1

Bell's Theorem: The Price of Locality

According to our naive, everyday conception, and even according to most of our refined theories, the physical world is composed of separate individually existing objects. The book on my desk sits apart from the glass, each constituted separate from the other and with its own intrinsic properties. The book has its mass, shape, number of pages, the marks of its history engraved on it. It is made up of atoms, each with its own physical constitution, tied together by chemical bonds. The glass similarly exists on its own, constructed from a separate complement of particles. There are, of course, relations between the book and the glass. The book is heavier and occupies more volume; there is a certain definite distance between them. Spatial separation plays a unique role: as an external relation it is not determined by any facts about the book and the glass taken individually. But once we have taken into account their intrinsic properties and their situation in space we appear to have exhausted the facts about the pair. All other facts about them are determined by these.

Each of the pair may influence the other. The glass, full of steaming tea, raises the temperature of the book which is in its proximity. But this interaction is mediated by other localized bits of matter. Air molecules around the glass are made more energetic through interactions with the tea, some wander off and communicate their energy with the book, heating it. The book exerts a slight gravitational pull on the glass and vice versa. This is a subtle matter, but we come to think of this too as a mediated interaction, an effect of a gravitational field.

The fields of classical physics are not so familiar as books or atoms but they too are local entities. Although an electric field may spread out and permeate the universe, the state of the field is determined entirely by its value at each point of space. Disturbances propagate through the field, but they do so by local interactions: changes in the field quantities induce other changes nearby and so ripple off to infinity. Like the transmission of heat, this process takes time as the vibrations of the field are passed along.

Einstein set great store by the idea that the physical state of the universe is determined by a set of locally defined physical magnitudes so that the state of any localized entity exists independently of all spatially separated systems. As he expressed it in a letter to Max Born:

> If one asks what, irrespective of quantum mechanics, is characteristic of the world of ideas of physics, one is first of all struck by the following: the concepts of physics relate to a real outside world, that is, ideas are established relating to things such as bodies, fields, etc., which claim "real existence" that is independent of the perceiving subject – ideas which, on the other hand, have been brought into as secure a relationship as possible with the sense-data. It is further characteristic of these physical objects that they are thought of as arranged in a space-time continuum. An essential aspect of this arrangement of things in physics is that they lay claim, at a certain time, to an existence independent of one another, provided these objects "are situated in different parts of space". Unless one makes this kind of assumption about the independence of the existence (the "being-thus") of objects which are far apart from one another in space – which stems in the first place from everyday thinking – physical thinking in the familiar sense would not be possible. It is also hard to see any way of formulating and testing the laws of physics unless one makes a clear distinction of this kind. This principle has been carried to extremes in the field theory by localizing the elementary objects on which it is based and which exist independently of each other, as well as the elementary laws which have been postulated for it, in the infinitely small (four-dimensional) elements of space.
>
> The following idea characterizes the relative independence of objects far apart in space (A and B): external influence on A has no direct influence on B; this is known as the "principle of contiguity," which is used consistently in the field theory. If this axiom were to be completely abolished, the idea of the existence of (quasi-) enclosed systems, and thereby the postulation of laws which can be checked empirically in the accepted sense, would become impossible. (Born 1971, pp. 170–1)

Bell's theorem addresses the implications, and ultimately the tenability, of this picture.

Given the extreme generality of the local conception of reality it is hard to imagine that it could, by itself, have any testable empirical consequences. No constraints have been put on the nature or complexity of the locally defined quantities. The locality condition allows, for example, that every particle in the universe could retain traces of every interaction it has ever undergone. It allows a system to be governed by laws which are deterministic or are probabilistic, placing no limit on the subtlety or sophistication of the laws. Nonetheless, Bell was able to show that some behavior of separated pairs of systems cannot be explained by *any* local physical theory if the systems do not interact. Although Bell's results can be derived in different ways and with great generality, we will begin by focussing on a singular fact about light.

Polarization

When one passes a beam of sunlight through a polarized filter, such as the material used in Polaroid sunglasses, two things happen. First, about half of the light is absorbed and half transmitted, as is immediately evident. Second, the light which is transmitted displays an entirely new and surprising characteristic: it shows a particular directionality. This directionality can be most easily observed if one passes the new beam through a second polarized filter. The effect of the second filter depends critically on its orientation with respect to the first. In one orientation the second polarizer will have no effect at all, allowing the entire beam to pass. But as it is rotated, the second filter allows less and less of the light through. By the time it has been turned 90°, it absorbs the beam entirely; as it is rotated further it permits ever more light to pass until, at 180°, the whole beam passes again.

The directionality that the sunlight acquires depends on the orientation of the first polarizer. When the first filter is rotated, the characteristic orientation at which the transmitted beam passes the second filter rotates with it. So light which has passed through a Polaroid filter acquires a new property, a polarization, which is associated with some direction perpendicular to its line of motion.

All that really concerns us is the behavior recounted above; the explanation of the phenomena will ultimately be irrelevant to our concerns. But to help fix our ideas it may help to recall the classical theory of polarization. The classical theory provides us with a simple picture of polarization which should, however, be taken *cum grano salis*, for it cannot be straightforwardly extended when quantum phenomena are taken into consideration.

According to classical physics, light is an electromagnetic wave, a propagating disturbance of the electric and magnetic fields. The fields that vary always point perpendicular to the direction of motion of the light. At any given moment the electric and magnetic fields are also perpendicular to each other, but as time goes on their direction and magnitude may change in any number of ways. For example, if we look at a ray of light head on as it comes toward us, the electric field may rotate in a circle, either clockwise or counterclockwise (circularly polarized light); or it may trace out an ellipse, rotating and varying in length; or it may simply oscillate back and forth without rotating, always remaining in a plane that points in a given direction. This last possibility, plane polarized light, is the case of interest to us. Plane polarized light has a characteristic direction, the direction of the plane in which the electric field vector always lies. Furthermore, for any direction θ we choose, light of any sort can be analyzed into a component plane polarized in that direction and a component polarized in the perpendicular (that is, $\theta + 90^\circ$) direction. Even circularly polarized light can be constructed from two such elements, if they are added together in the right way, with the right phase relations.

The phenomena recounted above are now easily explained. A Polaroid filter in effect analyzes all incoming light waves into two parts: one plane polarized in the direction of the filter's polarization, the other perpendicular to that direction. It then absorbs the perpendicular component, allowing only the plane polarized remainder to pass through. If the incoming light is unpolarized, this means that on average half of it will pass through and half be absorbed. The effect of the second filter then depends crucially on its orientation relative to the first. If they are perfectly aligned, the light which passes the first is already polarized in the direction of the second and so all gets through. If the second is misaligned by 90°, then exactly the component which passes the first will be absorbed by the second, and none will get through.

What if the two filters are misaligned by some angle α between 0° and 90°? We can represent the light coming through the first filter by a vector pointing in the θ direction whose length A represents the maximum amplitude of the electric field. The second filter resolves this vector into two components, one parallel to $\theta - \alpha$, the other perpendicular (see figure 1.1). The perpendicular component is absorbed by the filter, so the amplitude of the transmitted light is $A \cos \alpha$.

We now must appeal to a seemingly minor but highly significant fact. The energy of plane polarized light is proportional to the *square* of the amplitude of its electric field vector. So if we measure the amount of light which passes the second polarizer by the energy of the beam, the proportion of the beam that gets through is $A^2 \cos^2 \alpha / A^2 = \cos^2 \alpha$. Figure 1.2 shows the proportion of the beam which passes the second filter as a function of the angle of misalignment α.

As expected, when $\alpha = 0^\circ$ and the filters are aligned, all the beam is transmitted. When the filters are misaligned by 90° none of the light gets

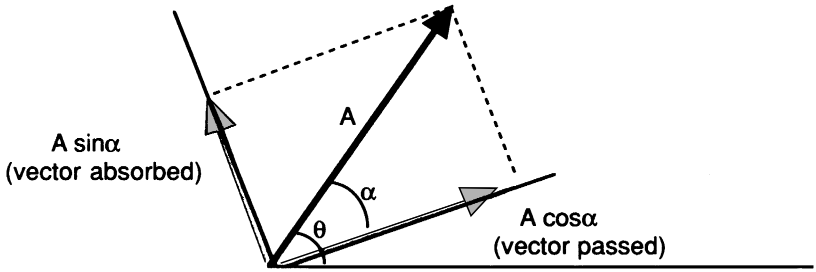

Figure 1.1 Resolving a Vector

through. But the most significant behavior is found between these extremes. For the moment we need only note that when $\alpha = 30^\circ$, $\cos^2 30^\circ = (\sqrt{3}/2)^2 = \frac{3}{4}$ of the beam gets through, while when $\alpha = 60^\circ$, $\cos^2 60^\circ = (\frac{1}{2})^2 = \frac{1}{4}$ of the light is transmitted.

Light Quanta

According to the classical conception light is a wave, spread out in space. Whenever a plane polarized beam impinges on a filter oriented at, say, 30° off of the polarization plane of the incoming beam the same thing happens: $\frac{3}{4}$ of the beam passes, and what gets through is polarized in the

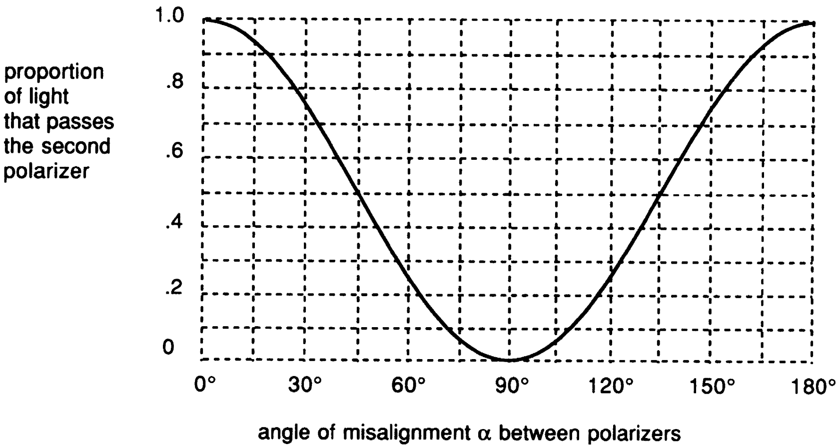

Figure 1.2 Proportion of Light Passing Second Polarizer

direction of the filter. A beam always comes out with its amplitude and energy reduced by a fixed proportion.

But as Einstein observed in 1905, light does not always behave like a wave. For example, when light falls on certain metals it can knock out electrons causing a current to flow, the so-called photoelectric effect. When one measures the energy of the electrons so liberated one finds that the energy of the incident light is not delivered uniformly over the surface of the metal as one would expect. The energy rather comes in small but discrete packets. At fine levels of analysis light behaves as if it is made up of particles. These light quanta, or photons, can be individually registered and counted by photomultiplier tubes.

The exact nature of the wave/particle duality of light need not detain us. We need only note two experimentally verifiable facts. First, light from certain sources has the effect of causing discrete, countable events in certain detection equipment. Second, if this light is passed through a polarizer, the resulting beam also behaves as if made up of photons and *the photons are each of exactly the same energy as those in the incoming beam*.

Nothing we have said so far could have prepared us for this new piece of information. It would have been plausible to guess instead that the photons coming through the polarizer would all have had their energy reduced by the same proportion as the energy of the beam as a whole. But in fact the transmitted photons are as energetic as the incoming ones, only the orientation of their polarization is changed by passage.

If the photons which survive the second polarizer each have the same energy as the incoming light quanta, how is the overall energy of the beam reduced? The only possibility is that the light coming out of the polarizer contains *fewer* photons than the light going in. Photons appear to be either transmitted complete through the filter or else swallowed whole.

It is worthwhile to note that all of this talk about light quanta need not be made precise. We could instead refer only to the observable behavior of pieces of laboratory equipment. When light from certain sources is directed at photomultiplier tubes discrete and countable events occur. When the light is passed through a filter fewer such events occur. A second filter again reduces the number, and the proportion of the reduction is the square of the cosine of the angle between the filters. These are the sorts of facts that we will be concerned to explain. The photon picture provides a convenient model of the underlying process, but the correctness of that model need not concern us. If the reader is puzzled by the particulate nature of light it may help to note that experiments similar to the ones we will describe can also be carried out on protons and electrons, archetypical particles. In those cases one measures the so-called "spin" of the particles by passing them through an inhomogeneous magnetic field.

If light behaves as if made up of quanta and if each such quantum which survives a filter has the same energy as it had coming in, then figure 1.2 takes on a new significance. The quantity $\cos^2\alpha$, which previously measured the proportion of the beam that passes the filter, now represents *the probability for each photon to pass*. If the energy of the beam is to be reduced to one quarter of its previous value when passed through a polarizer oriented at 60° it must be that only one quarter as many photons will compose the passed beam. And if we can create individual photons, it must be that their individual probability for surviving the polarizer is one out of four. As we turn the second polarizer from perfect alignment to perfect misalignment, the likelihood of each photon to get through the second polarizer drops in accord with the graph in figure 1.2.

The Entangled State

So far nothing much mysterious has happened. Polarization phenomena are not particularly strange, and the quantization of light, if unexpected, seems perfectly comprehensible. But one final observation, also apparently rather pedestrian, turns out to be enough to destroy our accustomed picture of physical reality.

When calcium vapor is exposed to lasers tuned to a certain frequency it fluoresces. As excited electrons in the atoms cascade down to their ground state they give off light. In particular, each atom emits a pair of photons which travel off in opposite directions. The polarization of the photons individually shows no preferred direction: for any randomly chosen direction θ the photons will pass a polarizer oriented in that direction half the time. But although the photons individually show no particular polarization, the pairs exhibit some striking correlations. Roughly, each member of a pair always acts as if it has the same polarization as its partner.

More precisely, the following can be observed.[1] Suppose that one photon, R, goes off to the right while its partner, L, goes off to the left. R and L will each eventually impinge on a filter which sits before a photomultiplier tube. If the two filters are set in the same direction then either both photons will pass the filter or both will be absorbed. When the filters are aligned, in whatever direction, the photons are *perfectly correlated*: each does what the other does. If the filters are misaligned, then the photons still behave as if they have the same polarization. That is, suppose R passes through its polarizer, which is oriented in direction θ. Then L will act as if it is polarized in direction θ. If the left polarizer is also oriented in direction θ then L will pass, as we have seen. If the left polarizer is oriented at $\theta + 90^\circ$ then L will be absorbed. And if the angle of misalignment θ is between 0° and 90° then L will pass the filter a proportion $\cos^2 \alpha$ of the time. Similarly, if R is absorbed by its polarizer, L will act as if it is polarized in the $\theta + 90^\circ$ direction. It will always pass a polarizer oriented at $\theta + 90^\circ$, always be absorbed by one at θ, and generally if the filter is oriented at $\theta + \alpha$ the photon will pass $\sin^2 \alpha$ of the time.

Let us say that a pair of photons *agree* if they are either both passed or both absorbed by their respective filters and *disagree* if one is transmitted while the other is not. Then if the two filters are aligned in the same direction the photons will always agree, half of the pairs being jointly passed, the other half jointly absorbed. If the polarizers are misaligned by 90° the photons always disagree, one being absorbed, the other not. And for any other angle of misalignment α the percentage of pairs which agree (in the long term) is $\cos^2 \alpha$, as shown in figure 1.2.

Note that when we set up a particular experiment we have two choices to make. First we must choose the angle θ of the right-hand polarizer. Then we choose the degree of misalignment α of the left-hand polarizer. If we decide to examine a case of perfect alignment ($\alpha = 0^\circ$) we are still at liberty to set the pair of filters in any direction θ we choose. The fact that we have two free variables, θ and α, is just a reflection of the fact that we have two decisions to make: the angle of the right polarizer and the angle of the left. But no matter how we set the two, the only relevant parameter for calculating the probability of agreement is α, the degree of misalignment (see figure 1.3). If $\alpha = 30^\circ$ the photons will agree $\frac{3}{4}$ of the

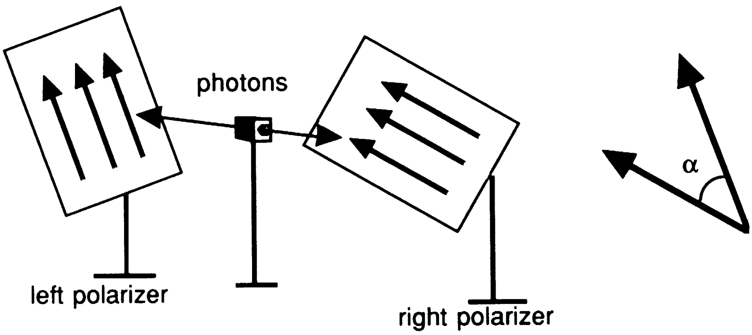

Figure 1.3 Experimental Set-up

time; if $\alpha = 60^\circ$ they will agree one time out of four. These simple facts about pairs of photons emitted by calcium vapor are enough to destroy any theory according to which physical reality is local.

How Do They Do It?

Suppose that you and a friend are set the task of reproducing the behavior of the photons: one of you will play photon L, the other photon R. These are the rules of the game: you and your friend start out together in a room (the "calcium atom"). You know that each of you will leave the room by a different door, and after some period of time you will each be asked a question. The question will consist of a number between 0 and 180 written on a piece of paper. Your answer must be either the word "passed" or "absorbed." Before you leave the room, you have no idea which question either of you will be asked. However, while in the room you and your friend are permitted to devise any strategy you please in order to coordinate your answers. Your aim is to ensure that after many repetitions of the game (you are permitted to adopt an entirely new strategy each time) your answers display exactly the same sorts of correlations as the photons show. That is, your strategies must ensure that, in the long run, when the question asked you differs from that asked your friend by an amount α, your answers agree $\cos^2 \alpha$ of the time.

For the moment, we will simplify your task even further. Unlike the photons, which have no information at all about which question will be $\cos^2 \alpha$ asked, you and your friend can know that only one of three possible questions, "0?", "30?" or "60?", will be asked. (We will eventually simplify the task even more, but it is easiest to begin here.) Of course, while you are in the room you still have no idea which of the three questions either of you will be asked. And once you leave the room, we suppose *you have no way of knowing what question has been asked (or will be asked) of your partner.* Your behavior may be determined by your agreed upon strategy and by the question you are asked, but not by the question which your friend happens to be asked. Once again, you must each respond either "Passed" or "Absorbed" when a question is asked.

Over a long run of this game you are aiming to reproduce the behavior of the photons in similar circumstances. That is, after a long series of plays you want to ensure that

Fact 1: When you and your friend happen to be asked the same question you always give the same answer.

Fact 2: When your questions differ by 30, that is, when one is asked "0?" and the other "30?" or one is asked "30?" and the other "60?", you and your friend agree $\frac{3}{4}$ of the time.

Fact 3: When your questions differ by 60, that is, when one of you is asked "0?" and the other "60?", your answers agree $\frac{1}{4}$ of the time.

After all, this is what the photons manage to do.

You and your friend are free to agree on any strategy you like, and you are free to vary your strategy from experiment to experiment. We may suppose that the questions to be asked are chosen at random, so that the pair of the questions "0?" to R and "30?" to L, for example, occurs $\frac{1}{9}$ of the time. It is not, however, important that the questions be asked equal amounts of the time, only that the choices be made at random, so that you can have no idea what is to come. How might you go about settling on a strategy?

The first obvious point is that there is no advantage, and much disadvantage, to using any sort of random element after you have left the room. For suppose your strategy demands that if asked the question "0?" you will decide your answer by a flip of the coin. Since you are unable to communicate with your partner there would be no way for your friend to know how you have answered the question, and so no way to be sure of matching your answer if asked the same question. In general, there is no possible way of satisfying Fact 1 above without deciding in the room *how each of you will answer each question if asked*. For without the knowledge of how your partner would answer a question you cannot act so as to ensure that your answers will match if you happened to be asked the same question.

Besides, no possible advantage can be gained by the introduction of random elements. If one of you may have to flip a coin when asked a question, why not flip it beforehand in the room and share the result with your partner? Or flip it three times, one for each possible contingency. Your partner would then have more information than would be available if you only appeal to the random element when actually asked the question. That excess information cannot possibly *degrade* your performance since, in the worst case, the information can just be ignored. Thus we have the simple result that *any strategy which involves local stochastic elements can do no better than a corresponding strategy where the random choices are made at the source*. A "local stochastic element" is a random process which takes place outside the room and whose outcome cannot be communicated to one's partner. In your case, "at the source" means "in the room"; for the photons it means "in the calcium atom." So in the first place, strategies utilizing local stochastic elements cannot ensure the perfect correlation when identical questions are asked, and in the second place, for every strategy using such elements an equally effective strategy which eschews them exists. If you have a penchant for flipping coins, you may as well flip them in the room. Given these two facts we may now narrow our search to strategies which involve no local stochastic elements. This means that when you leave the room each of you knows exactly what the other will do in each possible situation.

Furthermore, not any such deterministic strategy will do. Since you always run the risk of being asked identical questions, you and your friend must resolve to give the same answer as each other to each question. Only in this way can you assure that when answering identical questions your answers will tally.

So our situation has been greatly simplified. Only eight possible strategies are available, corresponding to the possible ways of answering the three questions. You might, for example, decide to answer "passed" no matter which of the three questions is asked. We will represent that strategy as ⟨P, P, P⟩, where the first slot represents the answer to "0?", the second to "30?" and the third to "60?". The eight possible strategies are then:

(1)	⟨P, P, P⟩	(2)	⟨A, A, A⟩	(A)
(3)	⟨A, P, P⟩	(4)	⟨P, A, A⟩	(B)
(5)	⟨P, A, P⟩	(6)	⟨A, P, A⟩	(C)
(7)	⟨P, P, A⟩	(8)	⟨A, A, P⟩	(D)

Since we are only interested in whether the answers given by you and your friend agree or differ, we can regard each of the corresponding mirror-image strategies above as equivalent. That is, if you choose either strategy (1) or strategy (2) you will agree no matter what pair of questions is asked, if you choose (3) or (4) you will disagree if one is asked "0?" and the other "30?" and agree otherwise, and so on. (Of course there are *other* facts, such as that in the long run approximately half the photons pass and half are absorbed, that would demand a judicious choice between the strategies in the right column and those in the left, but those facts have been omitted from our list.) So we may lump together strategies (1) and (2) calling each "strategy (A)," either (3) or (4) will be "strategy (B)," (5) or (6) "strategy (C)," (7) or (8) "strategy (D)." In order to ensure the strict correlations of Fact 1, you and your friend must choose among strategies (A), (B), (C), and (D) every time a new experiment is run. The only real option that is left open to you, then, is what proportion of the time each strategy will be chosen.

Let us suppose that your decisions over the long run result in choosing strategy (A) a proportion α of the time, strategy (B) β of the time, strategy (C) γ of the time, and strategy (D) δ of the time. α, β, γ, and δ must all be positive numbers (or zero), and of course $\alpha + \beta + \gamma + \delta$ must equal unity.

You and your friend must make your choice of which strategy to adopt in complete ignorance of what questions you are to be asked. Further, we may assume that the choice of questions is determined by a process which is random with respect to your choice of strategy. The experimenters, however they decide which question to ask, do not do so by predicating their choice on your predetermined strategy. In these circumstances, the long-run results of many repetitions of these experiments will depend solely on the values of α, β, γ, and δ. For example, suppose we wish to know how often the pair of question "0?", "60?" will receive answers which disagree. They will do so exactly when you have chosen strategy (B) or strategy (D), as can be verified by inspection. In the long run, you choose those strategies $\beta + \delta$ proportion of the time. And since the selection of experiments in which that pair of questions is asked constitutes a random selection from the sequence of strategies you choose, in the long run that pair of questions will receive disagreeing answers $\beta + \delta$ of the time.

By only selecting among the eight strategies we have ensured that Fact 1 will be satisfied. What of the other Facts? Fact 2 states that when the pair of questions "0?" and "30?" or "30?" and "60?" are asked, your answers will agree $\frac{3}{4}$ of the time. Another way of putting this is that your answers will disagree $\frac{1}{4}$ of the time. Similarly, Fact 3 states that when asked the pair of questions "0?" and "60?" you must disagree $\frac{3}{4}$ of the time. We have already seen that the proportion of the "0?" – "60?" experiments which yield disagreeing answers is $\beta + \delta$. By similar reasoning the proportion of "0?" – "30?" experiments which yield disagreement is $\beta + \gamma$, and the proportion of "30?" – "60?" experiments which yield disagreements is $\gamma + \delta$. To recover the correlations of the photons, then, you and your friend must arrange things so that

$$\gamma + \delta = 0.25$$
$$\beta + \gamma = 0.25$$
$$\beta + \delta = 0.75.$$

But now the rub becomes apparent. For on the one hand, the first two equations together imply that $(\beta + \gamma) + (\gamma + \delta) = 0.25 + 0.25 = 0.5$. But on the other hand, $(\beta + \gamma) + (\gamma + \delta) = 2\gamma + (\beta + \delta) = 2\gamma + 0.75$ (by the last equation). These results together imply that $0.5 = 2\gamma + 0.75$

or $2\gamma = -0.25$, so that $\gamma = -0.125$. But γ must be a positive number: it is not among your options to choose strategy (C) -12.5 percent of the time. In sum, there is no possible long-term selection of strategies that you and your friend can adopt which will ensure that your answers will display the same correlations as those of the photons.

Bell's Theorem(s)

The result just obtained can be generalized in many ways, all of which address themselves to variations of the question: given collections of two or more particles and a choice of observations that can be carried out on each, what sorts of constraints on the correlations among results can be derived if the observation carried out on one particle cannot influence the result of observations carried out on the others? We have just seen that in the case of two particles with three possible observations on each particle, if the results when the same experiments are carried out on both wings are perfectly correlated then (proportion of disagreement when experiments 1 and 2 are chosen) + (proportion of disagreements when experiments 2 and 3 are chosen) $\geq$ (proportion of disagreements when 1 and 3 are chosen). We can abstract from the exact nature of the experiments which are carried out, for in any such case the same reasoning leads to $(\gamma + \delta) + (\beta + \gamma) \geq (\beta + \delta)$.

John Stewart Bell inaugurated this line of investigation in his 1964 paper "On the Einstein–Podolsky–Rosen Paradox" (1987, ch. 2). Bell's result is couched in terms of expectation values, that is, long-term averages for observed quantities, and so takes on a slightly different form from ours. Bell also considers a case of perfect anti-correlation (disagreement when the same quantities are measured) rather than perfect correlation. Bell's result is:

$$|P(a,b) - P(a,c)| \leq 1 + P(b,c)$$

Where $P(a, b)$ is the expectation value of the product of the two observed results (the observations always yield the values ± 1).

Of course the assumption of perfect correlation or anti-correlation is an idealization relative to actual experimental situations: real laboratory conditions at best allow some approximation of perfect agreement or disagreement. Bell's result was further generalized by Clauser, Horne, Shimony and Holt (1969) to deal with imperfect correlations. In the case of our polarized photons, it is immediately clear that a small relaxation of the perfect correlation condition would not solve the difficulty. Even if you and your friend are lax enough to allow disagreement to occur 20 percent of the time when the same quantities are measured, the remaining correlations (i.e. 25 percent and 75 percent disagreement in the other experimental set-ups) cannot be recovered. The observed values are so far from the constraints imposed by the locality condition that no small perturbation will bring us back into the allowable range.

Of all the variations on Bell's theorem the most useful pedagogically is that of Greenberger, Horne, and Zeilinger (1989) (GHZ). The GHZ scheme involves three particles rather than two and has the advantage that all of the probabilities involved are either 0 or 1. Since there have been no experimental tests of the GHZ scheme, though, the exposition of it has been relegated to appendix A.

All of the Bell-type results express restrictions on the correlations that can be expected among various experimental observations. The only assumption needed to derive the restrictions is, as we have seen, that the experiment carried out on one particle can have no influence on the outcomes of observations on the other particle. You and your partner can have no advance knowledge of the questions that are to be asked, nor can you later acquire any information about what has been asked your partner so that you could adjust your own answer accordingly. If this condition were violated then the problem could be solved trivially. You need only first agree how your partner will answer each question. Once you know the question which has been asked her, you can decide how to answer the question asked of you. If the same question is asked both of you, you give the matching answer. If the questions differ by 30, you disagree 25 percent of the time; if they differ by 60, you disagree 75 percent of the time. Only information about the question asked need be transmitted since you could agree beforehand how your partner will respond.

A final remark on Bell's theorem is in order. Bell himself derived the result as part of an examination of so-called local hidden-variables theories. Such theories attempt to eliminate the stochastic element of orthodox quantum theory by adding extra parameters to the usual quantum formalism, parameters whose values determine the results of the experiments. Bell's results are therefore sometimes portrayed as a proof that local deterministic hidden-variables theories are not possible.

This is a misleading claim. It suggests that the violation of the inequalities may be recovered if one just gives up determinism or hidden variables. But as we have seen, the only assumption needed to derive the inequalities is that the result of observing one particle is unaffected by the experiment is carried out on the other. Subject to this restraint, no deterministic or stochastic theory can give the right predictions, no matter how many or how few variables are invoked.

Since a natural method of attempting to ensure the isolation of the two particles from one another is to carry out the relevant observations in distantly separated places, the isolation condition is generally called "locality". The assumption involved is that observations made on one photon can in no way alter the dispositions of the other photon to pass or be absorbed by its polarizer. Adopting this terminology uncritically for a moment, we have shown that Bell's inequality must be obeyed by any local theory of any sort. Adding stochastic elements does not help the situation at all, as noted above. So experiments verifying the violation of Bell's inequality would doom locality *tout court*.

Aspect's Experiment

To carry out exactly the experiment outlined above we would construct two polarization analyzers which could each be quickly set to any of three different positions: 0°, 30° or 60°. On reflection, though, we can see that the situation can be simplified even further. Suppose that each analyzer can be set at only one of two positions. We could then arrange things so that the right polarizer could be set at either 0° or 30° while the left can be set at 30° or 60°. The statistics to be reproduced are the ones we have already derived: when both analyzers are set at 30° on the photon always agree; when we have 0° on the right and 30° on the left or 30° on the right and 60° on the left they agree 75 percent of the time; when the right is set at 0° and the left at 60° they agree only 25 percent of the time. The analysis of this situation goes just as before, with exactly the same strategies available. Strategy $\langle P, A, A \rangle$ represents the decision that the right-hand photon will pass if measured at 0°, both will be absorbed if measured at 30°, and the left-hand photon will be absorbed if measured at 60°.

Indeed, the impossibility of satisfying these conditions is even more obvious now. The photons must agree on how they will both act if measured at 30°. In order to achieve the 75 percent agreement rate for the 0°–30° possibility, 25 percent of the time the right-hand photon must choose a strategy in which 0° differs from 30°. To achieve the 75 percent agreement for 30°–60°, the left-hand photon can only allow its 60° value to deviate from the common 30° value 25 percent of the time. But if the right-hand photon lets the 0° value deviate from the 30° value only 25 percent of the time, and the left-hand photon only allows 60° to deviate from 30° 25 percent of the time, then at least 50 percent of the time neither will so deviate. But then at least 50 percent of the time 0° and 60° will agree with each other (since neither deviates from the common 30° value) and the observed 0°–60° disagreement rate of 75 percent cannot be recovered.

The analysis again depends on the primary assumption: the setting of the polarizer on one side cannot be communicated to or have an effect on the photon on the other side. Experimentally we can try to ensure this condition in two ways. First, we want to separate the two analyzers from one another in space. Second, we want to choose the setting of the polarizer at the last possible moment. If the setting is chosen just before the measurement is made then the second photon could not adjust its strategy on the basis of prior knowledge of the question to be asked its partner. That is, if while you and your partner are in the room it has not yet even been decided which questions will be asked, then you cannot agree on a successful strategy while still in the room. Hence an ideal experimental condition will have two polarizers each of which can be set to one of two settings, well separated in space, with quick choices of the settings being made.

Such an experimental situation was realized in 1982 by Alain Aspect and his collaborators (Aspect, et al. 1982). Since physically rotating a polarizing filter cannot be done quickly, Aspect hit on a clever means of choosing between the two possible experiments on each side. Two polarizers and detectors were set up on each side of the experiment, with a very fast optical switch which could send the photons to either one (figure 1.4). Each of the optical switches alternated the beam between the two polarizers every 10^{-8} seconds. The right-hand apparatus was about 12 meters away from the left-hand one.[2] (These details will be of some importance in the coming chapters.) For the moment we can only note that it is in no way obvious how the result on the right-hand side could depend on which detector the left-hand photon is sent to. And if no such dependence exists then Bell's inequality cannot be violated: the quantum correlations could not be reliably produced.[3] In Aspect's experiment the quantum predictions were confirmed and Bell's inequality was violated.

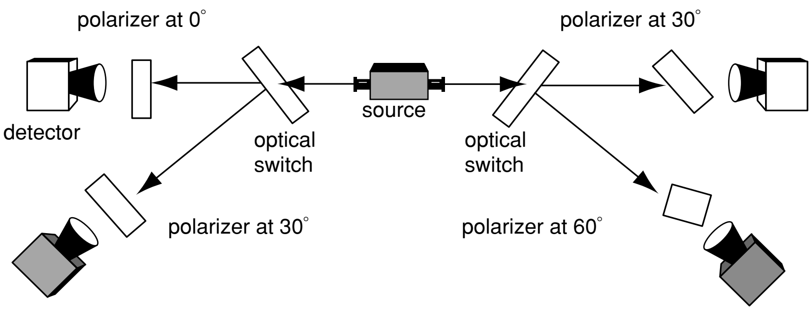

Figure 1.4 Aspect's Experiment

What is Weird About the Quantum Connection?

Aspect's experiment and other such experiments have produced observable data which cannot be predicted by any theory which disallows influence of the career of one particle on the behavior of the other once they separate. Somehow the particles must remain in communication, the observable behavior of one being determined, in part, by the nature of the observations carried out on its twin. After being created together the pair of particles remain interconnected.

This interaction among distantly separated particles presents profound interpretive difficulties. But one might initially be surprised that this behavior should elicit any concern at all. After all, classical physics is shot through with such causal connections among distant particles. Newtonian gravitational theory, for example, postulates that every massive particle in the universe exerts a gravitational force on every other, a force of magnitude Gm_1m_2/r^2. When a sparrow falls in Yugoslavia it has effects in New Brunswick and on Saturn and in the most distant galaxy. Some small gravitational tug will register in the smallest parts of the most far-flung stars. In the face of this sort of interconnectedness the quantum connection looks rather modest.

But there are at least three features of the quantum connection which deserve our close attention. All of them are, to some extent, surprising. The first two prevent our assimilation of these quantum effects to those of a force like gravitation. The last presents problems for reconciling the results of experiments like that of Aspect with the rest of our physical picture.

1. The Quantum Connection is Unattenuated

The fall of a sparrow in Yugoslavia may have its effects in New Brunswick and on Saturn and beyond, but the effect becomes progressively smaller the farther away one goes. Since the gravitational force drops off as the square of the distance it eventually becomes negligible if one is concerned with observable effects. The gravitational pull of the sparrow plays no noticeable role in affairs in New Brunswick, much less in the affairs of extra-galactic societies.

The quantum connection, in contrast, appears to be unaffected by distance. Quantum theory predicts that exactly the same correlations will continue unchanged no matter how far apart the two wings of the experiment are. If Aspect had put one wing of his experiment on the moon he would have obtained precisely the same results. No classical force displays this behavior.

2. *The Quantum Connection is Discriminating*

When the sparrow falls in Yugoslavia I feel a slight gravitational tug in New Brunswick. So does the computer on my desk, and the cat asleep on the bed. Every inhabitant of Princeton is jostled slightly, and to nearly the same extent as the population here. The effects of the sparrow's fall ripple outward, diminishing as distance increases, jiggling every massive object in its way. Equally massive objects situated the same distance from the sparrow feel identical tugs. Gravitational forces affect similarly situated objects in the same way.

The quantum connection, however, is a private arrangement between our two photons. When one is measured its twin is effected, but no other particle in the universe need be. If we create a thousand such correlated pairs and send the right-hand members all off in a group, each particle still retains its proprietary connection with its partner. A measurement carried out on one member of the right-moving hoard will influence only one member of the left-moving group, one particle situated in the midst of a thousand seemingly identical comrades.

The quantum connection depends on history. Only particles which have interacted with each other in the past seem to retain this power of private communication. No classical force exhibits this kind of exclusivity.

3. *The Quantum Connection is Faster than Light (Instantaneous)*

Of all the peculiarities of the particle communication, this might seem to be the most benign. For although no classical forces are unattenuated or discriminating, all were at least originally described as instantaneous. Classical gravitational and electrical forces were described as being determined by the contemporaneous global distributions of matter or of electric charge. Any change in that global distribution would therefore immediately have effects on the forces felt everywhere.

Although instantaneousness was a feature of the first theories of gravitation and electricity, it was not an essential feature. Newton thought that gravitation must be the effect of some subtle particles, about which he famously framed no hypotheses. He would therefore have expected a perfected theory of gravitation to take the speed of these particles into account. In such a final theory one would expect some delay to intervene between the sparrow's fall and the slight jostle it causes in New Brunswick. Of course, in the classical regime no a priori constraint could be put on the velocity of the gravitational disturbance, but one might reasonably expect it not to be infinite.[4]

But the modern theory of space and time differs radically from the classical view. The revolution has come in two stages, both initiated by Einstein: the Special and General Theories of Relativity. The Special Theory confers upon light, or rather upon the speed of light in a vacuum, a unique role in the space-time structure. It is often said that this speed constitutes an absolute physical limit which cannot be broached. If so, then no relativistic theory can permit instantaneous effects or causal processes. We must therefore regard with grave suspicion anything thought to outpace light.

The quantum connection appears to violate this fundamental law. Aspect's experiment was so contrived that the setting of the equipment at one side could not be communicated, even by light, in time to influence the other side. All three of these weird aspects of the quantum connection are related to spatial structure. Classical forces depend on spatial separation while the quantum connection does not. The effects of classical forces are determined by spatial dispositions: two electrons near one another will be (nearly) identically effected by distant gravitational or electrical sources. The quantum connection discriminates even among identical sorts of particles which are in close proximity to one another. Finally, the speed of the quantum communication appears to be incompatible with relativistic space-time structure.

Our concern will almost entirely be with the last of these three features. It is surprising that the communication between particles is unattenuated and discriminating, but often our best counsel is simply to accept the surprising things our theories tell us. The speed of the communication is another matter. We cannot simply accept the pronouncements of our best theories, no matter how strange, if those pronouncements contradict each other. The two foundation stones of modern physics, Relativity and quantum theory, appear to be telling us quite different things about the world. To understand, and perhaps resolve, that conflict we must consider carefully just what Relativity tells us about space and time.

APPENDIX A: THE GHZ SCHEME

If both you and your partner are uninformed about the question being asked the other, there is no strategy for answering questions which will reliably reproduce the quantum correlations in the long run. But this trouble matching the behavior of photons only appears in the long run: in every individual "experiment" you and your partner can be assured that your responses *in that particular experiment* are responses which the photons might have given. If, for example, you decide during one particular game that both players will answer "passed" no matter which question is asked, you can be assured, no matter which questions are asked, that your responses will not in themselves violate any quantum-mechanical predictions. For the only ironclad constraint imposed by the quantum correlations is that both partners give the same answer if they are both asked the same question. So long as you have agreed to a common response to each question you are safe on that run: the answers you give will certainly be quantum-mechanically permissible. Only after many games will your failure to match the target correlations emerge.

This state of affairs is frustratingly equivocal. Non-communicating partners can be certain that in no particular game will they diverge from what the photons might do, but can be equally certain that over time they must diverge in their cumulative behavior. There is something a bit ephemeral or ghostly about the problem, in that it lies entirely in long-term averages. Is this an indication of something deep about the nature of the quantum predictions?

A recent discovery by Daniel Greenberger, Michael Horne and Anton Zeilinger (1989) dispels this suspicion. Greenberger, Horne and Zeilinger (GHZ) found that in some instances quantum theory makes predictions about correlations between particles which are so strong that no local (i.e. non-communicating) strategy can be assured of matching the quantum predictions on any single run of the experiment. Although the GHZ scheme has not been tested in any actual experiment, it merits our attention as an indication of the strongly non-local character of quantum mechanics.[5]

The GHZ scheme uses three particles rather than two, and measures spin rather than polarization. The three particles can be created and allowed to separate to arbitrary distance, at which time a spin measurement is made on each. As in the Bell case, we can model the situation as a game. In this one, you and two partners begin together in a room. Some time after you depart the room, travelling in different directions, each of you will be asked one of two questions. We will denominate the questions "X?" and "Y?"; they correspond to measuring the spin of the particles in either of two orthogonal directions. To each question you must answer either "up" or "down" (in the literature, the responses of the particles are also often represented as 1 and -1). Since each particle can be asked one of two questions, there are eight distinct possible experimental arrangements, but of these only four will ever be used. Either two of the players will be asked "Y?" and the last one "X?" or all the players will be asked "X?". Using an obvious notation, we will represent these four experimental arrangements as $X_1Y_2Y_3$, $Y_1X_2Y_3$, $Y_1Y_2X_3$ and $X_1X_2X_3$.

GHZ noted that for one particular quantum state of a triple of particles, the following predictions can be made *with certainty*. If any of the first three experimental arrangements is chosen then the particles will, among them, respond "up" an *even* number of times. But if X measurements are made on all of the particles, they will, among them, respond "up" an *odd* number of times.

The quantum state does *not* fix exactly which particle will respond "up" and which "down" in any case. Nor does it predict whether the odd number of "up"s will be 1 or 3, or the even number 0 or 2. But according to quantum theory, the responses to the first three experimental arrangements will always include an even number of "up"s and the response to the fourth an odd number. Can you and your colleagues manage to duplicate this feat if each of you remains ignorant of the questions asked the others?

First, it is obvious that only a deterministic strategy will do. In any given experimental arrangement, once two of the partners have answered, the last partner will have to give a particular response in order for the number of "up" responses to come out right. If anyone is using stochastic mechanisms (and is unable to report the result to the other players) then he runs the risk of given an unacceptable answer.

While in the room, then, you must decide on how each partner will answer each question. The problem situation is vividly illustrated by the graph in figure A.1. Your task, in effect, is to write either "up" or "down" in each of the empty circles in such a way that an even number

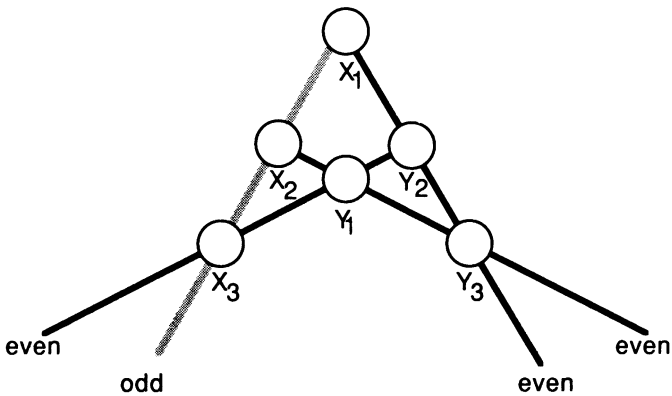

Figure A.1 The GHZ Problem

of "up"s lie in the triples linked by solid lines while an odd number lie in the remaining triple.

But this task is impossible. To solve it, the set of answers to $X_1Y_2Y_3$ must contain an even number of "up"s, as must the answers to $Y_1X_2Y_3$ and $Y_1Y_2Y_3$, while $X_1X_2X_3$ contains an odd number. Adding these up, $X_1Y_2Y_3+Y_1X_2Y_3+Y_1Y_2X_3+X_1X_2X_3 = \text{even} + \text{even} + \text{even} + \text{odd} = \text{odd}$. But $X_1Y_2Y_3 + Y_1X_2Y_3 + Y_1Y_2X_3 + X_1X_2X_3 = 2(X_1X_2X_3Y_1Y_2Y_3)$, that is, in taking the sum, the answer to each possible question was counted *twice*. The numbers of "up"s in the sum, then, must be even, not odd.

Since there is no solution to the task, the GHZ scheme brings home the problem for locality all the more sharply. No matter how you and your friends decide to answer the questions, there will always be at least one experimental arrangement for which you are bound to give the wrong answer. If each experimental arrangement is chosen with equal probability, then on average you must give a set of responses which is quantum mechanically *forbidden* at least one time in four. Every time the game is played you must run a risk of failing to do what the real particles are certain to do.

NOTES

1 When I write that this behavior can be observed, I am making certain idealizations about detector efficiencies, which do not approach 100 percent. What is correct to say is, first, that quantum mechanics predicts these correlations if the detectors were perfectly efficient. Second, what is observed is exactly in accord with the quantum mechanical predictions if the actual detector efficiencies are taken into account in the usual ways.

2 A very useful review of Aspect's experiment, as well as other experimental tests of violations of Bell's inequality can be found in Redhead 1987, pp. 107ff.

3 The qualifier "reliably" refers to the fact that the photons could, as it were, guess what the polarizer settings on each side will be, adjusting their strategies accordingly, and their guesses *could*, by pure chance, be right. But absent of any causal or informational connection, such accurate guessing would be a miraculous coincidence. The chance of a local system violating the Bell inequalities becomes arbitrarily small as the experiment collects more data.

4 It is interesting to note that such a perfected theory of Newtonian gravitation which incorporates a finite speed for the propagation of gravitational effects would necessarily lead to slightly different predictions than Newton's theory. Paul Gerber used the assumption that gravitational influence travels at the

speed of light to derive a corrected equation for planetary orbits. In this entirely Newtonian milieu, Gerber derived exactly the equation which Einstein recovered from the General Theory of Relativity. In particular, Gerber showed that this theory predicts the anomalous advance of the perihelion of Mercury. This is particularly remarkable since Gerber derived his result 17 years *before* Einstein, before the discovery even of the Special Theory of Relativity. Gerber's result is derived in a somewhat more perspicuous way by Petr Beckmann (1987, pp. 170–5). Robert Weingard (p. c.) has confirmed Beckmann's derivation, but remarks that the delayed-propagation version of Newtonian gravitation does not give the General Relativistic prediction for bending of light. Gerber's work was unfortunately seized upon by rabid anti-Relativists and has fallen into disrepute; the notion of a delayed-propagation Newtonian gravitational theory, though, is natural enough to warrant study.

5 For an extensive, but somewhat technical, discussion of GHZ-type schemes, see Clifton, Redhead and Butterfield (1991a). But see also Jones (1991) and Clifton, Redhead and Butterfield (1991b).