using Φ_0 at t_0 or t_1 as the pivot, one concludes that had S_{ax} been measured instead, S_{bx} would have been $-1/2$ for the outcome X_a^+, and $+1/2$ for the outcome X_a^-.

This result seems encouraging, for we have found a consistent family in which both S_{bz} and S_{bx} values appear, correlated in the expected way with S_{az} and S_{ax} measurements. However the S_{bz} states $z_b^\pm$ and the S_{bx} states $x_b^\pm$ in (24.5) are *contextual* properties in the sense of Ch. 14: z_b^+ and z_b^- both depend on Z_a°, and x_b^+ and x_b^- both depend on X_a°. This means—see the discussion in Ch. 14—that when using (24.5), one cannot think of S_{bz} and S_{bx} as having values independent of the quantum coin toss. Only if the toss results in Z_a° is it meaningful to talk about S_{bz}, and only if it results in X_a° can one talk about S_{bx}. And since the two outcomes of the quantum coin toss are mutually exclusive possibilities, one and only one of which will occur in any given experimental run, we have again failed to establish E4, and for basically the same reason pointed out in Sec.24.1 when discussing the family with two initial states that combines (23.22) and (23.28). Indeed, in the latter family S_{bz} and S_{bx} are contextual properties which depend upon the corresponding initial states—something we did not bother to point out in Sec.24.1 because dependence (in the technical sense used in Ch. 14) on *earlier* events never poses much of an intuitive problem. But does this contextuality mean that there is some mysterious long-range influence in that a last minute choice to measure S_{ax} rather than S_{az} would somehow determine whether particle b has a definite value of S_x rather that S_z? No, for dependence or contextuality in the technical sense used in Ch. 14 denotes a logical relationship brought about by choosing a framework in a particular way, and does not indicate any sort of physical causality. Thus there is no contradiction with the arguments presented in Sec.24.1 in support of E3.

The reader with the patience to follow the analysis in this and the previous section may with some justification complain that the outcome was already certain at the outset: if E4 really does contradict the basic principles of quantum theory, as asserted by E5, then it is evident that it can never be obtained by an analysis based upon those principles. True enough, but there are various reasons why working out the details is still worthwhile. First, there is no way to establish with absolute certainty the consistency of the basic principles of a physical theory, as it is always something more than a piece of abstract mathematics or logic; one has to apply these principles to various examples and see what they predict. Second, it is of some interest to find out where and why the seemingly plausible chain of arguments from E1 to E5 comes apart, for this tells us something about the difference between quantum and classical physics. The preceding analysis shows that it is basically violations of the single framework rule which cause the trouble, and in this respect the EPR paradox has quite a bit in common with the paradoxes discussed in previous chapters. But the non-classical behavior of contextual events can also play a role, depending on how one analyzes the paradox.

Third, the analysis supports the correctness of the basic *locality* assumption of EPR as expressed in E3, an assertion which is confirmed by the analysis in Ch. 23. Given that the EPR paradox has sometimes been cited to support the claim that there are mysterious nonlocal influences in the quantum world, it is worth emphasizing that the analysis given here does not show any evidence of such influences. On the other hand, certain *modifications* of quantum mechanics in which the Hilbert space is supplemented by "hidden variables" of a particular sort will necessarily involve peculiar nonlocal influences if they are to reproduce the spin correlations (23.9) of standard quantum theory, and these are the subject of the remaining sections in this chapter.

24.3 EPR and Hidden Variables

A *hidden variable* theory is an alternative approach to quantum mechanics in which the Hilbert space of the standard theory is either replaced by or supplemented with a set of "hidden" (the name is not particularly apt) variables which behave like those one is accustomed to in classical mechanics. One of the best-known examples was proposed in 1952 by Bohm, using an approach similar to one employed earlier by de Broglie, in which at any instant of time all particles have precise positions, and these positions constitute the new (hidden) variables.

The simplest hidden variable model of a spin-half particle is one in which the different components of its spin angular momentum simultaneously possess well-defined values, something which is not true if one uses a quantum Hilbert space, for reasons discussed in Sec. 4.6. A measurement of some component of spin using a Stern-Gerlach apparatus will then reveal the value that the corresponding (hidden) variable had just before the measurement took place. More complicated models are possible, but the general idea is that measurement outcomes are determined by variables that behave classically in the sense that they simultaneously possess definite values. John Bell pointed out in 1964 that hidden variable models of this kind cannot reproduce the correlation function $C(\boldsymbol{a}, \boldsymbol{b})$, (23.9) or (23.37), for spin-half particles in an initial singlet state, if one makes the reasonable assumption that no mysterious long-range influences link the particles and the measuring apparatuses. This result led to a number of experimental measurements of the spin correlation function. Most of the experiments have used the polarizations of correlated photons rather than spin-half particles, but the principles are the same, and the results are in good agreement with the predictions of quantum mechanics. Note that one can think of this correlation function as referring to particle spins in the absence of any measurement when one uses the framework (23.5), or as the correlation function between outcomes of measurements of the spins of both particles, (23.37). In line with most discussions of Bell's result, we shall think of $C(\boldsymbol{a}, \boldsymbol{b})$ as referring to measurement outcomes.

Before exhibiting one version of Bell's argument in Sec. 24.4 below, it is useful to look at a specific setup discussed by Mermin. Imagine two apparatuses, one to measure the spin of particle a and the other the spin of particle b, each of which can measure the component of spin angular momentum in one of three directions in space, u, v, and x, lying in the x, y plane, with an angle of 120° between every pair of directions, Fig. 24.1. The component of spin which will be measured is determined by a switch setting on the apparatus, and these settings will also be denoted by u, v, and x. Let $\alpha(w) = \pm 1$ denote the two possible outcomes of the measurement when the switch setting of the a apparatus is w: $+1$ if the spin is found to be in the $+w$ direction, $S_{aw} = +1/2$, and -1 if it is in the opposite direction, $S_{aw} = -1/2$. Let $\beta(w) = \pm 1$ be the possible outcomes of the b apparatus measurement when its switch setting is w. In any given experiment these results will be random, but if they are averaged over a large number of runs, the averages of $\alpha(w)$ and of $\beta(w)$ will be zero for any choice of w, whereas the correlation function (23.37) will be given by:

$$C(w_a, w_b) = \langle \alpha(w_a)\beta(w_b) \rangle = \begin{cases} -1 & \text{if } w_a = w_b, \\ +1/2 & \text{if } w_a \neq w_b, \end{cases} \tag{24.6}$$

since if the switch settings w_a and w_b for the a and b apparatuses are unequal, the angle between the two directions is 120°, and the inner product of the two corresponding unit vectors is $-1/2$.

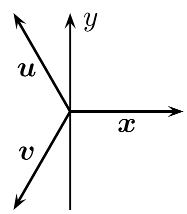

Figure 24.1: Directions u, v, x in the x, y plane.

Let us try and construct a hidden variable model which can reproduce the correlation function (24.6). Suppose that particle a when it leaves the source which prepares the two particles in a singlet state contains an "instruction set" which will determine the outcomes of the measurements in each of the three directions u, v, and x. For example, if the particle carries the instruction set $(+1, +1, -1)$, a measurement of S_{au} will yield the result $+1/2$, a measurement of S_{av} will also yield $+1/2$, and one of S_{ax} will yield $-1/2$. Of course, only one of these measurements will actually be carried out, the one determined by the switch setting on the apparatus when particle a arrives. Whichever measurement it may be, the result is determined ahead of time by the particle's instruction set. One can think of the instruction set as a list of the components of spin angular momentum in each of the three directions, in units of $\hbar/2$. This is what is called a "deterministic hidden variable" model because the instruction set, which constitutes the hidden variables in this model, determines the later measurement outcome without any extra element of randomness. It is possible to construct stochastic hidden variable models, but they turn out to be no more successful than deterministic models in reproducing the correlations predicted by standard quantum theory.

There are eight possible instruction sets for particle a and eight for particle b, thus a total of sixty-four possibilities for the two particles together. However, the perfect anticorrelation when $w_a = w_b$ in (24.6) can only be achieved if the instruction set for b is the *complementary* set to that of a, obtained by changing the sign of each instruction. If the a set is $(+1, +1, -1)$, the b set must be $(-1, -1, +1)$. For were the b set something else, say $(+1, -1, +1)$, then there would be identical switch settings, in this case $w_a = w_b = u$, leading to $\alpha(u) = \beta(u)$, which is not possible. Similarly, perfect anticorrelations for equal switch settings means that the instruction sets, once prepared at the source which produces the singlet state, cannot change in a random manner as a particle moves from the source to the measuring apparatus.

We will assume that the source produces singlet pairs with one of the eight instruction sets for a, and the complementary set for b, chosen randomly with a certain probability. Let $P_a(++-)$

denote the probability that the instruction set for a is $(+1, +1, -1)$. The correlation functions can be expressed in terms of these probabilities; for example,

$$\begin{aligned} C(u, v) = C(v, u) = &-P_a(+++) - P_a(++-) \\ &+ P_a(+-+) + P_a(-++) + P_a(+--) \\ &+ P_a(-+-) - P_a(--+) - P_a(---). \end{aligned} \tag{24.7}$$

Consider the following sum of correlation functions calculated in this way:

$$\begin{aligned} C(u, v) + C(u, x) + C(x, v) = \\ -3P_a(+++) - 3P_a(---) + P_a(++-) + P_a(+-+) \\ + P_a(-++) + P_a(+--) + P_a(-+-) + P_a(--+). \end{aligned} \tag{24.8}$$

Since the probabilities of the different instruction sets add to 1, this quantity has a value lying between -3 and $+1$. However, if we use the quantum mechanical values (24.6) for the correlation functions, the left side of (24.8) is 3/2, substantially greater than 1. Thus our hidden variable model cannot reproduce the correlation functions predicted by quantum theory. As we shall see in the next section, this failure is not an accident; it is something which one must expect in hidden variable models of this sort.

24.4 Bell Inequalities

The inequality (24.10) below was derived in 1969 by Clauser, Horne, Shimony, and Holt. As it is closely related to Bell's original result in 1964, this CHSH inequality is nowadays also referred to as a "Bell inequality", and by studying it one can learn the essential ideas behind such inequalities. We assume that when the a apparatus measures a spin component in the direction w_a, the outcome is given by a function $\alpha(w_a, \lambda) = \pm 1$ which depends both on w_a and a hidden variable, or collection of hidden variables, denoted by λ. Similarly, the outcome of the b measurement for a spin component in the direction w_b is given by a function $\beta(w_b, \lambda) = \pm 1$. In the example in Sec. 24.3, w_a and w_b can take on any of the three values u, v, or x, and λ should be thought of as the pair of instruction sets for *both* particles a and b. Hence λ could take on 64 different values, though we argued in Sec. 24.3 that the probabilities of all but 8 of these must be zero. For the purpose of deriving the inequality, one need not think of w_a as a direction in space; it can simply be some sort of switch setting on the a apparatus, which, together with the value of the hidden variable λ associated with the particle, determines the outcome of the measurement through the function $\alpha(w_a, \lambda)$. The same remark applies to the b apparatus and the function $\beta(w_b, \lambda)$. Also, the derivation makes no use of the fact that the two spin-half particles are initially in a spin singlet state.

The source which produces the correlated particles produces different possible values of λ with a probability $\rho(\lambda)$, so the correlation function is given by

$$C(w_a, w_b) = \sum_\lambda \rho(\lambda)\, \alpha(w_a, \lambda)\, \beta(w_b, \lambda). \tag{24.9}$$

(If λ is a continuous variable, $\sum_\lambda \rho(\lambda)$ should be replaced by $\int \rho(\lambda)\, d\lambda$.) Let a, a' be any two possible values for w_a, and b and b' any two possible values for w_b. Then as long as $\alpha(w_a, \lambda)$ and

$\beta(w_b, \lambda)$ are functions which take only the two values $+1$ or -1, the correlations defined by (24.9) satisfy the inequality

$$|C(a,b) + C(a,b') + C(a',b) - C(a',b')| \leq 2. \tag{24.10}$$

To see that this is so, consider the quantity

$$\begin{aligned} &\alpha(a,\lambda)\,\beta(b,\lambda) + \alpha(a,\lambda)\,\beta(b',\lambda) + \alpha(a',\lambda)\,\beta(b,\lambda) - \alpha(a',\lambda)\,\beta(b',\lambda) \\ &\quad = \big[\alpha(a,\lambda) + \alpha(a',\lambda)\big]\,\beta(b,\lambda) + \big[\alpha(a,\lambda) - \alpha(a',\lambda)\big]\,\beta(b',\lambda). \end{aligned} \tag{24.11}$$

It can take on only two values, $+2$ and -2, because each of the four quantities $\alpha(a,\lambda)$, $\alpha(a',\lambda)$, $\beta(b,\lambda)$ and $\beta(b',\lambda)$ is either $+1$ or -1. Thus either $\alpha(a,\lambda) = \alpha(a',\lambda)$, so that the right side of (24.11) is $2\alpha(a,\lambda)\beta(b,\lambda)$, or else $\alpha(a,\lambda) = -\alpha(a',\lambda)$, in which case it is $2\alpha(a,\lambda)\beta(b',\lambda)$. If one multiplies (24.11) by $\rho(\lambda)$ and sums over λ, the result of this weighted average is

$$C(a,b) + C(a,b') + C(a',b) - C(a',b'). \tag{24.12}$$

A weighted average of a quantity which takes on only two values must lie between them, so (24.12) lies somewhere between -2 and $+2$, which is what (24.10) asserts.

Consider the example in Sec. 24.3, and set $a = u$, $b = v$, $a' = b' = x$. If one inserts the quantum values (24.6) for these correlation functions in (24.12), the result is $3 \times 1/2 + 1 = 2.5$, which obviously violates the inequality (24.10). On the other hand, the hidden variable model in Sec. 24.3 assigns to the sum $C(u,v) + C(u,x) + C(x,v)$, see (24.8), a value between -3 and $+1$, and since $C(x,x) = -1$, the inequality (24.10) will be satisfied.

If quantum theory is a correct description of the world, then since it predicts correlation functions which violate (24.10), one or more of the assumptions made in the derivation of this inequality must be wrong. The first and most basic of these assumptions is the *existence of hidden variables* with a mathematical structure which differs from the Hilbert space used in standard quantum mechanics. This assumption is plausible from the perspective of classical physics if measurements reveal pre-existing properties of the measured system. In quantum physics it is also the case that a measurement reveals a pre-existing property *provided* this property is part of the framework which is being used to construct the quantum description. If S_{az} is measured for particle a, the outcome of a suitable (ideal) measurement will be correlated with the value of this component of spin angular momentum before the measurement in a framework which includes $|z_a^+\rangle$ and $|z_a^-\rangle$. However, there is no framework which includes the eigenstates of *both* S_{az} *and* S_{aw} for a direction w not equal to z or $-z$.

Thus the point at which the derivation of (24.10) begins to deviate from quantum principles is in the assumption that a function $\alpha(w_a,\lambda)$ exists *for different directions* w_a. As long as only a single choice for w_a is under consideration there is no problem, for then the "hidden" variable λ can simply be the value of S_{aw} at some earlier time. But when two (excluding the trivial case of w_a and $-w_a$) or even more possibilities are allowed, the assumption that $\alpha(w_a,\lambda)$ exists is in conflict with basic quantum principles. Precisely the same comments apply to the function $\beta(w_b,\lambda)$.

Of course, if postulating hidden variables is itself in error, there is no need to search for problems with the other assumptions having to do with the nature of these hidden variables. Nonetheless, let us see what can be said about them. A second assumption entering the derivation of (24.10) is that the hidden variable theory is *local.* Locality appears in the assumption that the outcome

$\alpha(w_a, \lambda)$ of the a measurement depends on the setting w_a of this piece of apparatus, but not the setting w_b for the b apparatus, and that $\beta(w_b, \lambda)$ does not depend upon w_a. These assumptions are plausible, especially if one supposes that the particles a and b and the corresponding apparatuses are far apart at the time when the measurements take place. For then the settings w_a and w_b could be chosen at the very last moment before the measurements take place, and it is hard to see how either value could have any influence on the outcome of the measurement made by the other apparatus. Indeed, for a sufficiently large separation, an influence of this sort would have to travel faster than the speed of light, in violation of relativity theory.

The claim is sometimes made that quantum theory must be nonlocal simply because its predictions violate (24.10). But this is not correct. First, what follows logically from the violation of this inequality is that hidden variable theories, if they are to agree with quantum theory, must be nonlocal or embody some other peculiarity. But hidden variable theories by definition employ a different mathematical structure from (or in addition to) the quantum Hilbert space, so this tells us nothing about standard quantum mechanics. Second, the detailed quantum analysis of a spin singlet system in Ch. 23 shows no evidence of nonlocality; indeed, it demonstrates precisely the opposite: the spin of particle b is not influenced in any way by the measurements carried out on particle a. (To be sure, in Ch. 23 we did not discuss how a measurement on particle a might influence the outcome of a *measurement* on particle b, but the argument can be easily extended to include that case, and the conclusion is exactly the same.) Hidden variable theories, on the other hand, can indeed be nonlocal. The Bohm theory mentioned in Sec. 24.3 is known to be nonlocal in a rather thorough-going way, and this is one reason why it has been difficult to construct a relativistic version of it.

A third assumption which was made in deriving the inequality (24.10) is that the probability distribution $\rho(\lambda)$ for the hidden variable(s) λ does not depend upon either w_a or w_b. This seems plausible if there is a significant interval between the time when the two particles are prepared in some singlet state by a source which sets the value of λ, and the time when the spin measurements occur. For w_a and w_b could be chosen just before the measurements take place, and this choice should not affect the value of λ determined earlier, unless the future can influence the past.

In summary, the basic lesson to be learned from the Bell inequalities is that it is difficult to construct a plausible hidden variable theory which will mimic the sorts of correlations predicted by quantum theory and confirmed by experiment. Such a theory must either exhibit peculiar nonlocalities which violate relativity theory, or else incorporate influences which travel backwards in time, in contrast to everyday experience. This seems a rather high price to pay just to have a theory which is more "classical" than ordinary quantum mechanics.