7

Bohm's Theory

There's an entirely different way of trying to understand all this stuff (a way of being absolutely deviant about it, a way of being polymorphously heretical against the standard way of thinking, a way of tearing quantum mechanics all the way down and replacing it with something else) which was first hinted at a long time ago by Louis de Broglie (1930), and which was first developed into a genuine mathematical theory back in the fifties by David Bohm (1952), and which has recently been put into a particularly clear and simple and powerful form by John Bell (1982), and that's what this chapter is going to be about.

Bohm's theory has more or less (but not exactly) the same empirical content as quantum mechanics does,[1] and it has much of the same mathematical formalism as quantum mechanics does too, but the metaphysics is different.

The metaphysics of this theory is exactly the same as the metaphysics of classical mechanics.

Here's what I mean:

This theory presumes (to begin with) that *every material particle in the world invariably has a perfectly determinate position.* And what this theory is about is the evolution of those positions in time. What this theory takes the *job* of physical science to be (to put it another way) is nothing other than to produce an *account* of those

1. Of course, there can't be *any* theory that has *exactly* the same empirical content as quantum mechanics does, since quantum mechanics (in its present unfinished condition, in the absence of any satisfactory postulate of collapse) doesn't *have* any exact empirical content!

evolutions; and the various *non*particulate sorts of physical things that come up in this theory (things like *force fields,* for example, and other sorts of things too, of which we'll speak presently) come up (just as they do in classical mechanics) only to the extent that we find we need to bring them up in order to produce the account of the *particle* motions.

And it turns out that the account which Bohm's theory gives of those motions is *completely deterministic.* And so, on Bohm's theory, the world can only appear to us to evolve *probabilistically* (and of course it *does* appear that way to us) in the event that we are somehow *ignorant* of its exact state. And so the very *idea* of probability will have to enter into this theory as some kind of an *epistemic* idea, just as it enters into classical statistical mechanics.

What the physical world consists of besides particles and besides force fields, on this theory, is (oddly) *wave functions.* That's what the theory requires in order to produce its account of the particle motions. The *quantum-mechanical wave functions* are conceived of in this theory as genuinely physical *things,* as something somewhat like force fields (but not quite), and anyway as something quite distinct from the particles; and the laws of the evolutions of these wave functions are stipulated to be precisely the linear quantum-mechanical equations of motion (always, period; wave functions never collapse on this theory); and the *job* of these wave functions in this theory is to sort of *push the particles around* (as force fields do), to *guide* them along their proper *courses;* and there are additional laws in the theory (new ones, *un*-quantum-mechanical ones) which stipulate precisely how they do that.[2]

2. Perhaps all this is worth spelling out in somewhat more pedantic detail. Here's the idea:

What *quantum mechanics* takes the wave function of a particle to be is merely a certain sort of mathematical representation of that particle's *state.*

What *this* theory takes the wave function of a particle to be, on the other hand, is a certain sort of genuinely physical *stuff.*

And the *physical properties* of such wave-functions-considered-as-stuff are (as with force fields) their *amplitudes* at every point in space.

And those amplitudes will invariably have determinate values (just as they invariably do, as purely mathematical objects, in quantum mechanics).

Setting Up

Let's start out by simply describing the mathematical formalism of the theory.

We'll begin with the case of a single structureless particle that's free to move around in only a single spatial dimension (since that case will turn out to be a particularly simple one, and a particularly instructive one), and then we'll build up from there.

The theory stipulates that the *velocity* of a particle like that at any particular time is given by the value of something called the *velocity function* $V(x)$ (which is a function of *position,* like the wave-function), evaluated at the point P at which the particle happens to be *located* at the time in question. Moreover, the theory stipulates that the velocity function for any particle at any time can be obtained (as a matter of law) from its *wave function,* at that *same* time, by means of a certain definite *algorithm.*

Let's set up a notation. Call the algorithm $V\{\ldots\}$, where the symbol $\{\ldots\}$ will serve as an empty *slot* into which any mathematical function of position can in principle be inserted.[3] What the theory is saying is that the velocity function $V(x)$ for any single structureless particle, moving around in a single spatial dimension,

---

It will frequently happen, of course (since the evolutions of wave functions here will invariably proceed in accordance with the dynamical equations of motion), that the wave function of a certain particle, at a certain instant, will fail to be an eigenfunction of one or another of the operators which the quantum-mechanical formalism associates with physical observables; but that will merely connote something about the *shape* of the wave-function-stuff here, something about its mathematical form, and *not* (as in quantum mechanics) that there is some *physical property of the world,* or some *potential* physical property of the world, which somehow *lacks a definite value!*

But of all this more later.

3. This algorithm, by the way, is stipulated to be precisely the same under all circumstances, irrespective of what the mass of the particle is, or what its charge is, or what forces it is subject to, or how those forces depend on space and time, or anything. All of those factors, of course, *do* play a role in determining the evolution of the *wave function* (since they all enter into the dynamical equations of motion); what they do *not* play any role in, however, is the way in which *velocity function* is obtained *from* the wave function.

at any particular time, will (as a matter of law) invariably be equal to the function $V\{\psi(x)\}$, where $\psi(x)$ is the wave function of that same particle at that particular time. And the theory further stipulates (as I mentioned above) that the *velocity* of that particle, at that time, will be equal to the number $V(P)$, where P is the *position* of the particle at that time.

And so if we're given the wave function of such a particle at some particular time t, and if we're also given its position at that time, then we can straightforwardly calculate its *velocity* at that time (and it's in this sense that the wave function can be said to push the particle around, to tell the particle where to go next).

And of course *that* will suffice, in principle, to calculate the position of the particle just an instant *after t* (since the velocity is just the rate of change of that position); and the *wave function* of the particle at that next instant can in principle be calculated too, with certainty, in the usual way, from the quantum-mechanical equations of motion; and so the *velocity* of the particle at that next instant can be obtained as well, and then the whole process can be repeated once again (in order to calculate the position of the particle and its wave function at the instant after *that*), and so on (as many more times as we like, up to whatever point in the future we choose).

And so the position of a particle and its wave function at any given time can in principle be calculated *with certainty,* on this theory, from its position and its wave function at any *other* time, given the external forces to which that particle is subject in the interval between those two times.

Suppose that the wave function of a particle at some particular initial time t is $\psi(x,t)$, and suppose that all of the forces are fixed in advance, and consider, in such circumstances (with the initial wave function held fixed, and with all of the external forces held fixed), how the motion of the particle will depend on its initial position.

Different initial positions will of course (in accordance with the laws of motion of this theory) pick out different *trajectories;* and each one of those different trajectories (in virtue of what it *means* to be a trajectory) will pick out some particular determinate position for the particle at every particular instant *later* than t.

Imagine a gigantic *swarm* of possible initial positions, and imagine the swarm of positions at some particular *later* time t', which that *initial* swarm (in accordance with the laws of this theory, given the initial wave function and given the external forces) evolves *into*.

And now imagine a very particular *sort* of swarm of possible initial positions. Remember (from Chapter 2) that the conventional quantum-mechanical prescription stipulates that the probability that a measurement of the position of a particle will find the particle to be located at any particular point in space will be equal to the absolute value of the square of the particle's *wave function*, at the time of the measurement, at that point. Imagine, then, that our swarm of possible initial positions happens to be distributed (with respect to the wave function) just as the quantum-mechanical probabilities are: imagine that the *density in space* of the possible positions in the initial swarm happens to be everywhere equal to the square of the absolute value of the initial wave function.

If that's so (and this is the punch line), then it can be shown to follow from the dynamical equations of motion for the wave function and from the form of the algorithm for the calculation of the velocity functions that the density in space of the positions in the *later* swarm (no matter *what* later time t' happens to be) will likewise be everywhere equal to the square of the absolute value of the wave function *then* (at the *later* time). That's what's depicted in figure 7.1.

Here's another way to put it. Consider the following fairy tale. Suppose that the form of a certain single-particle wave function at a certain initial moment t is $\psi(x,t)$; and suppose that at just that particular moment God places the *particle* associated with that wave function (which was *absent* before, in this fairy tale) somewhere in the world, and suppose that God makes use of some genuinely *probabilistic* procedure for deciding precisely where to *put* that particle, and suppose that that procedure happens to entail that the probability that the particle gets put at any particular point is equal to the square of the absolute value of the particle's wave function at that point. And suppose that thereafter God does no

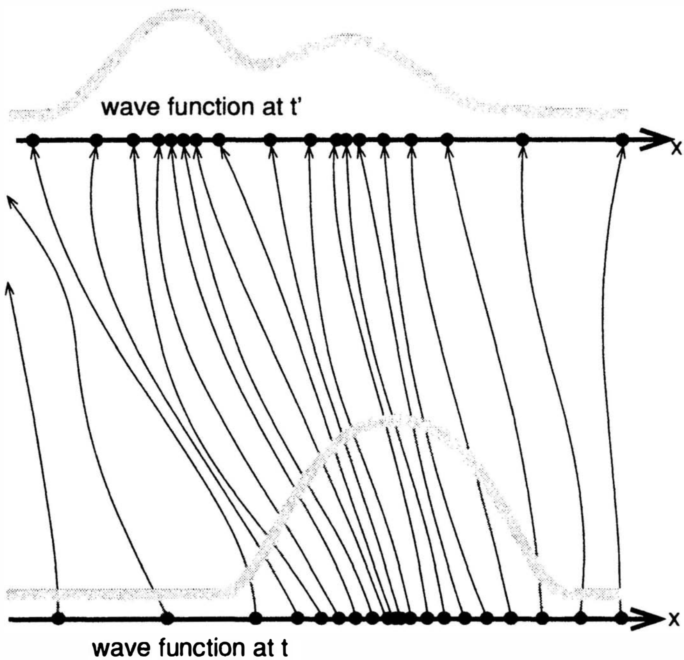

Figure 7.1

more meddling and allows everything to evolve strictly in accordance with the deterministic laws of Bohm's theory. Then it will be the case that the probability that the particle will be located at any particular point in space at any *later* time (given the way things were initially set up) will be equal to the square of the absolute value of the particle's wave function at that point at that later time.

Here's a more compact way to put it: What happens (and this is what the algorithm was explicitly cooked up in order to *guarantee;* and, as a matter of fact, this is a way of uniquely specifying precisely what that algorithm *is*) is that *the particle gets carried along* with the *flows of the quantum-mechanical probability amplitudes in the wave function,* just like (say) a cork floating on a river.

Here's how to parlay that into something useful:

There is what amounts to a fundamental postulate of Bohm's theory (let's call it the *statistical* postulate) which stipulates that if you're given the present wave function of a certain particle, and if you're given no information whatever about the present *position* of that particle, then, for the purposes of making calculations about motions of the particle in the future, what ought to be supposed, what ought to be plugged in to the formalism, is (precisely as it is in the fairy tale) that the *probability* that the particle is presently located at any particular point in space is equal to the square of the absolute value of its present *wave function* at that point.

That will guarantee (because of how the velocity algorithm works, because of how the particles always get carried along with the probability flows) that such calculations, under such circumstances, will necessarily entail that the probability that the particle will be located at any particular point in space at any particular *future* time will invariably be equal to the square of the absolute value of its wave function *then* (at the *future* time) at that point. And so in the event that we're given the present wave function of a single isolated particle, and we're given the external forces to which that particle is going to be subject, and we're given no information whatever (over and above what can be inferred, by means of the statistical postulate, from the wave function) about the present *position* of the particle (and it will turn out that that's more or less *all* of the information about the position of this sort of a particle that we can *ever* be in possession of, but of that more later), then Bohm's theory will entail precisely what principles C and D of Chapter 2 entail (that is: it will entail precisely what we know by experiment to be *true*) about the probability of *finding* that particle (if we should happen to *look* for it) at any particular point in space at any particular future time.

And so now we're apparently beginning to get somewhere.

The technical business of generalizing these laws, so as to be able to apply them to more complicated systems, goes (briefly) as follows:

(I) Three Space Dimensions. This part is trivial. The wave function (of a single isolated structureless particle, still) will now take on

values at every point in *three*-dimensional space (it will have the form $\psi(x,y,z)$), and the algorithm for calculating the velocity function will now become three algorithms (each of which looks much like the one-dimensional algorithm) for calculating three functions (one of which gives the velocity in the x direction, another of which gives the velocity in the y direction, and the third of which gives the velocity in the z direction), and each of those three functions will now be a function of three spatial positions. So the formalism will look like this:

(7.1)
$$V_x\{\psi(x,y,z)\} = V_x(x,y,z)$$
$$V_y\{\psi(x,y,z)\} = V_y(x,y,z)$$
$$V_z\{\psi(x,y,z)\} = V_z(x,y,z)$$

The evolution here will be just as deterministic as it was in the one-dimensional case. And the structure of these algorithms will entail that the particle will now get carried along with the flows of the quantum-mechanical probability amplitudes in the wave function in three-dimensional space. And there will be a straightforward generalization of the statistical postulate to the three-dimensional case, with precisely the same sorts of consequences as were described above.

(II) Spin. This part is pretty simple too. Spin properties, to begin with, are taken here to be mathematical properties of the *wave functions,* and the idea is more or less that *those* properties, *those* parts of the wave function, play no direct role whatever in the determination of the velocity functions. The idea is that the velocity functions are to be determined here *precisely* as in (7.1), from the *coordinate-space* wave function of the particle *alone.*

That will entail, for example, that the velocity functions for a particle whose quantum-mechanical state vector is

(7.2)
$$|\text{black}\rangle|\psi(x,y,z)\rangle \quad \text{or} \quad |\text{white}\rangle|\psi(x,y,z)\rangle$$
$$\text{or} \quad |\text{whatever you like}\rangle|\psi(x,y,z)\rangle$$

will be

$$V_i(x,y,z) = V_i\{\psi(x,y,z)\} \qquad (i = x, y, z) \tag{7.3}$$

The rule for handling states like

$$a|\text{black}\rangle|\psi(x,y,z)\rangle + b|\text{white}\rangle|J(x, y, z)\rangle \tag{7.4}$$

in which (since (7.4) is *nonseparable* between its coordinate-space parts and its spin-space parts) there isn't any such thing as *the* coordinate-space wave function of the particle, is that *each* of its various *different* coordinate-space wave functions gets to *contribute* to determining the velocity-functions in proportion to what you might call its "weight"; i.e.,

$$V_i(x,y,z) = \frac{|a\psi(x,y,z)|^2 V_i\{\psi(x,y,z)\} + |bJ(x,y,z)|^2 V_i\{J(x,y,z)\}}{|a\psi(x,y,z)|^2 + |bJ(x,y,z)|^2} \tag{7.5}$$

(III) Multiple-Particle Systems. This part is a little more complicated.

The first thing to talk about is the generalization of the conventional quantum-mechanical formalism of wave functions to cases of multiple-particle systems. That's pretty straightforward. The idea is just to write down something analogous to equation (2.27).

Suppose, then, that $|\psi_{1,2}\rangle$ represents an arbitrary quantum state of a two-particle system, and suppose that $|X_1 = x, X_2 = x'\rangle$ represents a state of that same system wherein particle 1 is localized at the (three-dimensional) point x, and wherein particle 2 is localized at the (three-dimensional) point x'. Then the two-particle *wave function* associated with the state $|\psi_{1,2}\rangle$ is defined to be

$$\psi(x,x') = \langle\psi_{1,2}|X_1 = x, X_2 = x'\rangle \tag{7.6}$$

considered as a function of x and x'.

And just as anything whatever that can be said of the quantum states of single particles can be translated into the language of single-particle wave functions, anything whatever that can be said of the quantum states of two-particle systems can be translated into the language of two-particle wave functions.

Note, by the way, that whereas the single-particle wave functions are functions of position in a three-dimensional space, the *two*-particle wave functions can be looked at as functions of position in a somewhat more abstract *six*-dimensional space. The first three of those six dimensions will refer to the space of possible locations of particle 1, and the *second* three of those six dimensions will refer to the space of possible locations of particle 2; and so picking out any particular *point* in that six-dimensional space will amount to picking out particular values for the locations of *both* of those particles.

The laws of Bohm's theory for two-particle systems which stipulate precisely how the two-particle wave functions push such pairs of particles around are formulated just as if it were a *single* particle that were being pushed around in a *six*-dimensional space. There will now be six algorithms (each of which looks much like the one-dimensional algorithm) for calculating six velocity functions (one for the velocity in each of the three spatial dimensions for each of the two particles), and (this is important) each of those six functions will be a function of position in the six-dimensional space. So the formalism will look like this:

7.7) $$V_i(x,y,z,x',y',z') = V_i\{\psi(x,y,z,x',y',z')\} \qquad (i = x,y,z,x',y',z')$$

And the way to get the *velocities* out of the velocity *functions* is to plug in the position of the *two*-particle system in the *six*-dimensional space.

The calculations of the wave functions and the positions of two-particle systems at arbitrary times from those wave functions and positions at any particular initial time (given the external forces to which the particles are subject in the interval between those two times) will proceed, as we shall see, very much as they do for single-particle systems, and (of course) with just as much certainty as they do for single-particle systems. It will now be the position of the two-particle system in the six-dimensional space that gets carried along with the flows of the quantum-mechanical probabilities, in that same six-dimensional space, in the wave function.

And what the statistical postulate will dictate here is that in the event that all we initially know of a certain two-particle system is its wave function (and that will turn out, once again to be all that we can *ever* know of such systems), then what ought to be supposed (for the purpose of making calculations about the future positions of the particles) is that the probability that the two-particle system was initially located at any particular point in the six-dimensional space is equal to the square of the absolute value of the wave function of the system, at that initial time, at that same particular point in the six-dimensional space (and note that this will reduce precisely to the *one*-particle statistical postulate in the event that the wave function of the two particles happens to be *separable* between them).

And that will guarantee (because of how these systems always get carried along with the probability flows) that such calculations, under such circumstances, will necessarily entail that the probability that the system will be located at any particular point in the six-dimensional space at any particular *future* time will invariably be equal to the square of the absolute value of its wave function *then* (at the *future* time) at that point. It will guarantee (to put it slightly differently) that such calculations, under such circumstances, will necessarily entail precisely the same things as the two-particle versions of principles C and D of Chapter 2 entail about the probability of *finding* the two particles (if we should happen to look for them) at any particular pair of points in the ordinary *three*-dimensional space at any particular future time.

And the reader will now have no trouble in constructing higher-dimensional formalisms, along precisely the same lines, for treating systems consisting of arbitrary numbers of particles, and even (in principle) for treating the universe as a whole.

And the statistical postulate, in a formalism like *that,* can be construed as stipulating something about the *initial conditions* of the universe; it can be construed (in the fairy-tale language, say) as stipulating that what God did when the universe was created was first to choose a wave function for it and sprinkle all of the particles into space in accordance with the quantum-mechanical probabilities, and then to leave everything alone, forever after, to evolve deterministically. And note that just as the one-particle postulate can be derived (as we saw above) from the *two*-particle postulate, *all* of the more specialized statistical postulates will turn out to be similarly derivable from *this* one.

The Kinds of Stories the Theory Tells

Let's start slow.

Let's look (to begin with) at some measurements with spin boxes.

Consider, for example, an electron, whose wave function is black; it is situated in coordinate space as pictured in figure 7.2, and it is (as shown in that figure) on its way into a *hardness* box.

Given all that, all of the *future* positions of this electron, on this theory, can in principle be determined, with certainty, from its *present* position (and so the *aperture* through which this electron will ultimately *exit* this box can in principle be determined from its present position, and so the *outcome* of the upcoming measurement

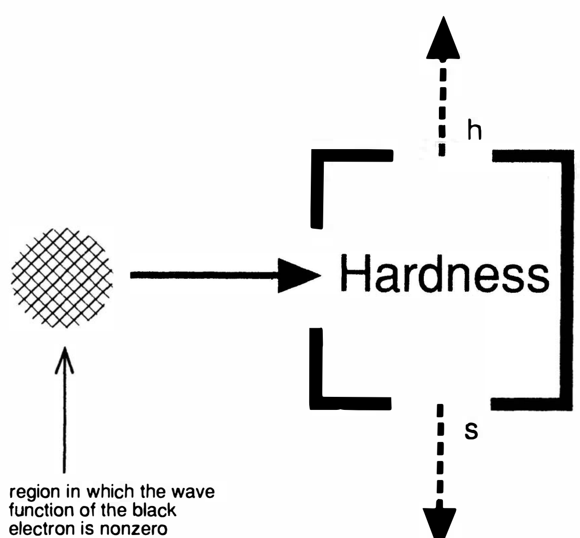

Figure 7.2

of the *hardness* of this electron can in principle be determined from its present position).

Let's see how that works. Suppose that the configurations of the hardness-dependent forces inside the hardness box are such as to cause *eigenfunctions* of the hardness to evolve as depicted in figure 7.3. Then the evolution of the wave function in question *here*, the *black* wave function, will of course be a linear *superposition* (with a plus sign and with equal coefficients) of the two evolutions

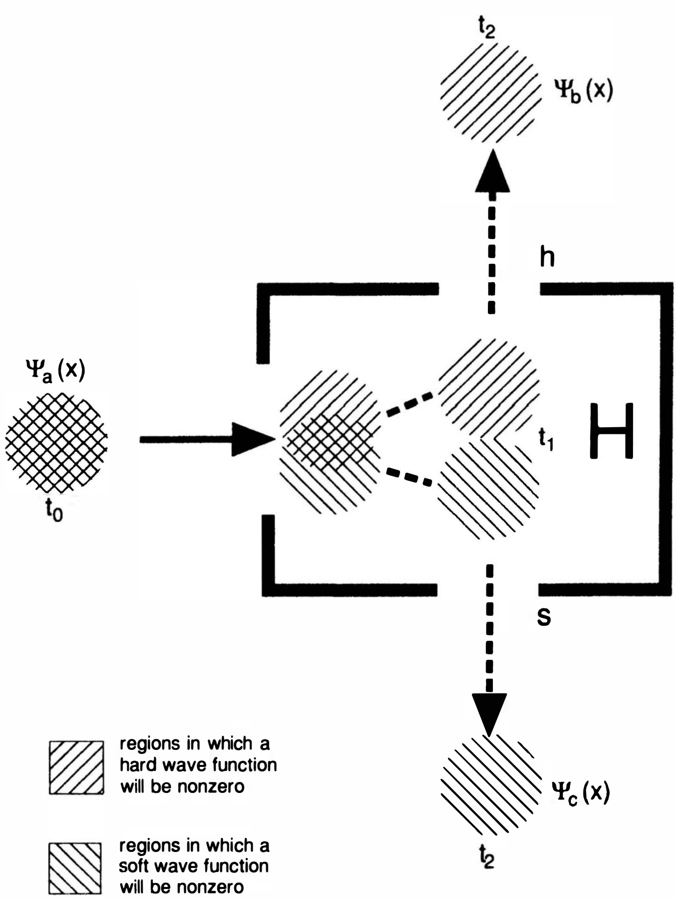

Figure 7.3

depicted there; so that (under these circumstances, in the time it takes for the wave function to pass all the way through the box):

(7.8) $$|\text{black}\rangle|\psi_a(x)\rangle \rightarrow 1/\sqrt{2}(|\text{hard}\rangle|\psi_b(x)\rangle + |\text{soft}\rangle|\psi_c(x)\rangle)$$

where $\psi_a(x)$ is a wave function which is nonzero only in the vicinity of the point *a* and which is headed to the right; $\psi_b(x)$ is a wave function which is nonzero only in the vicinity of the point *b* and which is headed *up;* and $\psi_c(x)$ is a wave function which is nonzero only in the vicinity of the point *c* and which is headed *down.*

Now consider how the motion of the electron is going to depend on its initial position. And keep in mind that what happens on this theory is that the electron, wherever it happens to be, invariably gets carried along with the local currents of the quantum-mechanical probability amplitudes. And note that the situation here has been cooked up so as to be perfectly symmetrical between the hard and the soft branches of the wave function. And so (for as long as the hard and the soft branches *overlap,* which will be up to time t_1 in figure 7.3) whenever the electron is located in the *intersection* of the two branches, then its velocity in the vertical direction will patently be *zero;* and whenever the electron is located exclusively in *one* or the *other* of those two branches (whenever, that is, the electron is located in a region of space in which the value of one of these two wave functions is *zero*), then it will invariably move perfectly along with that particular branch (the one that *isn't* zero there).[4] And all of that (if you think about it) will entail that in the event that the electron starts out in the *upper* half of the region where $\psi_a(x)$ is nonzero, then it will ultimately emerge from the *hard* aperture of the box; and in the event that the electron starts out in the *lower* half of that initial region, then it will ultimately emerge from the *soft* aperture of the box. It's as simple as that.

And note that in the event that the electron emerges from the hardness box through the *hard* aperture and is subsequently fed

4. The statistical postulate will guarantee that there will be no probability whatever of the electron's ever being located in a region of space in which the values of *both* of those wave functions will be zero.

into *another* hardness box (as in figure 7.4), then it will with certainty emerge from that *second* box through the hard aperture as well; and in the event that the electron emerges from the hardness box through the *soft* aperture and is subsequently fed into another hardness box (as in figure 7.5), *then* it will with certainty emerge from that second box through the *soft* aperture. Here's why: Once the electron gets out of the first box, if it emerges, say, through the

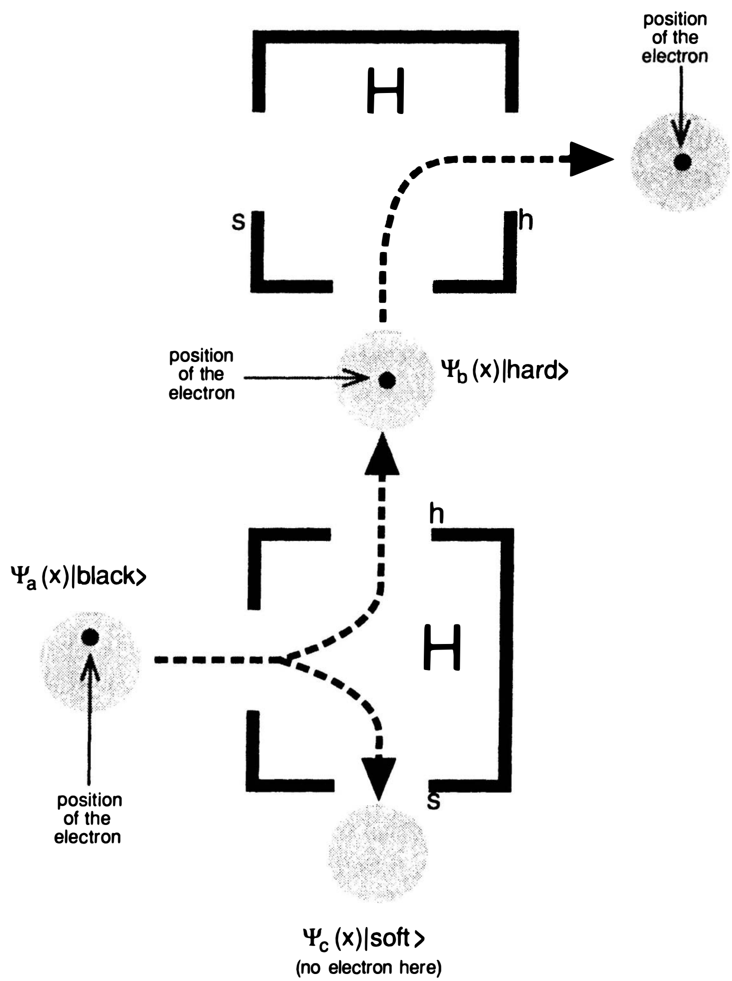

Figure 7.4

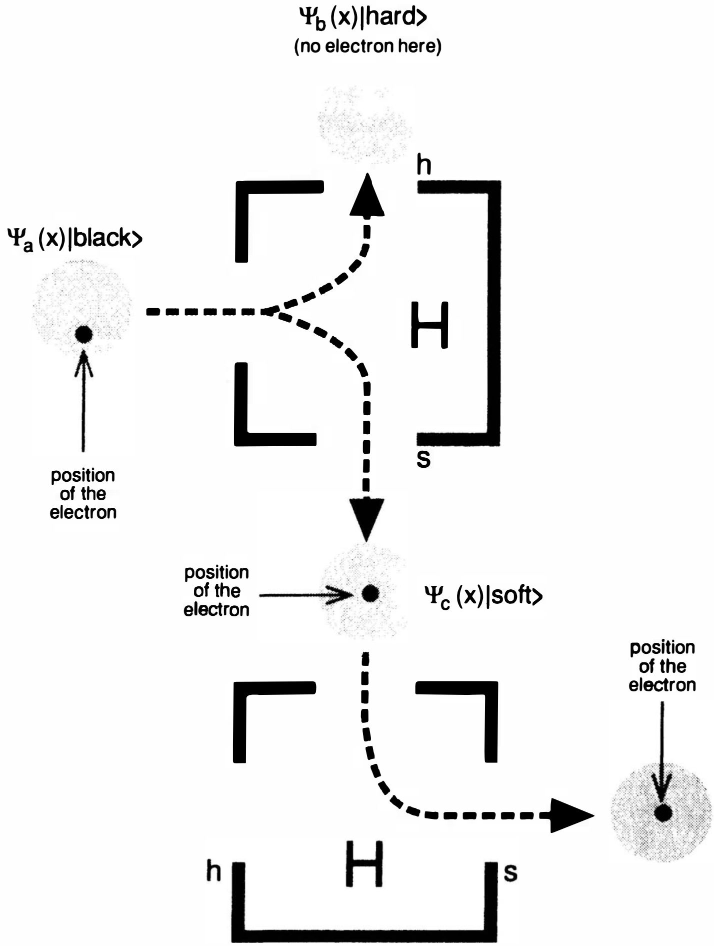

Figure 7.5

hard aperture, then (unless the hard and soft branches of the wave function are *reunited* in space *later on,* by means of reflecting walls, perhaps; but we'll talk about that in a minute) it will subsequently be moving in regions of space in which the amplitude of the *soft* part of its wave function is *zero;* and so that soft part will have *no effect whatever* on its subsequent motions; and it will simply be carried along through the second hardness box (and through whatever else it may encounter thereafter) with the *hard* part of its wave

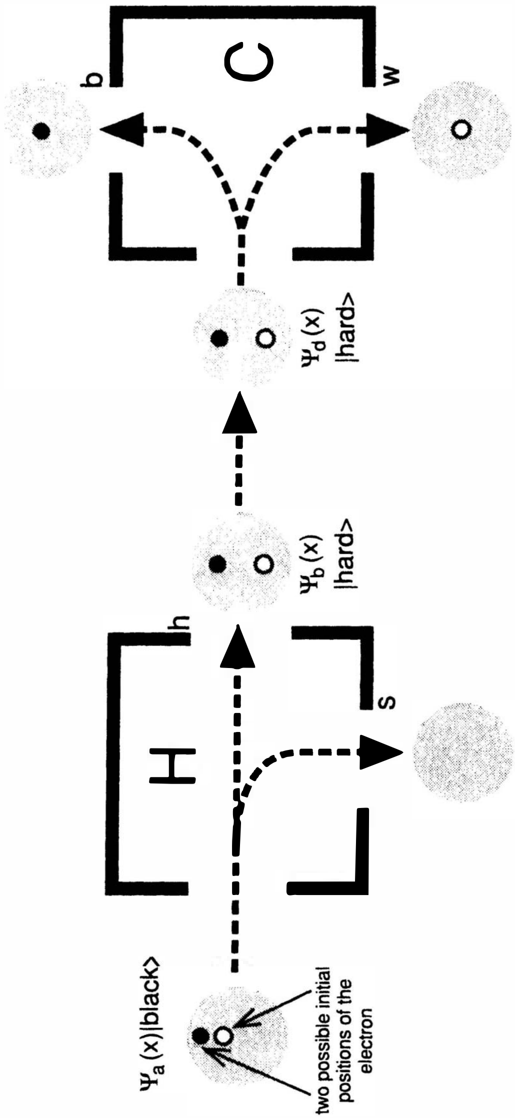
C
b
w
$\Psi_d(x)$ |hard>
$\Psi_b(x)$ |hard>
H
h
s
$\Psi_c(x)$|soft>
$\Psi_a(x)$|black>
two possible initial positions of the electron

Figure 7.6

function. And of course there is an analogous story to tell in the event that the electron emerges through the *soft* aperture of the first box.

And the reader will now be able to confirm for herself that in the event that the electron emerges from the hardness box (the *first* one) through, say, the hard aperture and is subsequently fed into a *color* box (as shown in figure 7.6), *then,* in the event that the electron is located in the *upper* half of the region in which $\psi_d(x)$ is nonzero, it will ultimately emerge from the *color* box through the *black* aperture, and in the event that it is located in the *lower* half of that region, it will ultimately emerge from the color box through the *white* aperture.

And in the event that the two branches of the wave function emerging from the two different apertures of the hardness box are ever *reunited,* by means, say, of an arrangement of reflecting walls and a "black box" like the one in figure 7.7, if the electron is *then* fed into a color box, *then,* no matter *what* position it initially had in the region in which $\psi_a(x)$ was nonzero, it will with certainty emerge from that color box through its *black* aperture. And if a wall is inserted somewhere along one of the two paths, then certain initial positions of the electron within $\psi_a(x)$ will entail that the electron never reaches the color box *at all,* and certain *other* such positions will entail that the electron finally emerges from the *white* aperture of the color box, and certain *other* such positions will entail that the electron will finally emerge from the *black* aperture of the color box.

And in the event that all that we initially happen to *know* of this electron is its *wave function* (which, again, will turn out to be all we *can* know of it, but of that more later), then the statistical postulate will straightforwardly reproduce all of the familiar quantum-mechanical frequencies of the various different possible outcomes of experiments like these (since the outcomes of experiments like these invariably come down to facts about the final *positions* of the *measured particles*). And that's going to turn out to be an instance of something a good deal more general: The fact that Bohm's theory gets everything right about the *positions* of things is going to entail that it *also* gets everything right about the out-

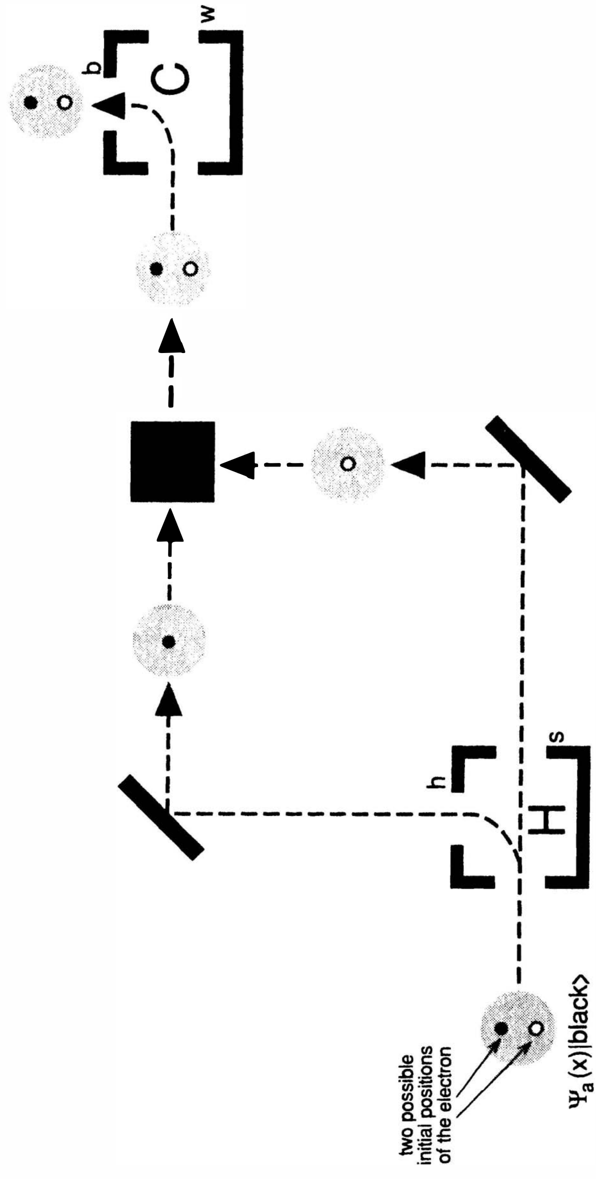
b
w
C
h
s
H
two possible
initial positions
of the electron
$\Psi_a(x)|\text{black}\rangle$

Figure 7.7

comes of any measurements of any quantum-mechanical observables whatever, so long as those outcomes get *recorded* (at one point or another) in the positions of things;[5] but a good deal more will need to be said, later on, about that too.

Here's something curious: What happens (as we saw before) when things are set up as in figure 7.2, and if the wave function of the electron is initially *black* and if the initial position of the electron is, say, within the *upper* half of the region where $\psi_a(x)$ is nonzero, is that the electron ultimately emerges from the *hard* aperture of the box.

Consider how all this depends on the *orientation* of the *hardness box*. Suppose that everything starts out just as described above (the wave function of the electron is black, the position of the electron is in the upper half of $\psi_a(x)$), except that the *hardness box* is *flipped over,* as in figure 7.8. Then the evolution of the wave function, as it passes through the box, will proceed as follows:

$$(7.9)\qquad |\text{black}\rangle|\psi_a(x)\rangle \rightarrow 1/\sqrt{2}(|\text{soft}\rangle|\psi_b(x)\rangle + |\text{hard}\rangle|\psi_c(x)\rangle)$$

5. And note (by the way) that that's a good deal *more* than can be said for the GRW theory (which is to say that it's a good deal more than can be said for *any* existing theory of the collapse).

The experiments that the GRW theory gets everything right about (after all) are just the ones whose outcomes get recorded in the position of something *macroscopic,* but the experiments that *Bohm's* theory gets everything right about are the ones whose outcomes get recorded in the position of *anything at all* (macroscopic or *microscopic*).

Bohm's theory, for example, is going to turn out to make the right predictions about the outcomes of experiments like the one depicted in figure 5.4, since the outcome of that sort of an experiment gets recorded in the *position* of the measured electron (the GRW theory, remember, turned out to entail that experiments like that, at least in their pre-retinal stages, don't have any outcomes at all); and Bohm's theory is also going to turn out to make the right predictions about the outcomes of experiments like the one depicted in figure 5.7, since the outcome of *that* sort of an experiment ultimately gets recorded in the position of the particle in the middle of John's head (and remember that *these* sorts of experiments turned out to be problematic for collapse theories *in general:* it turned out that there can't be *any* theory of the collapse on which *this* sort of an experiment has an outcome).

Nonetheless, it remains to be seen (to say the least) whether or not Bohm's theory makes the right predictions about the outcomes of absolutely all experiments whatsoever; but of that more later.

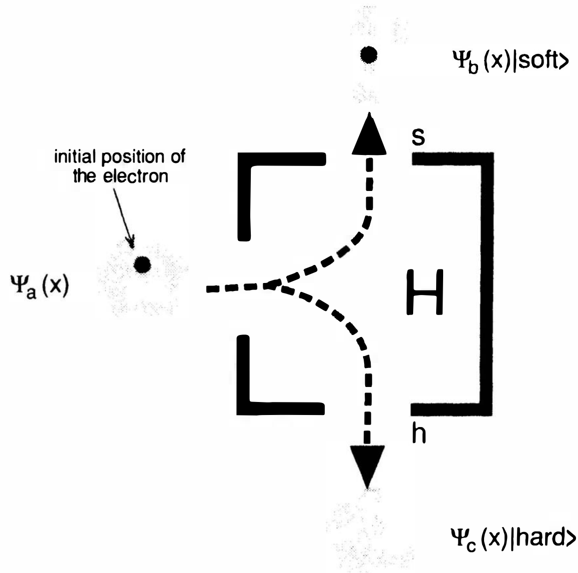

Figure 7.8

and the path of the particle through space (if you think it through) will patently be precisely the *same* as it would have been if the box *hadn't* been flipped, and so in *this* case the electron is ultimately going to leave the box through the *soft* aperture. And in the event that the electron starts out in the *lower* half of $\psi_a(x)$, with everything else as above (with the box flipped), *then* the electron will ultimately emerge through the *hard* aperture.

And so (even though this is a completely deterministic theory) the outcome of this sort of a "measurement" of the hardness of an electron will in general *not* be pinned down in this theory even by means of a complete specification of the electron's position and its wave function, which is (after all) *everything there is* to be specified about that electron. The outcome of such a hardness measurement is in general going to depend, on this theory, on precisely *how* and under precisely what *circumstances* the hardness gets measured, even down to the orientation of the hardness box in space.

And so it doesn't quite make sense, in general, on this theory, to think of hardness as an intrinsic property of electrons or of their wave functions (or of any combination of the two) at all. Properties like that (properties which, for these sorts of reasons, can't quite be thought of as intrinsic to the systems in question) have come to be referred to in the literature (for obvious reasons) as *contextual.*

And it turns out that color and gleb and scrad and momentum and energy and *every one* of the traditional quantum-mechanical observables of particles other than *position* are contextual properties on this theory too. As a matter of fact, there are theorems in the literature to the effect that *any* deterministic replacement for quantum mechanics whatever will invariably have to treat certain such observables as contextual ones.[6]

Note, however, that in the event that the wave function of an electron happens to be an *eigenfunction* of the hardness, then the outcome of a hardness measurement carried out on that electron will patently *not* depend on the orientation of the hardness box, and not on any *other* particulars of the condition of the hardness measuring device either, so long as whatever device that *is* satisfies the requirements for being a "good" device for measuring the hardness.

Similarly, if an electron is fed through one hardness box and then directly through *another,* it will with certainty emerge from both of those boxes through the *same* aperture, *irrespective* of the *orientations* of those two boxes or of anything else about them (so long as they're both *hardness boxes*), unless the hard and the soft branches of the electron's wave function have had a chance to overlap *in between* the two boxes.

And of course the same sorts of things are true of color as well, and analogous things are true of every other quantum-mechanical observable too.

For systems consisting of more than a single particle, Bohm's laws become explicitly *nonlocal.*

6. See, for example, Gleason, 1957, and Kochen and Specker, 1967.

Here's how that happens: Consider a two-particle system. The idea is that *each* of the *six* velocity functions for a two-particle system is going to depend, in general (as I mentioned before), on the location of the *entire* two-particle system in the *six-dimensional space*. And so (for example), in a system consisting of particle 1 and particle 2, once the wave function is fixed, the velocity in the x-direction of particle 1, at some particular moment, is in general going to depend not only on the position of particle 1 at that moment but *also* on the position of particle 2, at that *same* moment, no matter how far away particle 2 may (at that moment) happen to be! And so the *motions* of particle 2 (which will have the effect of changing the location, in the *six*-dimensional space, of the entire *two*-particle system) will in general play a very direct role, *instantaneously* (no matter how far apart the two particles may happen to be, or what may lie between them), in determining the velocities of particle 1. And (of course) vice versa.

Now, it turns out that in the event that the two-particle wave function is *separable* between the two particles, then these sorts of nonlocalities won't arise. In the event that the wave function is separable between the two particles, then it turns out that (once the wave function is given) the velocity of particle 1 always depends only on the position of particle 1 and the velocity of particle 2 depends only on the position of particle 2, and (more generally) the two-particle theory reduces completely, in that event (and if the particles don't interact with one another by means of ordinary *force fields*), to a pair of one-particle theories, just as in ordinary quantum mechanics. And if *that* weren't so (come to think of it), the *one*-particle theory wouldn't make any sense at all!

But things get more interesting if the wave function of the two particles is *non*separable.

Consider, for example, a pair of electrons; and suppose that the coordinate-space part of the wave function of one of those electrons (electron 1) is initially nonzero only in the vicinity of point a, and suppose that the coordinate-space part of the wave function of the *other* electron (electron 2) is initially nonzero only in the vicinity of point f (see figure 7.9); and suppose that the *spin-space* part of

Figure 7.9

the wave function of this system is of the (nonseparable) EPR type, so that the wave function *as a whole* looks like this:

$$(7.10) \qquad 1/\sqrt{2}(|\text{hard}\rangle_1|\psi_a(x)\rangle_1|\text{soft}\rangle_2|\psi_f(x)\rangle_2 - |\text{soft}\rangle_1|\psi_a(x)\rangle_1|\text{hard}\rangle_2|\psi_f(x)\rangle_2)$$

And suppose that electron 1 happens to start out in the upper half of the region in which $\psi_a(x)$ is nonzero and that it is now passed through a right-side-up hardness box, as in figure 7.10, so that the wave function of the two-particle system becomes:

$$(7.11) \qquad 1/\sqrt{2}(|\text{hard}\rangle_1|\psi_b(x)\rangle_1|\text{soft}\rangle_2|\psi_f(x)\rangle_2 - |\text{soft}\rangle_1|\psi_c(x)\rangle_1|\text{hard}\rangle_2|\psi_f(x)\rangle_2)$$

Electron 1 (for precisely the same reasons as in the single-particle case) will end up in the vicinity of point b. And once that's the case (and this is the punch line), the two-electron system will be located at a point in the six-dimensional space at which the value of the second term in (7.11) is 0. And so, thereafter (unless, or until, the two branches of the wave function subsequently come again to overlap in the six-dimensional coordinate space), that second branch will have *no effect whatever* on the motions of *either* of these electrons: electron 1 will behave, under all circumstances, as if its wave function were purely hard, and *electron 2 will behave, under all circumstances, as if its wave function were purely soft.* Electron 2, for example, will now emerge from any hardness box it gets fed through by the *soft* (as in figure 7.11) aperture, no matter *what* the orientation of that box or the initial position of that

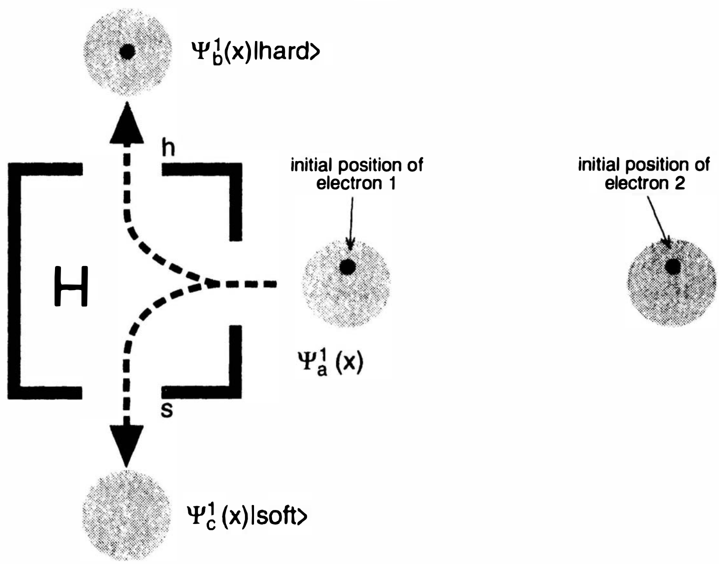

Figure 7.10

electron may happen to be. And so, even though the wave function evolves here entirely in accordance with the linear dynamical equations of motion, the passage of electron 1 through the hardness box brings about (as it were) an *effective collapse* of the wave function of the entire two-electron system, *instantaneously*, no matter how far apart they may happen to be or what may happen to lie between them.

And in the event that all we initially know of electrons like this one is that their wave function is the one in equation (7.10), then (as the reader can now easily confirm for herself) the statistical postulate and the Bohm-theoretic equations of motion will straightforwardly reproduce all of the standard quantum-mechanical predictions about the outcomes of measurements with spin boxes on EPR systems. And of course *that* entails (by means of Bell's theorem) that *some* sort of nonlocality was going to *have* to come up in this theory. But note that the particular sort of nonlocality that *does* come up in this theory turns out (compared, say, to the nonlocality in *quantum mechanics*) to be quite astonishingly *con-*

crete; it turns out (that is) to be a genuinely physical sort of *action at a distance.*

Suppose, for example, that electron 1 in figure 7.10 starts out (as above) in the upper half of the region in which $\psi_a(x)$ is nonzero, and suppose that electron 2 starts out in the upper half of the region in which $\psi_f(x)$ is nonzero. If that's so, and if electron 1 gets sucked through a right-side-up hardness box, as in figure 7.10, then (as we've just seen) electron 1 will end up effectively hard, and electron 2 (whatever *its* initial position is) will instantaneously become effectively *soft.* But of course in the event that electron *2* gets sucked through a hardness box first, and if *that* hardness box happens to be right-side-up, then (since electron 2 is in the upper half of $\psi_f(x)$) electron *2* will end up effectively hard, and electron *1* will instantaneously become effectively *soft* (and so if electron 1 *subsequently* gets sucked through a hardness box, with *any* orientation, then *it* will with certainty emerge by the *soft* aperture).

And if everything is initially as I just described, and electron 1

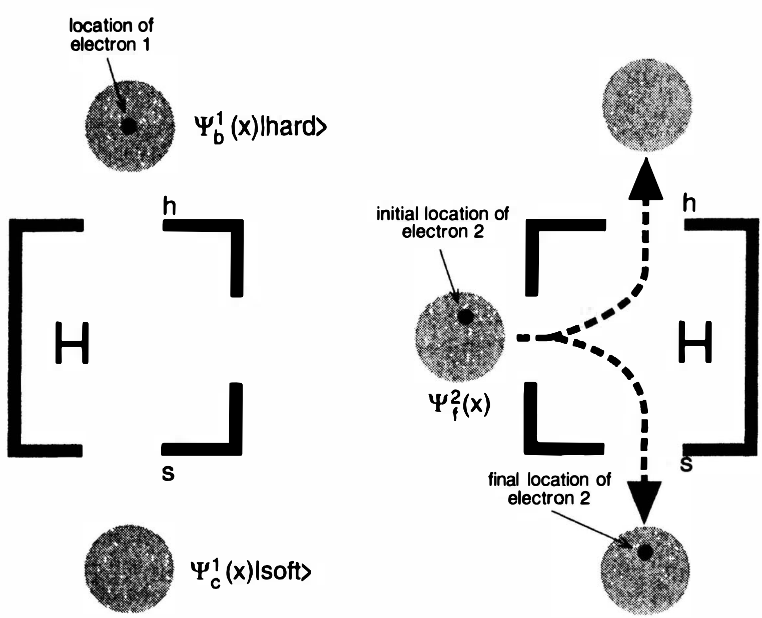

Figure 7.11

gets sucked through an *upside-down* hardness box, *then* electron 1 will end up effectively *soft* and electron 2 will instantaneously become effectively *hard*.

And so, if we're given the (nonseparable) wave function of such a system, and if we're given the positions of both of its constituent particles, then all this will patently afford a means of transmitting *discernible information,* instantaneously, over any distance, no matter what may lie in the way.

And so there can't possibly be any such thing as a Lorentz-covariant relativistic extension of this sort of a theory. This sort of theory will invariably require a *preferred frame;* this sort of theory (to put it another way) will invariably require an *absolute standard of simultaneity.*

But of course in the event that all we know of such a system is its *wave function,* then (as I mentioned above) all of the familiar quantum-mechanical probabilities for the outcomes of experiments with spin boxes on that sort of system will reemerge, and the sort of nonlocal information transmission just described will become impossible, and it will become impossible to determine (by means of experiments on a system like that, in any *non*covariant relativistic extension of Bohm's theory) *which* Lorentz-frame is the *preferred* one.[7]

The statistical predictions of any relativistic extension of Bohm's theory, then (supposing that the universe initially got started in the fairy-tale way), will turn out to be fully Lorentz-covariant, even though the underlying theory *won't* be;[8] and so taking Bohm's

7. The idea (once again) is that the outcomes of these sorts of measurements are recorded in the final *positions* of the two *particles;* and we know that in the event that all the information we initially have about such a pair of particles is their *wave function,* then the Bohm-theoretic probability distributions for those positions are going to be precisely the same as the quantum-mechanical probability distributions for those positions.

8. Of course, none of this talk will matter much unless it turns out that some relativistic extension of Bohm's theory (that is: some Bohm-type replacement for relativistic quantum field theory) can actually be cooked up.

It isn't clear whether or not that can be done. It turns out to be hard (for example) to figure out what sorts of field variables can possibly stand in for the *positions of particles* in a theory like that; it turns out to be hard (that is) to figure out what

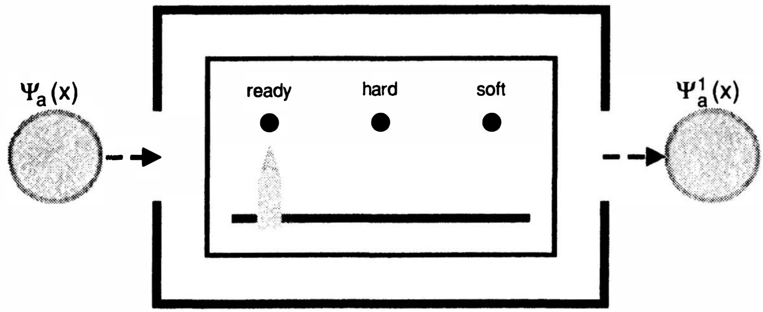

Figure 7.12

theory *seriously* will entail being *instrumentalist* about special relativity.[9]

Consider what happens on this theory in the event that the outcome of a measurement ultimately gets recorded in something macroscopic.

Consider, for example, the case of a hardness measuring device equipped with a macroscopic pointer, like the device depicted in figure 7.12. We've dealt with that sort of thing before.

---

sorts of field variables ought to get picked out by a theory like that as the *noncontextual* ones.

Bohm and Bell have both thought a good deal about these matters, but what they've come up with so far is only somewhat encouraging. Bohm (1952) has techniques for cooking up thoroughly Bohm-type replacements for a few particular field theories (but there are other such theories that those techniques *won't* work for); and Bell (1984) has a method which is a good deal more generally applicable but which generates theories that aren't completely *deterministic.*

Here's what the fundamental trouble is: Bohm's theory (as it presently stands) is quite deeply bound up with a very particular sort of *ontology;* the trouble (that is) is that this theory isn't a replacement for quantum theory *in general* (like the many-minds interpretation is), but only for those quantum theories which happen to be theories of *persistent particles;* and so the business of cooking up Bohm-type replacements for quantum theories of *other* sorts of systems (*field* systems or *string* systems or what have you) always has to proceed, without any guarantee of eventual success, case by case, from the ground up.

9. Bell (1976) began to explore what it's like to do that in a nice article called "How to Teach Special Relativity."

Suppose that a black electron is fed into that device, when the device is in its ready state. The wave function will evolve like this:

$$(7.12)\qquad |\psi_r\rangle_m(|\psi_a\rangle_e|\text{black}\rangle_e) \rightarrow$$
$$1/\sqrt{2}(|\psi_h\rangle_m(|\psi_b\rangle_e|\text{hard}\rangle_e) + |\psi_s\rangle_m(|\psi_b\rangle_e|\text{soft}\rangle_e))$$

where $|\psi_r\rangle_m$ is a state of the billions of particles in the pointer in which all of those particles are sitting more or less directly underneath the word "ready" on the dial (and similarly for $|\psi_h\rangle_m$ and $|\psi_s\rangle_m$), and $|\psi_a\rangle_e$ and $|\psi_b\rangle_e$ are the coordinate-space states of the electron depicted in figure 7.12.[10]

Once the electron passes all the way through the box and the wave function is the one on the right-hand side of equation (7.12), and the particles in the pointer are either under the "hard" on the dial or under the "soft" on the dial,[11] then the wave function of this composite system will have been effectively *collapsed* onto one or the other of the two terms on the right-hand side of equation (7.12), and (unless or until the two branches of the coordinate-space wave function of the particles in the pointer somehow *drift back together*) the electron will subsequently have an effectively determinate hardness.

And suppose that the two branches of the coordinate-space wave function of the particles in the pointer *do* eventually drift back together but that there happens to be an *air molecule* in the vicinity of the dial (as in figure 5.1), and that the *position* of that air molecule ends up *correlated* to the *hardness* of the electron (as in equation (5.4)). *Then* (unless or until the two branches of the coordinate-space wave function of the *air molecule* somehow drift

10. Note, to begin with, that the outcome of *this* sort of a hardness measurement is going to depend on the precise initial positions of the particles in the *pointer;* it's the *pointer* (and *not* the measured particle itself, as in figure 7.3) whose coordinate-space wave function gets split here; it's the pointer (and not the measured particle) whose position gets correlated to the hardness.

11. And note that either *all* of those particles will be under the "soft" or all of them will be under the "hard"; the form of the wave function (together with the statistical postulate) will entail that the probability that *some* of those particles end up in *one* of those places and some of them end up in the *other* will be *zero.*

back together) the electron will *still* have an effectively determinate hardness.

And even if the various branches of the *air molecule's* wave function were to drift back together *too,* but records of the outcome of the hardness measurements were still to survive, say, in the positions of ink molecules on the pages of a lab notebook, or in the positions of a few ions in a few neurons in some experimenter's brain, or even in so much as the position a single subatomic particle anywhere in the universe at all (as in the EPR example), *then* (unless or until we can manage to arrange that all that altogether ceases to be true) the electron will *still* have an effectively determinate hardness.

And so the business of *undoing* the effective determinateness of the hardness of this electron, or of empirically confirming that as a matter of fact that determinateness *is* merely effective, and that nothing has actually *collapsed,* will be (to say the least) quite fantastically difficult.[12]

And so Bohm's theory is going to make things look (for all practical purposes) as if wave functions *do* collapse when we do measurements with instruments with macroscopic pointers; and as a matter of fact, Bohm's theory is going to make things look as if wave functions collapse whenever we do measurements with any instruments whatever which (by one means or another) leave *records* of the *outcomes* of those measurements in the *positions of particles* in their *environments;*[13] and those collapses are going to appear to occur in more or less precisely the way we're *used* to, the

12. The difficulty here is of course precisely the same as the one we encountered on pages 88–92. The trouble is that the *environment* of a pointer like the one we're talking about here will act as a gigantic collection of extremely effective measuring devices for the pointer's *position;* and the business of confirming that as a matter of fact the Bohm wave function *hasn't* collapsed will involve either avoiding or reversing or somehow taking account of *every last one* of those measurements.

13. Note that these recordings need not be in the positions of anything *macroscopic;* what's important is merely that those recordings be in the positions of *something* and that they be (practically speaking) difficult to undo, or difficult to take account of. Consider, for example, the case of the measurement depicted in figure 5.4.

way we tried to *force* them to do in Chapter 5, even though it can *in principle* be empirically confirmed (on this theory) that in fact they don't occur at all.

And in the event that all we know of the systems involved are their *wave functions* at the moment just before the measurement interaction takes place, or in the event that all we know of those systems are their *effective* wave functions at the moment just before the measurement interaction takes place, then the statistical postulate will straightforwardly entail that the epistemic probabilities of the various possible effective collapses which that interaction may *bring about* will all be precisely the same as the *ontic* probabilities of the corresponding *actual* collapses (that is: the probabilities of principle D of Chapter 2) in *quantum mechanics.*

And it will turn out (as I've already mentioned) that this theory entails that all that we ever *can* know of the present states of such systems are their wave functions (or perhaps their *effective* wave functions); and *that's* how it happens (and I've already mentioned this too) that this theory has more or less the same empirical content as quantum mechanics does; that's how it happens (more particularly) that this theory entails that (even though the fundamental laws of the world are absolutely deterministic) we can never put ourselves in a position to predict any more about the outcomes of future experiments than the conventional quantum-mechanical *uncertainty relations* allow us to.

Let's see (finally) how that works.

Let's think through a simple example. Consider (say) an electron whose wave function happens to be $|\psi_a(x)\rangle|\text{black}\rangle$ and which is on its way into a hardness box, like the one in figure 7.2. Suppose that we initially know nothing of this electron other than its wave function; so that all that we're initially in a position to predict, with certainty, about the outcomes of spin measurements on this electron, is that (just as in quantum mechanics) any upcoming measurement of the electron's color will necessarily find it to be black.

Consider how things would stand if we were now somehow able to find out the *position* of this electron, or if we were merely able to find out, say, whether this electron is in the upper or in the lower half of the region where $\psi_a(x)$ is nonzero, without (in the course of finding that out) *changing* the electron's wave function. Then, of course, we would still be in a position to predict the outcome of any future measurement of the electron's color (since the wave function of the electron would still be the same), but we would *also* now be in a position to predict, with certainty, the aperture through which the electron would ultimately emerge from, say, a right-side-up *hardness* box; and so we would be in a position to violate the quantum-mechanical uncertainty relations; we would be in a position to predict *more* about this electron than quantum mechanics *allows* us to; we would be in a position to predict more about it than (as a matter of empirical fact) we find we *can* predict about it.

Let's figure out what's going on. Think, to begin with, about what it will involve (think, that is, about what sorts of *physical acts* it will involve) to find out whether the electron in the above scenario is located in the upper half or the lower half of the region in which $\psi_a(x)$ is nonzero. What we shall need to do is to bring some physical property of a measuring device (the position of its pointer, say) into the appropriate sort of *correlation* with the location of the electron. And what will need to be done in order to accomplish *that* (as a little reflection on the laws of this theory will show) is to bring about the analogous sorts of correlations in the *wave functions* of those two systems; what will need to be done (to put it another way) is to bring into being a wave function of the *composite* system which is *zero* in all those regions of the many-dimensional coordinate-space in which those correlations *fail* to obtain.

And so finding out whether the electron in the above scenario is located in the upper or the lower half of the region in which $\psi_a(x)$ is nonzero, *without* (in the process of finding that out) *changing the wave function,* is going to be (as a matter of *fundamental principle*) completely out of the question.

Suppose, for example, that we were to measure whether the electron is in fact in the upper or the lower half of the region in which $\psi_a(x)$ is nonzero (let's call that "region *a*" from here on), using a measuring device (m) like the one depicted in figure 7.13. Then (in accordance with the dynamical equations of motion) the

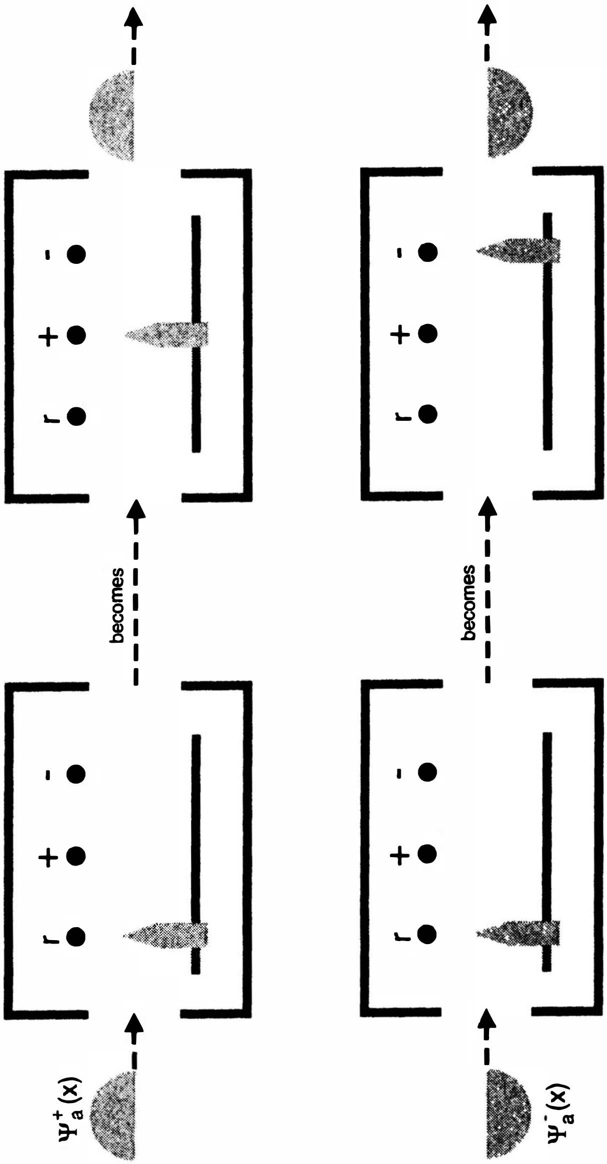

Figure 7.13

wave functions of the electron and the measuring device will evolve as follows:

$$(7.13)\quad |\mathrm{r}\rangle_m|\psi_a(x)\rangle_e|\mathrm{black}\rangle_e \rightarrow |+\rangle_m|\psi_a^+(x)\rangle_e|\mathrm{black}\rangle_e + |-\rangle_m|\psi_a^-(x)\rangle_e|\mathrm{black}\rangle_e$$

where $\psi_a^+(x)$ is equal to $\psi_a(x)$ in the upper half of region a and is 0 elsewhere, and $\psi_a^-(x)$ is equal to $\psi_a(x)$ in the *lower* half of region a and is 0 elsewhere, and of course $|\mathrm{r}\rangle$ and $|+\rangle$ and $|-\rangle$ are the states of the *pointer* in which it's pointing at "r" or "+" or "−" on the dial, respectively.

And so in the event that the electron is located in the *upper* half of region *a*, then the pointer will with certainty end up pointing at "+" and the coordinate-space *wave function* of the electron will be effectively *collapsed* onto $\psi_a^+(x)$, and the statistical postulate will entail (by means of a straightforward conditionalization) that the probability that the electron is located at any particular point in space is equal to $|\psi_a^+(x)|^2$; and in the event that the electron is located in the *lower* half of region *a*, *then* the pointer will with certainty end up pointing at "−" and the coordinate-space wave function of the electron will be effectively collapsed onto $\psi_a^-(x)$, and the statistical postulate will entail that the probability that the electron is located at any particular point in space will be equal to $|\psi_a^-(x)|^2$.

And all of that will ineluctably *change* the way in which the *future motions* of this electron depend on its present position. In the event, for example, that the electron's coordinate-space wave function gets effectively collapsed onto $\psi_a^+(x)$, then (as the reader can easily confirm) the outcome of an upcoming measurement with a hardness box will depend on whether the electron is located in the top *quarter* (not half!) or the next-to-the-top quarter of the region *a*. Of course the outcome of our position measurement will give us no idea whatever which of *those* two is the case (that is: the epistemic probabilities will be precisely fifty-fifty); and so everything perversely conspires together here so as to insure that that position measurement will have done us precisely *no good at all,* in so far as the violation of the uncertainty relations is concerned. And of course the same sort of thing happens in the event that the electron's wave function gets effectively collapsed onto $\psi_a^-(x)$.

And as a matter of fact, this sort of thing turns out to constitute an absolutely general *law:* it turns out (and the reader will now be in a position to construct an argument for this for herself) that this theory entails that if we initially know nothing of a system (*any* system) other than its wave function (if, that is, we know nothing more of the system's *location* in its many-dimensional coordinate space than can be *inferred* from its wave function by means of the *statistical postulate*), then all we shall possibly be in a position to know at any particular future time of that system's location in that space (by means of measurements, for example, or by any physical means whatever) will be what follows (by means of the statistical postulate) from that system's *effective wave function* at that future time.

And if we apply all that to the universe as a whole, and if we presume that the universe was initially created in accordance with the fairy tale described earlier, then this theory will entail that all that can ever be found out by any observer whatever about the present Bohm-state of any particular physical system can always be completely *summed up* (with the help of the statistical postulate) in a *wave function* (sometimes this will turn out to be the *actual* wave function of the system in question, but more often it will merely be an *effective* one). And so the Bohm theory (even though it's a completely deterministic theory) will systematically and invariably and unavoidably *prohibit* us from ever predicting the outcomes of future measurements of the positions of particles any more accurately than those *wave functions* allow us to do; it will (that is) prohibit us from ever predicting those outcomes any more accurately than the *uncertainty relations* allow us to do.

That's how this theory manages to clean up after itself: that's what entails (for example) that we can't ever empirically discover that the outcomes of measurements with spin boxes depend on the *orientations* of those boxes (even though they *do* depend on those orientations, if this theory is right); and that's what entails that we can't ever empirically discover that the outcomes of measurements with spin boxes can *also* depend, *nonlocally* (when states like (7.10) obtain), on the orientations of fantastically *distant* spin boxes (even though they *can* depend on the orientations of distant boxes, if this theory is right); and that's what entails that we can't ever find out what the natural standard of absolute *simultaneity* is (even though there *is* a natural standard of absolute simultaneity, if this theory is right); and so on.

And so if this theory is right (and this is one of the things about it that's cheap and unbeautiful, and that I like), then the fundamental laws of the world are cooked up in such a way as to systematically *mislead* us about themselves.

Here's what's so cool about this theory:

This is the kind of theory whereby you can tell an absolutely low-brow story about the world, the kind of story (that is) that's about *the motions of material bodies,* the kind of story that contains nothing cryptic and nothing metaphysically novel and nothing ambiguous and nothing inexplicit and nothing evasive and nothing unintelligible and nothing inexact and nothing subtle and in which no questions ever fail to make sense and in which no questions ever fail to have answers and in which no two physical properties of anything are ever "incompatible" with one another and in which the whole universe always evolves *deterministically* and which recounts the unfolding of a perverse and gigantic conspiracy to make the world *appear* to be *quantum-mechanical.*

And that conspiracy works (in brief) like this: Bohm's theory entails everything that quantum mechanics entails (that is: everything that principles C and D of Chapter 2 entail, everything that we empirically know to be *true*) about the outcomes of measurements of the positions of particles in isolated microscopic physical systems; and moreover it entails that whenever we carry out a measurement of *any* quantum-mechanical observable *whatever,* then (unless or until there somehow ceases to be any record of the outcome of that measurement in the position of even so much as a single subatomic particle anywhere in the universe; and *that* can presumably almost *never* come to pass, if the outcome of the measurement ever gets recorded in anything macroscopic) the measured system (or rather: the positions of all of the particles which *make up* that system) will subsequently evolve just as if that system's wave function has been *collapsed,* by the measurement, onto an *eigenfunction* (the one corresponding to the measured eigenvalue) of the measured observable, even though as a matter of fact it *hasn't* been; and it also entails that the *probabilities* of those "collapses" will be precisely the familiar quantum-mechanical ones.

Mentality

But there are interesting questions (which I want to just set up here, and then leave to the reader) about whether or not all that turns out to be enough.

Consider (for example) whether Bohm's theory guarantees that every sort of measurement whatsoever even *has* an outcome. Bohm's theory does a good deal *better* at that than any theory of the collapse of the wave function can (that's what we found out in note 5); but there are problem cases for Bohm's theory too.

Here's a science-fiction story (along the lines of the story about John, in Chapter 5) about one of those.

This story is going to involve a device (like the one depicted in figure 7.14) for producing a correlation between the *hardness* of a certain microscopic particle *P* and the hardness of an incoming electron (which is slightly different from what happens in the story in Chapter 5); a sort of measuring device for the hardness of electrons in which the "pointer" is the *hardness* of *P*. Here's how the device works: if the device starts out in its ready state, and if (say) a hard electron is fed through the device, then the hardness of that electron is unaffected by its passage through the device, and the device is unaffected too, except that *P* ends up, once the electron has passed all the way through, *soft;* and things work similarly (or rather, analogously) if a *soft* electron is fed through the device, if the device is initially in its ready state (in that case *P* ends up *hard*). Once the hardness of the electron gets correlated to the hardness of *P*, then *P*'s hardness can be measured, if we want to, by means

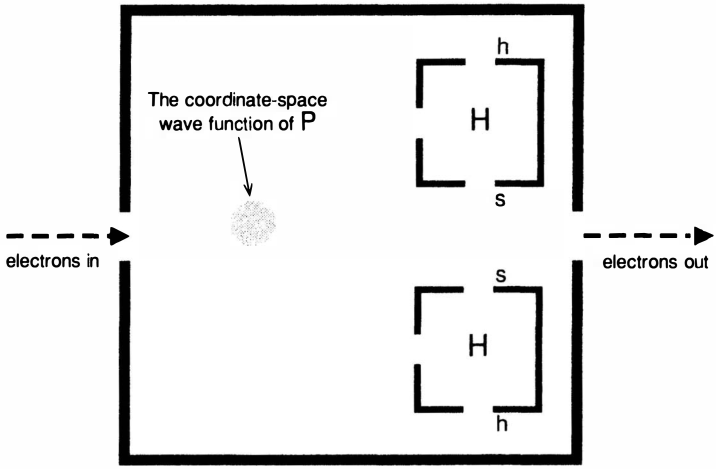

Figure 7.14

of either one of the two little hardness boxes (note that one of them is right-side-up and the other is up-side-down, but of that more later) on the right-hand side of the device.

A device like the one I just described is (according to the story) sitting in the middle of John-2's brain, and the particular way in which that device is now hooked up to the rest of John-2's nervous system makes him function as if his *occurrent beliefs* about the hardness of electrons that happen to pass through it are determined *directly* by the hardness of *P* (just as the occurrent beliefs of the *original* John about the hardness of those sorts of electrons were determined directly by the *position* of *his P*).

Here's what that means: Suppose that John-2 is presented with an electron which happens to be hard and is requested to ascertain what the value of the hardness of that electron is. What John-2 does (just as the original John did) is to take the electron into his head through his right door, pass it through his surgically implanted device (with the device initially in its ready state), and then expel it from his head through his left door. And when that's all done (that is, when *P* is soft but the value of the hardness of the electron is not yet recorded anywhere in John-2's brain *other* than in the hardness of *P*), John-2 announces that he is, at present, consciously aware of what the value of the hardness of the electron is, and that he would be delighted to tell *us* what that value is, if we would like to know.[14]

Now, the way that John-2 behaves, subsequent to all that, in the event that we *do* ask him to tell us what that value is, is (according to this story) as follows:

The rules of the game (to begin with, to keep things simple) are that John-2 can tell us about that either *verbally* or *in writing* (or in both ways, but of that more in a minute).[15]

And what the story stipulates is that in the event that we ask John-2 to tell us *verbally* what the value of the hardness of the electron is, then what he does (what he's programmed to do, what he's wired up to do) is to pass *P* through the upper hardness box (the one that's right-side-up), which is equipped with *P* detectors at its two exit apertures, which are in turn connected with certain of John-2's neurons, which are themselves ultimately connected with his mouth muscles and his throat muscles and his tongue muscles in such a way as to end up generating (in the case where John-2 is initially presented with a *hard* electron) the utterance "hard" (and of course in the event that John-2 is initially presented with a *soft* electron, then these same procedures and these same connections will end up producing the utterance "soft").

And (on the other hand) the story stipulates that in the event that what we ask John-2 to do is to report *in writing* what the value of the hardness of the electron is, *then* what he does (what, once again, he's wired up to do) is to pass *P* through the *lower* hardness box (the one that's *upside-down*) which is similarly (or rather, analogously) equipped with *P* detectors and hooked up to the rest of

14. There was a good deal of talk in Chapter 5 (which the reader would do well to bear in mind here too) about precisely how we ought to *take* announcements like that.

15. Of course, there are almost certainly going to be any number of other ways (*besides* speaking or writing) in which anybody like John-2 could manage to communicate information like that, but those other ways need not concern us right now; all that will turn out to be important for our present purposes is that there can be at least two ways for him to communicate it.

John-2's brain in such a way as to end up generating (in the case where John-2 is initially presented with a *hard* electron) the *inscription* "hard" (and of course, as above, in the event that John-2 is initially presented with a *soft* electron, then these same procedures and these same connections will end up producing the inscription "soft").

Good; now (and here comes the interesting part of the story) consider how John-2 is going to behave in the event that he is presented with (say) a *white* electron and requested to measure the hardness of that electron, and to remember the outcome of that measurement. When John-2 is done with all that (that is: when the electron has passed all the way through John-2's head, as depicted in figure 7.15; and when John-2 informs us that he now knows what the value of the hardness of that electron is, that he is now consciously *aware* of what the value of the hardness of that electron is, and that he can *tell us* that value, if we like, either verbally or

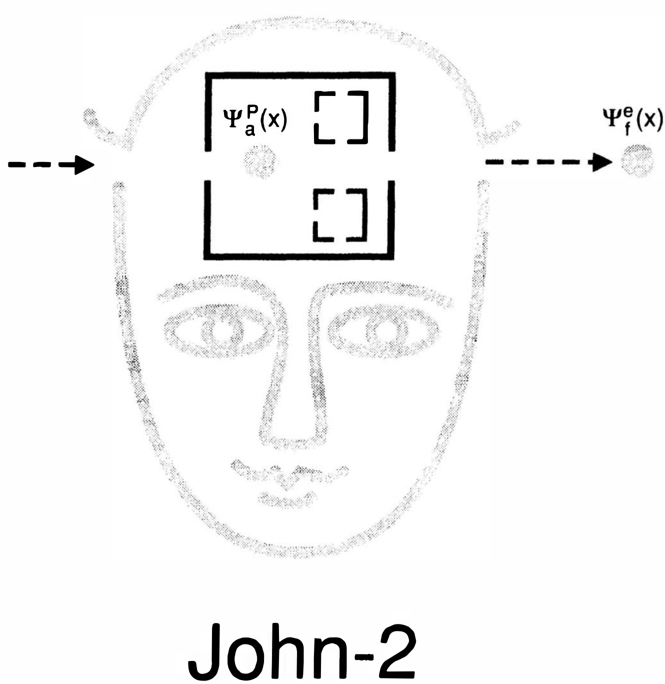

Figure 7.15

in writing), then (as the reader can easily confirm, since we've already had a good deal of practice with this sort of thing) the state of the composite system that consists of P and the measured electron is going to be:

$$(7.14)\quad 1/\sqrt{2}(|\text{soft}\rangle_P|\psi_a(x)\rangle_P|\text{hard}\rangle_e|\psi_f(x)\rangle_e - |\text{hard}\rangle_P|\psi_a(x)\rangle_P|\text{soft}\rangle_e|\psi_f(x)\rangle_e)$$

where $\psi_a(x)$ is a wave function which is nonzero only in the region a in figure 7.15, and $\psi_f(x)$ is a wave function which is nonzero only in the region f in figure 7.15. And of course the state in (7.14) is precisely the same state as the one in (7.10), which we discussed in considerable detail above.

And so (and all of what follows now can just be read off from what we found out before about (7.10)) if we were to ask John-2 to tell us, verbally (when (7.14) obtains), what the hardness of the measured electron is, then he would utter either "hard" or "soft,"[16] and of course the procedures that go on inside of John-2's head which *lead up* to one or the other of those two utterances will produce an *effective collapse* of the wave function in (7.14) onto (respectively) either its first or its second term, and so whichever of those two utterances John-2 ends up making will necessarily be *confirmed* to be correct by *any* subsequent measurement of the hardness of the electron *itself*.

And if, instead, we were to ask John-2 to tell us *in writing* (when (7.14) obtains) what the hardness of the measured electron is, then he would *write down* either "hard" or "soft."[17] The procedures that go on in John-2's head which lead up to one or the other of those *inscriptions* will produce the same sort of effective collapse as above, and so whichever of those two inscriptions John-2 ends

16. Note, by the way, that there is invariably going to be an objective and unambiguous matter of fact about *which* of those two utterances John-2 *makes*, since the act of uttering "soft" can be distinguished from the act of uttering "hard" in terms of the positions of all sorts of things (certain parts of John-2's tongue, for example, or his lips, or his throat).

17. And here again, for the same sorts of reasons, there will invariably be an objective matter of fact about which one of those he *does* write down.

up making will *also* necessarily be confirmed to be correct by any subsequent measurement of the hardness of the electron itself.

And moreover, if we were to ask John-2 for a verbal report on the hardness of the measured electron, when (7.14) obtains, and if, once that report is produced, we were to ask him for a *written* report on the hardness of that same electron, then the content of that written report will invariably coincide with the content of the earlier verbal report; and the content of any *second* verbal report on the hardness of that electron will also invariably coincide with the content of that first one; and the content of any later *written* report on that hardness will *also* invariably coincide with the content of any earlier one, and so on.

And so whatever plausible observable or functional criteria there may be for having a belief, or for having a true belief, or for having a justified true belief about the hardness of that measured electron, *all* of those criteria will plainly be satisfiable by anybody wired up as John-2 is, when a state like the one in (7.14) obtains.

And yet (and this is the punch line) if Bohm's theory is right, then John-2 can't *possibly* be a genuine "knower" of the hardness of the measured electron, when (7.14) obtains.

Here's why: Suppose (for example) that *P* happens to be in the upper half of region *a*, when (7.14) obtains, and suppose that John-2 is first requested to produce a verbal report about the hardness of the electron. Then the content of that report, and of any subsequent report, of either kind, about that hardness is going to be that the electron is *soft*. But in the event that John-2 is in precisely the condition described above and he is first requested to produce a *written* report about the hardness of the electron, *then* the content of *that* report, and of any subsequent report, of either kind, about that hardness is going to be that the electron is *hard*. And of course all of that will get *reversed* in the event that *P* happens to be in the *lower* half of region *a*, when (7.14) obtains.

And so, when (7.14) obtains (that is: before any report has been requested of John-2, but when John-2 claims to be a *knower* of the hardness of the electron, and when John-2 is in fact in a position to produce a *report* about that hardness, of *either* sort, which will with certainty be confirmed by any subsequent measurement of that hardness), John-2's *dispositions* are *completely unlike* those of any genuine rational *knower* of that hardness; since the *content* of whatever reports John-2 produces about that hardness will depend (as a matter of *fact,* but *not* in any *observable* way) on what *sort* of report is requested *first.*

The Incommensurability of Bohm's Theory and Many-Minds Theories

The business of comparing the empirical contents of Bohm's theory to the empirical contents of a *many-minds* theory turns out to be very peculiar.

To begin with, if *either one* of those two theories happens to be *the true theory of the world,* then the question of *which one* of them is the true one will be undeterminable by any imaginable sort of empirical evidence. It's pretty clear how that works: Each one of those two theories entails (as I've already mentioned) that the linear dynamical equations of motion are always exactly right, and each one of them entails that the probabilities of the outcomes of all of the sorts of experiments that *have* any outcomes (we'll talk more about that soon) are precisely the ones laid down in principle D of Chapter 2, and so (as a matter of fundamental principle) *there can't be any purely experimental means of deciding between them.*[18]

18. Any theory of the *collapse of the wave function,* by the way, is clearly going to *differ* from Bohm's theory and from a many-minds theory (and presumably from *any* theory in which there *isn't* any such thing as a collapse of the wave-function), in terms of its predictions about the outcomes of certain experiments.

Here's how that will work. Any theory of the collapse of the wave function (just in virtue of *what it is* to be a theory of the collapse of the wave function) is going to entail that there are certain superpositions of macroscopically different states of the world which, whenever they arise, collapse (more or less immediately) onto *one or another* of those states. Take any such theory (let's call it T). Design a hardness measuring device for which the "indicates-that-hard" state and the "indicates-that-soft" state are (as it were) macroscopically different *insofar as T is concerned* (that is: design the device in such a way as to guarantee that T entails that a state like the one in equation (5.2) will more or less immediately collapse onto the state in equation (5.1)). Prepare the device (either genuinely or effectively) in its ready state.

So, suppose it were to turn out that what our empirical experience entails is that the dynamical equations of motion are always exactly right (which is how I figure it probably will turn out, once the results are in); and suppose it were also to turn out that there aren't any purely theoretical reasons why one or the other of these two theories is somehow manifestly out of the running (which is maybe a little less certain). What that would mean is that questions about the structure of space and time, and questions about whether or not the world is deterministic (which were supposed to be the two central questions of the physics of this century, and which both happen to be questions on which these two theories radically disagree with one another), are the kinds of questions which there can't ever be scientific answers to. Period.

But there's something else that's weird too.

Consider the question (the one that I postponed a few paragraphs back) of *which* experiments *do* have any outcomes. Bohm's theory and many-minds theories disagree about that.

Consider, for example, a measurement of the hardness of a black electron that gets carried out by an automatic device with a pointer, like the one in figure 5.1. On Bohm's theory, there will be a determinate matter of fact about where that pointer ends up, and where that pointer ends up will be the sort of thing that counts as a piece of absolutely raw empirical data, the sort of thing (that is) that physics is fundamentally in the business of making predictions about. And of course on a *many-minds* theory, there *won't* be any determinate matter of fact about where that pointer ends up, and (consequently) where that pointer ends up can't *possibly* be the sort of thing that counts as a piece of absolutely raw empirical data, and (consequently) where that pointer ends up can't *possibly* be the

---

Feed a black electron into it. Let the interaction between the device and the electron run its course. Then measure the observable of the composite system consisting of the device and the electron that we talked about on pages 87–92, the one we called zip – color there, the one that's extremely hard (but not impossible) to measure. If T is true, then there will be nonzero probabilities of a number of different outcomes of the zip – color measurement; but if Bohm's theory is true, or if a many-minds theory is true, *then* the outcome of that measurement will invariably and with certainty be *zero*. And that's that.

sort of thing that physics is in the business of making any *predictions* about. And of course the conviction of any adherent of *Bohm's* theory to the effect that there *is* some matter of fact about where that pointer ends up will be perfectly explicable in the context of a *many-minds* theory as a simple *delusion.* And (consequently) there *can't* be any means of *resolving* that dispute.

And consider a measurement of the hardness of a black electron that gets carried out by a sentient observer whose brain is wired up as John-2's brain is. On a many-minds theory, there will be a determinate matter of fact about what each one of John-2's minds ends up thinking about the hardness of that electron, and those *thoughts* will be the sorts of things that count as pieces of absolutely raw empirical data and as the sorts of things that physics is fundamentally in the business of making predictions about. And of course on *Bohm's* theory, there *won't* be any matter of fact about what John-2 ends up thinking about the hardness of that electron, and (consequently) what he ends up thinking about that can't *possibly* be the sort of thing that counts as a piece of absolutely raw empirical data, or as the sort of thing that physics is fundamentally in the business of making predictions about. And of course the conviction of any adherent of a *many-minds* theory to the effect that there is some matter of fact about what any one of John-2's minds ends up thinking about that will likewise be perfectly explicable (because of the stuff we talked about in the first section of Chapter 6), in the context of *Bohm's* theory, as a delusion. And (consequently) there can't be any means of resolving *that* dispute either.

And so the upshot of all this is that it doesn't capture what's going on to say of these two theories that (in virtue of the impossibility of experimentally *telling them apart*) they're empirically *equivalent* to one another. What these two theories *are* is (in an extraordinarily radical way) empirically *incommensurable* with one another: What a many-minds theory takes physics to be ultimately about is *what observers think;* and it entails that there will frequently not even be *matters of fact* about *where things go.* And what *Bohm's* theory takes physics to be ultimately about (as I mentioned before) is *where things go;* and *it* entails that there will sometimes not even be matters of fact about *what observers think.*

Of course, there may happen to be sentient observers in the world whose thoughts supervene entirely on the *positions* of things. *Human* observers (the ones whose brains are wired up in the *natural* way, that is) are probably (most of the time) like that. And of course Bohm's theory and many-minds theories will *agree* that there are determinate matters of fact about the outcomes of any measurements that get carried out by observers like *that;* and (moreover) they will agree (as I mentioned above) that the *probabilities* of those outcomes will be precisely the ones laid down in principle D of Chapter 2. But that kind of thing (if it happens) will amount to a rather special case.