... 4 ...

The Measurement Problem

Now I want to begin to worry in earnest about whether or not the dynamics says the same thing as the postulate of collapse says about what happens to the state vector of a physical system when the system gets measured. Here's what looked worrisome back in Chapter 2: the dynamics (which is supposed to be about how the state vectors of physical systems evolve *in general*) is fully deterministic, but the collapse postulate (which is supposed to be about how the state vector of a system evolves when it comes in contact with a measuring device) isn't; and so it isn't clear precisely how the two can be consistent.

Let's figure out what the dynamics says about what happens when things get measured.

Suppose that everything in the world always evolves in accordance with the dynamical equations of motion. And suppose that we have a device (which operates in accordance with those equations, just like everything else does) for measuring the hardness of an electron; and suppose that that device works like this: The device has a dial on the front, with a pointer; and the pointer has three possible positions. In the first position the pointer points to the word "ready," and in the second position it points to the word "hard," and in the third it points to the word "soft." Electrons are fed into one side of the device and come out the other, and in the course of passing through (if the device is set up right, with the pointer initially in its "ready" position) they get their hardnesses measured, and the outcomes of those measurements get recorded in the final position of the pointer (figure 4.1).

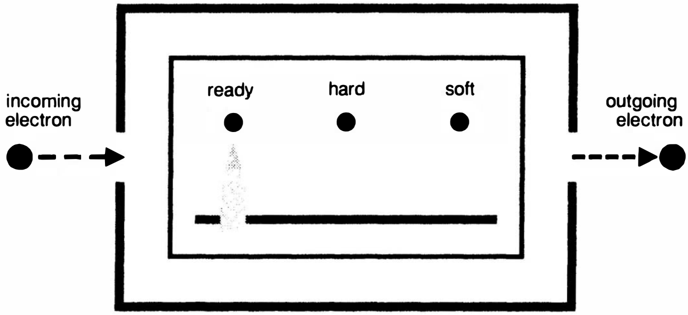

Figure 4.1

If the device is set up right, and if the dynamics is always true, then (to put all this another way) the dynamical equations of motion entail that it behaves like this:

(4.1) $|\text{ready}\rangle_m|\text{hard}\rangle_e \rightarrow |\text{“hard”}\rangle_m|\text{hard}\rangle_e$

and

(4.2) $|\text{ready}\rangle_m|\text{soft}\rangle_e \rightarrow |\text{“soft”}\rangle_m|\text{soft}\rangle_e$

That is: if the device (whose state vector is labeled with subscript *m*) is initially in the ready state, and if an electron (whose state vector is labeled with subscript *e*) that is hard gets fed through it, then the device ends up in the state wherein the pointer is pointing at "hard"; and if the device is initially in the ready state, and if a *soft* electron gets fed through it, then the device ends up in the state wherein the pointer is pointing at "soft." That's what it *means* for a measuring device for hardness to be a good one and to be set up right.

Now (still supposing that the dynamics is always true), consider what happens if this device (the one for measuring *hardness*) is set up right, and is in its ready state, and a *black* electron is fed into it. It turns out that (4.1) and (4.2), and the fact that the dynamical equations of motion are invariably linear, suffice by themselves to figure that out: The initial state of the electron and the measuring device is

$$(4.3) \qquad |\text{ready}\rangle_m|\text{black}\rangle_e = |\text{ready}\rangle_m\{1/\sqrt{2}|\text{hard}\rangle_e + 1/\sqrt{2}|\text{soft}\rangle_e\}$$
$$= 1/\sqrt{2}|\text{ready}\rangle_m|\text{hard}\rangle_e + 1/\sqrt{2}|\text{ready}\rangle_m|\text{soft}\rangle_e$$

which is precisely $1/\sqrt{2}$ times the initial state in (4.1) plus $1/\sqrt{2}$ times the initial state in (4.2). So, since (by hypothesis) the dynamical equations of motion entail that $|\text{ready}\rangle_m|\text{hard}\rangle_e$ evolves as in (4.1) and that $|\text{ready}\rangle_m|\text{soft}\rangle_e$ evolves as in (4.2) (that is: since this device is set up right, and since it's a good measuring device for hardness), it follows from the linearity of those equations that the state in (4.3), when the measuring device gets switched on, will necessarily evolve into

$$(4.4) \qquad 1/\sqrt{2}|\text{“hard”}\rangle_m|\text{hard}\rangle_e + 1/\sqrt{2}|\text{“soft”}\rangle_m|\text{soft}\rangle_e$$

That's how things end up, with certainty, according to the dynamics.[1]

And the way things end up according to the postulate of *collapse* (when you start with (4.3)) is

$$(4.5) \qquad \textit{either} \quad |\text{“hard”}\rangle_m|\text{hard}\rangle_e \ \text{(with probability } 1/2\text{)}$$
$$\textit{or} \quad |\text{“soft”}\rangle_m|\text{soft}\rangle_e \ \text{(with probability } 1/2\text{)}$$

1. This result often strikes people as mysteriously easy. The intuition is that measuring devices for hardness like the one described here must be extremely complicated contraptions (especially if you look at them on the level, say, of their constituent atoms) and must have extremely complicated equations of motion, the solution of which must be an extremely complicated matter. All of that is true. What simplifies things here is the fact that however complicated those equations may be, (4.1) and (4.2) are surely solutions of them (since the contraption we're dealing with is, by hypothesis, and whatever *else* it may be, a good measuring device for the hardness of an electron), and they (the equations) are surely *linear;* and those two facts are enough by themselves to insure that if a black electron is fed into the device, then those equations will entail that things will end up in the state in (4.4).

And the trouble is that (4.4) and (4.5) are measurably different situations.[2]

The state described in (4.5) is the one that's right; it is (as a matter of empirical fact) how things *do* end up when you start with (4.3).

The state described in (4.4) is *not* how things end up;[3] (4.4) is something very strange. It's a superposition of one state in which the pointer is pointing at "hard" and another state in which the pointer is pointing at "soft"; it's a state in which (on the standard way of thinking) *there is no matter of fact* about where the pointer is pointing.[4]

Let's make this somewhat sharper. Suppose that a human observer enters the picture, and *looks* at the measuring instrument (when the measurement is all done) and *sees* where the pointer is pointing. Let's figure out what the dynamics will say about that.

Suppose, then (just as we did before), that literally every physical system in the world (and this now includes human beings; and it

2. Perhaps this ought to be expanded on a bit. The point is that there are (in accordance with the postulates of quantum mechanics that were laid out in Chapter 2) necessarily measurable properties of the state in (4.4) whereby it can, in principle, be experimentally distinguished from either of the states in (4.5) and, as a matter of fact, from any other state whatever. There will be a good deal to say, later on, about precisely *what* those properties *are* (they're complicated ones, and their measurement will in general be extremely difficult); what's important for the moment is simply that those properties exist.

3. It isn't, at any rate, according to the conventional wisdom about these matters; but there will be much more to say about this later on.

4. Something ought to be mentioned in passing here, something that will turn out to be important later on.

What we've just discovered is that there is a certain fundamental effect of the carrying out of a measurement (namely: the emergence of some definite *outcome* of the measurement; the emergence of some *matter of fact* about precisely *what* the outcome of the measurement *is*) which is not predicted by the dynamical equations of motion.

But consider another effect of the carrying out of a measurement, one which we first described in the course of our discussions of hardness and color in Chapter 1: The carrying out of a measurement is disruptive of the values of observables of the measured system which are incompatible with the observable that gets measured. It turns out that the dynamical equations of motion *do* predict *that.*

includes the brains of human beings) always evolves in accordance with the dynamical equations of motion; and suppose that a black electron is fed through a measuring device for hardness that's set up right and that starts out in its ready state (so that the state of the electron and the device is now the one in (4.4)); and suppose that somebody named Martha comes along and looks at the device; and suppose that Martha is a competent observer of the positions of pointers.

Being a "competent observer" is something like being a measuring device that's set up right: What it means for Martha to be a competent observer of the position of a pointer is that whenever Martha looks at a pointer that's pointing to "hard," she eventually comes to *believe* that the pointer is pointing to "hard"; and that whenever Martha looks at a pointer that's pointing to "soft," she eventually comes to believe that the pointer is pointing to "soft" (and so on, in whatever direction the pointer may be pointing). What it means (to put it somewhat more precisely) is that the dynamical equations of motion entail that Martha (who is a physical system, subject to the physical laws) behaves like this:

(4.6) $$|\text{ready}\rangle_o|\text{ready}\rangle_m \rightarrow |\text{“ready”}\rangle_o|\text{ready}\rangle_m$$

and

$$|\text{ready}\rangle_o|\text{“hard”}\rangle_m \rightarrow |\text{“hard”}\rangle_o|\text{“hard”}\rangle_m$$

and

$$|\text{ready}\rangle_o|\text{“soft”}\rangle_m \rightarrow |\text{“soft”}\rangle_o|\text{“soft”}\rangle_m$$

In these expressions, $|\text{ready}\rangle_o$ is that physical state of Martha's brain in which she is alert and in which she is intent on looking at the

---

Look, for example, at the evolution from (4.3) to (4.4). Equation (4.3) is an eigenstate of the *color* of the electron whose hardness is about to be measured; but (4.4) (which is the state following the interaction of the electron with a good measuring device for the hardness, according to the equations of motion) *isn't*. If (4.4) is written out in terms of eigenstates of the color of the electron (which the reader can now easily do), *it* turns out to be a superposition of two states (with equal coefficients), in one of which the electron is black and in the *other* of which the electron is white.

pointer and finding out where it's pointing; $|\text{“ready”}\rangle_o$ is that physical state of Martha's brain in which she believes that the pointer is pointing to the word “ready” on the dial; $|\text{“hard”}\rangle_o$ is that physical state of Martha's brain in which she believes that the pointer is pointing to the word “hard” on the dial; and $|\text{“soft”}\rangle_o$ is that physical state of Martha's brain in which she believes that the pointer is pointing to the word “soft” on the dial.[5]

Let's get back to the story. The state of the electron and the measuring device (at the point where we left off) is the strange one in (4.4). And now in comes Martha, and Martha is a competent observer of the position of the pointer, and Martha is in her ready state, and Martha looks at the device. It follows from the linearity of the dynamical equations of motion (if those equations are right), and from what it means to be a competent observer of the position of the pointer, that the state when Martha's done is with certainty going to be

$$1/\sqrt{2}\,|\text{“hard”}\rangle_o|\text{“hard”}\rangle_m|\text{hard}\rangle_e + 1/\sqrt{2}\,|\text{“soft”}\rangle_o|\text{“soft”}\rangle_m|\text{soft}\rangle_e \tag{4.7}$$

That's what the dynamics entails.

And of course what the postulate of *collapse* entails is that when Martha's all done, then

$$\begin{array}{lll} \textit{either} & |\text{“hard”}\rangle_o|\text{“hard”}\rangle_m|\text{hard}\rangle_e & \text{(with probability } 1/2) \\ \textit{or} & |\text{“soft”}\rangle_o|\text{“soft”}\rangle_m|\text{soft}\rangle_e & \text{(with probability } 1/2) \end{array} \tag{4.8}$$

is going to obtain.

And (4.7) and (4.8) are empirically different. The state described in (4.8) is the one that's right; (4.7) is unspeakably strange. The state described in (4.7) is at odds with what we know of ourselves

5. It hardly needs saying that this is an absurdly oversimplified description of Martha's brain, and that this is an absurdly oversimplified account of the ways in which mental states are generally supposed to supervene on brain states; but all that turns out not to make any difference (not at *this* stage of the game, anyway). We can fill in the details whenever we want, to whatever extent we want. They won't change the arguments.

by *direct introspection.* It's a superposition of one state in which Martha thinks that the pointer is pointing to "hard" and another state in which Martha thinks that the pointer is pointing to "soft"; *it's a state in which there is no matter of fact about whether or not Martha thinks the pointer is pointing in any particular direction.*[6]

And so things are turning out badly. The dynamics and the postulate of collapse are flatly in contradiction with one another (just as we had feared they might be); and the postulate of collapse seems to be right about what happens when we make measurements, and the dynamics seems to be bizarrely *wrong* about what happens when we make measurements; and yet the dynamics seems to be *right* about what happens whenever we *aren't* making measurements; and so the whole thing is very confusing; and the problem of what to do about all this has come to be called "the problem of measurement."

We shall be thinking about that for the rest of this book.

6. This isn't anything like a state in which Martha is, say, *confused* about where the pointer is pointing. *This* (it deserves to be repeated) is something *really strange.* This is a state wherein (in the language we used in Chapter 1) it isn't right to say that Martha believes that the pointer is pointing to "hard," and it isn't right to say that Martha believes that the pointer is pointing to "soft," and it isn't right to say that she has *both* of those beliefs (whatever *that* might mean), and it isn't right to say that she has neither of those beliefs.