Interpreting spontaneous collapse theories

Peter J. Lewis*

Department of Philosophy, University of Miami, Coral Gables, FL 33124-4670, USA

Abstract

Spontaneous collapse theories of quantum mechanics require an interpretation if their claim to solve the measurement problem is to be vindicated. The most straightforward interpretation rule, the fuzzy link, generates a violation of common sense known as the counting anomaly. Recently, a consensus has developed that the mass density link provides an appropriate interpretation of spontaneous collapse theories that avoids the counting anomaly. In this paper, I argue that the mass density link violates common sense in just as striking a way as the fuzzy link, and hence should not be regarded as a problem-free alternative to the fuzzy link. Hence advocates of spontaneous collapse theories must accept some violation of common sense, although this is not necessarily fatal to their project.

Keywords: Spontaneous collapse; Spontaneous localization; Dynamical reduction; GRW theory; Fuzzy link; Mass density link; Counting anomaly

1. Introduction

In 1986, Ghirardi, Rimini and Weber showed how to construct a spontaneous collapse theory of quantum mechanics as a solution to the measurement problem. But in 1990, Albert and Loewer argued that the theory is in need of interpretation if its claim to solve the measurement problem is to be vindicated. Since then, a sizeable literature has developed, proposing, attacking and defending various potential interpretations of spontaneous collapse theories. Recently, a consensus has emerged around one particular interpretation, known as the mass density link. My goal in this paper is to challenge this consensus.

My main argument is that the mass density link entails a violation of common sense that I call the *location anomaly*. That is, the mass density link violates the common-sense assumption that if a number of ordinary macroscopic objects are located somewhere, then all of them can typically be found there. Before I make this case, I first give a brief account of the origin of spontaneous collapse theories, and the problem that gives rise to the need for an interpretation (section 2). I then outline the debate over possible interpretations, and the arguments that have recently led to consensus around the mass density link (section 3). I then turn to a general discussion of what it takes to provide an adequate interpretation of a theory, and on that basis construct my main argument against the mass density link (section 4). In section 5, I consider a recent proposal for construing the mass density link that doesn't fall prey to my argument, and show how it succumbs to a closely related argument. Finally, in section 6 I assess the acceptability of spontaneous collapse theories in light of these arguments.

2. Problems

Let us start at the beginning. The *measurement problem* in quantum mechanics arises due to a conflict between the wave-like and particle-like behavior of matter. The fundamental dynamical law of quantum mechanics—the Schrödinger equation—is a wave equation. It describes the evolution of a function over time, and this function spreads out like a wave and

* *Email address:* plewis@miami.edu (P. Lewis).

exhibits interference effects like a wave. The natural temptation, then, is to treat this so-called wavefunction as literally descriptive of the fundamental stuff of the world—to conclude that ordinary physical objects are not configurations of discrete particles, as classical mechanics would have it, but instead are distributions of wave amplitude.

However, our observations of individual physical systems seem to be inconsistent with the hypothesis that the fundamental stuff of the world is wave-like. One might hope to recover the behavior we attributed to particles by supposing that a particle is really a *wave packet*—a wave that is localized within a small volume of space. However, according to the Schrödinger equation, such a wave will rapidly spread out over the whole of space. The difficulty here is that our observations of particles and of the objects made up of them reveal them to be well localized in space, not spread out. A secondary difficulty is that the wavefunction, unlike a classical wave, is a wave in configuration space—a space of particle configurations. Consequently, the wavefunction of an n-particle object is a function of $3n$ parameters rather than 3. So a more accurate way to express the measurement problem is to say that the wavefunction of an n-particle object is typically spread over the whole of $3n$-dimensional configuration space, whereas our observations of the object reveal it to be well localized within this space.[1]

An early and influential solution to the measurement problem was the suggestion that quantum mechanics is in fact governed by two incompatible dynamical laws (von Neumann, 1932). On this proposal, the Schrödinger equation governs the evolution of the state between measurements, whereas a collapse dynamics takes over during a measurement. The collapse dynamics entails that when the location of an object is measured, the wavefunction becomes instantaneously localized within the configuration space region in which the object is found. This solution allows us to retain the reduction of particles to localized wave packets, but it also makes the laws of physics dependent on whether or not a physical process counts as a measurement. Not only is this distinction hopelessly vague, it is also hard to imagine how it could possibly be relevant to fundamental law (Bell, 1987, p. 117). Clearly a better solution is needed.

Many such solutions have been proposed, and most involve adding something to the wavefunction representation of physical systems. For example, Bohm's (1952) hidden variable theory adds particle positions; hence particles are not, after all, reducible to wave packets, but are irreducible entities that are pushed around by the waves. The solution that is perhaps closest in spirit to Schrödinger's original wave dynamics, in that it seeks to retain the wavefunction as the complete representation of the state of a physical system, is the spontaneous collapse theory of Ghirardi, Rimini and Weber (1986).[2] It is this theory and its variants that form the subject matter of this paper.

The core of all spontaneous collapse theories is a collapse dynamics that is independent of measurement. Recall that in the wavefunction representation of a system, each classical "particle" corresponds to three configuration space coordinates. According to the original Ghirardi, Rimini and Weber theory, for each "particle" there is a small, constant chance per unit time that the wavefunction will undergo an instantaneous, discontinuous collapse process in which it becomes highly localized in the corresponding three coordinates. For small, isolated systems, the chance of a collapse occurring during any reasonable period of time is negligible,

---

[1] One might object that we don't observe objects to be well localized in configuration space at all; we observe them to be well localized in three-dimensional space. But it is arguable that we observe configurations of things directly, just as we observe the positions of things directly. I address this issue further in section 5.

[2] Not all proponents of spontaneous collapse theories would agree that they seek to retain the wavefunction as the complete representation of a physical system. I will take up this controversy in section 5.

and this is sufficient for the theory to generate the known wave-like behavior of such systems. But for large, well-correlated systems, like ordinary macroscopic objects, the chance of a collapse occurring for some constituent "particle" becomes overwhelmingly likely, even over very short periods of time. The correlations between the coordinates mean that a collapse in any three coordinates localizes the whole object in configuration space, and hence the theory can account for the fact that the objects we directly experience are always well localized. Later versions of this theory incorporate the collapse mechanism as a small, continuous correction to the Schrödinger equation, rather than postulating a separate, discontinuous process (Ghirardi, Pearle and Rimini, 1990).

While these spontaneous collapse theories eliminate the reference to measurement from physical law, it is not immediately clear whether they successfully solve the measurement problem. The difficulty, first pointed out by Albert and Loewer (1990), is that the collapse process does not fully localize the wavefunction of either a particle or a macroscopic object within any finite region of the corresponding subspace of configuration space. Although the vast majority of the post-collapse wave intensity is concentrated in a small region, the wavefunction retains low-intensity "tails" that spread over the whole of the subspace; this has become known as the *tails problem*. If part of the measurement problem consists in the fact that the wavefunction for an ordinary object is spread over the whole of its corresponding subspace, then the measurement problem remains even for spontaneous collapse theories.

Still, there seems to be a relevant difference between a wavefunction that is almost entirely localized within a marble-sized region of configuration space and one which is spread evenly (or even lumpily) over a much larger region. Intuitively, at least, the former seems like the kind of thing that could ground our experience of a marble, whereas the latter does not. Perhaps what the tails problem indicates is not that the spontaneous collapse approach to the measurement problem fails, but that it requires some interpretation. In particular, what is needed is a precise specification, in wavefunction terms, of what it takes for an object to have a particular location.[3] Several such interpretations have been developed, but unfortunately they raise difficulties of their own.

3. Interpretations

The traditional rule for specifying the properties of objects in wavefunction terms is the *eigenstate-eigenvalue link*, according to which an object has a particular property if and only if the wavefunction is an eigenstate of the corresponding operator. The eigenstates for an object being located in a particular configuration space region are those in which all of the wavefunction intensity is concentrated in that region. Since spontaneous collapse theories cannot deliver these states, the eigenstate-eigenvalue link must be replaced if spontaneous collapse theories are to solve the measurement problem.

The alternative interpretation rule advocated by Albert and Loewer (1996, p. 87) has been dubbed the *fuzzy link* (Clifton and Monton, 1999, p. 699). In essence, it is simply a less stringent version of the eigenstate-eigenvalue link; rather than requiring that *all* the wavefunction intensity is located in the relevant configuration space region, it is enough that *most* of it is so located. More precisely, an object is located in a given spatial region if and only if the integral of the wavefunction intensity over that region in the coordinates of each of the particles in the object, and over the whole of space in the coordinates of every other particle, is at least $1 - p$, where p is

[3] Indeed, this is Albert and Loewer's own position on the import of the tails problem (Albert and Loewer 1996).

a parameter of the rule.[4] For suitable values of *p*, the fuzzy link entails that the states produced by spontaneous collapse theories for ordinary macroscopic objects are such that those objects have the locations we ordinarily take them to have. Hence the fuzzy link provides a dissolution of the tails problem, and spontaneous collapse theories interpreted according to the fuzzy link seem to successfully solve the measurement problem.

However, the fuzzy link itself raises a new difficulty for spontaneous collapse theories. Consider a spatial region, such as a box, and a large number *n* of independent macroscopic objects, such as a set of marbles. Suppose that for each marble, the wavefunction distribution is such that the marble is in the box according to the fuzzy link. Call this wavefunction distribution *state S* for ease of future reference. Now we can also ask whether state S is such that the composite system made up of all *n* marbles is in the box according to the fuzzy link—expecting the answer "yes". But note that the region of configuration space over which the integral is taken gets smaller the more objects we include in the compound system; the integral is restricted to a limited region in the coordinates of the particles included in the system, but ranges over the whole of space in the coordinates of every other particle, so the more particles we include, the smaller the region becomes. As a consequence, for a sufficiently large value of *n*, the integral for the compound system will become less than $1 - p$ (Lewis, 1997). Hence the fuzzy link entails that each marble individually is in the box, but the compound system made up of all the marbles is not. This phenomenon has become known as the *counting anomaly*.

This consequence of the fuzzy link is certainly surprising, and has generated a good deal of literature. Some argue that the counting anomaly is not a problem for spontaneous collapse theories, either because it doesn't arise (Frigg, 2003), or because it disappears when one tries to observe it (Clifton and Monton, 1999). However, since these arguments have been addressed elsewhere (Lewis, 2003b and Lewis, 2003a respectively), I will not consider them here. The more common response to the counting anomaly is that it does present spontaneous collapse theories with a problem, but that the problem can be solved by adopting an interpretation other than the fuzzy link (Bassi and Ghirardi, 1999; Bassi and Ghirardi, 2001; Parker, 2003; Monton, 2004). The interpretation that is universally appealed to is the *mass density link* of Ghirardi, Grassi and Benatti (1995).

Whereas the fuzzy link extracts the location of an object directly from the wavefunction distribution in configuration space, the mass density link proceeds by first constructing a mass density distribution in ordinary three-dimensional space. The mass density function $\mathcal{M}$ is defined for a point **r** in three-dimensional space as the expectation value of mass density at that point. In other words, since the wavefunction is a distribution over particle configurations, and each particle configuration determines a mass density at point **r**, the value of $\mathcal{M}$ for point **r** can be calculated as the mean value of the mass density at point **r,** weighted according to the squared wavefunction amplitude. In an analogous way, a variance function $\mathcal{V}$ can be constructed, whose value is the variance of the mass density at point **r** over the possible particle configurations. Ghirardi, Grassi and Benatti (1995) call the mass density at a point "accessible" (or "objective") when the ratio $\mathcal{V}/\mathcal{M}^2$ is much less than 1. An ordinary object like a marble is associated with a contiguous three-dimensional region of relatively high accessible mass density.[5]

[4] Which particles should count as being constituents of the object is itself a fuzzy matter, but this kind of vagueness should be familiar from classical mechanics.

[5] Monton (2004) drops the requirement that the mass density be accessible, for what strike me as good reasons. However, this change does not protect the mass density link from the argument in the next section.

According to the mass density link, what it takes for there to be a marble in the box is that there is a region of high accessible mass density in the box, and what it takes for there to be *n* marbles in the box is that there are *n* regions of high accessible mass density in the box. Hence the counting anomaly cannot arise for the mass density link; the composite system made up of *n* marbles is in the box if and only if each marble individually is in the box, no matter what the value of *n*. So it looks like the mass density link solves all our problems; the counting anomaly is avoided, the tails problem goes away, and spontaneous collapse theories provide a solution to the measurement problem.[6]

4. Interpretation as translation

One might quite reasonably take this to be the end of the story; indeed, Bassi and Ghirardi claim that the mass density interpretation of spontaneous collapse theories is "now universally accepted" (1999, p. 723). But the apparent victory of the mass density link here is spurious. The problem with the fuzzy link is that it violates a deeply held intuition concerning the locations of composite systems, namely that they are located where their parts are located. The mass density link shows that it is possible to vindicate this intuition. But I shall argue that this intuition is saved at the cost of violating another intuition concerning composite systems, namely that they are typically found where they are located.

To make my case, I first need to give an account of what it takes to provide an interpretation of a theory. The most straightforward answer is that an interpretation provides a rule for translating between the language in which the theory is couched and the language in which we couch our everyday claims about macroscopic objects (Lewis, 2003c). In the case of spontaneous collapse theories, the theory is couched in terms of the distribution of the wavefunction over configuration space. The interpretation rules we have looked at—the eigenstate-eigenvalue link, the fuzzy link and the mass density link—all stipulate conditions that the wavefunction must satisfy in order for it to be appropriate to say, for example, "There is a marble in the box". This is the understanding of interpretation that I will presuppose in the present section.

An analogy with more familiar scientific territory may be helpful at this point. In a famous passage, Eddington claims that the scientific description of a table is in conflict with the common-sense description, since according to common sense the table is solid, whereas according to classical atomic physics it is mostly empty space (1928, p. ix–x). Furthermore, Eddington claims that the scientific description must take precedence, since it is the product of rigorous testing; hence the common-sense description of tables as solid must be rejected (1928, p. xii). But as Stebbing (1944) forcefully argues, the alleged conflict dissolves on closer inspection. The language of physics doesn't contain terms like "table" and "solid", so physics by itself doesn't have the linguistic means to contradict the common-sense claim that the table is solid. To make the case that there is a conflict, one first needs to provide an interpretation of the relevant physical theory—a rule for translating between claims about the motion of electrons and atomic nuclei on the one hand, and claims about tables and solidity on the other.

The translation rule that Eddington seems to have in mind is something like this: An object counts as solid if and only if its volume is mostly taken up by the particles that compose it. Interpreted in that way, atomic physics *does* contradict the common-sense claim that the table is solid. But note that this does not entail that you can poke your finger through the table, or that a

[6] At least, the measurement problem is solved for non-relativistic quantum mechanics. Whether a spontaneous collapse theory can be made consistent with relativity remains to be seen.

cup placed on it will fall through. Ordinarily, claims like this constitute part of what we *mean* when we deny that the table is solid. In other words, Eddington's proposed interpretation rule requires that we give up the everyday entailment relations between claims about solidity and claims about supporting cups, resisting fingers and the like.

This is a high price to pay for adopting a particular translation rule. But why pay the price? Interpretations come cheap; it is easy enough to stipulate a new one according to which tables are typically solid. A massive reconfiguration of the entailment relations in everyday language, on the other hand, is surely a path of last resort, if indeed it is possible at all. So while one *could* interpret atomic physics as saying that tables are not solid, it is not necessary to do so, and certainly not desirable. In general, then, a good interpretation rule is one that respects the entailment relations between the claims of everyday language.

Eddington's example does not provide a serious candidate for a conflict between science and common sense. But the interesting thing about spontaneous collapse theories is that they apparently provide a *genuine* example of an Eddington-style conflict. The tails problem provides the initial statement of the problem; since the state of an object like a marble is never precisely an eigenstate of location, spontaneous collapse theories entail that marbles never have locations. This is in flat contradiction with common sense. But of course, just as atomic physics by itself says nothing about the solidity of tables, so spontaneous collapse theories as such say nothing about the locations of marbles. Hence the statement of the tails problem above must tacitly assume an interpretation rule for translating between the language of the theory and everyday language; indeed, it assumes the traditional eigenstate-eigenvalue link. Spontaneous collapse theories, coupled with the eigenstate-eigenvalue link, entail that marbles are never located in finite-sized regions. Given the traditional interpretation of quantum states, then, there is a prima facie conflict between spontaneous collapse theories and common sense.

But the lesson of the Eddington example is that, when faced with such a conflict, we should look for other ways of interpreting the theory. Indeed, we do have other options for interpreting spontaneous collapse theories, namely the fuzzy link and the mass density link. Both of these rules are an improvement over the eigenstate-eigenvalue link, since spontaneous collapse theories interpreted according to either rule allow that marbles can have locations. But recall the constraint that any adequate interpretation rule must respect the entailment relations between the claims of everyday language. One such entailment relation is that when each of n marbles is individually located in a box, then all n marbles taken together are in the box. That is, the counting anomaly is a violation of the interrelations between concepts that constitute common sense. So because of the counting anomaly, the fuzzy link is not an adequate interpretation rule, and the conflict between spontaneous collapse theories and common sense remains.

Hence the only remaining prospect for recovering common sense is the mass density link. The mass density link is apparently our best interpretation rule so far; it allows that marbles can have locations, and it does not exhibit the counting anomaly. However, I shall argue that it too fails to respect the interrelations between the concepts of common sense. To make this case, though, I first need to outline some further details of the spontaneous collapse process.

Recall that for each particle there is a small, fixed probability per unit time that the wavefunction will suddenly become highly localized in the coordinates of that particle. The point around which the collapse process is centered is chosen at random with a probability distribution given by the wavefunction intensity; it is this feature that ensures that spontaneous collapse theories reproduce the statistical predictions of standard quantum mechanics. But as mentioned above, the wavefunction intensity in the coordinates of any particle has tails that extend over the whole of space. So even if the wavefunction is currently well-localized right here in the coordinates of some particle, it is possible for it to jump to a state in which it is well localized ten miles away in the coordinates of that particle. And since the particles in a macroscopic object are well correlated in configuration space, a collapse to a point ten miles away in the coordinates of one particle in the object simultaneously localizes the wavefunction around that point in the coordinates of all the other particles in the object.

This behavior is anomalous; both the fuzzy link and the mass density link entail that the object in question has instantaneously jumped to a point ten miles away. The reason that such predictions do not result in empirical disconfirmation of spontaneous collapse theories is that the theories themselves predict that such jumps will be incredibly rare. For any macroscopic object, the wavefunction has tails in the coordinates of the particles in the object that extend over the whole of space, but the collapse dynamics ensures that the intensity of the tails at a point even a millimeter from the object is so small that one should never expect to witness a jump to that point.

However small the probability of such a jump for an individual object, though, it can become significant for a sufficiently large collection of independent objects, such as the system of *n* marbles that generates the counting anomaly.[7] Recall that for state S, the counting anomaly arises because the integral of the wavefunction intensity inside the box region in the coordinates of all the particles in the system becomes small; conversely, the integral outside the box region becomes large. But this is precisely the condition under which the probability of a jump to a radically different state becomes significant. Hence the state that generates the counting anomaly for the fuzzy link is highly unstable.

The mass density link claims of this state that it is a state in which all *n* marbles are in the box, and hence does not generate the counting anomaly. But there are further entailments between everyday-language claims that we need to take into account, namely between claims about where objects are *located* and claims about where those objects can be *found*.[8] For example, any adequate interpretation rule should say that there is a marble in the box if and only if you will typically find it in the box when you look.[9] The mass density link satisfies this constraint; the wavefunction distributions which it counts as ones in which the marble is in the box are also ones in which you are almost certain to find it there.

But similarly, any adequate interpretation rule should say that there are *n* marbles in the box if and only if you will typically find *n* marbles in the box when you look. This is just another instance of the same entailment between everyday-language claims. However, the mass density link fails to respect this entailment; for state S, the mass density link says that all *n* marbles are in the box, yet there is almost no chance of finding them all there when you look. Let us call this violation of common sense the *location anomaly*.

---

[7] The reason that the objects have to be independent—that is, at most weakly interacting—is that otherwise the correlations between the positions of the objects suppress both the anomalous jumping behavior and the counting anomaly.

[8] Note that the entailment is between two *claims*, not between a claim and its empirical basis. That is, the entailment is no different in kind from the one that is violated by the counting anomaly.

[9] The caveat "typically" is necessary to prevent the constraint from being too strong; spontaneous collapse theories cannot absolutely guarantee that you will find the marble in the box, due to the small probability of an anomalous "jump" to a different location. But the situation here is no different from the case of solidity; atomic physics cannot absolutely guarantee that a cup placed on a solid table will be supported, although typically it will be.

The location anomaly means that the mass density link is in exactly the same position as the fuzzy link; both involve violations of the web of entailments between everyday-language claims, and hence both require a reconfiguration of those entailments. Admittedly, the location anomaly only occurs for large values of *n*, but this is equally the case for the counting anomaly. One could argue about which anomaly requires the more radical revision of common-sense entailments, but the point is that one does not want to get into the business of revising these entailments *at all* if one can help it. Certainly the location anomaly means that the mass density link cannot claim to be a problem-free solution to the difficulties facing spontaneous collapse theories.

There are at least two serious objections to the above argument that need to be addressed at this point. The first objection is that the behavior of state S according to the mass density link does not constitute a violation of everyday-language entailments, precisely because state S is unstable. For an unstable state, it is only to be expected that it will be found in a different state from the one it is in right now. So surely the right claim to make about state S is that all *n* marbles *were* in the box, momentarily, but due to the instability of the state they are *no longer* in the box when you look for them (Bassi and Ghirardi, 1999, p. 729; Monton, 2004, p. 413). Thus the apparent violation of common sense dissolves.

Before responding to this objection, it is interesting to note that the fact that there is a very low probability of finding all the marbles in the box has nothing to do with the speed of human observers; even if we were able to examine the contents of the box extremely rapidly, we would still almost certainly not find all *n* marbles there. Spontaneous collapse theories, by design, reproduce the statistical predictions of standard quantum mechanics; the probability of an outcome is proportional to the square of the wavefunction amplitude associated with that outcome. Since only a small proportion of the wavefunction amplitude for the *n*-marble system is in the box, then any reliable device for determining the locations of the marbles, no matter how fast it works, will almost certainly return the result that fewer than *n* of them are in the box (Clifton and Monton, 1999). In other words, state S doesn't even support the counterfactual claim that all the marbles would be found in the box if we could examine the contents fast enough.

But even though *no* observation method will reliably yield the result that all *n* marbles are in the box, one might still insist that they are all there *before* the observation. My response to this objection is that the claim that all the marbles are in the box in such a situation still violates the entailments of everyday language. Think again about the analogy with solidity. Suppose a theory, together with an interpretation rule, entails that water is solid at room temperature, but only for isolated moments of extremely short duration. Suppose that it is a further consequence of the interpreted theory that the solidity vanishes whenever an object is placed on the surface of the water. I think it would be quite reasonable to complain that this is an abuse of the everyday concept of solidity. Solidity, in the everyday sense, can't be manifested for a moment, because what it takes for an object to be solid, such as supporting coffee mugs, can't be manifested for a moment. One might object that a coffee mug *can* be supported for a moment, though this support wouldn't be directly observable, and hence that water can be solid in the everyday sense even for isolated moments. But in the case under consideration, this move is blocked by the stipulation that the solidity vanishes whenever an object is placed on the surface. So if someone claims that water is solid for isolated moments, and only when nothing is placed on it, then they are using the word "solid" in other than its everyday sense. Certainly their use of the word doesn't support the entailments we ordinarily make based on claims of solidity.

But this is exactly the situation we are in with regard to the marbles in the box. It is claimed that all *n* marbles are in the box, but only for a moment of extremely short duration, and only if nobody tries to determine that they are in the box. Again, the locution "in the box" is not being used in its everyday sense. In the everyday sense, the marbles cannot be in the box for an moment, because part of what it takes for them to be in the box—namely reliably being found in the box on observation—cannot be manifested in a moment. One might object that it *is* possible for the marbles to be in the box for a moment, even in the ordinary sense, in the same way that a speeding bullet can be in a box for a moment. But note that the speeding bullet at least supports the counterfactual claim that if we could examine the contents of the box fast enough, we would find it there. Under the current proposal, even this move is blocked, since no detection method will reliably find the marbles in the box. So the claim that all the marbles are in the box must be using the locution "in the box" in other than its everyday sense, one that does not support the entailments we typically make based on such assertions. Hence an interpretation rule, considered as a translation between the language of physics and everyday language *as it is*, simply cannot say that all the marbles are in the box in this case.

The second potential objection to my argument concerns the size of the *n*-marble system required for state S to become problematic. As Bassi and Ghirardi (1999) have pointed out, for state S to exhibit the counting anomaly, *n* must be very large indeed—so large that it would be impossible in practice to construct such a state, and even if we could, it would be impossible to directly observe the number of marbles in the box. The same goes for the location anomaly; the size of system required is way beyond what it is possible for us to construct, or to directly observe.

There are two possible ways this observation could be taken to undermine my argument. The first is that everyday language use does not entail anything about composite systems that are so large as to be beyond our possible experience; hence the fact that the mass density link says of state S that all *n* marbles are in the box, even though they would almost certainly not all be found in the box, does not show that the mass density link conflicts with everyday language. However, this amounts to admitting that "in the box" for state S does not have its ordinary meaning, and hence that it is not the case for state S that all *n* marbles are in the box *in the same sense* that a single marble is in the box. While it might be acceptable to insist that all *n* marbles are "in the box" in this new sense, the mass density link can no longer claim to solve the counting anomaly; each individual marble is in the box in a sense that all *n* are not.

The second way to construe this objection is that, since we will never run across a state like S, the fact that such a state violates the entailments of everyday language need not concern us. To an extent, I think this form of the objection has some merit. It is certainly true that, as a practical matter, we would not have to modify our actual language use if we were to adopt the mass density link. This makes the mass density link much less problematic than Eddington's interpretation of atomic physics, for example, which *would* have radical consequences for actual language. But even if the mass density link leaves our actual language unscathed, it still violates the *entailments* between our claims. The objection perhaps shows that the mass density link may be tenable *despite* the location anomaly, but it does not make the anomaly go away. Finally, it is important to note that this objection applies equally well as a defense of the fuzzy link against the counting anomaly; to the extent that it shows that the mass density link might perhaps be tenable, it shows that the fuzzy link might perhaps be tenable too.

My thesis, then, is that spontaneous collapse theories with the mass density link are no less problematic than with the fuzzy link. Interpreted using the fuzzy link, spontaneous collapse theories violate the common-sense entailment that if each of *n* ordinary, macroscopic objects is individually in some spatial region, then all *n* objects are in that region. Interpreted using the mass density link, spontaneous collapse theories violate the entailment that if *n* objects are located in a spatial region, then all *n* objects can typically be found in that region.[10] Hence however they are interpreted, spontaneous collapse theories conflict with common sense.

For the purposes of this section, I have been assuming that the role of interpretation rules like the mass density link is to provide a translation between the language of physics and everyday language. But this is not the only possible way in which interpretation rules have been viewed; Monton (2004) takes a much more ontologically serious approach to the mass density link, which apparently bypasses the arguments of this section. It is to Monton's approach that I now turn.

5. Interpretation as ontology

Thus far, I have assumed that spontaneous collapse theories are "wavefunction only" theories—theories according to which the wavefunction is the basic entity of quantum mechanics. But this is not a universal point of view. Monton provides two lines of argument against what he calls the *wavefunction ontology*, both based on the fact that the wavefunction is an object in a 3*N*-dimensional configuration space, where *N* is the number of particles in the universe. The first line of argument is that since there is nothing intrinsic to the configuration space that determines which configuration-space coordinate corresponds to which three-dimensional particle coordinate, there is no way that the wavefunction by itself can underpin the claims we make about three-dimensional objects (Monton, 2002). The second line of argument is that the wavefunction ontology is radically revisionary of common sense, since it entails that the claims we make about three-dimensional objects are literally false (Albert, 1996; Monton, 2004). Monton's solution to both these problems is to reject the wavefunction ontology, and to endorse instead an ontology of genuine three-dimensional objects.

The shift in ontology changes the terms of the debate over the interpretation of spontaneous collapse theories. The output of interpretation rules like the fuzzy link and the mass density link, rather than being a manner of speaking about the distribution of wavefunction intensity in configuration space, is now to be taken as literally descriptive of the distribution of physical objects in three-dimensional space. In fact, calling the two links *interpretation* rules looks inappropriate from this point of view. According to the wavefunction ontology, the wavefunction distribution in configuration space provides a description of physical reality, and the link translates this description into the everyday language of marbles, boxes and such. But according to Monton's three-dimensional ontology, the wavefunction distribution is a convenient yet unphysical mathematical formalism, and the job of the link is to extract from it the true physical description of three-dimensional objects. On this latter view, the fuzzy link and the mass density link are perhaps better described as integral parts of competing physical theories, rather

[10] This argument addresses not only the original appeal to the mass density link by Bassi and Ghirardi (1999; 2001), but also the more recent defense of the mass density link by Parker (2003). Parker argues that Lewis' original (1997) critique of the mass density link illegitimately presupposes something like the fuzzy link; otherwise, there would be no reason to deny that state S can be one in which the composite system is in the box. But the above argument shows that the mass density link is problematic in its own right, irrespective of the status of the fuzzy link.

than as competing interpretations of the same physical theory; indeed, Monton describes them this way (2004, p. 411).

Monton's three-dimensional ontology makes the counting anomaly look even more problematic for the fuzzy link that it does already. Under the wavefunction ontology, one can perhaps convince oneself that the counting anomaly is not so bad. If one treats the fuzzy link as a rule for extracting everyday language descriptions from physical states, then one can perhaps countenance there being a state that is truly described by "Each marble individually is in the box", but is not truly described by "All *n* marbles is in the box". After all, there is no inconsistency at the level of the wavefunction description of the system, and one should probably not expect complete consistency from everyday language. However, if one treats the fuzzy link as a means of extracting the true physical description of reality from an abstract mathematical representation, then one expects the same rigor from the resulting description as from any other part of physics. But the resulting description is plainly inconsistent; there are *n* three-dimensional physical objects present, and each of them is in the region of three-dimensional space defined by the box, but it is not the case that the composite system made up of all *n* of them is in this region (Monton 2004, p. 413). Where an inconsistency in everyday language can perhaps be tolerated, and inconsistency in one's basic physical theory surely cannot.

Furthermore, Monton's three-dimensional ontology makes the mass density link look less problematic than it did before. Recall the argument of the previous section; the mass density link fails to provide a translation into everyday language that respects the common-sense entailments between our claims. But according to Monton, the mass density link is not intended to provide a translation between the language of physics and everyday language, so the fact that it fails to do so can hardly be held against it. Rather, the mass density link translates between the configuration-space description of a system, expressed in terms of a wavefunction distribution, and the three-dimensional representation, expressed in terms of the distribution of regions of accessible mass density. Since regions of accessible mass density are not objects of everyday-language discourse, they are not subject to the constraints imposed by the interrelations of everyday-language concepts. Hence under Monton's version of the mass density link, we are free to claim that state S is a state in which there are *n* regions of accessible mass density in the box, because such a description doesn't entail that one should find *n* marbles in the box when one looks.

Hence the mass density link looks much more attractive than the fuzzy link under Monton's three-dimensional ontology. However, I see several reasons not to adopt Monton's approach. First, Monton relies on arguments that the wavefunction ontology is untenable, and I don't think these arguments are conclusive. Regarding Monton's first line of argument, a case can be made that there *is* an intrinsic correspondence linking particular configuration-space coordinates and particular three-dimensional particle coordinates; a configuration space is, after all, a space of particle configurations (Lewis, 2004). This case, if cogent, also reduces the force of Monton's second line of argument. If points in configuration space can, contra Monton, be associated with particle configurations in three-dimensional space, then claims about the locations of objects in three-dimensional space need not be regarded as false, but can be taken as partial descriptions of the actual configuration-space situation.

Furthermore, if the wavefunction ontology is tenable, then I think there are strong reasons to prefer it. The spontaneous collapse approach to solving the measurement problem in quantum mechanics is usually taken as a rival to the hidden variable approach (Bell, 1987, p. 201). A major advantage that hidden variable theories have over spontaneous collapse theories is that they have no need for an empirically risky and potentially ad hoc collapse dynamics. But on the other hand, a major advantage that spontaneous collapse theories apparently have over hidden variable theories is that they have no need for extra representational machinery; the wavefunction by itself can account for our observations of the world. But if one accepts Monton's three-dimensional ontology, this advantage is lost; spontaneous collapse theories postulate a collapse dynamics *and* extra representational machinery. If one is going to add "hidden variables" anyway, why not use them to avoid the extra complications of the collapse dynamics?

But my most serious objection to Monton's defense of the mass density link is that it raises the problem of interpretation all over again. Suppose we grant Monton the three-dimensional ontology. Then the mass density link is immune to the argument of the previous section precisely because it does not act as a link between the language of physics and everyday language, but only as a link between one level of theory and another. However, if we take the output of the mass density link to be theoretical claims about regions of mass density, then we need to be told how these claims are related to everyday-language claims about marbles and boxes. The obvious answer, and the one that is presupposed in all presentations of the mass density link, is that there is a marble in the box exactly when there is an appropriate-sized region of relatively high accessible mass density in the box. But putting this translation rule together with the link between the wavefunction and regions of mass density yields precisely the translation rule between wavefunction talk and everyday language criticized in the previous section. That is, the resulting translation rule says of state S that all the marbles are in the box, and hence entails the location anomaly. One could adopt a different interpretation of the regions of mass density, one that doesn't say of state S that all the marbles are in the box, but such a rule would entail the counting anomaly. Hence Monton's ontological defense of the mass density link only puts off the interpretive problem.

6. Conclusion

The counting anomaly seems at first glance like a simple technical problem that can be solved by a suitable choice of interpretation rule. But an acceptable solution turns out to be remarkably hard to come by. The mass density link, the favored way of avoiding the counting anomaly, fails to rescue common sense, since it violates the intuition that if a number of ordinary macroscopic objects are located somewhere, then all of them can typically be found there. This argument is quite general; any interpretation rule for spontaneous collapse theories that avoids the counting anomaly thereby runs into the location anomaly. And while it is possible to disguise the anomaly by taking an ontological approach to mass density, the anomaly reappears as soon as one tries to connect the new ontology to everyday language.

If the foregoing is correct, then neither the fuzzy link nor the mass density link can rescue spontaneous collapse theories from a conflict with common sense, and no alternative rule one might construct could do better. If one chooses to interpret spontaneous collapse theories according to the fuzzy link, then there are states in which each of n marbles individually is in the box, and yet it is not the case that the composite n-marble system is in the box. In other words, the common-sense entailment that says that a whole is located where its parts are located holds only as an approximation for relatively simple systems. Admittedly, this includes all the systems we are ever likely to run into, since the systems that violate the entailment in a noticeable way are impractically large. Still, the position that this entailment is an approximation is counter-intuitive enough.

Alternatively, if one chooses to interpret spontaneous collapse theories according to the mass density link, there are states in which *n* marbles are in the box, and yet it is not the case that one will typically find *n* marbles in the box. In this case, the common-sense entailment that says that ordinary macroscopic objects can typically be found where they are located holds only as an approximation for relatively simple systems. Again, it will hold for any of the systems we are likely to run into, but the position that this entailment is an approximation is counter-intuitive enough.

None of this constitutes a knockdown argument against spontaneous collapse theories. The anomalies do not constitute an empirical refutation of spontaneous collapse theories; although the behavior of state S cannot be reconciled with common sense, it does not conflict with our experience, since we have never encountered a state like S, and nor are we likely to. So it is possible to retain spontaneous collapse theories if we are willing to modify common sense accordingly. The advocate of spontaneous collapse theories faces a choice between the fuzzy link and the counting anomaly on one hand, and the mass density link and the location anomaly on the other. I do not think there is a correct answer here; it is a matter of choosing a convention for talking about things. It seems to me that simplicity favors the fuzzy link, since it works within the standard wavefunction formalism, rather than supplementing it with a mass density distribution. However, perhaps I could be convinced that the location anomaly does less damage to common sense than the counting anomaly, and hence that the mass density link is to be preferred despite its somewhat baroque structure.

Acknowledgements

I would like to thank Bradley Monton, Daniel Parker and two anonymous referees for helpful comments on earlier versions of this paper.

References

Albert, D. Z. (1996). Elementary Quantum Metaphysics. In J. Cushing, A. Fine and S. Goldstein (eds.), *Bohmian mechanics and quantum theory: An appraisal* (pp. 277–284). Dordrecht: Kluwer.

Albert, D. Z. and B. Loewer (1990). Wanted dead or alive: Two attempts to solve Schrödinger's paradox. In A. Fine, M. Forbes and L. Wessels (eds.), *Proceedings of the Philosophy of Science Association, Volume 1* (pp. 277–285). Michigan: Philosophy of Science Association.

Albert, D. Z. and B. Loewer (1996). Tails of Schrödinger's cat. In R. Clifton (ed.), *Perspectives on quantum reality* (pp. 81–91). Dordrecht: Kluwer.

Bassi, A. and G. C. Ghirardi (1999). More about dynamical reduction and the enumeration principle. *British Journal for the Philosophy of Science*, 50, 719–734.

Bassi, A. and G. C. Ghirardi (2001). Counting marbles: Reply to Clifton and Monton. *British Journal for the Philosophy of Science*, 52, 125–130.

Bell, J. S. (1987). *Speakable and unspeakable in quantum mechanics*. Cambridge: Cambridge University Press.

Bohm, D. (1952). A suggested interpretation of the quantum theory in terms of 'hidden' variables. *Physical Review*, 85, 166–193.

Clifton, R. and B. Monton (1999). Losing your marbles in wavefunction collapse theories. *British Journal for the Philosophy of Science*, 50, 697–717.

Eddington, A. S. (1928). *The nature of the physical world*. New York: Macmillan.

Frigg, R. (2003). On the property structure of realist collapse interpretations of quantum mechanics and the so-called 'counting anomaly'. *International Studies in the Philosophy of Science*, 17, 43–57.

Ghirardi, G. C., R. Grassi and F. Benatti (1995). Describing the macroscopic world: Closing the circle within the dynamical reduction program. *Foundations of Physics*, 25, 5–38.

Ghirardi, G. C., P. Pearle and A. Rimini (1990). "Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles. *Physical Review* A, 42, 78–89.

Ghirardi, G. C., A. Rimini and T. Weber (1986). Unified dynamics for microscopic and macroscopic systems. *Physical Review* D, 34, 470–91.

Lewis, P. J. (1997). Quantum mechanics, orthogonality and counting. *British Journal for the Philosophy of Science*, 48, 313–328.

Lewis, P. J. (2003a). Counting marbles: reply to critics. *British Journal for the Philosophy of Science*, 54, 165–170.

Lewis, P. J. (2003b). Four strategies for dealing with the counting anomaly in spontaneous collapse theories of quantum mechanics. *International Studies in the Philosophy of Science*, 17, 137–142.

Lewis, P. J. (2003c). Quantum mechanics and ordinary language: The fuzzy link. *Philosophy of Science*, 70, 1437-1446.

Lewis, P. J. (2004). Life in configuration space. *British Journal for the Philosophy of Science*, forthcoming.

Monton, B. (2002). Wavefunction ontology. *Synthese*, 130, 265–277.

Monton, B. (2004). The problem of ontology for spontaneous collapse theories. *Studies in History and Philosophy of Modern Physics*, 35, 407–421.

Parker, D. (2003). Finding your marbles in wavefunction collapse theories. *Studies in History and Philosophy of Modern Physics*, 34, 607–620.

Stebbing, L. S. (1944). *Philosophy and the physicists*. New York: Penguin.

Von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer.