Quantum mechanics, emergence, and fundamentality

Peter J. Lewis

June 6, 2016

Abstract

Quantum mechanics arguably provides the best evidence we have for strong emergence. Entangled pairs of particles apparently have properties that fail to supervene on the properties of the particles taken individually. But at the same time, quantum mechanics is a terrible place to look for evidence of strong emergence: the interpretation of the theory is so contested that drawing any metaphysical conclusions from it is risky at best. I run through the standard argument for strong emergence based on entanglement, and show how it rests on shaky assumptions concerning the ontology of the quantum world. In particular, I consider two objections: that the argument involves Bell's theorem, whose premises are often rejected, and that the argument rests on a contested account of parts and wholes. I respond to both objections, showing that, with some important caveats, the argument for emergence based on quantum mechanics remains intact.

A strong form of holism or emergence is often taken to be a consequence of quantum mechanics (Teller 1986; Hawthorne and Silberstein 1995; Wallace and Timpson 2010; Healey 2016). I think this is basically right: quantum mechanics provides us with good (though not incontrovertible) evidence that physical systems have emergent properties. But the standard argument for this conclusion in terms of entanglement is only part of the story. This is because entanglement is a purely formal property of the mathematical representation of certain quantum systems, and the interpretation of that mathematical representation in the case of quantum mechanics is notoriously contested.

There are two related issues here. The first is that the most direct argument from entanglement to emergence is via Bell's theorem, but the major interpretations of quantum mechanics all deny the assumptions of Bell's theorem in one way or another. The second is that arguments for emergence assume a model of wholes built up out of smaller parts that, while intuitive, is controversial in the quantum mechanical context. I think these issues can be addressed, but at the cost of adding some caveats to the claim that quantum mechanics entails emergence.

The term "emergence" is used in a bewildering variety of senses, so let me begin by specifying how I will use them here. I am interested in strong rather than weak emergence, in Chalmers' (2006) usage. That is, my concern is whether there are high-level phenomena that are not deducible even in principle from truths in the low-level domain, rather than truths that are merely unexpected. In particular, I am interested in whether there are cases in which this "in principle" epistemic block arises because of irreducible high-level ontology. That is, by "emergence" I mean what Silberstein and McGeever (1999, 182) call "ontological emergence", namely the possession by systems or wholes of "causal capacities not reducible to any of the intrinsic causal capacities of the parts nor to any of the (reducible) relations between the parts".

Emergence in this sense is closely related to what Healey (2016) calls "physical property holism", namely the existence of physical objects "not all of whose qualitative intrinsic physical properties and relations supervene on qualitative intrinsic physical properties and relations in the supervenience basis of their basic physical parts." Given some plausible assumptions, emergence and holism in the above senses amount to the same thing. That is, if causal capacities are a kind of property, and irreducibility entails lack of supervenience, then a system that exhibits emergence also exhibits holism. And if every physical property entails at least one unique causal capacity, and lack of supervenience entails irreducibility, then a system that exhibits holism also exhibits emergence. So from here on I use the two terms interchangeably. Glossing over complications concerning the nature of reduction and of intrinsic properties, we can give a rough characterization of the target phenomenon as follows:

Emergence (holism): A physical system exhibits emergence (holism) iff it has properties that are not reducible to the intrinsic properties of its parts.

1 The case for emergence

Let us begin by examining the standard case for emergence (in the above sense) based on quantum mechanics. An electron has a property called *spin*: relative to a specified direction, the state of the electron can be spin-up, written $|\uparrow\rangle$, or spin-down, written $|\downarrow\rangle$. The electron's state can also be a *superposition* $a|\uparrow\rangle + b|\downarrow\rangle$ of the spin-up and spin-down states in any proportions a and b, where $|a|^2 + |b|^2 = 1$. If the spin of an electron in such a state is measured, one obtains the result spin-up with probability $|a|^2$ and spin-down with probability $|b|^2$.

Now consider a pair of electrons. Their joint state can be written e.g. $|\uparrow\rangle_1 |\downarrow\rangle_2$, in which case the first is spin-up and the second is spin-down. Similarly, each electron can be in a superposition state $(a|\uparrow\rangle_i + b|\downarrow\rangle_i)$, so that the joint state of both electrons is $(a|\uparrow\rangle_1 + b|\downarrow\rangle_1)(a|\uparrow\rangle_2 + b|\downarrow\rangle_2)$, which can be expanded to $a^2|\uparrow\rangle_1|\uparrow\rangle_2 + ab|\uparrow\rangle_1|\downarrow\rangle_2 + ab|\downarrow\rangle_1|\uparrow\rangle_2 + b^2|\downarrow\rangle_1|\downarrow\rangle_2$. In this case, if the spin of both electrons is measured, then for each electron there is a probability $|a|^2$ of obtaining spin-up and $|b|^2$ of obtaining spin-down, where these probabilities are entirely independent of each other.

It is also possible to prepare a pair of electrons in the state $|S\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle_1|\downarrow\rangle_2 - \frac{1}{\sqrt{2}}|\uparrow\rangle_1|\downarrow\rangle_2$. The notable thing about $|S\rangle$ is that it is not factorizable into a state of electron 1 and a state of electron 2. The lack of factorizability is reflected in the fact that the probabilities of obtaining spin-up and spin-down for each electron are no longer independent of each other: the outcome for electron 1 is spin-up iff the outcome for electron 2 is spin-down. States like $|S\rangle$ are called *entangled*.

It is natural to regard entanglement as evidence for the existence of emergence. The entangled state looks like a property of the pair of electrons, a causal capacity responsible for the correlated outcomes we observe. But the entangled state by definition can't be factored into a state of electron 1 alone and a state of electron 2 alone: it is an irreducible state of the pair. Quantum mechanics does allow us to assign states to the individual electrons—such states are called density operators—but the individual electron states don't entail the correlations between the measurement outcomes. So it looks like the explanation for the correlations is in terms of an irreducible joint property of the pair: the system composed of the two electrons has a property that is irreducible to the properties of its constituent parts.

So far we have assumed that quantum mechanics constitutes a complete description of the relevant causal capacities of the system, but this assumption has been challenged (Einstein, Podolsky and Rosen 1935). That is, one might suspect that even if quantum mechanics, as it stands, doesn't represent the properties of the individual electrons responsible for the correlations we observe, nevertheless such properties must exist. Bell's theorem (Bell 1964) apparently rules out such a completion. Bell proves that, subject to some plausible assumptions, no assignment of properties to the individual particles can generate all the correlations we observe for pairs of entangled particles. The entangled states that quantum mechanics assigns to such pairs of particles correctly predict all these observations. So it looks like the explanation of the results we observe requires the existence of properties of pairs of particles that are not reducible to properties of the particles taken individually.

This is arguably the best case we have for the existence of (strong, ontological) emergence.[1] But drawing any metaphysical conclusion on the basis of quantum mechanics is a fraught enterprise, precisely because the interpretation of the theory of quantum mechanics is so unsettled. This general concern applies in particular to the case of emergence.

There are at least two aspects of the argument for emergence one might worry about. First, it relies on the conclusion of Bell's theorem. However, many interpreters of quantum mechanics, including Bell himself (2004, 59), regard Bell's argument as a reductio of his assumptions rather than a direct argument for his conclusion. Indeed, the three major realist interpretations of quantum mechanics (Bohm's theory, the GRW theory and the many-worlds theory) all violate Bell's assumptions in one way or another, and so avoid the force of his conclusion. Second, the argument for emergence takes for granted that the entangled state is ascribed to a compound system with two individual electrons as parts. But the common-sense analysis of wholes into smaller parts is also frequently challenged on the basis of quantum mechanics. I consider these difficulties separately, starting with the latter.

[1] Lancaster and Pexton (2015) argue for (strong, ontological) emergence based on the fractional quantum Hall effect, but since the heart of their argument is the role of entangled states in the production of the effect, this is not a separate line of argument. Chalmers (2006) maintains that consciousness provides the only clear example of strong emergence, but the conceptual issues surrounding consciousness are arguably even more contested than those surrounding quantum mechanics.

2 Configuration space realism

The spin of an electron is usually represented as a vector—a quantity with magnitude and direction. Vector quantities are quite familiar in physics: we have no trouble understanding the velocity of an object as a vector-valued property of that object. So we are initially inclined, I think, to think of spin the same way—as a vector-valued property of an individual electron.

But the mathematics of spin is less familiar than this analogy makes it sound. The state $|\uparrow\rangle_1$ of a single electron is a vector in a two-dimensional space—not a three-dimensional space, as a direct spatial reading of spin might lead you to expect. Still, one might be tempted to regard this vector as representing a property of the electron. However, the state $|\uparrow\rangle_1 |\downarrow\rangle_2$ of a pair of electrons is a single vector in a four-dimensional space. For factorizable states like this one, the vector can be decomposed into a two-dimensional vector for electron 1 and a two-dimensional vector for electron 2. But for non-factorizable states like $|S\rangle$ the four-dimensional vector cannot be so decomposed.

So far, it looks like I have just been restating the argument for emergence. But some have seen it as pointing towards a way to avoid emergent properties. Realists typically regard the mathematical structure of a theory as a reflection of the structure of the world. So if the mathematical representation of the two-electron system inhabits a four-dimensional vector space, then we should think of the fundamental ontology of the system as inhabiting a four-dimensional space too. But then rather than thinking of the system as two objects each bearing a two-dimensional vector property, we should think of it as a single object bearing a four-dimensional vector property. So this isn't a compound system after all (despite appearances), and the argument for emergence fails before it gets started.

This view of fundamental ontology is usually expressed in terms of position properties rather than spin properties. The position of a single particle (at a time) can be represented using a wave function—a complex-valued function of three spatial dimensions. It is natural to think of the wave function of a single particle as analogous to a field, defined by an amplitude and phase at each point of three-dimensional space. But the position state of two particles is represented by a complex-valued function of six dimensions, and in general the position state of N particles is represented by a complex-valued function of $3N$ dimensions. That is, the wave function of a compound system is represented in a *configuration space*—a space in which each point corresponds to a specific *configuration* of all the particles in the system. When the wave function is entangled, it cannot be factorized into separate wave functions for each of the particles in the system, and hence cannot be represented in a three-dimensional space.

Albert (1996; 2013) argues that since the wave function inhabits a multi-dimensional configuration space, we should regard properties ascribed to points of this space as the fundamental ontology described by quantum mechanics.[2] Configuration space points are the basic physical parts. But in that case, even entangled states supervene on the properties of the basic physical parts of the system. Loewer (1996, 104) claims that it is an advantage of this configuration space realist approach that it requires no emergent properties—that it vindicates David Lewis's Humean supervenience, in the sense that every property of a system supervenes on the properties of its point-like parts. Those point-like parts, however, are points of configuration space, not points of ordinary three-dimensional space.

If Loewer is correct, then quantum mechanics does not entail emergence after all, because we have misidentified the parts of the system in the argument of the previous section. Rather than a compound system consisting of two fundamental entities (the electrons), we have a simple system consisting of a single point-like entity. Hence it is irrelevant that we cannot factor state $|S\rangle$ into a state of electron 1 and a state of electron 2: state $|S\rangle$ can simply be ascribed *in toto* to the single point-like entity. We can still talk in terms of electrons if we like, but they aren't fundamental. The individual electrons supervene on the fundamental ontology, in the sense that their individual (density operator) states can be derived from state $|S\rangle$.

There is something odd about this as an argument against holism, though. According to the configuration space realist, the most fundamental ontological part—a single point in configuration space—represents a possible configuration of particles over the entire universe. This sounds like holism with a vengeance, and seems to avoid holism on a technicality.

Let me try to diagnose more precisely the sense of fishiness here. What configuration space realism suggests is that our ordinary sense of what is a part and what is a whole is mistaken. A configuration space point is a basic part, even though it represents what we take to be the whole universe.

[2]Note that this is a controversial way of interpreting the quantum state; see the papers in Ney and Albert (2013) for opposing views. But it is a contender, and hence might be thought to provide a potential escape from emergence.

And a single electron is a compound object, in the sense that its state can contains contributions from many distinct configuration space points. And the reason this is fishy, I contend, is that "part" and "whole", as they appear in the definition of emergence at the beginning of this paper, are not technical terms of an interpretation of quantum mechanics, but precisely terms used in their ordinary sense.

That is, when we describe emergence as the possession by wholes of properties that are not reducible to the properties of their parts, we mean "part" and "whole" in an intuitive sense, the sense in which a tine is part of a fork and an atom is part of a molecule. In investigating holism, we want to know whether something like a classical reductivist understanding of the physical world is correct—according to which the properties of partridges supervene on the properties of particles, for example. This understanding presupposes a picture of wholes and parts localized in ordinary three-dimensional space, and a failure of this presupposition such that the basic "parts" are spread over three-dimensional space is tantamount to holism. We can introduce a technical vocabulary if we like, a vocabulary in which what we ordinarily call a part becomes a whole, and what we ordinarily call a whole becomes a part, and this technical vocabulary might have some theoretical usefulness. But a terminological choice like this can't undermine an argument for emergence.

A configuration space realist might insist that we should follow the physics in deciding what to call a part and what to call a whole, since physics is a better guide to fundamental reality than ordinary language. That's fine, but I think it misses the point. Under such a proposal, all the properties of a partridge would indeed supervene on the properties of the basic physical parts. In fact, the instantiation of one basic physical part—one configuration space point—would be sufficient to instantiate the partridge and all its properties, so in a sense the partridge has no proper parts. Perhaps this is what physics tells us. But if so, then physics tells us that electrons, partridges and galaxies are on a par as far as fundamentality goes. To conclude from this that physics tells us that holism is false would be rather absurd.

What is needed, I think, is a small modification in the definition of emergence/holism:

Emergence (holism): A physical system exhibits emergence (holism) iff it has properties that are not reducible to the intrinsic properties of its *spatially local* parts.

If configuration space realism is correct, and objects have no spatially local parts, then holism is trivially true. This seems exactly as it should be.

3 Priority monism

Even if you accept the argument of the previous section that radical redefinition of "part" and "whole" should not affect whether quantum mechanics entails emergence, you might still think that quantum mechanics challenges the metaphysical priority of parts over wholes. This is the position defended by Schaffer (2010). Schaffer contends that what we should take as fundamental is whatever can act as the relevant supervenience base. It looks like individual particles or individual space-time points cannot function as a supervenience base in quantum mechanics: for an entangled state, the state of the whole does not supervene on the states of the point-like parts. But if we turn our ontology on its head, supervenience is unproblematic. For the two-electron entangled spin system, the state of each electron individually supervenes on the state of the pair (in the sense that the density operator for each electron can be derived from the entangled state $|S\rangle$ of the pair). Similarly, the (density operator) position state of each of a system of N particles can be recovered from the wave function of the N-particle system. In general, we should take the physical state of the entire universe as fundamental: everything else supervenes on that.

Schaffer calls this view *priority monism*.[3] Note that priority monism does not challenge the ordinary conception of parts and wholes: a tine is still a part of a fork, even if the fork is more fundamental. So priority monism is not a direct challenge to the argument for emergence based on quantum mechanics; indeed priority monism is *motivated* by the failure of supervenience of the properties of wholes on the properties of their parts. But nevertheless there is a sense in which priority monism might be thought to make emergence beside the point.

The sense is this. Emergence can be understood as a failure of dependence: the properties at one level fail to depend only on the properties at a more fundamental level. While quantum mechanics may entail a failure of supervenience of the properties of wholes on the properties of parts, this is

[3]Like configuration space realism, priority monism is controversial: see responses by Bohn (2012) and Calosi (2014). But again like configuration space realism, it is a contender, and hence might be thought to pose an indirect challenge to the argument from quantum mechanics to emergence.

not a genuine failure of dependence according to the priority monist, because the dependence between the properties goes the other way up. The properties of the parts depend on the properties of the whole. If we understand metaphysical dependence correctly, then there is no failure of dependence, and the failure of supervenience of wholes on parts is not what we should be concerned about.

If priority monism is correct, then there are no "free-floating" higher-level properties: every higher level property is tethered by dependence to properties at a more fundamental level. So priority monism might make you feel better about emergence. One might even go further and argue that the absence of free-floating higher-level properties means that there is no real holism or emergence here. But it is hard to see how this could work. Defining holism as the failure of dependence of parts on wholes is fairly clearly absurd, since the existence of fundamental wholes *is* holism (as argued in the previous section). Attempting to define holism in terms of fundamentality rather than parthood courts triviality: if "x is more fundamental than y" means "y depends on x", then trivially everything depends on that which is more fundamental. And even if the triviality can be avoided, the absurdity remains, since one ends up asserting that irreducible wholes do not constitute holism. Priority monism treats the symptoms of emergence, but it is not a cure.

Earlier, I noted two worries you might have regarding the standard argument for emergence based on quantum mechanics. The second of these is the one we have been considering—that the argument presupposes a commonsense analysis of wholes into smaller parts. My contention has been that this presupposition is entirely appropriate: the sense of "part" and "whole" appearing in definitions of emergence and holism is the ordinary sense, not some technical sense that one might construct to overcome some of the interpretive difficulties of quantum mechanics. So emergence is safe from any redefinition of "part" and "whole" such as is implicit in configuration space realism. Further, any restructuring of dependence relations between parts and whole, such as priority monism, leaves the argument for emergence unscathed. Priority monism might be an attractive way to understand a world with emergent properties, but it doesn't undermine the existence of emergence.

However, the first worry I noted above remains: the standard argument for emergence depends on Bell's theorem, but most interpretations of quantum mechanics reject the assumptions of Bell's theorem. Let us turn to that now.

4 The case for emergence, revisited

The assumptions on which Bell's theorem is based can be formulated in various ways, but the following formulation is quite standard. Consider a pair of particles, passing through space-like separated space-time points x_1 and x_2 respectively. Suppose that one of a set of distinct measurements can be performed on one or both particles, where each measurement has a set of distinct outcomes. Then Bell's theorem follows from the following three assumptions:

Locality: The outcome of a measurement on particle 1 at x_1 cannot affect the properties of particle 2 at x_2.

Independence: The properties of particle 1 at x_1 cannot affect the choice of measurement performed on particle 2 at x_2.

Uniqueness: Every (good) measurement has exactly one outcome.

These assumptions are easy to motivate. According to special relativity, there is no fact of the matter about which of x_1 and x_2 occurs earlier in time, so on the plausible assumption that a cause must occur earlier than its effect, it is impossible for a cause at x_1 to produce an effect at x_2. This ensures the truth of Locality and Independence. And Uniqueness just seems to express a truism about good measurements: if your measurement somehow results in more than one of a set of distinct outcomes, then it wasn't a good measurement! Given these assumptions, Bell's theorem follows: when the particles are prepared in an entangled state, no assignment of properties to the individual particles can generate the correlations we observe between measurement outcomes.

But despite the plausibility of the premises, there are reasons independent of the emergence debate to think that the conclusion of Bell's theorem is unacceptable. The trouble is that we don't know how to explain the correlations we observe *except* by appeal to properties of the individual particles. Holistic properties of the pair of particles don't help in this regard. The entangled state can be regarded as a holistic property of the pair of particles. Such a property entails conditionals such as "if the outcome of a spin measurement on particle 1 in some direction is spin-up, then the outcome of a spin measurement on particle 2 in the same direction is spin-down". But it doesn't entail any *unconditional* measurement outcomes for either particle; it doesn't explain why the outcome of the spin-measurement on particle 1 was spin-up rather than spin-down (say). To accomplish the latter, we need an intrinsic property of particle 1. So even if we avail ourselves of holistic properties, accepting the conclusion of Bell's theorem leaves the outcomes of measurements unexplained.

Given the unacceptability of Bell's conclusion, many interpreters of quantum mechanics, including Bell himself, prefer to reject one of his premises. This restores the possibility of explaining measurement results on entangled states in terms of the properties of the individual particles, but it also undermines the direct argument for emergence. The surprising thing, perhaps, is that even though the major interpretations of quantum mechanics violate one of Bell's assumptions, they do not thereby avoid the need to postulate emergent properties. Let us briefly see why.[4]

Consider first the Locality assumption. If it fails, then it is straightforward to arrange that the properties of the individual particles explain the measurement results we observe: a measurement on particle 1 at x_1 changes the properties of particle 2 at x_2, thereby bringing about a correlation between the measurement results we obtain for particle 1 and particle 2. Two of the leading interpretations of quantum mechanics, Bohm's theory (Bohm 1952) and the GRW theory (Ghirardi, Rimini and Weber 1986), violate the Locality assumption. Hence, one might think, they allow for an explanation of the results of spin-measurements on entangled particles that does not appeal to holistic properties of the pair.

But in fact this thought is short-lived: a causal theory that exploits nonlocal influences to explain entanglement correlations is hard to come by. The reason has to do with the particular nature of entanglement. In order to explain the correlations exhibited by entangled particles, an intervention on one particle (e.g. a measurement) has to affect the *particular* distant particle with which it is entangled. But law-like causal influences don't work this way: they affect any particle of a particular *type*. For example, consider Newtonian gravitation: the motion of one massive body instantaneously affects the motion of every other massive body, no matter how distant. In order for a causal influence to affect precisely one distant particle, we need to equip

[4]See Lewis (2016) for more details.

our theory with an irreducible relation between the two particles concerned. That is, non-local theories like Bohm and GRW can't do without emergence.

In each case, the way this emergence is instantiated is that the entangled quantum state is retained as descriptive of the two-particle system: the entangled state represents a holistic property linking the two particles together. Then in addition, the theory proposes some new causal machinery connecting this holistic property to the results we observe. In the case of Bohm's theory, the new machinery is a law that dictates how the quantum state "steers" the positions of the particles. The law is non-local, in that the velocity of each particle depends on the quantum state evaluated at the position of the other particle—so a measurement on one particle can instantaneously affect the motion of the other. And the dependence of the motion of each particle on the entangled quantum state shows how a holistic property enforces the correlations we observe between this particular pair of particles: if one particle is steered towards a spin-up result, then the holistic connection between them means that the other is steered towards a spin-down result.

In the case of the GRW theory, the extra machinery is a law that causes occasional collapses of the quantum state. Again, this law is non-local: a collapse triggered by a measurement on one particle instantaneously affects the state of the other. And again, the entangled state of the pair of particles is essential in explaining why the measurement outcomes for the two particles are correlated: entanglement means that a collapse centered on a spin-up result for the first particle is also a collapse centered on a spin-down result for the second. So in both GRW and Bohm's theory, emergent properties play an essential role, and the violation of Locality does not threaten the need for emergence.

Now consider the Uniqueness assumption. If it fails, then we don't need a special intrinsic property of particle 1 to explain why the outcome of a spin-measurement on it was spin-up rather than spin-down: instead, we can say that both outcomes are equally instantiated, and say that the particular result you see is a consequence your place in the branching structure of reality. This is the many-worlds approach (Everett 1957; Wallace 2012). Again, one might think that the violation of Uniqueness allows for an explanation of the results of measurements on entangled pairs that does not appeal to holistic properties of the pair.

But again the thought is short-lived, and for the same reason: although the many-worlds approach obviates the need for intrinsic properties of each particle, it still requires a special link between precisely this pair of particles to explain the correlations observed in each branch of reality. If the spins of the particles are measured in the same direction, then the branches in which particle 1 is spin-up are also branches in which particle 2 is spin-down. Again, the many-world theory retains the entangled quantum state as descriptive of the two-particle system, representing a holistic property linking the two-particle system together.[5] The many-worlds approach does not add any additional causal machinery, but instead interprets the quantum state as describing a branching reality. Even so, the holistic properties are necessary to provide the requisite structure in the branches to produce the observed correlations between spin measurement outcomes within each branch.

So all three of the major research programs in the interpretation of quantum mechanics embody emergence. In this sense, perhaps, it doesn't matter whether there is a direct argument for emergence from the formalism of quantum mechanics itself. Our best theories of the quantum world all exhibit emergence, and that is enough.

5 The case against emergence

But first a word of caution. We have not considered violations of the Independence assumption. This oversight can be justified: there is no fully-developed interpretation of quantum mechanics that succeeds via violating Independence. But there is an ongoing research program in this direction. The trick is to find a way that the properties of one particle can affect the choice of measurement on the other. Since the measurement can be chosen *freely*, you might think that any such theory would threaten free will, or at least require a kind of global physical conspiracy (Lewis 2006).

But Price (1994) suggests that instead one can regard the free choice of measurement on the second particle as causing the properties of the first. Since the measurement can be chosen after the two particles have been produced (and to avoid the need for non-local causal influences),[6] this means that the causal influence involved runs backwards in time. This is the *retro-*

[5] How that holistic property is instantiated in the world is an interesting question. See Wallace and Timpson (2010) for a concrete proposal.

[6] That is, according to the retrocausal approach, all causal influences propagate along timelike curves, in either the forward or the reverse temporal direction. So there is no need for direct causal influences between spacelike separated locations, such as appear in Bohm and GRW.

causal approach to quantum mechanics.

The promise of this approach is that it can yield a coherent interpretation of quantum mechanics without the problematic non-locality of Bohm and GRW, and without the problematic branching reality of many-worlds. The reason that non-locality in Bohm and GRW is considered problematic is that it involves a prima facie conflict between quantum mechanics and special relativity. The primary reason that many-worlds branching is considered problematic is that it is hard to square the probabilistic predictions of quantum mechanics with a theory in which every outcome occurs (on some branch of reality). Much has been written to try to mitigate these problems, but still, it would be nice to be able to avoid them altogether.

How does the retrocausal approach affect the argument for emergence? The important feature of a retrocausal theory in this regard is that there is no need to postulate a special holistic property connecting the two particles involved in the entangled state. There is a direct casual chain connecting the measurement on the second particle to the properties of the first, mediated by the second particle itself. That is, just as a particle can carry traces of earlier measurements performed on it, so in a retrocausal theory a particle can carry traces of later measurements performed on it. Entangled particles are typically produced at a common source, so the second particle can carry these traces to this common source, and thereby affect the first particle.[7] The fact that the particles themselves carry the causal influences means that the special connection between these two particles is built in to the causal story via their common origin; there is no need to postulate a holistic property of the pair to do the job.

Put another way, the retrocausal approach opens up the possibility of a genuinely epistemic interpretation of the quantum state. The three major interpretive strategies all take the quantum state as playing a representational role, and when the state is entangled, what is represented is an irreducible joint property of the entangled pair. But in the retrocausal approach, it may be possible to regard the quantum state as a reflection of an agent's state of knowledge about a quantum system. When the state is entangled, this just means that the agent's knowledge includes irreducible conditionals—e.g. knowledge that the second particle is spin-down in a given direction *if* the

[7]The phenomenon of entanglement swapping complicates this story a little, as the two entangled particles at the end of the experiment do not have a common souce (Ma et al. 2012). However, retrocausal theories can easily handle such cases too (Price and Wharton 2015).

first particle is spin-up in that direction. But each particle can still carry its own intrinsic spin property, unknown to the agent.

The retrocausal approach vividly illustrates why there is no direct argument from the quantum formalism to the existence of emergence. Since denying the Independence assumption is possible, there is no reason in principle that the correlations exhibited by entangled pairs could not be explained without emergent properties. But it should be stressed that, while this explanatory strategy is perfectly coherent, it cannot at present be embedded within a fully developed retrocausal interpretation of quantum mechanics. The research program is not at a stage yet where it can deliver measurement outcomes that are demonstrably equivalent to those of standard quantum mechanics.[8]

6 Conclusion

Quantum mechanics is the best place to look for emergence, and it is the worst place to look for emergence. It is the best place to look insofar as the phenomenon of entanglement provides measurement outcomes that resist explanation in terms of the intrinsic properties of the particles involved. But it is the worst place to look insofar as constructing explanations in quantum mechanics is strongly interpretation dependent, and the interpretation of quantum mechanics is still a matter of controversy.

Still, we can say this much. First, interpretive strategies that challenge the intuitive division of systems into wholes made up of smaller parts do not undermine arguments for emergence based on quantum mechanics, because the sense of "part" and "whole" appearing in the definition of emergence are the ordinary ones, not the technical notion that might be the result of such a challenge. Second, there is no direct argument from the formalism of quantum mechanics to emergence: the possibility of violating the Independence assumption provides a counterexample to such an argument. But third, there is an indirect argument for emergence, in that all the well-developed interpretations of quantum mechanics involve emergent properties as a crucial part of their explanation of measurement outcomes for entangled systems.

[8] Retrocausal research programs include the transactional interpretation of Cramer (1986) and Kastner (2012), the two-vector formalism of Aharonov and Vaidman (1990), the retrocausal Bohmian approach of Sutherland (2008), the classical wave approach of Wharton (2010), and the particle-based approach of Price (2012).

In this sense, quantum mechanics gives us good reason to think that the world contains emergent properties.

Acknowledgements

I would like to thank the participants in the Durham Emergence Project and an anonymous referee for useful comments on earlier incarnations of this paper.

References

Aharonov, Yakir, and Lev Vaidman (1990), "Properties of a quantum system during the time interval between two measurements," *Physical Review A* 41: 11–20.

Albert, David Z. (1996), "Elementary Quantum Metaphysics", in J. Cushing, A. Fine, and S. Goldstein (eds.), *Bohmian Mechanics and Quantum Theory: An Appraisal.* Dordrecht: Kluwer, 277–284.

Albert, David Z. (2013), "Wave function realism", in Ney and Albert (2013), 52–57.

Bell, John S. (1964), "On the Einstein-Podolsky-Rosen Paradox," *Physics* 1: 195–200. Reprinted in Bell (2004), 14–21.

Bell, John S. (2004), *Speakable and Unspeakable in Quantum Mechanics, Second Edition.* Cambridge: Cambridge University Press.

Bohm, David (1952), "A suggested interpretation of the quantum theory in terms of" hidden" variables, parts I and II," *Physical Review* 85: 166–193.

Bohn, Einar D. (2012). Monism, emergence, and plural logic. Erkenntnis, 76, 211–223.

Calosi, Claudio (2014). Quantum mechanics and priority monism. Synthese, 191, 915–928.

Chalmers, David J. (2006), "Strong and weak emergence," in P. Davies & P. Clayton (eds.), *The Re-Emergence of Emergence*. Oxford: Oxford University Press, 244–256.

Cramer, John G. (1986), "The transactional interpretation of quantum mechanics," *Reviews of Modern Physics* 58: 647–687.

Einstein, Albert, Boris Podolsky, and Nathan Rosen (1935), "Can quantum-mechanical description of physical reality be considered complete?," *Physical Review* 47: 777–780.

Everett III, Hugh (1957), ""Relative state" formulation of quantum mechanics," *Reviews of Modern Physics* 29: 454–462.

Ghirardi, Gian Carlo, Alberto Rimini, and Tullio Weber (1986), "Unified dynamics for microscopic and macroscopic systems," *Physical Review D* 34: 470–491.

Hawthorne, James, and Michael Silberstein (1995), "For whom the Bell arguments toll," *Synthese* 102: 99–138.

Healey, Richard (2016), "Holism and nonseparability in physics," *Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2016/entries/physics-holism/>.

Kastner, Ruth E. (2012), *The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility*. Cambridge: Cambridge University Press.

Lancaster, Tom, and Mark Pexton (2015), "Reduction and emergence in the fractional quantum Hall state," *Studies in History and Philosophy of Modern Physics* 52: 343–357.

Lewis, Peter J. (2006), "Conspiracy theories of quantum mechanics," *British Journal for the Philosophy of Science* 57: 359–381.

Lewis, Peter J. (2016), *Quantum Ontology: A Guide to the Metaphysics of Quantum Mechanics*. Oxford: Oxford University Press.

Loewer, Barry (1996), "Humean supervenience", *Philosophical Topics* 24: 101–127.

Ma, Xiao-Song, Stefan Zotter, Johannes Kofler, Rupert Ursin, Thomas Jennewein, Časlav Brukner, and Anton Zeilinger (2012), "Experimental delayed-choice entanglement swapping," *Nature Physics* 8: 479–484.

Ney, Alyssa and David Z. Albert (2013), *The Wave Function*. Oxford: Oxford University Press.

Price, Huw (1994), "A neglected route to realism about quantum mechanics," *Mind* 103: 303–336.

Price, Huw (2012), "Does time-symmetry imply retrocausality? How the quantum world says "Maybe"?," *Studies in History and Philosophy of Modern Physics* 43: 75–83.

Price, Huw, and Ken Wharton (2015), "A live alternative to quantum spooks," arXiv preprint arXiv:1510.06712.

Schaffer, Jonathan (2010), "Monism: The priority of the whole", *Philosophical Review* 119: 31–76.

Silberstein, Michael, and John McGeever (1999), "The search for ontological emergence," *Philosophical Quarterly* 49: 182–200.

Sutherland, Roderick I. (2008), "Causally symmetric Bohm model," *Studies in History and Philosophy of Modern Physics* 39: 782–805.

Teller, Paul (1986), "Relational holism and quantum mechanics," *British Journal for the Philosophy of Science* 37: 71–81.

Wallace, David (2012), *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation*. Oxford: Oxford University Press.

Wallace, David, and Christopher G. Timpson (2010), "Quantum mechanics on spacetime I: Spacetime state realism," *British Journal for the Philosophy of Science* 61: 697–727.

Wharton, Ken (2010), "Time-symmetric boundary conditions and quantum foundations," *Symmetry* 2: 272–283.