Chapter 8: The Collapse of the Quantum State

Jeffrey A. Barrett

May 2, 2020

1 Wigner’s solution

Eugene Wigner did not present his version of the Friend story as a criticism of quantum mechanics. Rather, it was part of his explanation for how one should properly understand the theory. In particular, he used it to argue that the quantum-mechanical state of a system must collapse when the system is apprehended by a conscious mind.

Wigner held that a consistent formulation of quantum mechanics requires one to endorse a strong variety of mind-body dualism, something that he believed most physicists had already accepted.

Until not many years ago, the ‘existence’ of a mind or soul would have been passionately denied by most physical scientists. ... There are [however] several reasons for the return, on the part of most physical scientists, to the Spirit of Descartes’ ‘Cogito ergo sum’ ... When the province of physical theory was extended to encompass microscopic phenomena, through the creation of quantum mechanics, the concept of consciousness came to the fore again: it was not possible to formulate the laws of quantum mechanics in a consistent way without reference to consciousness. (1961, 168)

The worry about consistency here comes, as we saw in the last chapter, when one tries to say when collapses occur. Wigner believed that one can avoid the quantum measurement problem by opting for a strong Cartesian dualism.¹

He argued for his proposed solution to the measurement problem as follows:

The important point is that the impression which one gains at an interaction may, and generally does, modify the probabilities with which one gains the various possible impressions at later interactions. In other words, the impression one gains at an interaction, called also the result of an observation, modifies the wave function of the system. ... [I]t is the entering of an impression into our consciousness which alters the wave function because it modifies our appraisal of the probabilities for different impressions which we expect to receive in the future. It is at this point that the consciousness enters the theory unavoidably and unalterably. (1961, 172–3)

¹Descartes famously held that the mind and body were metaphysically distinct objects. This is one of the central philosophical conclusions of his Meditations on First Philosophy (1641). The main argument is that since one can be certain that one exists as a thinking being (I am thinking, therefore I exist) while remaining uncertain whether one’s body is an illusion (produced perhaps by an evil demon), one’s mind and one’s body cannot be the same thing. While it is unlikely that most physicists believed that quantum mechanics committed them to this sort of mind-body dualism, it is significant that Wigner thought they did. He was a co-recipient of the Nobel Prize in Physics two years later in 1963 for his work in quantum mechanics and in particular the application of symmetry principles to the theory.
But it is not just our appraisal of the probabilities for different future impressions that changes when the impression one gains enters our consciousness. Wigner required that a physical collapse of the state occurs whenever a conscious mind apprehends the state of a measured system and gains the impression of the measurement result. It is this change in the physical state that requires one to change one’s appraisal of future probabilities.

On Wigner’s understanding of the theory, the dynamical laws of the standard collapse formulation of quantum mechanics are sharpened as follows:

4. Laws of motion (Wigner):

I (Wigner). Linear dynamics: when no conscious mind apprehends the state of a physical system $S$, it evolves in the standard deterministic, linear way: $\hat{U}(t_0, t_1)\langle \psi(t_0) \rangle_S$, where $\hat{U}$ is a unitary operator.

II (Wigner). Nonlinear collapse dynamics: when a conscious mind apprehends the state of a physical system $S$, it will instantaneously collapse to an eigenstate of the observable being measured $|\chi \rangle_S$ with the standard quantum probabilities $|\langle \chi | \psi \rangle|^2$, where $|\psi \rangle_S$ is the initial state of the system.

He believed that this sharpening of the laws was “required” for the consistency of the theory, and he considered it to be the “simplest way out” of the quantum measurement problem (1961, 180).²

Wigner believed that both the interaction between the measuring device $M$ and the object system $S$ and the interaction between the friend’s physical body $F$ and the measuring device $M$ in the Wigner’s Friend story were both perfectly ordinary physical interactions and, hence, correctly described by the usual deterministic linear dynamics. The composite physical system, then, evolves to the state

$$|\text{super}\rangle_{FMS} = \frac{1}{\sqrt{2}} |\uparrow_x \rangle_F |\uparrow_x \rangle_M |\uparrow_x \rangle_S + \frac{1}{\sqrt{2}} |\downarrow_x \rangle_F |\downarrow_x \rangle_M |\downarrow_x \rangle_S$$

But, Wigner argued, this cannot be the final state of the composite system. Rather, in order for there to be a determinate result, the composite system $FMS$ must be either

$$|\text{up}\rangle_{FMS} = |\uparrow_x \rangle_F |\uparrow_x \rangle_M |\uparrow_x \rangle_S$$

or

$$|\text{down}\rangle_{FMS} = |\downarrow_x \rangle_F |\downarrow_x \rangle_M |\downarrow_x \rangle_S$$

Wigner believed that were he to ask his friend the question “What did you feel about the result of your measurement before I asked you?”, the friend would certainly reply, “I told you already, I got the result [“$\uparrow_x$” or “$\downarrow_x$”]” as the case may be. That is, the friend would report that the result of her measurement “was already decided in [her] mind” before she was asked (1961, 176). Hence Wigner concluded:

If we accept this, we are driven to the conclusion that the proper wave function immediately after the interaction of friend and object was already either [state $|\text{up}\rangle_{FMS}$] or [state $|\text{down}\rangle_{FMS}$] and not the linear combination [state $|\text{super}\rangle_{FMS}$]. . . .It follows that the being with a consciousness must have a different role in quantum mechanics than the inanimate measuring device . . . .(1961, 176–7).

²This is a modification to rule 4 of the standard formulation. We will characterize alternative formulations of quantum mechanics by saying how they might be understood as modifications of the standard theory. Unless otherwise noted, the rest of the theory stays the same.
While it is not logically inconsistent to deny that the friend is right in reporting that she already had a determinate measurement result before she was asked, Wigner took such an option to be unacceptable. He argued that to deny that the friend has the same sort of determinate experiences that we do "is surely an unnatural attitude, approaching solipsism, and few people, in their hearts, will go along with it" (1961, 177–8). Further, by a basic principle of charity, it must be when the friend herself apprehends the state, and not just when Wigner asks her what her result was, that the composite system collapses to a state where she has a determinate measurement record. Given that everything else in the story is just ordinary physical systems interacting in ordinary ways, the only principled place to put the collapse is where a nonphysical entity, the friend’s mind, forms an impression of the result. Here the friend simultaneously creates a determinate physical record, makes that record accurate by causing a collapse of the correlated measuring device and object system, and apprehends the value of the record. And all this happens in a way that generates measurement records that satisfy the standard quantum statistics.

In some sense, this proposal immediately solves the measurement problem by saying when collapses occur. But one only really knows precisely when collapses occur here if one knows precisely which physical systems are associated with minds and the conditions under which these minds apprehend measurement outcomes.

For his part, Wigner advertised this as a feature of his theory and not a flaw. Since minds affect the quantum-mechanical states of physical systems, they also affect the objective, observable properties of those systems. Namely, if it turns out that one’s friend is not in fact conscious and hence does not cause a collapse, then the state of the composite system $FMS$ would be

$$\frac{1}{\sqrt{2}}| \uparrow_x \rangle_F | \uparrow_x \rangle_M | \uparrow_x \rangle_S + \frac{1}{\sqrt{2}}| \downarrow_x \rangle_F | \downarrow_x \rangle_M | \downarrow_x \rangle_S$$

so an $A$ measurement of the composite system would certainly yield the result +1. But if the friend caused a collapse, then the final state of the composite system will be $| \text{up} \rangle_{FMS}$ or $| \text{down} \rangle_{FMS}$, so an $A$ measurement of the composite system would yield result +1 or −1 each with probability 1/2. The upshot is that, while they would be extraordinarily difficult to perform, there are at least in principle experiments that would determine what systems cause collapses, and hence what systems are conscious. Wigner took the fact that one now has an empirical way for determining what systems are conscious to be a virtue—the theory and our intuitions concerning which systems are associated with conscious minds are inter-testable.

By saying how he thought the Friend story should go, Wigner provided a sharpening of the standard collapse theory and a proposed resolution to the measurement problem. His proposal has the virtue of saying more precisely when the linear dynamics and the collapse dynamics obtain. But it does not really pin this down until one knows which systems are in fact conscious. One could appeal to one’s intuitions about which systems are conscious, but this is arguably not much better than appealing to one’s intuitions concerning which systems should count as measuring devices on the standard theory.

The most salient problem is that Wigner’s theory requires a strong variety of mind-body dualism where minds do not supervene on physical states but cause physical events. An observer’s mental state is not determined by her physical state. Rather, her mental state causes her physical state to collapse whenever an impression enters into her consciousness, whatever that might mean. Wigner’s proposal is a return to the spirit of Descartes, but insofar as one is committed to mental processes supervening physical processes this is not an attractive prospect.

Finally, one would like one’s solution to the quantum measurement problem to suggest how one might reconcile quantum mechanics with the dynamical constrains of special relativity. But,
of course, Wigner’s proposal for when and how collapses occur is as incompatible with relativity as the standard formulation of quantum mechanics, and for precisely the same reasons.

2 GRW

Giancarlo Ghirardi, Alberto Rimini, and Tullio Weber (GRW) (1985; 1986) showed how to formulate a purely physical theory that says when and why collapses occur without requiring a commitment to a strong variety of mind-body dualism. The thought is that one wants macroscopic systems to have determinate classical properties so that one can explain things like tables with determinate locations and the motion of the moon but one also wants microscopic systems to obey the linear dynamics so that one can explain things like interference effects and the stability of matter. GRW accomplish this by changing the dynamical laws of the theory. In order to understand the main idea behind their proposal, we will start by considering a toy version of their theory we will call GRW*.

GRW* supposes that physical systems are composed of fundamental particles, and it describes the particles in terms of their wave functions. Specifically, the state of a particle $p$ is given by a complex-valued function $|\psi(r)\rangle_p$ over the possible positions where the particle might be found. The state of a collection of $N$ particles is given by their wave function over $3N$-dimensional configuration space.3

GRW* is just like the standard collapse formulation of quantum mechanics except that one replaces the linear dynamics and the collapse dynamics by a single hybrid dynamical law

4. (GRW*) Law of motion: The state of every physical system $S$ evolves in the standard deterministic, linear way $|\psi(t)\rangle_S = \hat{U}(t_0,t_1)|\psi(t_0)\rangle_S$ except that each particle in the system has a small probability $\lambda$ per unit time of collapsing randomly to an eigenstate of position. The probability of a particle collapsing to a position in a specified region is given by the standard quantum probabilities.

It is as if an external observer measures the position of a randomly selected particle at a randomly selected time and collapses it to an eigenstate of position, except that it happens spontaneously without requiring an observer of any sort. GRW chose a collapse rate $\lambda$ of about $10^{-16}$/sec, or about once every $10^8$ years. That makes collapses rare for microscopic systems containing just a few particles but frequent for macroscopic systems containing something on the order of Avogadro’s number of $6.022 \times 10^{23}$ of particles. The idea is that microscopic systems will behave quantum-mechanically most of the time, and macroscopic systems will behave almost classically most of the time.

Consider the Wigner’s Friend story in the context of GRW*. Suppose that the initial state of the friend’s object system $S$ is

$$\alpha|\uparrow_x\rangle_S + \beta|\downarrow_x\rangle_S$$

On the hybrid dynamics, the composite system evolves linearly until one of its constituent particles collapses. Suppose that time $t_1$ is just before the first particle in the composite system collapses.

3See the discussion of the wave function and $3N$-dimensional configuration space in section 5.7.
The system then starts by evolving as follows:

\[ |\psi(t_0)\rangle_{FMS} = |\alpha^{-}\rangle_F|\alpha^{+}\rangle_M(|\alpha\uparrow_x\rangle_S + |\beta\downarrow_x\rangle_S) \]

\[ \downarrow \]

\[ |\psi(t_1)\rangle_{FMS} = \alpha|\alpha^{-}\rangle_F|\alpha^{+}\rangle_M|\alpha\uparrow_x\rangle_S + \beta|\alpha^{-}\rangle_F|\alpha^{+}\rangle_M|\beta\downarrow_x\rangle_S \]

Suppose the measuring device pointer is a macroscopic system (so that one can read it) that indicates the measurement result by the direction it is pointing. Since it is macroscopic, it contains something on the order of Avogadro’s number of particles. Hence, it is very likely that one of the particles in the pointer, whose position is now correlated to the \(x\)-spin of system \(S\), will collapse to an eigenstate of position. Consequently, GRW predicts that the state \(|\psi(t_1)\rangle_{FMS}\) will be extremely unstable. In order to see the sense in which it is unstable, we will rewrite it in terms of the \(n\) particles that make up \(M\)’s pointer:

\[ |\psi(t_1)\rangle_{FMS} = \alpha|\alpha^{-}\rangle_F|\alpha^{+}\rangle_M|\alpha\uparrow_x\rangle_{m_1} |\alpha^{+}\rangle_{m_2} \cdots |\alpha^{+}\rangle_{m_n} |\uparrow_x\rangle_S + \beta|\alpha^{-}\rangle_F|\alpha^{+}\rangle_M|\alpha\downarrow_x\rangle_{m_1} |\alpha^{+}\rangle_{m_2} \cdots |\alpha^{+}\rangle_{m_n} |\downarrow_x\rangle_S \]

where \(m_k\) is the \(k\)th particle in the pointer and \(|\alpha^{+}\rangle_{m_k}\) is the state where \(m_k\) is in the region where the pointer indicates the result “\(x\)-spin up” and \(|\alpha^{-}\rangle_{m_k}\) is the state where it is in the region where the pointer indicates the result “\(x\)-spin down.” Since \(n\) is very large, while the collapse rate for each particle is small, it is highly likely that some particle will randomly collapse to an eigenstate of position in the next fraction of a second. The hybrid dynamics tell us that the probability of it collapsing to an eigenstate of position in the region where the pointer indicates the result “\(x\)-spin up” is \(|\alpha|^2\) and the probability of it collapsing to an eigenstate of position in the region where the pointer indicates the result “\(x\)-spin down” is \(|\beta|^2\).

Suppose that the first particle to collapse is particle \(m_{217}\) (as in figure 1). The probability that it will collapse to an eigenstate of position in the region where the pointer indicates the result “\(x\)-spin up” is \(|\alpha|^2\). Suppose it does. Since particle \(m_{217}\) is now determinately in the “\(x\)-spin up” pointer region, the probability of finding it in the “\(x\)-spin down” region is zero. If one assumes that the positions of all of the particles that constitute the pointer are perfectly correlated, one can think of the collapse as multiplying the second term of the state \(|\psi(t_1)\rangle_{FMS}\) of the composite system by zero. After renormalizing, the resultant state of the composite system is

\[ |\psi(t_{1.00001})\rangle_{FMS} = |\alpha^{-}\rangle_F|\alpha^{+}\rangle_M|\alpha\uparrow_x\rangle_{m_1} |\alpha^{+}\rangle_{m_2} \cdots |\alpha^{+}\rangle_{m_n} |\uparrow_x\rangle_S \]

\[ = |\alpha^{+}\rangle_F|\alpha^{+}\rangle_M |\uparrow_x\rangle_S \]
In this case, the friend will see the pointer determinately indicating \( x \)-spin up because its center of mass is in this region of ordinary space. Hence the friend will end up in the state

\[
|\text{up}\rangle_{FMS} = |\uparrow_x\rangle_F |\uparrow_x\rangle_M |\uparrow_x\rangle_S
\]

Similarly, the probability that particle \( m_{217} \) will collapse to an eigenstate of position in the pointer region \( "x\)-spin down" is \( |\beta|^2 \) in which case, the resultant state of the composite system is

\[
|\psi(t_{1.00001})\rangle_{FMS} = |\uparrow_x\rangle_F |\downarrow_x\rangle_{M1} |\downarrow_x\rangle_{M2} \cdots |\downarrow_x\rangle_{Mn} |\downarrow_x\rangle_S
\]

In this case, the friend will end up in state

\[
|\text{down}\rangle_{FMS} = |\downarrow_x\rangle_F |\downarrow_x\rangle_M |\downarrow_x\rangle_S
\]

and see the pointer determinately indicating \( x \)-spin down because its center of mass is in this region of ordinary three-dimensional space.

So GRW* predicts the standard quantum statistics for the position of the macroscopic measuring-device pointer, and hence predicts determinate position records that satisfy the standard quantum statistics. And it does so without requiring anything nonphysical to cause the collapse of the composite system. GRW* predicts the right quantum statistics for the determinate position of \( M \)'s pointer regardless of whether any conscious entity observes it.

On the hybrid dynamics, microscopic systems with relatively few particles will evolve linearly most of the time and will hence behave quantum mechanically, and the frequent collapses of the particles to eigenstates of position in macroscopic systems will keep their centers of mass close to eigenstates of position. And since the expectation values for position evolve classically under the linear dynamics, the expected positions of these centers of mass will evolve in an approximately classical way.

This is how GRW* explains the appearance of the classical, macroscopic world. But one only gets determinate physical measurement records, and hence determinate appearances, under special conditions.

Suppose one has an \( x \)-spin measuring device \( M^\dagger \) that records the result of its measurement of the system \( S \) in the position of a single particle \( p \) instead of the position of a macroscopic pointer. One might think of \( p \) as a microscopic pointer. Assuming that \( M^\dagger \) perfectly correlates the position of \( p \) with the \( x \)-spin of \( S \), the linear dynamics predicts that the composite system \( M^\dagger S \) will end up in the state

\[
\alpha|\uparrow_x\rangle_{M^\dagger} |\uparrow_x\rangle_S + \beta|\downarrow_x\rangle_{M^\dagger} |\downarrow_x\rangle_S
\]

But since the position of only the particle \( p \) is correlated with the \( x \)-spin of \( S \), GRW* predicts that this state will be extremely stable. Hence, it is highly unlikely that the interaction will produce a determinate measurement record. For this reason, GRW* would predict that there is likely no collapse of the state in the two-path experiment where we tried to record the position of the electron in the position of a single particle. Consequently, it would explain the destruction of the two-path interference effect \textit{in that case} as the result of decoherence effects and not a collapse of the state of the electron.

There is a good argument that a measuring device with a pointer consisting of a single particle should not count as a measuring device at all. If one wanted to know where \( M^\dagger \)'s pointer was, one
would need to read its position with a second measuring device \( M^\dagger \) with a pointer one could see. Suppose that \( M^\dagger \) perfectly correlates the position of its macroscopic pointer, and hence one that can be seen, with the position of \( p \). By the linear dynamics, the resultant state will be

\[
\alpha |\uparrow_x^\prime\rangle_{M^\dagger} |\uparrow_x\rangle_S + \beta |\downarrow_x^\prime\rangle_{M^\dagger} |\downarrow_x\rangle_S
\]

Inasmuch as \( M^\dagger \)'s pointer contains on the order of Avogadro’s number of particles, GRW* predicts that this state will be extremely unstable since it is very likely that a collapse of one of the particles in \( M^\dagger \)'s macroscopic pointer will collapse to an eigenstate of position leaving the composite system in either state

\[
|\uparrow_x^\prime\rangle_{M^\dagger} |\uparrow_x\rangle_S
\]

or state

\[
|\downarrow_x^\prime\rangle_{M^\dagger} |\downarrow_x\rangle_S
\]

with the standard quantum probabilities \(|\alpha|^2\) and \(|\beta|^2\) respectively. And this second interaction will very likely count as a successful measurement on the theory yielding a perfectly determinate measurement record. In each of the collapsed final states it will look like \( M^\dagger \) had a determinate measurement record all along and that the result that it recorded accurately represented the \( x \)-spin of \( S \). But, of course, \( M^\dagger \) had no determinate measurement record and \( S \) had no determinate \( x \)-spin before \( M^\dagger \)'s measurement.

This is something characteristic of measurement in GRW*. A measurement does not reveal the physical properties of an object system—it creates them.\(^4\) When \( S \) starts in the state

\[
\alpha |\uparrow_x\rangle_S + \beta |\downarrow_x\rangle_S
\]

it does not have any determinate \( x \)-spin whatsoever. But when the position of a measuring device’s macroscopic pointer becomes correlated to \( S \)'s \( x \)-spin and a particle in that pointer collapses to an eigenstate of position, that gives \( S \) a determinate \( x \)-spin by giving the pointer’s center-of-mass a determinate position.

That a pointer be macroscopic and well-correlated to the property of the object system being measured is, however, not by itself sufficient for there to be a determinate measurement record on GRW*. Suppose that one had a measuring device \( M^\dagger \) that recorded the result of its \( x \)-spin measurement of \( S \) in the \( z \)-spins of each particle in a system \( P \) consisting of an Avogadro’s number of spin-1/2 particles. Here the \( z \)-spins of the particles in \( P \) act as a sort of pointer—but not a very good one given how GRW* works. Suppose that the measurement interaction between \( M^\dagger \) and \( S \) perfectly correlates \( M^\dagger \)'s pointer \( P \) with \( S \)'s \( x \)-spin. The linear dynamics, then, predicts that the state of the composite system \( M^\dagger + S \) will be

\[
\alpha |\uparrow_x\rangle_{M^\dagger} |\uparrow_x\rangle_S + \beta |\downarrow_x\rangle_{M^\dagger} |\downarrow_x\rangle_S
\]

But here, since \( M^\dagger \)'s record is in terms of the \( z \)-spins of the particles of \( P \) and since collapses to eigenstates of position do nothing to spin, this state is perfectly stable under the GRW* dynamics. And it remains stable as long as no particle positions become correlated with the \( x \)-spin of \( S \) or the \( z \)-spins of the particles in \( P \).

\(^4\) Note that this is also true for other collapse theories. On Wigner’s theory, for example, there is typically no matter of fact about the value of the observable being measured until it is measured.
That said, inasmuch as $M^\dagger$’s record involves the $z$-spins of many particles, it may be difficult to keep the positions of particles in $M^\dagger$’s environment from becoming correlated with the $z$-spins of the particles in its pointer. If the $z$-spins of the particles in $P$ become correlated with the positions of enough particles, then the entangled superposition of the composite system consisting of $S$, $M^\dagger$, and its environment will be unstable. A collapse of any particle whose position is well-correlated with any of the $z$-spins of the particles in $P$ will collapse both the $z$-spin of that particle and the $x$-spin of $S$. This will consequently yield a determinate measurement record, and it will accurately represent the newfound and new made $x$-spin of $S$.

Here environmental interactions that would just produce decoherence effects under the linear dynamics alone, may also produce collapses on the GRW* hybrid dynamics. And these collapses would then yield determinate physical records—something that environmental decoherence cannot provide under the linear dynamics.

The measurement records that ultimately determine whether or not a theory is empirically adequate are those records that determine the experiences of human observers. For his part, Wigner simply stipulated that mental states are always determinate and cause collapses and then end up in states that agree with the physical records produced by the collapses. This immediately guaranteed determinate mental records on which one’s experience might supervene. Further, given the details of the collapse dynamics, these records will exhibit the standard quantum statistics. The directness of this explanation, however, is also what renders it ad hoc. To the credit of the theory, since GRW* does not stipulate nonphysical minds with always determinate states, it must account for the determinate experience of observers indirectly by appealing to standard physical facts.

Suppose that $Q$ is a physical property on which the mental records of an observer $F$ in fact supervene. GRW* will very likely make $F$’s mental records determinate and ensure that they are distributed in the standard quantum mechanics way if the value of $Q$ is either determined by or correlated with the positions of a large number of particles after a measurement interaction. That is, one likely gets a determinate mental record in the theory whenever the physical record on which the mental record supervenes is in fact correlated with the positions of many particles. Note that GRW* does not treat all physical observables the same way—position plays a special role in explaining how the theory seeks to account for our experience.

That GRW* requires that one’s measurement records ultimately supervene on the positions of particles is a version of the preferred basis problem. Since the Hilbert space formalism represents every physical observable of a system in the same way, there is at least an ad hoc feel to a dynamics that singles out one observable to make determinate, especially when position was chosen precisely because its being determinate also plausibly makes determinate those measurement records that we in fact take ourselves to have. Put the other way around, GRW did not choose to make the $z$-spin of a particular neutrino near the center of the spiral galaxy NGC 5457 determinate even though this was a perfectly coherent theoretical option—that would have done nothing whatsoever to explain our determinate experience of coffee cups and baseballs. Rather, they chose to make something determinate that arguably provides determinate measurement records given how we in fact record measurement outcomes and the sorts of interactions we believe typically obtain between those measurement records and their environments. This particular ad hoc selection of a preferred observable has the virtue of plausibly providing determinate measurement records. And choosing position as a physically preferred quantity is arguably less ad hoc than Wigner’s proposal of simply stipulating that mental states are always determinate.
There are, however, two significant problems with choosing position as the dynamically preferred physical quantity. The first problem is technical. GRW⋆ requires particles to collapse to eigenstates of position, but there are no such eigenstates in the standard Hilbert space representation. The reason is that Hilbert space needs to be separable, which means it can have at most a countably-infinite dimension. But since there are a continuously infinite number of possible positions, the Hermitian operator representing exact position would have to have a continuously infinite number of mutually orthogonal eigenvectors. This is only possible in a continuously-infinite dimensional space, and such spaces are not suitable for representing quantum-mechanical states since there is typically no unique linear decomposition of a vector with respect to a given basis.

The second problem is physical. GRW⋆, like the standard theory and Wigner’s theory, is not compatible with the constraints of relativity. Consider two particles $A$ and $B$ and two $x$-spin measuring devices, one particle and one measuring device at each end of the experiment $M_A$ and $M_B$, that record their measurement results in the positions of macroscopic pointers. The composite system starts in the stable state

$$\left| \text{"r"} \right>_{M_A} \left| \text{"r"} \right>_{M_B} \left[ \frac{1}{\sqrt{2}} (|\uparrow_x\rangle_A |\downarrow_x\rangle_B + |\downarrow_x\rangle_A |\uparrow_x\rangle_B) \right]$$

and, given the properties of good measuring devices, evolves by the linear dynamics to the unstable state

$$\frac{1}{\sqrt{2}} (|\text{"u"} \rangle_{M_A} |\text{"d"} \rangle_{M_B} |\uparrow_x\rangle_A |\downarrow_x\rangle_B + |\text{"d"} \rangle_{M_A} |\text{"u"} \rangle_{M_B} |\downarrow_x\rangle_A |\uparrow_x\rangle_B)$$

When the first particle in the pointer of either $M_A$ or $M_B$ collapses to an eigenstate of position, the composite system will collapse to either

$$|\text{"u"} \rangle_{M_A} |\text{"d"} \rangle_{M_B} |\uparrow_x\rangle_A |\downarrow_x\rangle_B$$

or

$$|\text{"d"} \rangle_{M_A} |\text{"u"} \rangle_{M_B} |\downarrow_x\rangle_A |\uparrow_x\rangle_B$$

with probability $1/2$ for each. The first pointer particle to collapse disentangles the state of the composite system and instantaneously gives both its own pointer and the pointer of the distant system determinate positions that correspond, respectively, to the determinate $x$-spins it instantaneously gives the two particles $A$ and $B$. Since there is no physical matter of fact concerning the order of spacelike-separated events, if $M_A$ and $M_B$ are sufficiently far apart and if they make their measurements at the nearly same time in the laboratory frame of reference, there will be no physical matter of fact concerning which pointer particle collapsed first. But that means that there can be no consistent dynamical story for how the systems disentangle if the two particles are described by a single wave function.

The incompatibility of GRW⋆ and relativity can also be seen by the role played by the new physical constants in the theory. Consider, for example, the collapse rate $\lambda$. Since relativistic clocks disagree in different inertial frames, one must choose a preferred inertial frame in order to even specify $\lambda$. But special relativity requires that there be no preferred inertial frame.

The incompatibility with relativity is particularly striking given the privileged dynamical role played by exact position in GRW⋆. If there were eigenstates of exact position, the expectation value of energy of such a state would have to be unbounded. The reason for this has to do with the quantum-mechanical relationship between position and momentum. The quantum state
representing the position of a particle $p$ is given by a complex-valued function $|\psi(r)\rangle$ over all possible positions. The value of this function given the probability of finding the particle in a spacial region $R$:

$$P(R) = \int_R |\psi(r)|^2 dr$$

Suppose that $|\psi(r)\rangle$ is a gaussian wave packet with a probability density given by $|\psi(r)|^2$. The same state can be expressed as a complex-valued function of momentum $|\phi(p)\rangle$ with a probability density given by $|\phi(r)|^2$. The two wave functions $|\psi(r)\rangle$ and $|\phi(p)\rangle$ are related to each other by a mathematical operation called a Fourier transform. This operation has two properties that matter here. First, if $|\psi(r)\rangle$ is gaussian, then so is $|\phi(p)\rangle$. Second, the narrower the gaussian wave packet $|\psi(r)\rangle$, the wider the gaussian wave packet $|\phi(p)\rangle$ and the other way around (as in figure 2). This means that the sharper particle $p$’s position, the less sharp its momentum. And

![Figure 2: wave functions in position and momentum space](image)

this means that the more determinate $p$’s position, the higher the probability of finding it with a large momentum. In the limit as $p$’s position becomes perfectly determinate, the expectation value of its momentum is unbounded, which means that the expectation value for its energy would also be unbounded. The upshot is that if any particle in a measuring-device pointer ever did collapse to a state were it had a perfectly precise position, the device would explode destroying the universe. This renders GRW⋆ incompatible with relativity insofar as one takes relativity to require the conservation of energy. Inasmuch as the universe is still here, it also means that GRW⋆ is not even close to being empirically adequate.

---

5The variable $p$ in the expressions here is particle $p$’s momentum, and $r$ is position.
3 GRW

GRW explicitly recognized that GRW* is empirically unacceptable. The theory that they ultimately presented, which we will just call GRW, is similar to GRW* except that they replaced the dynamics of the standard formulation of quantum mechanics with a more complicated hybrid law

4. Law of motion (GRW): The state of every physical system $S$ evolves in the standard deterministic, linear way $|\psi(t_1)\rangle_S = \hat{U}(t_0, t_1)|\psi(t_0)\rangle_S$ except that each particle has a small probability $\lambda$ per unit time of randomly collapsing to a narrow gaussian wave packet in position with width $1/\sqrt{\alpha}$. The probability of the wave packet being centered in a specified region $R$ is given by the standard quantum probabilities.

In contrast with the GRW* dynamics, the state here evolves as if an observer made an unsharp quantum measurement of the position of a randomly selected particle at random times and caused a corresponding unsharp collapse to a narrow gaussian state that is close to an eigenstate of position. But, again, on the hybrid dynamics this happens at a particular moment to the wave function in $3N$-dimensional configuration space randomly and without the necessity of an observation.

The hybrid dynamics still fails to conserve energy, but GRW chose the width of the collapsed gaussian wave packet $1/\sqrt{\alpha}$ to be wide enough, given the particle collapse rate, to keep the violation of the conservation of energy low enough to explain why we have not seen it yet. They also chose the width of the gaussian to be narrow enough so that a collapsed system in that state might plausibly be said to have an almost determinate position (specifically, they took $1/\sqrt{\alpha} \approx 10^{-5}$ cm). Their choice of $\lambda \approx 10^{-16}$/sec makes collapses rare for those systems we know behave linearly and common for those systems that we take to have determinate positions. And it works with the proposed width of the collapsed wave packet to make the violation of the conservation of energy small enough to explain why we have not seen it. While all this is clearly ad hoc, it works pretty well.

Consider this hybrid dynamics in the context of a measurement of the $x$-spin of a system $S$. Suppose $S$ starts in the symmetric superposition

$$\frac{1}{\sqrt{2}}(|\uparrow_x\rangle_S + |\downarrow_x\rangle_S)$$

The linear dynamics, then, predicts that the state of the composite system $M + S$ will be

$$\frac{1}{\sqrt{2}}(|\uparrow_x\rangle_M|\uparrow_x\rangle_S + |\downarrow_x\rangle_M|\downarrow_x\rangle_S)$$  \hspace{1cm} (3.1)

If $M$ records its result in the position of a macroscopic pointer, then this state is unstable under rule 4(GRW) because of the choice of collapse rate. But this time, when the first pointer particle collapses, it collapses to a narrow gaussian wave packet instead of an eigenstate of position with the center of this wave packet determined by the usual quantum probabilities. Suppose that the first pointer particle to collapse is $m_{17}$. Its wave packet has equal chance here of ending up centered in region “$\uparrow_x$” and centered in region “$\downarrow_x$”. Suppose it collapses to a state that is centered in region “$\uparrow_x$”. Since the resultant state of $m_{17}$ is a gaussian wave packet, it is still in a superposition of being in region “$\uparrow_x$” and in region “$\downarrow_x$”. It is just that the norm-squared of the amplitude associated with it being in region “$\downarrow_x$” is now very small. The effect of $m_{17}$’s collapse, then, is to
multiply the second term of the state above by a factor that is just slightly greater than zero. After
renormalizing, the resultant state is

$$|\text{grw}\rangle_{MS} = \alpha |\uparrow_x\rangle_M |\uparrow_x\rangle_S + \beta |\downarrow_x\rangle_M |\downarrow_x\rangle_S$$

where $|\alpha|^2 \gg |\beta|^2$. Note that $|\text{grw}\rangle_{MS}$ is not an eigenstate of $M$’s pointer indicating the result “$\uparrow_x$”
or anything else. On the standard interpretation of states, this is a state where $M$ is close to
recording a determinate measurement result but doesn’t, $S$ is close to having a determinate $x$-spin
but doesn’t, and the two systems are close to being disentangled but aren’t. So this is a state
where $M$ fails to record or indicate any determinate measurement outcome whatsoever on the
standard interpretation of states.

This is the tails problem. Under plausible assumptions, $|\text{grw}\rangle_{MS}$ is close (in the Hilbert space
metric determined by the inner product) to a state where the measurement record is determinately
$x$-spin up. But if one wants to say that $|\text{grw}\rangle_{MS}$ is a state where $M$ determinately indicates
the result $x$-spin up, then one needs a new way to interpret quantum-mechanical states.

To this end, one might further tune the theory by replacing the standard interpretation of states
with a new rule 3(GRW)

3. Interpretation of states (GRW): A system $S$ has a determinate property if and only if it is
sufficiently close to being in an eigenstate of having that property.

Here one can either choose a concrete standard of what one means by sufficiently close once and
for all or, as with classical treatments of vague predicates, allow the standard to depend on the
explanatory context at hand. One advantage of the latter is that it does not involve adding another
physical parameter to the theory. If one wants to explain why one’s measuring device does not
explode, one might note that neither the pointer nor any of the particles that constitute it are
in fact in eigenstates of position. But if one wants to explain why the pointer appears to record
the determinate outcome $x$-spin up, one might argue that its center of mass is close enough to an
eigenstate of position that it should count as having the corresponding determinate position.

That even classical predicates are typically vague for macroscopic systems gives some plausibility
to the proposed revision of the standard interpretation of states. Consider the claim that the Eiffel
Tower is in Paris. No one would take this to mean that every atom of iron that was present in
the original structure in 1889 is now in Paris. Some of the original Eiffel Tower atoms are almost
certainly in New York, which does precisely nothing to undermine the every-day truth that the
Eiffel Tower is in Paris. The predicate “is in Paris” is vague, but still useful. The fact that the
Eiffel Tower is in Paris (even when some of its original atoms are in New York) explains why one
sees the Eiffel Tower when one looks north-west down the Champ de Mars. The thought is that
the vague claim that the pointer on a particular measuring device records the result $x$-spin up is
similarly useful. The fact that the pointer indicates the result $x$-spin up (even when it is not in an
exact eigenstate of position) explains why one sees it pointing at the result $x$-spin up.

One might, however, object that there is an important disanalogy between the two cases. The
Eiffel Tower appears to be in Paris on classical grounds because most of its (microscopic) parts
are simply and determinately in Paris. But GRW does not characterize any of the parts of the
pointer as simply and determinately pointing at any particular measurement result. Because of the
infinite tails of the gaussian wave packet, one’s finest-grained description will always characterize the
pointer and its parts as being in an entangled superposition of recording all possible measurement
outcomes, being part of the motor of a taxi in Manhattan, being in orbit around $\alpha$ Centauri, etc.
The real issue, one might insist, is not the language of vague predicates. Rather, it is whether a pointer with an almost determinate position in the sense provided by GRW looks like a pointer with a fully determinate position. It concerns whether the theory correctly predicts our determinate experience of seeing macroscopic objects with perfectly determinate positions.

For its part, the theory says that a physical observer will typically end up in an entangled superposition of having recorded every possible measurement outcome. One of the records will get most of the quantum-mechanical amplitude, but that does not mean that it will feel like seeing a pointer with a fully determinate position. All it means formally is that the norm-squared of one coefficient on one term in the composite state is much larger than the other. The empirical adequacy of the theory presumably depends on our being able to say what one would experience in such a state. And that apparently has nothing to do with the language of vague predicates.

That said, it is not at all clear that we need the center of mass of a system to be exactly determinate in order to explain why it appears to have a determinate classical position. Indeed, inasmuch as the expectation value for energy would be unbounded under standard quantum assumptions, we have never seen a pointer or any other physical system in an exact eigenstate of position, and we never want to see it. Since the standard formulation of quantum mechanics does not even allow one to represent such states, if that theory ever explained the apparent determinate positions of macroscopic objects, then GRW arguably does as well. Further, it is not at all clear that we need a physical theory to tell us what it feels like to be in a particular physical state. It is, one might argue, enough for there to be something physical on which our determinate experience might plausibly be taken to supervene. And GRW arguably provides physical records that do that.

There is more to say about what one should want from a physical theory that purports to be empirically adequate and hence to explain our experience. This will be a recurring theme throughout the rest of the chapter and book. That it need only predict something physical on which our determinate experience might be taken to supervene is arguably too weak. We want the supervenience relation needs to be one that meshes with our beliefs about the sorts of physical facts that might in fact matter to experience. Starting with the next section, we will see how this plays out in a number of concrete formulations of quantum mechanics.

Like GRW*, GRW has the salient virtue of treating all physical processes, including measurement, in precisely the same way, but, unlike GRW*, it has the added virtue of predicting a relatively stable universe. GRW conclude that their theory “reproduces in a consistent way quantum mechanics for microscopic objects and classical mechanics for macroscopic objects, and provides the basis for a conceptually appealing description of quantum measurement” (Ghirardi, et al. 1986, 491). But it also exhibits a number of unsatisfactory features.

Instead of the position being the preferred physical quantity as in GRW*, here it is almost position with gaussian width $1/\sqrt{\alpha}$. Inasmuch as choosing position as a preferred physical quantity and specifying a just-right collapse rate of $\lambda$ was ad hoc, this choice makes the theory look yet more ad hoc. Indeed, insofar as GRW explicitly constructed their theory and chose the collapse rate $\lambda$ and the gaussian width $1/\sqrt{\alpha}$ so that microscopic systems will almost always behave linearly and hence exhibit distinctive quantum behavior and the centers-of-mass of macroscopic systems will almost always remain close to eigenstates of position, exhibit roughly classical behavior, and not heat up so much that we would have already noticed, the theory is manifestly ad hoc from the start.

In addition to being hoc, there is nothing subtle about the incompatibility of this version of GRW and relativity. Since clocks in different inertial frames disagree, one must choose a preferred
inertial frame to specify the collapse rate $\lambda$; but here, since different inertial frames do not agree on measurements of length either, one must also choose a preferred inertial frame to specify the width $1/\sqrt{\alpha}$ of the collapsed wave packet.

Consider the EPR experiment again. A collapse of a single particle in either of the two pointers in GRW will instantaneously collapse the state of the composite system just as in the standard theory. Here it will give both pointers almost determinate positions, both particles almost determinate, almost opposite spins, and almost disentangle the systems. The statistical distribution of such records will be the same as predicted by the standard collapse theory. But like both the standard theory and GRW*, one must be able to specify the temporal order in which spacelike separate events occur in GRW in order to tell a consistent dynamical story, and special relativity says that there is simply no physical matter of fact concerning the temporal order of spacelike separated events. The incompatibility of GRW and relativity can also be seen in the violation of the conservation of energy, something that relativity arguably requires. In short, if one believes that special relativity is right, then this version of GRW can’t be.

4 GRWr, GRWm, and GRWf

There are various ways of understanding GRW. The ontological commitments associated with each provide different accounts of experience.

Inasmuch as they take macroscopic systems to be constituted by particles, one might imagine that Ghirardi, Rimini, and Weber are committed to an ontology that includes particles. As they put it, a composite system is “a system of $N$ particles” and the localization process (the spontaneous collapse of a particle) occurs “independently for each constituent of a many-particle system” (1986, 476). But this is at least somewhat misleading. The theory characterizes the state of the particles in terms of the wave function of the composite system in configuration space. Consequently, one cannot appeal to the particle-like properties of the particles to account for the experience of observers. Rather, the talk of particles ultimately just serves to characterize the degrees of freedom of the system and hence the structure of the state space. Physically, the theory provides nothing beyond the wave function of the composite system and how it evolves to account for our experience.\(^6\) As a result, the most straightforward way to interpret GRW is arguably in terms of a direct wave function realism.

In parallel with the options discussed below we will call this interpretation GRWr.\(^7\) Here one takes the wave function in configuration space to be a concrete physical object. As David Albert describes the view

```
The sorts of metaphysical objects that wave functions are, on this way of thinking, are (plainly) fields— which is to say that they are the sorts of of objects whose states one specifies by specifying the values of some set of numbers at every point in the arena in which they live, the sorts of objects whose states one specifies (in this case) by specifying the values of two numbers . . . at every point in the configuration space of the universe. (2013, 53)
```

\(^6\)On this view, particles cannot be understood as providing a primitive ontology for the theory. We will discuss the significance of this as we go.

\(^7\)Inasmuch as this view might be understood as simply sharpening the ontological commitments of GRW themselves, one might take GRW and GRWr to be the same theory. Calling it GRWr here just marks the explicit interpretation of the wave function as a physically real field in configuration space.
The theory is just GRW understood as describing a field in configuration space as the only physically real thing. As Albert puts it, the world consists of “exactly one physical object—the universal wave function” (Albert 2013, 54).

This view may seem to be closely analogous with the Eiffel Tower example. Here one might imagine that a measuring device pointer indicates a particular result because most of the field stuff that constitutes the pointer in fact indicates that result after a measurement-like interaction. There is something right about this, but the situation is more subtle than this quick characterization might suggest.

The wave function \( |\psi(r)\rangle_S \) is a complex-valued function over points \( r \) in a \( 3N \)-dimensional configuration space, where \( N \) is the number of particle degrees of freedom in system \( S \).\(^8\) As a result, the wave function representing the world is a function in \( 3N \)-dimensional configuration space, where \( N \) is the total number of particles in the universe, something many orders of magnitude greater than Avogadro’s number. So if one takes \( |\psi(r)\rangle_S \) to directly represent a physical object, the only physical object, then it is an object that lives in an unimaginably high-dimensional space, not in ordinary three-dimensional space. It is this object that we are seeing when we observe pointers, baseballs, coffee cups, and such. Indeed, we are ourselves also just manifestations of this single, high-dimensional object. One is left, then, with the task of providing a dynamical story for how this object gives rise to observers who seem to experience everyday three-dimensional macroscopic objects situated in ordinary three-dimensional space. That one can do precisely that should seem plausible. Given how the wave function evolves in GRW, there is clearly something on which the standard quantum measurement records might supervene. But the story one tells here requires some care.

In order to account for an observer’s experience of seeing an ordinary three-dimensional object where she sees it, one needs a special sort of supervenience relation. Specifically, one needs to suppose that one’s experience roughly supervenes in a particular way on the observer’s brain-record degrees of freedom of the full wave function in \( 3N \)-dimensional configuration. One needs a supervenience relation that says something like that whenever the quantum state is within some cutoff distance \( \epsilon \) of a brain-record eigenstate in Hilbert space, then one has a determinate experience corresponding to that eigenstate. This special supervenience relation provides an interpretational principle that ties physical states to experience. What one is seeing when one looks at an ordinary extended object here is coarse-grained features of the wave function in \( 3N \)-dimensional configuration space that arise from the correlations between the physical degrees of freedom.

On GRWR all one has to explain one’s experience is an evolving field in a high dimensional space. It is the correlations in the degrees of freedom of the \( 3N \)-dimensional wave function that produce the illusion of ordinary three-dimensional objects in ordinary three-dimensional space behaving as we find them. Space seems to be three-dimensional because of the structure of these correlations. Specifically, it is because the Hamiltonian treats the degrees of freedom in \( 3N \)-dimensional configuration space as if they described a collection of \( N \) particles interacting in three-dimensions.\(^9\) The determinate experience of a three-dimensional world inhabited by ordinary three-dimensional objects in GRWR results from (1) the precise structure of correlations in the \( 3N \)-dimensional wave function under the GRW dynamics and (2) the special supervenience relation that relates mental

\(^8\) As discussed in section 5.7.

\(^9\) The Hamiltonian gives the energy properties of a physical system—in this case the entire physical world. In so doing, it determines the linear operator that shows up in the standard quantum dynamics by describing how the particle degrees of freedom are dynamically related.
states and the $3N$-dimensional wave function. But there are other ways to understand GRW.

On GRWm one starts with the wave function in configuration space under the GRW dynamics but understands it as coding for a continuous distribution of matter in ordinary three-dimensional space. Specifically, one uses the wave function in configuration space to calculate the marginal mass distribution associated with each particle in three-dimensional space then sums over the mass densities for all $N$ particles. This gives a description of the physical world in terms of a continuous distribution of matter at each time. We and everything we see are the result of this continuous matter sloshing about in ordinary three-dimensional space.

Inasmuch as this view involves a matter field in ordinary space, it might also seem to be closely analogous to the Eiffel Tower example. Here one might imagine that sees the measurement pointer where one does because it is just a pointer-shaped bump in the distribution of matter and one simply sees the location of the bump when one looks for the pointer. But again, the situation is more subtle than it might at first appear.

Here one needs to spell out the details of the dynamical story that ties one’s brain record of the position of the pointer to the mass density that constitutes the pointer. Inasmuch as this story will involve details of how the wave function evolves in $3N$-configuration space, it is unclear that the account of experience one ends up with is significantly more intuitive than on GRWr. The point is that one is not directly seeing the pointer on GRWm. Observation is mediated either by the universal wave function in $3N$-dimensional configuration space or by something like a brute-fact law for the evolution of the matter density that codes for the dynamical information in the wave function. In neither case is this particularly intuitive. We will return to this issue of intuitiveness in connection with GRWf since it arguably has all of the virtues of GRWm plus a few extra.

On GRWf the wave function in configuration space determines the probability densities of flashes in ordinary 3-space, and it is constellations of such flashes that explain the appearance of macroscopic objects. One might calculate the location of flashes by evolving of the wave function in configuration space under the standard GRW dynamics, then associating a flash with the spacetime point at the center of the gaussian wave packets that are generated by the random collapses. Since flashes just occupy spacetime locations, they are not the sort of things that have energies. Hence, one might argue, there is in fact no actual violation of conservation of energy.

Allori, Goldstein, Tumulka, and Zanghi describe how GRWf accounts for our experience:

[F]or a reasonable choice of the parameters of the GRWf theory, a cubic centimeter of solid matter contains more than $10^8$ flashes per second. That is to say that large numbers of flashes can form macroscopic shapes, such as tables and chairs. That is how we find an image of our world in GRWf. (2012, 6)

But if one takes the flash ontology seriously, flashes are more than just the manifest image of the world—there is a sense in which they are the world. On this account, physical objects are

\footnotetext{10}{See Albert (2013) and (2015) for discussions regarding how one might account for experience in GRW-type theories.}

\footnotetext{11}{This view was proposed by Benatti, et al. (1995). See Allori (2013) for a description of the approach and a comparison with other options.}

\footnotetext{12}{See Allori, Goldstein, Tumulka, and Zanghi (2012) for a prescription for how to do this.}

\footnotetext{13}{The inspiration for this view can be traced to John Bell’s (1987, 181–195) discussion of quantum interference phenomena. See also Tumulka (2007) and Allori (2013, 67–8) for descriptions of this approach.}

\footnotetext{14}{That said, insofar as the flashes constitute pointers that indicate the same measurement outcomes as pointers in GRW, it would nevertheless appear that there was a small violation of conservation of energy—precisely the violation predicted by standard GRW.}
constituted by constellations of flashes. There is, however, an important sense in which we never see the flashes that constitute an extended object.

The flashes that constitute an object are not themselves the manifest image of the object. Inasmuch as the flashes do not show up in the theory’s dynamics, they do not do anything. Specifically, they do not cause an observer’s experience of the positions, shapes, or anything else of the extended three-dimensional objects they constitute. Flashes just occupy spacetime points, distributed with the standard quantum statistics.

To be sure, the theory predicts that there will be a statistical correlation between the flashes that constitute an extended object and the flashes that constitute a good observer’s brain records of that object. So if one explains one’s determinate experience by supposing that it supervenes on the value of one’s brain-record flashes, one might infer probabilities concerning the position or shape of an extended object from one’s experience. But one’s experience does not supervene on the flashes that constitute the extended objects one takes oneself to see. Rather, it supervenes on the flashes that constitute one’s brain records, assuming that one can make good sense of that. Again, the correlation between the two is mediated by either the wave function in 3N-dimensional configuration space or by something like a brute-fact law that codes for the dynamical information in the wave function and determines the correlated locations of the flashes.

That flashes have no dynamical role in GRWf is an important feature of the theory. It allows for a sort of dynamical compatibility between GRWf and relativity. Since flashes just occupy spacetime points and do not do anything dynamically and since there is nothing non-relativistic about the statistical distribution of flashes, one just needs a relativistic description of the evolution of the wave function to get a relativistic formulation of GRWf.

Roderick Tumulka (2007) has shown how to do this. Getting a relativistic description of the evolution of the wave function involves a trick. Consider a situation where we want to represent a pair of distant particles in something like an EPR state. Rather then there being a simple matter of fact regarding their composite state or the state of either particle alone, the trick is to allow the wave function of the composite system to depend on the choice of a spacelike hypersurface. Since the wave function is hypersurface-dependent, there is typically no simple matter of fact regarding whether a GRW collapse has occurred at a time. While the state of the composite system depends on one’s choice of inertial frame, the states associated with different inertial frames (associated with different families of hyperplanes) can be related in such a way that they are all compatible with the standard quantum statistics for the location of flashes under the GRW dynamics. This is possible because the standard quantum statistics do not themselves select a preferred inertial frame. The collection of hyperplane-dependent wave functions does not determine a preferred inertial frame precisely because they are hyperplane-dependent, and the statistical distribution of flashes does not determine a preferred inertial frame either. Since all one has are the hyperplane-dependent wave functions and the spacetime flashes, and since neither of these determines an inertial frame, such a formulation of GRWf might be understood as satisfying the dynamical constraints of relativity.

There are two things to note. First, this does not mean that GRWf is local. There is a sense in which the standard quantum statistics themselves characterize nonlocal correlations. Specifically, while a hypersurface-dependent formulation of GRWf is dynamically compatible with relativity,

\[\text{15}\] The distribution of flashes is described by the standard quantum statistics which do not allow one to send superluminal signals nor do they select any preferred inertial frame. See Bell (1987, 52–62) for one discussion.

\[\text{16}\] This idea has a long tradition among people who want to formulate a collapse theory that does not require one to select a preferred inertial frame. See for example Aharonov and Albert (1981). Note that the wave function remains a function in 3N-configuration space—it is just (typically) a different function for each choice of spacelike hyperplane.
the distribution of flashes still violates the Bell inequalities. In this sense, a sense weaker than dynamical incompatibility with relativity, the theory is nonlocal just by dint of being empirically adequate. Second, since there is typically no simple matter of fact concerning the wave function of a system at a time, this means that the flashes must determine the values of the local measurement records. On this view, one might argue, the collection of hypersurface-dependent wave functions just amounts to a part of the statistical recipe for where to put flashes. And again, any particular distribution of flashes over spacetime is perfectly compatible with the dynamical constraints of relativity.

While such a theory is dynamically compatible with relativity in the sense just described, there is also a sense in which it is not a relativistic theory at all. One might expect a truly relativistic formulation of quantum mechanics to predict the phenomena that are taken to be characteristic of relativistic quantum mechanics, phenomena like particle creation and annihilation, how photons are found to scatter from a single free electron, etc. One should want novel relativistic quantum phenomena to result naturally from imposing relativistic constraints on quantum mechanics. One should not have to put such phenomena in later by hand.17 The explanation of the phenomena should follow from the relativistic dynamics in the context of quantum mechanics. Settling for compatibility with relativistic constraints by simply stipulating hypersurface-dependent states is nothing like that.18

5 empirical ontology and experience

We have seen how GRW might be formulated in terms of a physically real field in configuration space, a matter distribution in ordinary three-dimensional space, and flashes at spacetime points. Such options represent different choices of an empirical ontology. One’s choice of empirical ontology specifies the basic structure of the physical world and hence the entities that one can appeal to in explaining experience.

A recurring issue over the next few chapters concerns the sort of empirical ontology that would allow for a satisfactory account of experience. One approach is to require that the empirical ontology be a primitive ontology—an ontology of ordinary objects or fields in three-dimensional space that directly explain our experience.19 The thought is that if a theory has a primitive ontology, then one can think of the objects of the theory as providing a manifest image of the world that simply corresponds to our experience of a three-dimensional world with ordinary three-dimensional objects.

17One would, for example, like to be able to derive something like the Klein-Nishina (1929) scattering formula from one’s relativistic formulation of quantum mechanics. This consideration also applies in the context of formulations of Bohmian mechanics that seek to account for relativistic phenomena by sticking them in by hand. Everettian formulations of quantum mechanics are arguably the best positioned to explain genuinely relativistic phenomena. But, as we will see over the next two chapters, it is nontrivial for such theories to explain even determinate measurement records.

18That said, it may be possible to make some progress here. Tumulka’s characterization of the state of a relativistic system is given by a multi-time wave function with one time variable for each of the N particle degrees of freedom. The evolution of the multi-time wave function is given by N linear equations that must satisfy a consistency condition. This condition is automatically satisfied for Hamiltonians describing non-interacting particles, but it is very difficult to satisfy for interacting particles in even the simplest of situations. But there is reason to suppose that the prospects are better for the state of a multi-time quantum field. In addition to allowing for a consistent account of particle-like interactions, such a formulation might allow one to capture something like particle creation and annihilation phenomena. See Lienert, Petrat, and Tumulka (2017) for a discussion.

19See Allori (2013) for a description and motivation of this approach.
in motion. Indeed, this is precisely how Allori, Goldstein, Tumulka, and Zanghi characterize how flashes in GRWf explain our experience (2012, 6). Here the primitive ontology is taken to be a suitable surrogate for specifying what is directly observable in the theory.

But we have seen that the situation is significantly more subtle than this might suggest. Given that we do not directly see the flashes that constitute everyday objects on GRWf one needs an account of how it is that the primitive ontology, being what it is (in this case flashes situated in spacetime), makes determinate our most immediately accessible records. That the ontology is primitive is supposed to help in providing this account, but insofar as the flashes that constitute the objects we take ourselves to see are themselves not directly observable and are mediated by either the 3N-dimensional wave function or something like a brute-fact law that codes for the same information, it is unclear why a primitive ontology of three-dimensional objects is automatically preferable to an empirical ontology more generally, perhaps one that like GRWr appeals to the wave function directly.

In both GRWr and GRWf one has something on which mental states might be taken to supervene, but in neither theory is the explanation particularly intuitive. Whether one prefers the account of experience provided by GRWr over the account provided by GRWf turns on whether one finds it more plausible that mental states roughly supervene on brain-record degrees of freedom of the wave function (as required by GRWr) or that they supervene on the positions of brain-record flashes (as required by GRWf). Primitive ontologists prefer GRWf because the flashes can be understood as objects with ordinary spacial locations. But one can easily imagine someone who prefers the dynamical and/or metaphysical economy of GRWr.

Ultimately, one wants an empirical ontology that provides records on which one’s experience might plausibly be taken to supervene. Unsurprisingly, there are competing intuitions regarding precisely what this should be and, hence, what the best formulation of quantum mechanics and the most plausible account of experience will look like. We will return to discuss competing empirical ontologies and their relative virtues again when we consider Bohmian mechanics and variants of that theory (chapters 11 and 12), but we will begin with the general issue of what it should even mean for a physical theory to be empirically adequate in the next chapter.

GRW also faces a less subtle empirical issue, one that we should get out of the way before moving on. If physical objects really do appear to warm up as predicted by the standard GRW dynamics, then that would be a remarkable empirical success for the theory. It would make the worries we have expressed along the way seem entirely beside the point. But we arguably already have good inductive reason to believe that the nonlinear behavior predicted by GRW never happens.

Whenever we have been technologically able to measure an interference effect that would determine whether a system evolves linearly or collapses, we have always found that it evolves linearly. This was true for the simplest physical systems, and it has proven true as we have learned how to measure interference effects involving increasingly more complex systems. We are in the possession then of a sort of inductive argument on the complexity of the physical systems we have observed that they always evolve linearly. This does not immediately rule out the GRW formulation. GRW themselves explicitly chose a collapse width and collapse rate so that their theory was not ruled out by what had been observed to that time. And, while the region for choosing these parameters has been significantly reduced by careful experimental work since then, there is still some room to make a choice. That the choice is clearly ad hoc is why I would bet against the novel predictions of the theory. It could be right, but I do not believe it.

There is also a significant philosophical issue at hand, one that concerns how one understands
physical theories more generally. When one goes from the standard collapse theory to GRW, one changes the dynamical laws of the theory, but, as we have seen, one must also change how one interprets the theory by giving up the standard eigenvalue-eigenstate link. This illustrates why it is virtually impossible to separate a physical theory from its interpretation.

The various GRW-type theories we have considered explain determinate measurement outcomes differently and their explanations are all radically different from the standard theory’s. The special dynamical role that position plays in GRW goes hand-in-hand with how states are interpreted in GRWr, GRWm, and GRWf. The mathematical specification of the physical theory and the physical interpretation of the mathematical objects work together to make predictions and provide explanations. Inasmuch as physical theories provide explanations and those explanations depend on how one understands the theory, one’s interpretational principles are an essential part of one’s physical theory.