Completeness, supervenience and ontology

2007 J. Phys. A: Math. Theor. 40 3151

(http://iopscience.iop.org/1751-8121/40/12/S16)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) 3151–3171 doi:10.1088/1751-8113/40/12/S16

Completeness, supervenience and ontology

Tim W E Maudlin

Department of Philosophy, Rutgers University, 26 Nichol Avenue, New Brunswick,
NJ 08901-1411, USA

Received 11 October 2006
Published 7 March 2007
Online at stacks.iop.org/JPhysA/40/3151

Abstract

In 1935, Einstein, Podolsky and Rosen raised the issue of the completeness of the quantum description of a physical system. What they had in mind is whether or not the quantum description is informationally complete, in that all physical features of a system can be recovered from it. In a collapse theory such as the theory of Ghirardi, Rimini and Weber, the quantum wavefunction is informationally complete, and this has often been taken to suggest that according to that theory the wavefunction is all there is. If we distinguish the ontological completeness of a description from its informational completeness, we can see that the best interpretations of the GRW theory must postulate more physical ontology than just the wavefunction.

PACS numbers: 03.65.Ta, 03.65.−q, 41.20.−q, 01.70+w

1. Informational completeness, supervenience and reduction

When Einstein, Podolsky and Rosen posed the fateful question 'Can quantum-mechanical description of reality be considered complete?' [1], they introduced an important new piece of terminology into the foundations of physics. The concept of the *completeness* of a description was exactly the concept needed to raise questions about the status of quantum theory—and the apparent non-locality inherent in the standard formulation of quantum theory—in a particularly sharp way. But the tools adequate for one problem can sometimes be misleading in other contexts, and I fear that the notion of completeness required for the EPR argument needs to be sharpened again if we are to make progress in understanding the physical accounts of the world provided by different versions of the quantum theory. In particular, recent discussions about various versions of the spontaneous localization theory come into better focus once we attend to different ways that the notion of completeness can be understood.

For the purposes of the argument of Einstein, Podolsky and Rosen, the key notion is that of the *informational completeness* of a physical description of a situation. We can say that a description is informationally complete if every physical fact about the situation can be recovered from the description. It follows from this definition that if a theory provides descriptions that are informationally complete and two physical situations are given the same description by that theory, then the situations are physically identical in all respects.

1751-8113/07/123151+21$30.00

A description can be informationally complete in this sense even though it is not, intuitively, a direct description of what is physically real. For example, consider classical electromagnetic theory. The usual (albeit naïve) understanding of Maxwell's equations is that they describe the dynamics of a fundamental physical entity, the electromagnetic field. In this naïve understanding, it is the field that is physically real. But it was soon understood as a purely mathematical consequence of Maxwell's equations that any field that obeys the equations could also be described by the use of the scalar and vector potentials. The description in terms of these potentials is informationally complete, in that two situations described by the same scalar and vector potentials would be identical in all physical respects (at least with respect to electromagnetics). Nonetheless, the scalar and vector potentials were *not* taken to be 'direct' descriptions of the physical reality: *different* scalar and vector potentials could be used to describe exactly the same situation, reflecting arbitrary choices of gauge in framing the mathematical description. Gauge degrees of freedom in the mathematics do not correspond to physical degrees of freedom in the world. So one would be making a serious error if one tried to read off the physical ontology directly from classical electromagnetic theory presented in terms of the vector and scalar potentials. One *could*, of course, interpret the theory as postulating the physical reality of the potentials rather than the fields, but the price would be acceptance of physically different situations (corresponding to what we think of as choice of gauge) that would display exactly the same observable behaviour. Since there were no grounds to consider such a possibility, the most reasonable understanding of classical electromagnetic theory was exactly that given in textbooks: what is real is the fields, and the potentials are merely mathematical conveniences whose ultimate physical credentials are secured because one can derive the field values from them.[1]

If one keeps in mind the example of the classical electromagnetic potentials, which were thought to be informationally complete but not physically real, the distinctions I want to focus on should become clear.

There are many other examples of classical descriptions that were considered informationally complete but were nonetheless not thought to directly represent the entire physical ontology. Consider the electromagnetic field and the charge density in classical theory. Given only a description of the field, one could recover full information about the charge density by simply taking the divergence, so the description of the field would, in this sense, contain full information about the charge density. And the situation here is not symmetrical: full information about the distribution of charge would not provide full information about the field, as the existence of multiple distinct vacuum solutions demonstrates. In the argot of philosophers, the charge distribution *supervenes* on the field values, since there cannot be a difference in charge distribution without a difference in the field, but the field does not supervene on the charge distribution. Even more exactly, the charge distribution *nomically supervenes* on the field values since one uses a physical law—Maxwell's equations—to derive the former from the latter.[2]

---

[1] It is often said that the Aharonov–Bohm effect forces us to recognize the vector potential as physically real rather than as a mere mathematical convenience. The situation is more complex than this: the right thing to say is that we now recognize that the appropriate mathematical representation of the electromagnetic field is neither the classical field nor the classical potentials, but rather the connection on a fiber bundle. The Aharonov–Bohm effect occurs because the global structure of such a bundle can change even though the curvature of the bundle (the 'field') changes locally only in a small region that is seemingly irrelevant to the experiment (the interior of the solenoid). The classical vector potential—which is a vector field *on space*—only takes a particular spatial form in virtue of an arbitrary convention ('choice of gauge') relating these spatial directions to structures on the fiber bundle. Given different such conventions, the spatial representation can differ.

[2] The fact that one appeals to a *physical law* to specify this form of supervenience shows up when one specifies the sense of 'can't' in 'there can't be a difference in this without a difference in that'. In nomic supervenience, one means

But even though everyone agrees that in classical theory the description of the field is informationally complete, and the charge distribution supervenes on the field values, it is still also the case that in the usual understanding of the classical theory *there is more to the physical world than just the field: there is also the charge distribution*. The supervenience is suggestive, and may motivate a project of trying to understand the charge distribution as somehow *nothing but* the field (think of attempts to understand point charges as *nothing but* singularities in the electromagnetic field), but the supervenience does not, by itself, show that such a project can succeed, or should be undertaken. Indeed, there are clear cases of nomic supervenience in which any such attempt to reduce the ontology of the theory would be crazy. In any deterministic theory, for example, the global physical state of the world at any moment nomically supervenes on the global state at any other moment: there cannot be a difference in one without a difference in the other. In this sense, given the dynamical laws, the state at any particular moment is informationally complete (as Laplace pointed out). But no one would suggest because of this that we think of the state at one moment as *all that exists*: indeed, it is the various different states at different times that the dynamical laws link to one another.

To sum up, informational completeness implies a form of supervenience, and supervenience is often taken to be a indication that there ought to be some form of *ontological reduction*: if there cannot be a difference in one thing without a difference in another, and if all the facts about one thing can be derived from facts about the other, why not suspect that the first thing is *nothing but* an aspect of the other? Lots of examples suggest this kind of reasoning: there cannot be a difference in the facts about the tables in the room without a difference in the distribution of atoms in the room, and all the facts about the tables could (in some decent sense of 'could') be derived from a complete description of the facts about the atoms, so ought we not to conclude that tables are *nothing but* complex collections of atoms? A physical description that includes all the facts about the atoms has not *left out* the tables; it has given a complete physical specification of the tables. Tables are nothing *over and above* the atoms. I hope that it is now clear that these claims—which are claims about the *ontological status* of tables, about their physical nature—are claims that go beyond the simple observation that 'table talk' supervenes on 'atom talk'. For the supervenience can hold in cases where no one thinks that the one thing ontologically reduces to the other.

Sometimes philosophers advert to Ockham's Razor to argue from supervenience to ontological reduction: after all, *entia non sunt multiplicanda praeter necessetatem*, and if one is in possession of an informationally complete description, then it cannot be strictly *necessary* to postulate anything else. Claims about the extra ontology could somehow be translated into claims about the informationally complete state. But for all the surface plausibility of this line of thought, the examples show it to be empty. If the laws of physics are deterministic, that in no way suggests that claims about the late stages of the universe are just fancy ways of making complicated claims about its initial state. Nor need claims about the charge distribution be understood as nothing but claims about the divergence of the electric field.

So we have three sorts of examples before us. In the case of atoms and tables, ontological reduction is clearly in order: tables are nothing but structured collections of atoms. In the deterministic universe, reduction is clearly not in order: the state at one time is something different from the state at another time. And the charge distribution/divergence of the field gives an intermediate case. The attempt to somehow reduce charges or charged particles to nothing but states of the field does not seem crazy, but neither does it seem inevitable. It is a reasonable sort of physical research program, to be judged, in the end, on the advantages and

---

that there can't be a difference in one without a difference in the other *supposing the physical laws to remain fixed*. Different forms of supervenience result from different readings of 'can't'.

disadvantages that come with the reduction. And there is probably little of a general nature that can be said about what those advantages and disadvantages might be.

Having specified what it is for a description to be informationally complete, it will be useful to introduce the somewhat vaguer notion of an *ontologically complete* description. An ontologically complete description of a physical situation should provide—in a relatively transparent way—an exact representation of all of the physical entities and states that exist. If the charge distribution is the distribution of some matter—not a field—then an ontologically complete description should directly specify both the field and the matter. In classical electromagnetic theory, an ontologically complete description need not mention the vector potential, since in that theory the potential is not physically basic. An ontologically complete description should say just what there is and no more. Although this is a not a perfectly sharp characterization, and there could be reasonable disputes in particular cases about exactly how to apply it, it should be clear enough for the ensuing discussion.

2. Completeness in the EPR argument and the measurement problem

Foundational discussions of quantum theory often pose the question ‘Is the wavefunction of a system complete?’. But in view of the distinctions just made, we should be wary about exactly what such a question portends. One issue is whether a specification of the wavefunction is informationally complete, whether it pins down, one way or another, all of the physical facts about an individual system. A quite different question is whether the description of the wavefunction is ontologically complete, in which case the theory would hold that the wavefunction is all there is. It is this latter question that will most directly concern us in our investigation of the theory of Ghirardi, Rimini and Weber and variants. But it was the former question, and the former question alone, that concerned Einstein, Podolosky and Rosen.

Let us see how the informational completeness of the quantum state of a system (as ascribed by the usual Cophenhagen rules) implies a radical non-locality in nature. The argument—which I believe to be the heart of EPR’s concerns—does not concern the Heisenberg uncertainty relations or anything like them. Rather, it runs as follows.

Create, by the usual means, a pair of electrons in the singlet state. Supposing that the wavefunction is informationally complete, every physical fact about the electrons—both singly and jointly—is implied by the wavefunction. But the wavefunction does not ascribe any particular spin in any direction to either particle. So *if the wavefunction is informationally complete* neither particle *has* a spin in any direction. Such spins are not among the physical features of the system.

Now measure the spin of one particle, and adjust the wavefunction according to the usual Copenhagen rules on the basis of the outcome. The wavefunction ascribed to the unmeasured particle changes, and in virtue of that change the unmeasured particle now *does* have a particular spin in a particular direction. So the physical state of the unmeasured particle has changed due to the measurement made on the twin. This is so no matter how far apart the twins are, whether there are intervening barriers, etc. The change in the state of the unmeasured particle in virtue of the measurement made on the twin is exactly the sort of ‘spooky action-at-distance’ that Einstein abhorred.

Of course, one could respond to this argument by insisting that *no* physical change occurred in the distant twin. If after the measurement it has a spin in a particular direction, then it already had that spin even before the measurement was made, and all the distant measurement does *provide information about a pre-existing state of affairs* (for surely we have *gained information*, we can make *better predictions* about the distant particle after local measurement than before). This is exactly the way Einstein responded to the argument.

But the response requires admitting that the original singlet state was *not* informationally complete: the particle had a particular spin in a particular direction even though one could not read that fact off from the wavefunction. *It is only if the description one has of a system is not informationally complete that one can interpret a seemingly non-local change in that description as* '*merely getting new information about the system without physically changing the system*'. For if the initial description is informationally complete, then one cannot merely find out some already existing fact: all the facts about the system are already reflected in the description.

The EPR argument forces a choice on interpreters of quantum theory: either the wavefunction is not informationally complete, or there is spooky action-at-a-distance. It was no part of the EPR argument to suppose that the wavefunction is ontologically complete. That issue never even needs to be raised. For if the wavefunction is not informationally complete, then it is clearly a further task of physics to postulate a sort of description that is informationally complete, in much the same way that the gross thermodynamic description of a classical system is not informationally complete while the exact phase point is. Notice also that it is no part of EPR's argument to suppose that one can experimentally determine the complete physical description, any more than the postulation of exact positions and velocities of the particles in a box of gas required specifying an experimental method for determining those exact positions and velocities.

Once one sees that EPR was arguing for the informational incompleteness of the quantum description, it becomes obvious that certain approaches to interpreting quantum theory are not responses to EPR's argument but concessions to it. For example, the 'statistical interpretation' is supposed to hold that the wavefunction is only a description of an *ensemble* of systems, not of any particular single system. From this point of view, an 'equal superposition of live cat and dead cat' is not a puzzle: it is just a description of a group of cats, half of which are alive and half dead. But this evidently concedes the point that the wavefunction is not a complete description of any particular cat in the ensemble: for any given cat, it is either a physical fact that it is alive or it is dead. The wavefunction does not contain this information.

Similarly, those quantum information theorists who interpret the wavefunction as merely a representation of some individual's information about a system tacitly suppose that the wavefunction is (in the sense defined) informationally incomplete. For either that individual knows all the physical features of the system or does not. If he knows all of its physical features, then anyone else who similarly knows all of its features will ascribe it the same wavefunction, so associating the wavefunction with the individual is doing no work. It is exactly because different people can have different *partial* accounts of the physical state of a system that it could be appropriate for them to use different representations of it. Maybe there is even a good physical reason that no one can *have* or *verify* a complete physical description of a system; nonetheless, the descriptions they do have—even if they are the best they *can* have—are incomplete.

The argument of the EPR paper, then, concerns the informational completeness of the quantum description, not its ontological completeness. But when the notion of completeness arises in other contexts, the question of its exact meaning must be addressed anew.

In statements of the measurement problem, the notion of informational completeness is similarly useful. In Bell's famous phase, the linearity of Schrödinger's equation implies that 'either the wavefunction, as given by the Schrödinger equation, is not everything, or it is not right' ([2], p 201). This leaves solutions to the measurement problem with two (not mutually exclusive) choices: either modify the Schrödinger evolution or assert that the wavefunction is not everything. But in what sense 'not everything'? Is the question whether the wavefunction is *ontologically* complete (all there is to the physical world is wavefunction) or *informationally* complete (one can recover all physical facts from the wavefunction)? In the context of Bell's argument, the issue is clearly only informational completeness: the argument is that it is either a plain physical fact about, e.g., an individual cat that it ends up alive (and not dead) or that it ends up dead (and not alive). Since the wavefunction as given by the Schrödinger equation does not favour one result over the other, it fails to reflect a plain physical fact about the cat. The wavefunction, as given by the Schrödinger equation, is *informationally* incomplete.

The wavefunction as given by the GRW dynamics, by contrast, appears to be informationally complete. Depending on exactly how the stochastic hits come out, the final wavefunction will either be distinctly cat alivish or cat deadish. As Bell puts it, 'The wavefunction commits itself very quickly to one pointer reading or the other' ([2], p 204). The *obvious* worry about considering the wavefunction informationally complete has disappeared, although there still may be lingering subtle difficulties about how *exactly* to read off the complete physical situation from the wavefunction[3]. From this point of view, the new dynamics provides the resources to solve the measurement problem by postulating a physical process that corresponds to von Neumann's collapse.

But what comes along almost inevitably with this observation is the assumption that in the GRW theory the wavefunction ought to be thought of as not only informationally but also *ontologically* complete, that *all there is* in this theory is the wavefunction[4]. Again, Bell provides an example. Immediately after the sentence quoted above, he begins a new section writing 'There is nothing in this theory but the wavefunction' ([2], p 204). One natural reading of this claim, I think, is not merely that the wavefunction is informationally complete but that it is ontologically complete as well. That may not have been Bell's intent—indeed the following discussion of the exact ontology of the theory is difficult—but it is the natural reading. And for quite a long time, most philosophers have assumed that the best understanding of the GRW theory is one in which the wavefunction directly represents, at the deepest level, all there is. Often, these philosophers take wavefunction monism to be a signal advantage for GRW over any dualistic theory, by appeal to Ockham's Razor. GianCarlo Ghiradi himself, though, has not committed himself to this view, and I think he is correct not to. So we are finally in a position to ask the central question: even granting that the wavefunction in the GRW theory is informationally complete, what grounds are there for accepting or rejecting the additional claim that it is ontologically complete as well?

3. Local beables and a space-time ontology

In 'The theory of local beables', Bell introduced a very nice piece of terminology, that of the local beable. A 'beable' is a (speculative) piece of ontology: something that a theory postulates as being physically real. Part of Bell's explication makes use of the same example as we used above:

> The word 'beable' will also be used here to carry another distinction, that familiar already in classical theory between 'physical' and 'non-physical' quantities. In Maxwell's electromagnetic theory, for example, the fields **E** and **H** are 'physical' (beables, we will say) but the potentials **A** and φ are 'non-physical'. Because of gauge invariance the same physical situation can be described by very different

---

[3] These difficulties arise if one tries to hold on to the rule that physical facts exactly correspond to facts about which operators the wavefunction is an eigenstate of, and with which eigenvalues. The so-called 'tails' problem undercuts the utility of this rule.

[4] I use the term 'wavefunction' ambiguously to refer to both a certain piece of physical reality and the mathematical object used to represent it. I hope the meaning in each use is clear.

> potentials. It does not matter that in Coulomb gauge the scalar potential propagates with infinite velocity. It is not really supposed to *be* there. It is just a mathematical convenience ([2], p 52).

What Bell calls the beables of a theory, then, are just what are also called the physical ontology of the theory: what the theory postulates to exist. What, then, are the *local* beables?

Bell again:

> We will be particularly concerned with the *local* beables, those which (unlike for example the total energy) can be assigned to some bounded space-time region. For example, in Maxwell's theory the beables local to a given region are just the fields **E** and **H,** in that region, and all the functionals thereof. It is in terms of local beables that we can hope to formulate some notion of local causality. Of course we may be obliged to develop theories in which there *are* no strictly local beables. That possibility will not be considered here (p 53).

Local beables do not merely exist: they exist *somewhere*. If a theory has local beables, then the distribution of those beables throughout all of spacetime deserves the name David Lewis gives it: a 'mosaic'. That is, one specifies the global distribution of these beables throughout spacetime by simply specifying the beables in each open spacetime region. Even more: one can specify the global distribution of the local beables by specifying the beables in a set of *arbitrarily small* regions of spacetime, so long as the set covers the spacetime.

In a famous letter to Born, Einstein noted that the progress of physics had tended towards the most radical possible version of local physics, in which not only are all fundamental physical quantities local, but the fundamental laws are as well:

> If one asks what, irrespective of quantum mechanics, is characteristic of the world of ideas of physics, one is first of all struck by the following: the concepts of physics relate to a real outside world, that is, ideas are established relating to things such as bodies, fields, etc., which claim 'real existence' that is independent of the perceiving subject—ideas which, on the other hand, have been brought into as secure a relationship as possible with the sense data. It is further characteristic of these physical objects that they are thought of as arranged in a space-time continuum. An essential aspect of this arrangement of things in physics is that they lay claim, at a certain time, to an existence independent of one another, provided these objects 'are situated in different parts of space'. Unless one makes this kind of assumption about the independence of the existence (the 'being-thus') of objects which are far apart from one another in space—which stems in the first place from everyday thinking— physical thinking in the familiar sense would not be possible. It is also hard to see any way of formulating and testing the laws of physics unless one makes a clear distinction of this kind. This principle has been carried to extremes in the field theory by localizing the elementary objects on which it is based and which exist independently of each other, as well as the elementary laws which have been postulated for it, in the infinitely small (four-dimensional) elements of space ([3], p 170).

Einstein notes that in classical field theory *all* of the beables are local, and local in the strongest sense: the entire physical situation is nothing but the sum of the physical situations in the infinitely small regions of space-time. One might imagine loosening this: perhaps one could not 'localize' a beable in regions smaller than, say, Planck scale. Einstein further notes that this extreme localization in field theory applied not just to the ontology but to the laws: one could determine whether or not the laws of electromagnetics hold throughout spacetime by checking each infinitely small open region and seeing whether the laws hold there. Such laws could not be formulated unless the entire physical ontology were, in this most extreme sense, local.

A rather fine philosophical distinction can be made here. According to Bell's definition, classical electromagnetic theory contains non-local beables: the total energy of the universe is an example. Nonetheless, we want to call the ontology of this theory completely local: in an obvious sense, there is nothing in this theory but completely local stuff. The point is that one could determine the total energy of the universe if one knew only the energies in arbitrarily small local regions that cover spacetime provided one also had two more pieces of information: how the small open regions overlap, and *that* they collectively cover the spacetime. The total energy of the universe satisfies this criterion, and we might call it, somewhat oxymoronically, a *global local* quantity[5].

There are theories with local ontologies but non-local laws. The Newtonian gravitational theory of point particles provides an example. Here the ontology is local (global mass distribution is determined by the mass distribution in the small regions), but one could not determine whether the dynamical laws were being satisfied by looking at the small regions individually. One might determine, for example, that a particle in a small region is suddenly accelerated, but the cause of the acceleration need not be evident in the region: it could be a change in a distant gravitating body. Complete locality of ontology need not imply locality of laws.

Finally, we should note that in a theory *some* but not *all* of the beables could be local. Bohm's theory provides an obvious example: the particle trajectories are local beables but the wavefunction, which is equally part of the ontology, is not. Bell clearly means to consider theories like this. What Bell leaves out of account is only theories that have no local beables at all. And he curiously provides no direct grounds for this omission, beside the plausible suggestion that a theory lacking any local beables could have no claim to be a theory with only local causation. So it might be worthwhile to consider the challenges that would confront a theory that posits no local beables at all.

Back in the bad old days of logical positivism, it was often asserted that physical theories only acquire meaning or content insofar as they have implications about *experience*. Taken at face value, this dictum implies that, e.g., Newtonian gravitational theory has no content. For although given appropriate initial conditions one can derive many seemingly physical predictions from the theory—about how fast a dropped object on the Earth will fall, or how the moon will orbit the Earth, or when eclipses will occur—one would look in vain for any derivable claims about experience. At the most mundane level, although the theory may make a prediction about when an eclipse will occur, it will not predict whether anyone will be *looking* at it. And at a much deeper level, given the ontology of the theory it could not

---

[5] Roderich Tumulka (p.c.) wonders why it is not simpler just to say that the total energy is the integral of the energy density, which is local, and, in general, why I write in terms of facts about the physical state of small open regions rather than facts about the physical state *at each point* in space-time. For many purposes, talk of the state at a point will seem simpler, but there are cases in which the relevant state of affairs properly belongs to an open region rather than a point. For example, as suggested in Einstein's letter to Born, local laws (represented by differential equations) can only be said to hold or fail to hold in a region, since they describe spatial and temporal variation. Further, physical states of affairs at individual points lead to certain worrisome puzzles. Consider, for example, two functions representing energy densities that differ only at a single point (at least one will evidently be discontinuous). Should those be thought to represent *different possible physical states*, even though they agree on the energy content of every measurable region? One might well want to identify the physical content of these states, and thereby reject, at a fundamental level, the notion of an energy density being a physical state of affairs at a point: the density 'at a point' can instead be *defined* as an ideal element through a limiting process, taking average energy densities of ever smaller regions.

have implications about experience without solving the mind-body problem, without, that is, having principles that connect the motion of matter (in brains) with conscious experience. Since Newtonian theory contains no such principles, no predictions about experience will be forthcoming.

Of course, in the bad old days it was sometimes thought that *all* meaningful claims were just oblique ways of making claims about experience: in a world with no experience, there could be no atoms or stars or eclipses. And since that claim seems on its face absurd, the even more desperate gambit was tried of making all meaningful claims to be about *merely possible* experience, as if counterfactuals about what conscious beings *would have* experienced in the first moments of the Big Bang provided a solid foundation for one's understanding of the claims of cosmology. But since this project was a dismal failure, there is no need to rehearse it in detail.

There was a reasonable concern behind all this foolery. In order to be of interest, physical theories have to make contact with some sort of evidence, some grounds for taking them seriously or dismissing them. And the acquisition of evidence by humans clearly does involve experience at some point. So it is not surprising that one might focus one how physical claims relate to experience in an attempt to get a handle on the problem of evidence. But for all that, it turns out to be the wrong handle to grasp since the connection between physical descriptions and experience has never been made precise enough to admit analysis.

Rather, in classical physics the evidential connection is made between the physical description and a certain class of *local beables*, such as the positions of macroscopic objects. For example, it is a straightforward prediction of Newtonian mechanics that, neglecting the resistance of the air, heavy bodies let go from the same point at the same time will hit the ground together. This is a claim that can (with the appropriate idealizations) be rigorously derived. So, one thinks, a natural way to test the theory is the method of Galileo: drop the bodies and see what happens. And this is indeed a good test *provided one can determine through observation when the bodies hit the ground*. But it is part of our pre-theoretical beliefs about the world that this is precisely the sort of thing one can determine, if the bodies are large enough. Our ability to reliably observe such facts is not itself derived from the physics: it is rather a presupposition used in testing the physics. So the contact between theory and evidence is made exactly at the point of some local beables: beables that are predictable according to the theory and intuitively observable as well.

The pre-theoretical intuition that certain physical states of affairs are unproblematically observable is not couched in the terminology of a physical theory: it is couched in everyday language. If Galileo drops rocks off the Leaning Tower, what is important is that we accept that it is observable *when the rocks hit the ground*. If the physical theory itself asserts that rocks are made up of atoms, then it will follow *according to the theory together with intuition* that we can observe when certain collections of atoms hit the ground, but this latter is clearly not the content of the observation. If the theory says instead that rocks are composed of fields, then it will follow that we can observe when certain fields hit the ground, or when the field values near the ground become high. It is easy enough to see how to translate the claim that we can see the rocks into the proprietary language of atomic physics or continuum mechanics or string theory. But the critical point is that *the principles of translation are extremely easy and straightforward when the connection is made via the local beables of the theory*. Collections of atoms or regions of strong field or regions of high mass density, because they are local beables, can unproblematically be rock-shaped and move in reasonably precise trajectories. If the theory says that this is what rocks really *are*, then we know how to translate the observable phenomena into the language of the theory, and so make contact with the theoretical predictions.

It is perhaps misleadingly narrow to insist that the existence of the right sorts of local beables in a theory makes the contact between the theory and its evidential base transparent: more directly, it makes the connection between the theoretical picture and the world as we pre-theoretically take it to be transparent. We take the world to contain localized objects (of unknown composition) in a certain disposition that changes through time. These are the sorts of beliefs we *begin with*. A physics that cannot somehow account for these beliefs is a physics that we would not have any use for. This is not to say that a physics with no local beables at all could not, in principle, account for those beliefs, but it is to say that understanding such a theory, and its relation to our pre-theoretical beliefs, is going to be a much, much more complicated business than understanding a theory with observable local beables.

It is worth noting here that a persistent abuse of terminology has helped to obscure these basic points. In discussions of quantum theory, the postulation of anything in the ontology beside the wavefunction is commonly called the postulation of *hidden variables*, such as the particles in Bohmian mechanics. Since the wavefunction itself is *not* a local beable, any version of quantum mechanics that has local beables at all will risk having those local beables denominated 'hidden'. But if they really *were* hidden, i.e. if we could not easily tell just by looking what they are, then the postulation would not help solve the problem of contact with evidence at all. It is exactly because the local beables are *not* hidden, because (according to the theory) it is easy to physically produce correlations between the disposition of those beables and the state of a 'measuring apparatus' (or the state of our brain), that they can play the right role in our epistemology. The local beables—at least some of them–had better be manifest rather than hidden. In Bohm's theory they are.

Because of the mediating place of local beables, classical physics could be tested without mention of the mind-body problem or the problem of connecting claims about experience with physical descriptions: the evidence, after all, was stated not in the language of experience but in the language of local physical facts (e.g., that the rocks hit the ground together). What, then, would be the situation of a theory that lacked local beables altogether? How could the connection between the theory and the world be made?

Since we have no such theory to hand, nothing definite can be said here. But some observations are in order. First, it is rather hard to see why a theory that lacks local beables altogether would bother to postulate anything like *spacetime*: after all, if there is nothing *in* any local region of spacetime, why think there is a spacetime? Why have the container if there is nothing contained? One would imagine that a theory that does away with local beables would also do away with the locations that the local beables might have inhabited. So the entire classical picture of local beables in a spacetime would have to be replaced.

And if spacetime goes, then *a fortiori* questions about whether the theory is relativistic go as well. A relativistic theory is exactly one whose laws can be framed using only the resources provided by the relativistic metric. But if there is no spacetime with a relativistic metric in the first place, the question of whether a theory can be framed using only the metric becomes idle.

Furthermore, if the local beables and the locations are removed from the physical ontology, it is hard to see how evidential contact with the world is to be made *except* at the level of conscious experience. Such a result may well warm the hearts of some, who want to derive from quantum theory the most radical consequences. But whatever one's attitude towards the result, it cannot be overemphasized how difficult this project would be to carry out. The physical description of the universe would have no local structure at all: nothing that corresponds to our intuitive picture of brains or neural activity. Nonetheless, some connection is supposed to be recognizable between that physical description and our conscious *experiences*, even though we have nothing like a precise vocabulary with which those experiences, as such, can be described. Once again, the mediating role of local beables obviated all these problems: there is no doubt a question about how, when we look at a rock with a certain shape, a conscious experience of certain sort arises, but classical physics could put that question off for another day (which perhaps would never come). All classical physics needs is the belief that experiences *as of* a rock of a certain shape typically *are* experiences of a rock with that shape, and the physics could take care of the rock. It is hard to see even how to begin if the physics has, in its own terms, to take care of the *experiences*.

Local beables also make transparent the explanation of the intersubjective character of physics. If there are local physical facts, such as the directions that pointers are pointing or whether cats are alive or dead, and if those facts are easily accessible to observers, then we see how many observers could easily come to share beliefs about the world. They all look at the same cat or the same pointer. If there are no local facts at all, what is the source of intersubjective agreement? What is there *outside of the various observers* that all the observers could independently become aware of, and hence agree on?

Once again, these puzzling questions might seem grist for some radical mills. What is 'black hole complimentaritary' except the claim that there is no common set of facts, even of the most seemingly evident kind, that all observers would come to agree on? What is the many-worlds theory but the claim that what we call 'observation' of a Schrödinger cat is not a process by which many people can come to agreement about the state of the cat, but rather a process by which many people all subdivide into many many many people, largely unaware of each other's presence, with the illusion that everyone who looked 'saw the same thing'? Perhaps these approaches could be made precise and clear, with the consequence that they can get the right claims about experience while jettisoning everything we ever thought we knew about the physical world. But this is a daunting task, and one that a theory with the appropriate sort of local beables can avoid.

In sum, a physics devoid of local beables would be a radically different kind of physics, a physics faced with problems of a completely different scale and sort than any theory in human history. It would have to make due without tables and chairs and rocks and trees and pointers and spacetime itself, in anything like the way we took them to be. It would be a change in the physical account of the world infinitely more staggering than, say, the addition of a few compactified dimensions to spacetime, or the admission of a physical foliation of spacetime, or the understanding of electrons as states of small vibrating strings, or the introduction of a discrete spacetime structure at the Planck scale.

As David Albert has pointed out (p.c.), even if one accepts the mediating role of local beables, a critical question still remains. One might, as in Bohmian mechanics, take both the spacetime and the local beables as ontological primitives. But one might also try instead to *derive* a physical structure with form of local beables from a basic ontology that does not postulate them. This would allow the theory to make contact with evidence still at the level of local beables, but would also insist that, at a fundamental level, the local structure is not itself primitive. The notion is that the dynamics of a very high-dimensional object in a high-dimensional space could somehow implicitly contain within it—as a purely *analytical* consequence—a description of local beables in a common low-dimensional space. This approach turns critically on what such a derivation of something isomorphic to local structure would look like, where the derived structure deserves to be regarded as *physically* salient (rather than merely mathematically definable). Until we know how to identify physically serious derivative structure, it is not clear how to implement this strategy. As we will shortly see, for example, the same GRW dynamics can be supplemented by *different* local beables: a mass density ontology and a flash ontology. Presumably, an account that treats these local beables as derived would determine whether one or the other (or neither) of these is the derivative ontology that is 'really' implicit in the GRW dynamics. At this point, I cannot see any principles that would tell us how to decide between these. So, at least until the principles of identifying derivative ontology are clarified, the only clear way for a theory to *have* an ontology of local beables is to directly *postulate* such an ontology. Since a theory that takes the wavefunction of the universe to be ontologically complete (not just informationally complete) makes no such postulate, it would seem to have no local beables at all.

3.1. Local beables and spontaneous collapse

Does the theory of Ghirardi, Rimini and Weber postulate any local beables? Although this appears to be a well-formed question, it turns out not to be, on account of a certain vagueness in the term 'the theory of Ghirardi, Rimini and Weber'. Some terminological conventions can help clear the matter up.

Let us call the following two principles the *core principles* of GRW: principle (1) the wavefunction evolves in accord with the GRW dynamics, and principle (2) the wavefunction is *informationally* complete. If one denies either of these principles, then it seems to me appropriate to say that one is rejecting GRW. But these two principles by themselves do not completely specify a theory, and in particular they do not specify an ontology. Again, an analogy might help. One could take as core principles of Maxwell's theory (1) the electromagnetic field evolves in accord with Maxwell's equations, and (2) the electromagnetic field is informationally complete. But these two principles are equally compatible with diametrically opposed views about the status of the charge density in the theory. One could hold (the usual view) that the electromagnetic field is *not* ontologically complete: there is more to the physical ontology than the field. There is, in addition to the field, the charge density, another local beable of the theory. The charges are there, distributed through space in a certain way. The charge density is related to the electromagnetic field by means of a *law*, in virtue of which one can determine the charge density by taking the divergence of the electric field. This allows the field to be informationally complete without being ontologically complete. On the other hand, one could take the view that the field is not only informationally complete but also ontologically complete: the field is really all there is. The charge density is then related to the field not by means of a law but by means of a *definition* (or perhaps an *ontological analysis*): the charge density just is the divergence of the electric field, whatever that happens to be. In both cases the charge distribution supervenes on the state of the field, but in the first case it is a matter of nomological supervenience and in the second logical (or definitional, or analytical) supervenience[6]. These two views postulate two different ontologies, and deserve to be called two different theories, even though they equally respect the core principles of Maxwell's theory. For convenience, we can call them two ontologically distinct *versions* of the theory.

---

[6] An exactly parallel issue comes up in some versions of analytical mechanics with respect to $\mathbf{F} = m\mathbf{A}$. Does this equation represent a *law*, as Newton thought, or merely a *definition* (of, e.g., $\mathbf{F}$)? Proponents of Ockham's razor may prefer the latter view, which reduces the ontology of the theory by eliminating forces as something over and above masses and accelerations. According to this view, the equation called 'Newton's law of motion' couldn't possibly be *false*, since it is just an implicit definition of the force. It could, of course, be a *pointless* or *unhelpful* definition. Newton, however, simply would reject this view. He took forces to have a real existence independent of masses and their accelerations, and took the equation to express a law relating these different things. He also thought that we had straightforward access to at least some facts about forces, masses and accelerations that made the law testable, and that observations—within experimental error—confirmed it. So the view that $\mathbf{F} = m\mathbf{A}$ is a definition is properly thought of as a revision of Newton's theory, not an explication or alternative presentation of it. There are other ways of considering Newton's law, and we have not touched on the complications that arise from the fact that $\mathbf{F} = m\mathbf{A}$ holds, in any case, only for the *net* force on an object, while various individual laws (e.g. the law of gravitation) specify only *component* forces. But the example still serves to distinguish treating an equation as an expression of a law as opposed to a definition.

Similarly, there can be different versions of GRW: theories with different ontologies that equally respect the core principles. There is, for example, a theory according to which the wavefuction is not only informationally complete but also ontologically complete. We can call this version *bare GRW*. But there are other versions that supplement the wavefunction with additional ontology. These versions are different physical accounts of the world than bare GRW, and differ from each other. As we noted above, the most common understanding of GRW among philosophers is as bare GRW: a monistic theory in which all there is the wavefunction. And one might think that this understanding is clearly the best-motivated one because of principle 2. For if the wavefunction is informationally complete, then one apparently *could* interpret any other 'ontology' as merely the consequence of *definitions*, as just misleading ways of talking about facts about the wavefunction. But there could also be countervailing arguments. Just as it could be most plausible to regard Maxwell's theory as postulating both the field and the charge density, with a law (rather than a definition) connecting them, and most plausible to regard Newton's theory as postulating forces, masses and accelerations, with law (rather than a definition) connecting them, so too there could be clear advantages to a version of GRW that has both the wavefunction and something else in its ontology, with a law (rather than a definition) connecting them.

The most obvious possible advantage of some non-bare GRW theory is that, unlike bare GRW, it could postulate local beables. It should be clear that bare GRW has no primitive local beables since by definition all it has is the wavefunction and the wavefuction is not a local beable. The wavefunction has no 'value' or 'state' in a small open region of space time, for it is not defined on space time but on a very high-dimensional space. And much of the burden of this essay has been to argue that local beables make the testability of a theory a relatively transparent affair.

The alert reader, though, may at this point detect an incoherence in the argument to this point. For I have claimed (1) the usual understanding of GRW has been bare GRW, (2) bare GRW contains no local beables (3) without local beables, it is difficult to understand how a theory could be testable. But it is also a plain historical fact that GRW has been taken, from the beginning, to be a theory with clear testable consequences. How could all these be simultaneously true?

(Indeed, the above set of claims could also be made for the 'standard' interpretation of quantum mechanics, however one understands it. The Copenhagenists clearly rejected the EPR argument that the quantum-mechanical description is (informationally) incomplete, and, as with GRW, the natural default position is that the wavefunction is not only informationally but ontologically complete: all there is the wavefunction. But then the theory has no local beables, and the question arises about how it can be testable and make contact with laboratory operations and observations. The standard answer is by the association of concrete laboratory operations with Hermitian operators, but that just raises the question of the principles by which *that* association is made: there is nothing in a piece of physical apparatus that naturally suggests a Hermitian operator!)

The answer to the puzzle might be that there is, after all, a direct route from the quantum description to claims about experience that does not pass through local beables in spacetime, so the argument I have been making is just wrong. But I think the correct answer is rather that there are quite natural ways of associating a description involving local beables in spacetime with wavefunctions, and that it is these associations that make the theory comprehensible, even though according to official doctrine the local beables do not really exist. Let us see how that can work.

Begin with the easiest case: a single electron in a two-slit apparatus. First of all, one always begins by at least conceptualizing the *apparatus* in terms of local beables in spacetime:

an opaque wall with two slits in it in the usual geometrical configuration. As a Copenhagenist might observe, we begin by conceptualizing at least *part* of the experimental situation in classical terms (in the sense of local beables, not in the sense of Newtonian physics), and we use that description when we come to treat the electron. It is, for example, this conceptualization of the wall that guides us as to the form of the potential to use in the Hamiltonian of the electron. If, at the end of the day, one comes to believe that there really is not any wall with a particular geometrical structure in spacetime, then the utility of this starting point will require some explanation, but let us leave that wrinkle aside.

For the single particle case, the configuration space of the system is isomorphic to physical space, so in this picture (which already has a physical space that contains the wall!) it is easy to add a wavefunction for the single electron and to picture it as a physical field on physical space. So it is also easy to associate local beables with the electron, at least in the imagination. One could even imagine the single-particle wavefunction *itself* to be a local beable—to have ontologically independent parts that exist in disjoint regions of spacetime—but since that conceit will clearly not survive even the transition to a two-particle state, we will not consider it. There are other possibilities for local beables that immediately recommend themselves.

What is normally said of the electron in this situation? Begin with the *location* of the electron. Since the wavefunction is informationally complete, and the wavefunction does not pick out any particular point of space, one cannot say that the electron is in some particular small region. The wavefunction is 'smeared out' over a large region of space, so it is usually said that the electron itself is 'smeared out'. Or, using the resources of wave-particle duality, one says in this circumstance that the electron 'acts like a wave rather than a particle', and clearly the right thing to say about (classical) waves is that they occupy a large region of space. More formally, the wavefunction provides, in the usual way, a probability distribution for the possible outcomes of a 'position measurement', and it is easy enough to devolve into talking about the electron itself being distributed over spatial regions in a way that matches the probability distribution: proportionally more of it in places where there would be proportionally higher chance of 'finding' it with a 'measurement'.

And since the electron has a mass, it is easy to fall into thinking of the mass itself as 'smeared out' in just this way, that is, to associate with the electron a local beable, a mass distribution, that is proportional to the squared amplitude of the wavefunction in position representation. It is strictly against the principles of the standard interpretation to take such a mass distribution *seriously*, to really think it is physically *there*, but the image of such a mass distribution inevitably presents itself when one tries to picture what is going on. And this *picture* of a local beable is enough to provide the normal sort of linkage between the theoretical description and concrete situations.

Indeed, in the case of the single-particle two-slit experiment, the picture of the evolution of the mass density in space is intuitively satisfying. The electron begins 'smeared out'—the mass density is significantly non-zero over a large area—and the smear does indeed go through both slits. This explains how the interference pattern can form only when both slits are open. The smear reaches the screen still smeared out, but the first GRW hit suddenly localizes it: almost all the mass density gets suddenly concentrated in a very small spatial region. What could more intuitively correspond to the particle-like appearance of the electron at a particular place on the screen? It is essential, of course, that in this picture the mass distribution is a local beable that inhabits a common ordinary space with the screen, so the clumping of the mass can occur in some region of the screen.

Once we go from one to two particles being treated purely quantum mechanically, the situation becomes more complicated. Again, we start by conceiving of the apparatus as having a certain geometrical structure in space time. But if we fire entangled pairs of particles at the slits, the wavefunction for the pair is no longer a function on something isomorphic to physical space: it is a function on a space of twice as many dimensions. So the wavefunction itself no longer can be directly interpretable as a local beable. As Bell says:

> [T]he wavefunction as a whole lives in a much bigger space (than physical space), of 3N dimensions. It makes no sense to ask for the amplitude or phase or whatever of the wavefunction at a point of ordinary space. It has neither amplitude nor phase nor anything else until a multitude of points in ordinary three space are specified ([2], p 204).

Nonetheless, the basic picture of an evolving mass density in ordinary space can be carried through. There is still a perfectly definite probability distribution for the results of ‘position measurements’ carried out on each particle. Each particle can be associated with an evolving mass distribution in this way, and each mass distribution thought of as a local beable. (The various mass distributions could all be combined into one big one, summing at each point in space, or each particle could be thought of as having its own proprietary mass, so there are many, distinct, coexisting local beables.) ‘Wave-like’ and ‘particle-like’ behaviour of the particles would again correspond to the mass distribution being smeared out and clumped up, respectively.

It must, I think, be something like this picture that Bell had in mind when he said that the GRW wavefunction ‘commits itself very quickly to one pointer reading or the other’. He cannot mean, of course, that the wavefunction of the pointer quickly evolves into an *eigenstate* of the position operator: that is not the case because of the tails. But if one associates a mass distribution in ordinary space with the wavefunction in this way, then in short order almost all of the mass density associated with the pointer will be either in one indicator position or another, depending on how the hits occur.

All of this strikes me as quite sensible, a reasonably clear method for associating a picture of local beables moving in ordinary space with the evolution of the wavefunction in a high-dimensional space. Thinking in this way, one would know what to expect to see (according to the GRW theory) upon entering a room where a certain kind of experiment was being conducted. The fundamental question left to us is whether one can consistently adopt this method of making sense of the theory *while still denying local beable status to the mass distribution*. Can one reasonably maintain that the theory is really ontologically monistic, that all of this talk of local beables in ordinary space is just a fiction, but still use this method to extract predictions from the theory?

In one sense of ‘can’ one can, but I do not think such a move is scientifically respectable. Clearly, the method does provide a sort of map from the quantum state to expectations about experience: one’s experiences will be *as of* a certain set of local beables disposed in ordinary space. But it is not legitimate to use this technique while denying the real existence of the beables (and the space!), without offering anything at all in its place. It is as if an opponent of the atomic theory of matter were to offer the following as a rival theory: matter is not, in fact, atomic, but it always behaves exactly as if it were *and I have no further account of why this should be so*. What the wavefunction monist has is just a field evolving in a very high-dimensional space, quite unlike what we take ordinary space to be. It is true that under a certain fictive mapping, one can associate that evolution with the motion of local beables in an ordinary-looking space, and that motion corresponds to what we see. But under other perfectly well-defined fictive mappings, it will correspond to a bizarre evolution of local beables in ordinary space, or a bizarre evolution in a bizarre space, and so on. Why should one pay any more attention to one fictive mapping than another?

There are situations where the practical effectiveness of a fiction can be explained: it is easy enough to derive from a heliocentric account of the Solar system an explanation of why certain predictions made by a rival geocentric account will be accurate. But here the comprehensibility of the heliocentric theory is essential. If we did not already know how to connect the heliocentric states with observations, we could not use the theory to explain the observational successes of the rival. But our original problem was how to understand what the monistic theory is claiming about the physical world, how it makes connection to experiment and empirical evidence in the first place. We cannot appeal to mere fictions to solve that problem.

For these reasons, I take bare GRW to be a very problematic theory. It only postulates as real a complicated evolution of something in a very high-dimensional space, yet has to account for observations that seem to be of particularly situated objects in a three-dimensional space. A merely fictive three-dimensional space populated with merely fictive local beables derived by choice of one out of an infinitude of possible ways to generate the fiction from the reality does not solve the problem.

But it is clear what would solve the problem: remove all the talk of fiction! If one believes that in addition to the wavefunction *there really is* an ordinary space that *really does* contain local beables that *really do* evolve in a specified way determined by the wavefunction, then you have something. The existence of an infinitude of other merely fictive beables in a merely fictive space definable from the wavefunction is neither here nor there: *they* obviously play no role in the explanation of our experience. This change from bare GRW to a non-bare theory that admits more than just the wavefunction need in no way impugn principle 2: if the distribution of the local beables is determined by law from the wavefunction, then the wavefunction can continue to be informationally complete. And the fact that the non-bare theory can have explanatory power that far outstrips the bare version is no more puzzling than the fact that a theory that postulates the real existence of atoms can have explanatory power unavailable to a theory that does no more than to deny their existence.

This is the importance of Ghiradi's emphasis on a mass-distribution function (see [4] and [5]). By restoring an uncontroversial local ontology, the fundamental physical picture becomes clear and questions about spacetime structure—particularly about the prospects for a fully Lorentz invariant theory—can be cleanly framed. And once one marks the distinction between a monistic theory (in which the wavefunction would automatically be informationally complete) from a non-monistic theory (in which, due to nomic supervenience, the wavefunction may turn out to be informationally complete), one also sees that there are distinct different *options* for the local ontology. Such options correspond to fundamentally different physical theories, all of which my share exactly the same (GRW) dynamics for the wavefunction alone.

Most famously, Bell himself suggested quite a different sort of local beable that one could add to the core principles of GRW. His explication is admittedly somewhat obscure, but his description suggests a theory in which the fundamental ontology includes ordinary space time and a set of physically distinguished space-time points. This passage follows on directly from the passage last cited.

> However, the GRW jumps (which are part of the wavefunction, not something else) are well localized in ordinary space. Indeed, each is centred on a particular space-time point $(\mathbf{x}, t)$. So we can propose these events as the basis of the 'local beables' of the theory. These are the mathematical counterparts in the theory of real events at definite places and times in the real world (as distinct from the many purely mathematical constructions that occur in the working out of physical theories, as distinct from things that may be real but not localized, and as distinct from the 'observables' of other formulations of quantum mechanics, for which we have no use here). A piece

> of matter then is a galaxy of such events. As a schematic psychophysical parallelism we can suppose that our personal experience is more or less directly of events in particular pieces of matter, our brains, which events are correlated with events in our bodies as a whole, and they in turn with events in the outer world ([2], p 205).

The obscurity of Bell's proposal, for me, lies in the claims that the jumps are 'well localized in ordinary space', given that he has just insisted that the jumps are part of the wavefunction and the wavefunction itself does not 'live' in ordinary space. There is, however, a straightforward *association* of each jump with a point of ordinary space (or, more exactly, space time): as Bell says, the Gaussian used to represent the jump is indexed to a particular place. And the key to Bell's suggestion—which is more than a matter of simply *noting* this association—is to invest those locations with a real, distinct, physical character. It my help to consider an analogy: in classical physics, every finite extended material system is *associated* with a particular point in space, viz its centre of mass. But that does not mean there is any local beable at that place: the centre of mass can be in a region of empty space time. Making the centres of the GRW jumps into local beables is a *physical* posit, one that Ghirardi's mass distribution ontology does not make, just as Bell invests Ghirardi's mass distribution with no physical reality, even though the corresponding *mathematical* object is perfectly well defined. This physical posit, which has come to be known as the *flash* ontology, is clearly a different one from the mass density ontology suggested by Ghirardi.

These alternative ontologies of local beables that can be appended to the core principles of GRW illustrate how much latitude the core principles leave open. Each of these theories postulates the same dynamics for the wavefunction, and in each of these theories the wavefunction is informationally complete, but the space-time pictures of physical reality could hardly be more different. If all one could see were space time and the local beables (but if one could see them in all details), the fine-grained pictures provided by the theories would have nothing in common. In Ghirardi's world, one sees an evolving continuous mass density that would usually change continuously but would sometime undergo discontinuous changes that result in the sudden clumping of the mass in small regions of space. In Bell's world, one would see almost nothing at all: just a relatively sparse distribution of flashes, one for each of the sudden jumps in the alternative theory. But squinted at from a distance, both Ghirardi's and Bell's world would present a recognizable macroscopic world: where we think there are tables we would find, in the one case, a table-shaped concentration of mass, and in the other a table-shaped constellation of flashes.

This macroscopic similarity, though, should not lull one into the false belief that these are just two 'pictures' or 'presentations' of the same physical theory. Accepting one or the other theory will have far-reaching theoretical consequences. As an obvious example, accepting the mass density theory suggests that one might profitably try to remove the jumpiness of the evolution of the wavefunction: a continuously evolving wavefunction would correspond to a continuously evolving mass density, rather than one that evolves continuously only most of the time. So it becomes a natural thing to investigate continuous spontanious localization models, as Ghirardi and his collaborators have done. The resulting behaviour of the local ontology would be more elegant, and perhaps more plausible.

On the other hand, eliminating the jumps from Bell picture would eliminate the local ontology altogether! Continuous spontaneous localization of the wavefunction is the *last* thing that one would pursue in this theory. So the research programmes that naturally arise from the alternative pictures are in some ways diametrically opposed.

And the difference between the theories runs even deeper. Given a clear local ontology, one is in a position to ask clear questions about matters such as Lorentz invariance in a possible relativistic version of the theory[7]. And again, the contrast between the two approaches could not be more stark.

It is easiest to begin with the non-relativistic versions, which already exist. Both versions make use of absolute simultaneity in their formulation, for a unique time parameter is used in specifying the dynamics of the wavefunction. But would that absolute simultaneity *manifest itself* at the level of the local beables? Supposing, again, one had full and unfettered access to the local beables in two different regions of space, could one somehow tell, by examination of the contents of one region, *when something happened* in a distinct region?

Here is a simple experiment. Put a single electron into an equal superposition of being on the right side of the room and one the left side of the room. According to the mass-distribution theory, half the mass will then be located on the right and half on the left, and this distribution will almost certainly persist for centuries if nothing more is done. This particular distribution of mass would not be, in Ghiradi's terminology, *accessible*, because no experiment could determine that that *is* the distribution, but that is not relevant to our concerns: the mass is *there* and we are pretending that we have access to all the local beables as they are.

Now suppose we make a 'position measurement' on the right side of the room. By means of the usual GRW dynamics, once we appropriately couple the electron to a large device, there will almost immediately be a collapse, with the result that the mass density on the *left-hand* side of the room will either suddenly double or suddenly be reduced to essentially zero. So if we could see that sudden jump, we could identify the exact moment that the distant measurement was made. We could determine that two distant events (the measurement and the jump in mass density) took place at the same moment of absolute time, just by keeping careful track of the local beables.

On the other hand, what would happen according to Bell's flash ontology? Once the superposition is prepared, *absolutely nothing* would locally exist on the left side of the room: there would be no flashes associated with the electron. The 'measurement' on the right would be associated with flashes, and the exact position of those flashes would be a matter of chance, but no matter which way the measurement came out *there would almost certainly be no flash on the left, since it is extremely unlikely that the electron itself receive a* '*hit*'. No matter when the right-hand measurement was made, or what its outcome is, the left-hand part of the room will be devoid of any local beables. Evidently, one could not use this ontology to determine distant absolute simultaneity in the way one can use the mass density ontology.

More generally, in the mass density version, if one had perfect epistemic access to the local beables then one could reliably send messages from one side of the apparatus to the other, while in the flash version, even granting such perfect access, this would not be possible. Facts like this suggest that a completely relativistic (but non-local!) version of GRW could be formulated, a hope that has been realized, at least to some extent, by Roderich Tumulka [7]. So the question of whether or not a version of GRW admits of fundamental Lorentz invariance (as opposed to merely phenomenal Lorentz invariance) turns in part on the exact local ontology of the theory. This ought not to be very surprising since relativity is a theory of spacetime structure, and hence has its most immediate consequences for physical magnitudes that have unproblematic locations in spacetime.

One more important conceptual consequence of having a clear ontology of local beables is worthy of note: it can help explicate the transformational properties of *non-local* ontology under a spacetime transformation. Again, the simple non-relativistic case can serve as an example.

---

[7] The observation that the ontology of local beables is crucial to the analysis of the symmetries of a theory was made by Sheldon Goldstein in [6].

We have already seen that the mathematical representation of the wavefunction of a single particle is just like the mathematical representation of a certain kind of local beable, which we call a scalar field. The scalar field as a local beable is a physical magnitude that exists at a point (or in an indefinitely small open region), whose possible physical states have one degree of freedom. We have to be careful to keep clearly separate here the scalar field as a *mathematical object* (an assignment of a scalar to each point in spacetime) and the scalar field as a kind of *physical entity*, a sort of local beable whose most transparent mathematical representation is a mathematical scalar field. For there can be other physical entities that are *represented* by mathematical scalar fields but are not physical scalar fields as just defined. But how can we tell the physical scalar field from other entities that just happen to be representable by mathematical scalar fields?

One clue lies in the transformational properties of the mathematical representation under a change in space-time reference frame. Suppose, for example, we have a physical scalar field, and the appropriate mathematical representation of it in a particular inertial frame. Then we know automatically what the appropriate mathematical representation is in any boosted frame: the value of the field at the point (x', t') in the new frame is just the value of the field at $(x - \mathbf{v}t, t)$ in the original. And from this, we can also say what should count, in a single fixed frame, as a boosted *field*: it would have the value at $(x - \mathbf{v}t, t)$ that the original field has at (x, t). Evidently, any spatial periodicity of the mathematical representation of the physical scalar field will not change under either transformation.

In contrast, the periodicity of the single-particle wavefunction does change under a Galilean boost. This shows already that even the single-particle wavefunction is not a local beable (or at least: not a physical scalar field) even though the usual mathematical representation used for it is the same as that used for a physical scalar field. Even the single-particle wavefunction does not 'live' in spacetime, even though one can profitably 'visualize' it as propagating through space time when analysing, e.g., the two-slit experiment.

How then *should* the (mathematical representation of the) wavefunction transform under a Galilean boost? In a theory with the right kind of local beables, this transformation can be *derived*, while in a theory without a clear local ontology, the transformation can only be *postulated*.

Take Bohmian mechanics as an example. Suppose one has the single-particle wavefunction specified in a particular frame of reference, and the guidance equation that indicates how the particle's velocity (in that frame) is a function of the wavefunction and the particle position. In particular, suppose only that the particle's velocity is proportional to the gradient of the wavefunction (the easiest way to generate a vector field from a scalar field). Then it is obvious that the (mathematical) wavefunction *cannot* transform as a physical scalar field under a Galilean boost since the gradient would not be changed, but the velocity of the particle would be. So even in the single-particle case, Bohmian mechanics implies that the wavefunction is not a (physical) scalar field, and one can even *derive* what the mathematical transformation of the wavefunction must be under the Galilean boost. In a theory that has both local beables and a non-local ontology, the space-time transformation properties of the (mathematical representation of) the local beables may be determined by *what the local beables are*, and the transformation properties of the (mathematical representation of) the non-local ontology may be determined, or at least constrained, by the transformations of the local ontology and the structure of the theory. But in a theory with *no* local beables at all, it is hard to see how one could derive, or motivate, or make plausible how to implement a space-time transformation like a Galilean boost.

Note that no mention has been made, in the Bohmian derivation of the Galilean transformation, of measurements or observations or anything like that. What we know the transformation properties of are the local beables and the functions of the local beables, such as particle velocities. The question of what the outcome would be if one tried to *measure* or *observe* a velocity never had to be raised. The 'standard' approach, however, might try to save the situation with an appeal to measurement as follows.

Suppose we have a wavefunction, even a single-particle wavefunction, expressed in a particular inertial frame, and we want to know how it ought to transform under a Galilean boost. We take the as fundamental purpose of the wavefunction to give probabilistic predictions for the outcomes of 'measurements' and nothing else: there is no associated local ontology at all. Still, one might argue as follows: the boosted wavefunction ought to be whatever produces the same predictions for a *boosted* laboratory (or boosted measurement set-up) as the original wavefunction yields for a laboratory at rest in the initial frame. This then allows one to work backward to the form of the boosted wavefunction. Unlike the derivation through the local beables, ineliminable reference must be made to 'measurements' and 'measurement outcomes', but the standard interpretation is already up to its neck in that commitment anyway.

But this sort of 'derivation' is a cheat if one tries to coherently maintain a stance of wavefunction *monism*. For what it presumes is that we have a clear account of what should count as the *boosted laboratory*, and hence a *boosted measurement interaction*. But the wavefunction monist has no principled way to specify how to boost the laboratory, as a physical transformation: that is just another instance of the problem we are trying to *solve*.

The reason that this problem does not come up in practice is because the 'standard' interpretation is a legacy of the Copenhagen view, and the Copenhagen view does not postulate wavefunction monism. Copenhagenism insisted on the *necessity* of having a classical description somewhere, the description of the 'measurement situation': the infamous Copenhagen 'cut' was exactly between a quantum realm and a classical realm. And the classical description would, of course, be in terms of local beables, so there is no problem applying a spacetime transformation to *it*. Within this sort of a dualistic picture the problem of spacetime transformations of the wavefunction can be approached. The problem, of course, is that this sort of dualistic ontology is impossible to take seriously: no one ever thought that there were *really* two different sorts of physical systems, the classical and the quantum, that somehow *interact*. If that were the view, then the 'cut' would be a matter of physical fact: somewhere the classical and quantum bits of ontology would actually *meet*. Furthermore, it is evident that the 'classical objects', measuring apparatus and so on, are *composed out of* 'quantum stuff' (electrons, protons, and so on), so this cannot really be a dualistc ontology.

In the confused morass of Copenhagenism, the observation that the 'cut' could, For All Practical Purposes, be moved about at will within a large range was taken to show that the cut itself corresponded not to a physical fact but to a convention, or something like that. But if the theory can be formulated without a cut at all, let it be so formulated. Having removed the cut and put everything in the quantum ontology, one would evidently remove all the local beables, and all the problems we have been discussing would return.

None of this constitutes a proof that one could not, in a coherent and systematic way, formulate a theory that eliminates all local beables. It does suggest that in such a theory, spacetime, and hence spacetime transformations, would not play any fundamental role. Questions about the Lorentz Invariance of such a theory, or compatibility with Relativity, could not even be framed in a clear manner. It is obscure what could be meant by saying that such a theory is Lorentz Invariant except at the gross observational level: nothing we would count as a laboratory operation could pick out a preferred reference frame. But *that* sort of Lorentz Invariance is already achieved, e.g., by Bohmian mechanics. If one wants something more serious, then one had better have spacetime, and local beables, in the theory in order for the possibility of serious Lorentz invariance to arise.

It has been a long hard struggle from the mysticism of Copenhagen back to a clear idea of what a physical ontology is. Once one appreciates the central methodological role that local beables play in the formulation and testing of theories, the straightforward introduction of local beables into the GRW theory and the rejection of ontological completeness of the wavefunction (even in a theory in which the wavefunction is informationally complete) can be recognized as a positive, and perhaps even necessary, step in the development of that theory.

Acknowledgments

I wish to thank Roderich Tumulka, Detlef Dürr, Nino Zanghi, David Albert, and especially Shelly Goldstein for trying to keep me honest. All lapses are my own.

References

[1] Einstein A, Podolsky B and Rosen N 1935 Can quantum mechanical description of reality be considered complete? *Phys. Rev.* **47** 777–80

[2] Bell J S 1987 *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press)

[3] Born M 1971 *The Einstein-Born Letters* ed Born I (New York: Walker) trans.

[4] Benatti F, Ghirardi G C and Grassi R 1995 Describing the macroscopic world: closing the circle within the dynamical reduction program *Found. Phys.* **25** 5–38

[5] Bassi A and Ghirardi G C 2003 Dynamical reduction models *Phys. Rep.* **379** 257–426

[6] Goldstein S 1998 Quantum theory without observers Part 1: *Phys. Today (March)* 42–46
Goldstein S 1998 Quantum theory without observers Part 2: *Phys. Today (April)* 38–42

[7] Tumulka R 2006 A relativistic version of the Ghirardi-Rimini-Weber model *J. Stat. Phys.* (at press)