

A prolegomenon to the ontology of the Everett interpretation

David Wallace

September 19, 2011

This article began as an attempt to explain, without assuming any technical physics or advanced mathematics, what the metaphysical (and in particular, the ontological) implications are of the Everett interpretation of quantum theory, often known as the many-worlds interpretation.

The difficulty is that physics is not, by and large, *gratuitously* technical. It is possible to give a non-technical impression of what Everettian quantum theory is like, but metaphysicians don't just need impressions: they need clear statements of just what a given physical theory is saying.

In my view — and this is an underlying theme of the article — this has led to trouble. Physical theories don't, straightforwardly, *say* anything: the theory, shorn of any interpretation, is a piece of mathematics. In the case of classical point-particle mechanics — that is, in the case of Newton's physics — that piece of mathematics is relatively easy to understand intuitively, at least at a surface level. But in the case of quantum theory, the mathematics is very abstract and its connection to observable facts is both indirect and controversial. There is a great temptation, in trying to communicate quantum physics to philosophers, to tacitly smuggle in controversial interpretative posits as if they were part of the formalism; there is a comparable temptation to present certain, relatively-simple-describable special cases as if they were general; there is, in short, a serious danger of miscommunication, and this is exacerbated by the very different styles of philosophy and modern physics, a difference not entirely recognised by philosophers whose model of physics remains something like Newton's Laws.

In this article, I do not pretend to present the mathematics of quantum theory more carefully, in a way that avoids any such communications failure; such a task lies far beyond its scope. My goal, instead, is to give some insight into the features of quantum mechanics, mathematically and interpretationally, that must be understood before any metaphysical discussion of the Everett interpretation can get off the ground. Much of what has been said so far on the subject, I believe, has failed to fully understand those features; as such, it has begun the discussion on the wrong foot. By the end of the article, I will have begun a positive discussion of the Everett interpretation's metaphysical consequences, but I can do no more than begin it without presuming a good

deal of technical physics. (As it is, I try to avoid assuming anything more complicated than complex numbers.)

The structure of the article is as follows. In section 1, I explain the concept of *state space*, a commonplace idea in physics since at least the 19th century but one which will be unfamiliar to most philosophers (though it has a family resemblance to the set of possible worlds). In section 2 I apply this discussion to quantum mechanics, and in doing so provide a perhaps-unfamiliar perspective on the quantum measurement problem. In section 3 I outline the Everett interpretation and its claims about higher-level ontology: about chairs, tables and the like. In sections 4–5 I consider the question of fundamental ontology in the Everett interpretation, first exhibiting what I think are the main misconceptions available, and then making some provisional comments on what positive story can be told.

1 The concept of state space

Perhaps the most familiar theory in physics is the theory of Newtonian point particles. Physically, what the theory is about is the movement of objects through space, where those objects are small enough relative to the distances between them that they can be treated as pointlike. (The mechanics of the planets in the solar system is the most important practical example). Mathematically, we represent the instantaneous state of the system of particles by a collection of points in three-dimensional space.

It's perfectly possible to do the mathematics of Newtonian mechanics in terms of this collection of points, but more advanced treatments of mechanics instead represent the instantaneous state of an N -particle system not by N points in three-dimensional space but by one point in $3N$ -dimensional space - with the first three coordinates encoding the position of particle 1, the second three encoding the position of particle 2, and so forth. This $3N$ -dimensional space is called the *configuration space* of the system — so-called because every point in it represents a possible instantaneous configuration of the system. And so it's possible to reexpress the equations of Newtonian physics — the equations for how those N particles move through space — as equations governing the movement of a single point through configuration space.

This is just one example of a general move made throughout physics, and more generally in the study of dynamical systems: the introduction of a *state space*, where each point in the space represents the state of the entire system being studied. A curve in such a space represents an entire history of the system (one point for each moment of time¹). A great many such state spaces are used in physics and elsewhere in science (indeed, multiple sorts of phase space are used even in classical mechanics). Not all of them regard the points in the state space as representing possible physical states of a system; sometimes, a point in

¹Mathematically speaking, this means that a history is really a parametrised curve, or path; equivalently, it's a map from whatever space represents the instants of time into the state space.

the state space needs to be understood as representing the various *probabilities* (be they chance or credence) that the actual physical system in question has of being in a given state.

It is important to realise that, mathematically speaking, state spaces are highly structured: they are not mere sets of points. *Some* of that structure is specified by the very general nature of the dynamical theory under study: the configuration spaces of *any* classical-mechanical theory, for instance, must be so-called “differential manifolds”. But *most* of the mathematical structure is fixed by the specific details of the system under study. In the theory of N Newtonian point particles, for instance, configuration space cannot be a *homogenous* $3N$ -dimensional space: it has to have enough structure that we can read off the positions of each separate particle from a single point in configuration space. (The mathematical details of this are unimportant: it suffices to observe that in a homogenous $3N$ -dimensional space, no structural features of the space distinguish one point from another.)

Indeed, while a physical system represented in the state-space formalism may have a huge amount of structure, that structure is not encoded at all in the state-space point (which, as a mere point, is entirely structureless). *All* the structure is encoded in the location of the point in state space, which in turn is only possible because of the highly structured nature *of* state space.

Metaphysically speaking, the state space concept may seem of little relevance: the metaphysics of a theory, presumably, is to be understood in terms of the actual properties and relations holding between the objects that make up the world according to the theory, and state space is just an abstract mathematical tool. Certainly in Newtonian mechanics (and, arguably, in classical mechanics more generally), this seems correct: N -particle point mechanics is a theory of N particles moving in homogenous three-dimensional space, not a theory of one particle moving in a highly-inhomogenous, very-high-dimensional space, and the fact that we can represent those N particles via the state-space formalism is mathematically useful but not metaphysically germane. Unfortunately, in quantum mechanics things are not so straightforward.

2 Quantum mechanics and the measurement problem

Quantum theory, unlike classical mechanics, is usually first presented in its state-space form. The particular state spaces used for quantum physics are called *Hilbert spaces*,² and they come equipped with a deterministic dynamics: through every point in Hilbert space there is exactly one dynamically permissible curve. Hilbert space also has a crucial property called *linearity*: if ψ and ϕ are

²Strictly speaking, the true state space is the so-called *projective Hilbert space*, not the full Hilbert space. In addition, the state space described here is only the space of so-called “pure states”. It can be argued — indeed, Chris Timpson and I argue (Wallace and Timpson 2010) — that the larger space of “mixed states” is more properly understood as the true state space. For the purposes of this article, though, such details just add unneeded complexity.

points in the state space, then it makes sense to talk about the state $\alpha\psi + \beta\phi$, where α and β are complex numbers satisfying $|\alpha|^2 + |\beta|^2 = 1$. Furthermore, the dynamics respects this linearity: if ψ evolves after t seconds into $\psi(t)$, and ϕ evolves after t seconds into $\phi(t)$, then $\alpha\psi + \beta\phi$ evolves after t seconds into $\alpha\psi(t) + \beta\phi(t)$.

One might expect that to understand the relation between quantum mechanics and physical reality (and in particular, to understand the way in which quantum mechanics has empirical significance), we would need to understand what physical goings-on are represented by each point in Hilbert space. But in fact this is not how the relation between quantum mechanics and the world is normally presented. Instead, textbooks generally take an explicitly instrumentalist line: to every quantum system is associated a set of possible *measurements* that can be carried out on it, some, though not all, of which are designated as measurements of classical quantities like position and momentum. Given any such measurement, each state in the Hilbert space uniquely determines a probability distribution over the possible outcomes of that measurement, which is interpreted as the probability that a given measurement, if actually performed, will produce a given outcome.

That this rule provides empirically adequate predictions is not in question: quantum theory is by far the most thoroughly empirically confirmed theory in science. What *is* in question is how to go beyond this rather mysterious *algorithm*, and understand quantum theory as a description of physical reality: as David Lewis puts it, the challenge is to “see how [quantum mechanics] looks when it is purged of instrumentalist frivolity and dares to say something not just about pointer readings but about the constitution of the world” (Lewis 1986, p.xi). This is (one way of describing) the infamous *quantum measurement problem*: to solve the problem is to either make sense of unmodified quantum mechanics as a physical theory in good philosophical standing, or to replace it with some equally-empirically-adequate theory in good standing. (Were this article aimed at a physics audience I would digress for some time as to just why any solution — any move beyond instrumentalist ‘frivolity’ — is needed; metaphysically-inclined philosophers, I take it, need no such digression.)

To see the severity of the problem (and, in doing so, to get clearer on the form of the measurement algorithm), let’s attempt to make sense of quantum-mechanical state space in the same way we do for classical-mechanical state space. Consider (say) the state space of a single electron. There are *some* states — so-called “wave-packet states” — such that the measurement rule predicts that the state will almost certainly have some value of position in some particular narrow range. Let ψ_x denote such a wave-packet state, one for which a position measurement will almost certainly give a result close to x . Suppose, then — waving aside any qualms about “almost certainly” and “close to x ” — that what ψ_x represents, physically, is a particle located at position x . And similarly, we might suppose, ψ_y represents a particle located at position y .

Here’s the problem. What does $\alpha\psi_x + \beta\psi_y$ represent? By the linearity principle, it too is a possible state of the electron system. Applying the measurement

algorithm to the state tells us³ that it has probability very close to $|\alpha|^2$ of being found to have position close to x , and probability very close to $|\beta|^2$ of being found to have position close to y . So $\alpha\psi_x + \beta\psi_y$ doesn't seem to represent an electron located at any particular place in space.

Well, since one thing we *do* know about $\alpha\psi_x + \beta\psi_y$ is that measuring the position of a system it represents sometimes gives x and sometimes gives y , its tempting to interpret the state space as a space of probability distributions, not as a space of actual physical states (as noted above, in some other contexts in science, that's how state space works). In other words (on this approach) to say "the state of the system is $\alpha\psi_x + \beta\psi_y$ " is to say something like "the electron has probability $|\alpha|^2$ of being at position x and probability $|\beta|^2$ of being at position y . (This was how Einstein hoped that quantum theory would turn out.⁴)

This probability strategy turns out not to be viable, though, because of the quantum-mechanical phenomenon called *interference*. Suppose, for instance that the dynamics of the system are such that ψ_x evolves after 5 seconds into $(1/\sqrt{2})(\psi_x + \psi_y)$, and ψ_y evolves into $(1/\sqrt{2})(\psi_x - \psi_y)$. (This is a perfectly possible form for the dynamics of real systems.) In that case, if we delay our position measurement by 5 seconds, the measurement algorithm tells us that, whether the system was initially in state ψ_x or in state ψ_y , there's a 50% probability of getting a position measurement of about x , and a 50% probability of getting a position measurement of about y . ($|1/\sqrt{2}|^2 = 0.5$.)

If that's the case, and if $\alpha\psi_x + \beta\psi_y$ represents a system that is definitely in either position x or position y , we'd predict that this state, too, leads to a 50% probability of each position when measured after 5 seconds. But in fact, if we apply linearity, we find that the *state* after 5 seconds is $(1/\sqrt{2})[(\alpha + \beta)\psi_x + (\alpha - \beta)\psi_y]$. If $\alpha = \beta$, for instance, the state after 5 seconds is just ψ_x , i.e. a position measurement after 5 seconds will almost certainly give result x . In physics jargon, there is interference between the ψ_x and ψ_y parts of the state, so that the x outcome is *reinforced* and the y outcome is *cancelled out*; interference phenomena like this are very general, and rule out the possibility of a probabilistic interpretation of the state space.

The only possibility left seems to be that $\alpha\psi_x + \beta\psi_y$ does represent something physical. But (a) it's not clear what — an electron in two places at once, or somehow in both places or neither, seems the only real possibility available, but it's not at all clear what that means; and (b) if electrons can be in two places at once, how is this compatible with the measurement algorithm, which says that we definitely *find* it in one place or another?

We can push this point further once we remember that measurement devices, too, are physical systems, made out of microscopic matter that is supposedly subject to the principles of quantum mechanics. As such, the process of measurement *itself* ought to be represented physically.

To see how that works, let's suppose we have a measurement device represented by a pointer, that can be in three states: pointing left, pointing right,

³Since I haven't actually stated how the measurement algorithm works, just take this on trust.

⁴See Einstein, Podolsky, and Rosen (1935).

and pointing nowhere. And suppose the measurement is set up so that if the electron is measured in position x the pointer moves so that it points left, and if it is measured in position y , it moves so that it points right. We can certainly find a state space suitable for such a pointer, and indeed can find wave-packet states ϕ_L (for the pointer pointing left), ϕ_R (for it pointing right), and ϕ_0 (for it pointing nowhere). The idea of these states, as with the electron, is that ϕ_L (say) is a state such that, if we measure where the pointer is — with the naked eye, or otherwise — we're pretty much guaranteed to get the result that it's in the pointing-left position.

Given state spaces for the electron and for the pointer, quantum theory gives us a recipe to construct a state space (the so-called “tensor product space”) for the combined system of electron-plus pointer. If ϕ is any state for the electron alone, and ψ any state of the pointer alone, there is then a combined state $\phi \otimes \psi$ of both together, which gives the same experimental predictions as ϕ for measurements of the electron and the same experimental predictions as ψ for measurements of the pointer.

If the measurement device works as intended, the dynamics of measurement must look something like this:

$$\begin{aligned}\psi_x \otimes \phi_0 &\longrightarrow \psi_x \otimes \phi_L \\ \psi_y \otimes \phi_0 &\longrightarrow \psi_y \otimes \phi_R.\end{aligned}$$

In other words, if the electron starts off in a state such that its position is always found to be x , the pointer must reliably end up in a state such that *its* position is always found to be on the left (and similarly for y). But now, the linearity of the dynamics causes trouble: what if we measure the electron's position when it is in the mysterious state $\alpha\psi_x + \beta\psi_y$? The dynamics in this case *have* to give

$$(\alpha\psi_x + \beta\psi_y) \otimes \phi_0 \longrightarrow \alpha\psi_x \otimes \phi_L + \beta\psi_y \otimes \phi_R.$$

So now there seems to be a contradiction between our measurement *algorithm* and the actual physical process of measurement. The algorithm tells us that the measurement should give x a fraction $|\alpha|^2$ of the time and y the rest of the time, and hence that the pointer should point left a fraction $|\alpha|^2$ of the time and right the rest of the time. But the actual physical process never gives ‘left’ or ‘right’ as pointer states at all, and is not indeterministic at all: instead, it deterministically gives the strange, indefinite state $\alpha\psi_x \otimes \phi_L + \beta\psi_y \otimes \phi_R$, in which the pointer seems to be pointing left and pointing right at the same time.

Notice that at *this* stage, the strategy of interpreting the state space as a space of *probabilities* looks very attractive. If what it *meant* for the state to be $\alpha\psi_x \otimes \phi_L + \beta\psi_y \otimes \phi_R$ was that the pointer had probability $|\alpha|^2$ of pointing left and probability $|\beta|^2$ of pointing right, such states would be unmysterious. Furthermore, as a practical matter (it turns out) the interference phenomena which previously prevented us from interpreting states probabilistically do not occur at the level of macroscopic states: the dynamical process called *decoherence* (the

details of which lie beyond this article⁵ guarantees that the effects of interference are invisibly weak for macroscopic systems like pointers on measurement devices.

To sum up: state spaces in non-quantum physics represent either (i) possible physical states of a system, or (ii) possible probability distributions over physical states of a system. At the microscopic level, interference means that it's impossible to adopt (ii), and (i) seems the only option. But at the macroscopic level, (i) seems flatly contradictory to our observations, and (ii) seems a much better fit — and furthermore, the interference phenomena that rule out (ii) at the microscopic level are undetectable at the macroscopic level because of decoherence.

What physicists do *in practice* is adopt (i) for microscopic systems like electrons and atoms, and adopt (ii) for macroscopic systems like measurement devices. The distinction between microscopic and macroscopic is rough-and-ready: basically, systems (for these purposes) count as “macroscopic” exactly when decoherence means that interference is undetectably small.

Pragmatically speaking, physics gets by fine with this dual interpretation of the state space; conceptually and philosophically speaking, though, it is profoundly unsatisfactory. Sometimes (these days, usually only in introductory textbooks) one talks about the “projection postulate” or the “collapse of the quantum state”, a mysterious, indeterministic physical process that is supposed to take over from the normal (linear) quantum-mechanical dynamics whenever a measurement is made, and so to effect a transition from (i) to (ii); but this is really just a way of applying sticking plaster to the problem, and essentially no-one regards it as a principled solution.⁶

More principled solutions to the measurement problem (here I restrict my attention to solutions which avoid instrumentalism and leave classical logic alone) by and large hold on to the idea that the quantum state space represents the state of something physical, essentially because interference leaves them no choice. There are then essentially three available moves:

1. The hidden-variable move: “quantum mechanics is not everything”.⁷ Points in quantum state space do represent the state of *something*, but they do not (at least, not alone) represent the complete state of the physical system under study. Additional information is required to do that, and so the state space of quantum mechanics needs to be supplemented with a different state space, the space of states of so-called “hidden variables”. Normally, points in this state space encode the positions of the point par-

⁵For those details, see Zurek (1991) or Halliwell (2010) for an introduction, Joos *et al* (2003) or Schlosshauer (2007) for a systematic study, and Bacciagaluppi (2005), Wallace (2008, pp.22–29) and Wallace (2011, chapter 3) for philosophical considerations.

⁶Philosophers sometimes call this the “Dirac-von-Neumann” interpretation of quantum mechanics. Physicists often call it “the Copenhagen interpretation”, but they use that term for at least two other approaches to quantum mechanics.

⁷Here and afterwards I borrow from (and adapt) J.S.Bell’s famous observation that “[e]ither the [quantum state together with the quantum dynamics] is not everything, or it is not right” (Bell 1987a).

ticles out of which macroscopic objects are assumed to be made.⁸

2. The dynamical-collapse move: “quantum mechanics is not right”. Points in quantum state space do represent the complete state of the physical system under study, but the linear dynamics are not actually correct, at least at the macroscopic level. Instead, they must be modified so that macroscopically indefinite states like $\alpha\psi_x \otimes \phi_L + \beta\psi_y \otimes \phi_R$ do not actually occur (or, if they do occur, rapidly “collapse” into definite states like $\psi_x \otimes \phi_L$).
3. The Everettian move: “quantum mechanics is everything, and it is right”. Points in quantum state space do represent the complete state of the physical system under study, and the linear dynamics *are* correct. Macroscopically indefinite states like $\alpha\psi_x \otimes \phi_L + \beta\psi_y \otimes \phi_R$ are physically reasonable after all, and should be understood as describing a *multiplicity*: a situation in which there are two pointer (or two sets of pointers), one pointing left and one pointing right, and with each dynamically isolated from the other.

My concern in this article, of course, is exclusively with the Everettian move.

3 The Everett interpretation

The immediate question one asks about the Everett interpretation — why do we only see one pointer, if actually there are two? — can be resolved by remembering that you too, dear reader, are a physical system, and if χ_L and χ_R are, respectively, states in your state space representing you seeing a pointer pointing left and you seeing it pointing right, then the same linearity argument used above predicts that the state of (you-plus-pointer-plus-electron), once you look at the pointer, will be

$$\alpha\psi_x \otimes \phi_L \otimes \chi_L + \beta\psi_y \otimes \phi_R \otimes \chi_R .$$

In other words, *you* will be in a state of seeing left and seeing right at the same time, and this state (according to the Everett interpretation) should also be understood as telling us that there are two yous, one seeing the pointer pointing left and one seeing it pointing right.

Notice — crucially — that although the state above is the sum of two macroscopically very different state, in each term in the sum the results of the two measurements are correlated (in each term the electron has a particular position, the pointer records it as having that position, and you observe the pointer as so recording it.)

Once a system gets above a certain size, it cannot help being measured constantly — by chance collisions with the atmosphere and with sunlight, if by

⁸For technical reasons (originating, for the most part, in relativistic quantum physics), this assumption is actually pretty questionable in modern physics.

nothing else. In doing so, the multiplicity spreads to more and more systems, while the correlations in each term in the state remain. In due course, the state (schematically) evolves into something like

$$\alpha(\text{Whole planet is as if electron was found in position } x) \\ + \beta(\text{Whole planet is as if electron was found in position } y).$$

When this, too, is understood as representing both states of affairs simultaneously, the “many-worlds” label for the Everett interpretation starts to sound apposite.

Of the three strategies given above, philosophers of physics — by and large — have strongly tended to prefer the first two. Physicists, on the other hand — by and large — show a general preference for the Everett interpretation (though at least as many reject the idea of a *realist* solution to the measurement problem altogether)⁹. Nor is it hard to see why physicists adopt this attitude. As I have been at pains to stress, we specify the mathematical formalism of quantum physics by (a) providing a state space, and (b) providing a dynamics for points in that state space. Both the hidden-variable and dynamical-collapse strategies requires us to change one or both of these, to replace quantum theory with a new physical theory, one way or another, and physicists are generally loath to replace successful scientific theories for purely philosophical reasons, especially where — as is the case for both hidden-variable and dynamical-collapse theories — no replacement is at present available.¹⁰

The Everett interpretation, on the other hand, is a *pure* interpretation of quantum mechanics. It leaves the quantum formalism, dynamics and state space alike, completely alone. The problems with it, if any, are entirely philosophical. From the point of view of this article, we might enumerate them as follows:

1. **The problem of microscopic ontology:** What are the physical properties of the possible states of a system represented by the various points of quantum-mechanical state space?
2. **The problem of macroscopic ontology:** By virtue of what is it the case that the “wave-packet” quantum states mentioned above — the states

⁹To be clear: I am not making the (relatively strong) claim that physicists prefer Everett to the various anti-realist or quasi-realist strategies on the table (Copenhagen, or physics-as-information, or some kind of operationalism, etc.). I am making the weaker claim that there is relatively little support for change-the-formalism strategies in mainstream physics, and plenty for the Everett interpretation. You can see this in practice just by looking at the physics literature: hardly any of it explores change-the-formalism strategies, and much of it uses Everettian language in more-or-less explicit ways. But I base it as much on conversations with colleagues, at conferences, etc. (It’s somewhat variable by community: string theorists and quantum cosmologists are mostly Everettians; quantum information people are mostly neo-operationalists.)

¹⁰There are versions of both theories that are adequate for *non-relativistic* physics, and there have been some encouraging signs in recent years (Dürr *et al* 2004, Struyve and Westman 2007; Colin and Struyve 2007), that at least *phenomenologically* successful versions of hidden-variable theories might be constructible in the relativistic regime. Less progress has been made for dynamical-collapse theories, though Tumulka (2006) makes some interesting first steps.

that physicists use to represent macroscopic systems with approximately determinate positions — should actually be interpreted as the states of classically-behaving macroscopic objects?

3. **The problem of macroscopic multiplicity:** By virtue of what is it the case (given (2)) that for macroscopic systems, *sums* of wave-packet states should be interpreted as representing *multiple, non-interacting* classically-behaving macroscopic objects?
4. **The problem of probability:** By virtue of what is it that processes, like measurement, which (given (3)) *objectively* cause the macroscopic world to split into multiple copies, are treated scientifically as *probabilistic* processes?

Problems 2-4 are unavoidably tied up in questions of *emergence*, of just how facts about large-scale systems like tables and chairs (and measurement devices) supervene on facts about microphysics.¹¹ Here the Everett interpretation may seem to fail at the first hurdle, for there is, I think, virtually no chance of understanding macroscopic objects in quantum mechanics as some kind of mereological agglomeration of small components. However, this view of emergence is (in my view) highly unattractive even setting aside quantum mechanics, bearing as it does almost no resemblance to the way emergence actually functions in science (in, say, zoology, or fluid dynamics, or psychology, or ...). There, typically, one finds that emergence is all about *structure*: theory A is emergent from theory B in some circumstance when the structural features of the world as described by theory A are instantiated in the structure of the world according to theory B. Or, at the level of objects rather than theories: higher-level objects are patterns in lower-level ones.

Granted this perspective on emergence, an answer to (2) becomes available. Wave-packet states represent classical objects in motion because their dynamical behaviour instantiates the dynamical behaviour of those classical objects; in other words, the structural features of the classical world which justifies our talk of “chairs” and “tables” and “pointers” are in fact represented — to a high degree of accuracy — by structural features of the dynamics of the wave-packet states. As for (3), it follows from (again) the dynamical process of decoherence that where the quantum state of a macroscopic system is a sum of wavepackets, the structural features of the macroscopic system are just the collection of all the structural features of the individually-evolving wave-packet states considered separately. So if a single wave-packet has the right structure to instantiate a classical system, a sum of such wave-packets has the right structure to instantiate a collection of independent systems. Let that “system” be something the size of a planet or solar system (or, indeed, of the Universe as a whole, in principle) and that “collection of independent systems” becomes a collection of independently evolving macroscopic worlds.

This also gets us part-way towards a solution to the problem of probability. While at the *fundamental* level the state space and dynamics of quantum theory

¹¹Here I draw heavily on my own earlier work; cf Wallace (2003, 2010a).

are those of a deterministic theory, at the *emergent* level, quantum mechanics has the structure of an indeterministic theory. Roughly speaking, what this means is that to every path (representing a possible history of the system) in the emergent state space, quantum theory assigns a number between 0 and 1, and those numbers satisfy the probability calculus: at the level of mathematical formalism (and skipping some technicalities), this is just what it *is* for a mathematical theory to be a theory of indeterministic, probabilistic dynamics.

The remaining problem in (4) is again philosophical: the numbers which quantum mechanics assigns to (emergent) histories are mathematically *suitable* for representing probabilities, but *do* they represent probabilities? The question really lies beyond this article.¹² Suffice it to say that the contemporary literature is largely concerned with two strategies: (i) that it is justifiable for it to be a *posit* of the theory that the numbers do represent probability (or at least, that it is no less justifiable than the equivalent posit in classical mechanics); (ii) that considerations from symmetry and decision theory allow us to deduce that rational agents in a quantum universe would act *as if* the numbers represent probability. (An optional extra move is to claim that if rational agents act *as if* the numbers represent probability, that's all there is to the claim that they *do* represent it.)

None of these problems, however, address the question of the *fundamental* ontology of quantum mechanics. Furthermore, and quite apart from the intrinsic interest of the question, it might be thought that we cannot be confident in any story of emergence unless we are confident what it is emerging *from*. (Maudlin (2010), in particular, criticises the Everett interpretation for having an inappropriate micro-ontology to appropriately ground macro-level facts; Hawthorne (2010) raises some similar concerns.) In particular, normally our concepts of space and time are treated as constant between higher-level and lower-level theories, so that for (e.g.) some higher-level object to exist in spacetime region K it must be instantiated not just by any old objects and properties in the lower-level theory, but by objects and properties themselves located in K . As such, getting some understanding of the relation between spacetime and the microscopic ontology might well be crucial for the larger Everettian project. This project will be my concern for the remainder of the article.

4 Microscopic ontology: problematic strategies

Both in informal conversation and in published writings, several approaches to thinking about the micro-ontology of (Everett-interpreted) quantum mechanics frequently recur. Most of these approaches, I think, are seriously flawed, due mostly to misconceptions about quantum theory rather than internal conceptual problems. In this section I identify and criticise three such approaches.

¹²For extensive discussion, see Saunders *et al* (2010), Greaves (2007), Wallace (2011), and references therein.

The eigenvalue-eigenstate approach

Philosophical discussions of quantum theory often refer to something called the “eigenstate-eigenvalue link”, or E-E link for short.¹³ In the language of this article, the E-E link says that a quantum state possesses a given property if and only if, according to the measurement algorithm (which, recall, associates to each quantum state a probability distribution over measurement outcomes), a measurement to detect that property would have 100% probability of so detecting it. If a system has nontrivial probability for each of several outcomes of a measurement of a given quantity, it is said, according to the E-E link, to possess an *indefinite value* of that quantity. (Monton (2006) explicitly constructs an ontology of quantum physics based on the E-E link.)

If the E-E link were correct, it would cause problems for the ontology of the Everett interpretation. In particular, no realistic macroscopic system (in the absence of wavefunction collapse) ever has a quantum state which would give any single macroscopic outcome 100% probability. So macroscopic systems would have indefinite values of ordinary quantities like position, and the onus would be on the Everettian to explain why we should re-interpret indefiniteness as multiplicity. Indeed, this might seem to undermine the previous section’s discussion of emergent multiplicity: there aren’t two pointers, there’s just a pointer with an indefinite position. (Albert’s “bare theory” (Albert 1992) explicitly considers the Everett interpretation from this perspective; see also chapter 4 of Barrett (1999).)

Actually, though, none of this is relevant to the Everett interpretation, because the E-E link has nothing much to do with quantum theory. One might be forgiven, if one learned one’s quantum mechanics exclusively from philosophy discussions, for thinking that the E-E link was part of quantum mechanics itself, but it isn’t. The measurement *algorithm* is a core part of the practice of quantum mechanics, but the E-E link isn’t part of that algorithm. It’s an interpretative assumption. Its motivation (I think) comes from the idea that measurement must be discovering some pre-existing measured value, in which case that value must be possessed by a system iff it is certain to give that value as a result of measurement. But this isn’t realised in any realist interpretation of quantum mechanics — not hidden-variable theories, not dynamical-collapse theories, not the Everett interpretation. It plays no part in the actual practice of physics (the term occurs just once in a search of Physical Review over the last century).¹⁴ And it is anyway incompatible with the actual physics of quantities with continuously many measurement outcome possibilities, like position and momentum.¹⁵

¹³For recent examples, see Albert and Loewer (1996), Barrett (1999), Lewis (2003), and Dickson (2007).

¹⁴In fact, the very framework used in modern physics to define the measurement algorithm — the “positive operator-valued measure” framework — does not even allow the E-E link to be satisfactorily defined. It only makes sense if we use the somewhat outdated “projection-valued measure” framework.

¹⁵On a technical level, this is because no particle’s wavefunction will remain localised in any spatial region smaller than all of space for any nonzero length of time, unless unphysical

For that reason, the E-E link has no particular significance for the Everett interpretation; no more, in my view, does it have any significance for quantum mechanics in general.

Reifying state space

Some discussions of the ontology of quantum theory seem to lose track of the nature of state space. Sometimes (largely, to be fair, in conversation) one hears talk of quantum mechanics as describing the evolution of systems in Hilbert space, just as classical mechanics describes the evolution of systems in ordinary three-dimensional space. But this fails to appreciate that Hilbert space is the *state* space of quantum mechanics. To reify it — to treat it as a space of physical points, just as substantialists do about ordinary space — is no more justified than to reify the configuration space of classical mechanics, and to regard classical mechanics as about the evolution of a point in $3N$ -dimensional space.¹⁶

Very few people are willing to defend Hilbert-space realism in print. There is, however, considerable discussion of — and some support for — a more sophisticated view that, I think, suffers from a subtler form of the same fallacy. In other writing (Wallace and Timpson 2010) Chris Timpson and I have called this view *wavefunction realism*; we could equally call it *configuration space realism*.¹⁷ To understand the view, it is necessary to recognise that in many (but not all) quantum systems, the quantum theory of that system is obtained by starting with a *classical* theory and then “quantizing” it. This quantizing process is mathematically at most heuristically defined, and has no real conceptual justification, but it has proved highly useful as a way to construct new quantum theories. Roughly speaking, the quantum state space of a quantized theory can be represented as the space of *complex functions* on the configuration space of the classical theory, so that for a quantum system of N point particles, for instance, the quantized theory’s state space can be represented as a space of complex functions on the $3N$ -dimensional configuration space of that theory. (I say “can be represented as” rather than “is” advisedly: this is only one way to represent the structure of the quantum state space, and mathematically speaking there is no particular reason to prefer it over other representations).

Wavefunction realism involves the reification of the configuration space as a physical space, so the quantum system is a complex field on this very high-

infinite potentials are used.

¹⁶An extreme proponent of structural realism might argue that it is not *wrong* to so regard classical mechanics, since the state-space description is isomorphic to the more conventional description. I confess to a certain weakness for this extreme view; however, even its proponents would acknowledge that it is highly *unhelpful* to think of classical mechanics in this way. (Indeed, more generally, the kind of methods that a conventional realist uses to ascertain the *true* ontology of a theory will be pretty much the same as those used by a structural realist to ascertain the most *perspicuous* way to think about that theory.)

¹⁷Tim Maudlin (2010) calls it *wavefunction monism*, though he treats it as essentially synonymous with the view that the state space of unaugmented quantum theory is the correct state space for a quantum system.

dimensional space. This has led Albert (1996) to claim that according to quantum mechanics, it is an illusion that space is 3-dimensional; really, it's $3N$ -dimensional, where N is the number of particles in the universe. This view of the ontology of (no-hidden-variable) quantum mechanics has probably been the most commonly assumed in the recent literature (authors who tacitly or explicitly assume it, other than Albert, have included Maudlin (2010), Hawthorne (2010), Lewis (2004) (though see also Lewis, this volume), and Monton (2006)).

For technical reasons (largely to do with relativistic quantum field theory; cf Wallace and Timpson (2010)), wavefunction realism is not really viable (or not without major changes) as an ontology of our universe. More importantly, though, it makes the same unmotivated *conceptual* move as Hilbert-space realism: it reifies a mathematical space without any particular justification. Recall that in classical physics, configuration space is not a physical space at all: it is a space of possible classical states, and as such is very high-dimensional and very non-homogenous. It does not play quite that role in quantum theory: states are represented not as points in configuration space but as functions over it, and some of their structure is manifest in the structure of the function, not implicit in the structure of the space. Nonetheless, it still has a considerable role in encoding state structure: wave-packet states, in particular, are featureless blobs in the configuration-space representation of state space, and the structural differences between, say, a wave-packet state representing a uniform gas of particles and a wave-packet state representing those particles arranged into a sculpture is encoded mathematically *entirely* by the position of the respective blobs in configuration space.

As such, configuration space is not a natural candidate for reification. A strong positive case for wavefunction realism would be required if that position is to be taken seriously as a candidate for quantum micro-ontology. To the best of my knowledge, such a case has nowhere been presented; the position is adopted, so far as I can see, largely on the erroneous assumption that quantum theory forces it on us.

Many worlds at the micro level

Multiplicity, in the Everett interpretation, is an emergent, high-level notion. The theory is a “many-worlds” theory in the same sense that modern astrophysics is a “many-stars” theory: in both cases, the objects being multiplied are not represented in the fundamental structure of the theory. Indeed, there appears to be nothing in quantum mechanics to say that a quantum state represents a universe which *fundamentally* consists of multiple classical components. (The very idea of classicality is emergent and high-level.)

However, popular accounts of the many-worlds theory — e.g., Deutsch (1997) — do sometimes talk as if the multiplicity was fundamental rather than emergent. Furthermore, the configuration space representation of state space *seems* to support the idea: don't each of the points in configuration space represent a classical world?

However, the idea does not work in practice. The configuration space points

themselves have no dynamics; they are simply points. The dynamics of the theory is encoded entirely in the distribution of complex numbers to the points: it is that distribution, not the points themselves, which evolve over time in accordance with the dynamical equations. Furthermore, as stressed previously, the configuration space representation is just one way to represent the quantum state space, and has no special status.

This might suggest that additional dynamics, and additional rules, need to be added (the rules to pick out the configuration representation — or another representation — as preferred; the dynamics to say how each individual world evolves). At one point, the philosophical literature on the Everett interpretation was quite focussed on this project (see Barrett (1999) and references therein for more details). But all such strategies abandon the idea of the Everett interpretation as a pure interpretation of quantum mechanics; instead, they become proposals for modifying the theory, and regarded as such, do not seem competitive with less ontologically expensive strategies. This kind of approach, therefore, never really found favour in physics; in philosophy, it has been moribund for some while. Contemporary defences of the Everett interpretation, almost exclusively,¹⁸ restrict multiplicity to the emergent level.

(I should note, however, that there is one way in which a *sort* of multiplicity does occur at the fundamental level in quantum theory: it occurs in the so-called “sum-over-histories” formalism developed by Feynman. There has been little exploration of this formalism in the philosophy literature (Arntzenius (2010) is an exception) and I have no observations to make here, except that Feynman’s histories are *not*, mathematically speaking, the same as Everett’s branches. The topic — though technical — might well repay further study.)

5 Microscopic ontology: positive observations

I have said much in the previous section about how *not* to understand the ontology of quantum mechanics, but little positive; I will finish this article with some provisional thoughts. To begin, recall why the Everett interpretation has a problem of microscopic ontology in the first place. The reason goes back to my original discussion of state space. Both classical and quantum theories can be formulated in state-space terms; however, in the classical case the state-space formalism is constructed some way into the process, once the actual metaphysical content of a “classical state” is already understood. In the quantum case, by contrast, the state space is given *first*, with even the question of *whether* points in that space represent physical states of the system being controversial.

Once this is recognised, though, it can be seen that specifying a micro-ontology for quantum mechanics *in general* makes no more sense than doing so for classical mechanics in general. In classical mechanics, points in state space can represent (depending on the classical system in question) the length of a

¹⁸The only recent exception of which I am aware is Allori *et al* (2009), though it is not entirely clear to me whether they actually advocate the position they present.

spring, or the positions of some particles in physical space, or the relative distances between some particles without any substantival concept of space, or the strength of some field at every point in space, or even the shape of space itself. The only things in common between these different classical-mechanical theories are some rather general, abstract features of their structure and dynamics. Similarly, it's not clear to me that *anything* very substantive can be said about the ontology of quantum theory in general, over and above similarly-general, similarly-abstract points.

So: what *can* be said at such an abstract level? The main distinctive feature of quantum theory in general is the relation between subsystems and the overall system. In classical physics, if systems A and B have state spaces S_A and S_B , the state space of the combined system $A + B$ is just the Cartesian product of S_A and S_B : that is, the set of ordered pairs of states of A and states of B . It follows that (i) any state of the combined system uniquely specifies states of the component systems; (ii) all there is to the combined system being in a given state is that its subsystems are in the appropriate states.

In quantum theory, the state space of $A + B$ is the *tensor*, rather than Cartesian, product of the separate state spaces. The mathematical details aren't crucial; what *is* crucial is that both (i) and (ii) fail. This can be seen directly by looking again at the electron-plus-pointer system when its state is

$$\alpha\psi_x \otimes \phi_L + \beta\psi_y \otimes \phi_R.$$

It's fairly clear that there are no states χ, δ of electron and pointer separately such that this state can be written as $\chi \otimes \delta$. In the terminology of quantum mechanics, the two systems are *entangled*. It is really entanglement, rather than linearity and interference, that gives rise to the weirdnesses of quantum mechanics (small waves on a pond have linearity and interference!) because it is entanglement that causes microscopic indefiniteness to be magnified up to the macro level during the measurement process.

To say anything general about the metaphysics of quantum theory, then, we have to say something about the metaphysical status of entanglement. Some provisional attempts have been made; Healey (1991, 1994), for instance, regards entanglement as a form of holism, and explores exactly what kind (partly through analogies with holism elsewhere in physics); cf also the extensive list of references in Healey (2009). So far, however, this discussion does not seem to have made much contact with the recent ontological discussions of quantum theory in the metaphysics literature.

But to say anything that goes beyond this generality, we need to look at specific quantum theories. From the metaphysician's perspective, the natural choice is *quantum field theory*, the general framework for our current most-successful quantum theories and (in particular) for the theory that underpins the Standard Model of particle physics.¹⁹ In quantum field theory (unlike earlier versions of

¹⁹It may seem odd that a *field* theory underpins *particle* physics. Details would take us too far afield; suffice it to say that (i) particles, too, are emergent entities in modern physics; (ii) the popular impression of particle physics as about the behaviour of lots of little point particles

quantum theory), ordinary classical space-time is explicitly represented in the formalism. Indeed (speaking loosely²⁰) the theory works by assigning a space of quantum states to every point of spacetime. It's pretty clear what that state space represents: the state of that spacetime point. But it's also necessary to remember that because of entanglement, the list of properties of each spacetime point does not remotely exhaust the list of properties of the theory as a whole. Chris Timpson and I Wallace and Timpson (2010) regard this as a major failure of Lewis's doctrine of Humean supervenience, the doctrine that all facts about the world supervene on monadic properties of spacetime points and the spacetime relations between them: in our view, the entanglement between (say) spacetime regions *A* and *B* should be understood precisely as encoding certain irreducible relations between *A* and *B*. Simon Saunders (1995, 1996, 1997) puts it differently (though compatibly): as he sees quantum field theory, spacetime events stand not only in spatial and temporal relations with one another, but in a different kind of relation which he calls 'modal' (though its relation to metaphysical modality is unclear.)

But from this point on, the devil is in the details, and it becomes impossible to discuss the topic further without delving deep into the mathematical structure of quantum theory in general and quantum field theory in particular. As this article has aimed to avoid such technical details, I take my leave here.

Acknowledgements

Thanks to Chris Timpson for helpful comments on this paper.

References

- Albert, D. Z. (1992). *Quantum mechanics and experience*. Cambridge, Massachusetts: Harvard University Press.
- Albert, D. Z. (1996). Elementary quantum metaphysics. In J. T. Cushing, A. Fine, and S. Goldstein (Eds.), *Bohmian Mechanics and Quantum Theory: An Appraisal*, Dordrecht, pp. 277–284. Kluwer Academic Publishers.
- Albert, D. Z. and B. Loewer (1996). Tails of Schrödinger's Cat. In R. Clifton (Ed.), *Perspectives on Quantum Reality*, pp. 81–92. Dordrecht: Kluwer Academic Publishers.
- Allori, V., S. Goldstein, R. Tumulka, and N. Zanghi (2009). Many-worlds and Schrödinger's first quantum theory. Forthcoming in *British Journal for the Philosophy of Science*; available online at <http://arxiv.org/abs/0903.2211>.
- Arntzenius, F. (2010). Quantum mechanics, narratability and relativity. Forthcoming.

whizzing about bears about as much relation to real particle physics as the earth/air/fire/water theory of matter bears to the Periodic Table.

²⁰To be less loose would also be to court controversy; cf Wallace (2010b).

- Bacciagaluppi, G. (2005). The role of decoherence in quantum mechanics. In the Stanford Encyclopedia of Philosophy (Summer 2005 edition), Edward N. Zalta (ed.), available online at <http://plato.stanford.edu/archives/sum2005/entries/qm-decoherence>.
- Barrett, J. A. (1999). *The quantum mechanics of minds and worlds*. Oxford: Oxford University Press.
- Bell, J. S. (1987a). Are there quantum jumps? In C. W. Kilmister (Ed.), *Schrödinger: Centenary of a Polymath*, Cambridge. Cambridge University Press. Reprinted in Bell (1987b), pp. 201–212.
- Bell, J. S. (1987b). *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.
- Colin, S. and W. Struyve (2007). A Dirac sea pilot-wave model for quantum field theory. *Journal of Physics A* 40, 7309–7342.
- Deutsch, D. (1997). *The Fabric of Reality*. London: Penguin.
- Deutsch, D. and P. Hayden (2000). Information flow in entangled quantum systems. *Proceedings of the Royal Society of London A* 456, 1759–1774.
- Dickson, M. (2007). Non-relativistic quantum mechanics. In *Handbook of the Philosophy of Science: Philosophy of Physics, Part A*, pp. 275–417. Amsterdam: Elsevier.
- Dürr, D., S. Goldstein, R. Tumulka, and N. Zanghi (2004). Bohmian mechanics and quantum field theory. *Physical Review Letters* 93, 090402.
- Einstein, A., B. Podolsky, and N. Rosen (1935). Can quantum-mechanical description of reality be considered complete? *Physical Review* 47, 777–780.
- Greaves, H. (2007). Probability in the Everett interpretation. *Philosophy Compass* 38, 120–152.
- Halliwell, J. J. (2010). Macroscopic superpositions, decoherent histories and the emergence of hydrodynamic behaviour. In S. Saunders, J. Barrett, A. Kent, and D. Wallace (Eds.), *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford: Oxford University Press.
- Hawthorne, J. (2010). A metaphysician looks at the Everett Interpretation. See Saunders, Barrett, Kent, and Wallace (2010), pp. 144–153.
- Healey, R. (1991). Holism and nonseparability. *Journal of Philosophy* 88, 393–421.
- Healey, R. (1994). Nonseparable processes and causal explanation. *Studies in History and Philosophy of Science* 25, 337–374.
- Healey, R. (2009). Holism and Nonseparability in Physics. In Edward N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy (Spring 2009 edition)*. Available online at <http://plato.stanford.edu/archives/spr2009/entries/physics-holism/>.

- Joos, E., H. D. Zeh, C. Kiefer, D. Giulini, J. Kupsch, and I. O. Stamatescu (2003). *Decoherence and the Appearance of a Classical World in Quantum Theory* (2nd Edition ed.). Berlin: Springer.
- Lewis, D. (1986). *Philosophical Papers, Vol. II*. Oxford: Oxford University Press.
- Lewis, P. J. (2003). Counting marbles: Reply to critics. *British Journal for the Philosophy of Science* 54, 165–170.
- Lewis, P. J. (2004). Life in configuration space. *British Journal for the Philosophy of Science* 55, 713–729.
- Maudlin, T. (2010). Can the world be only wavefunction? In S. Saunders, J. Barrett, A. Kent, and D. Wallace (Eds.), *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford: Oxford University Press.
- Monton, B. (2006). Quantum mechanics and 3N-dimensional space. *Philosophy of Science* 73, 778–789.
- Saunders, S. (1995). Time, Decoherence and Quantum Mechanics. *Synthese* 102, 235–266.
- Saunders, S. (1996). Time, Quantum Mechanics and Tense. *Synthese* 107, 19–53.
- Saunders, S. (1997). Naturalizing Metaphysics. *The Monist* 80(1), 44–69.
- Saunders, S., J. Barrett, A. Kent, and D. Wallace (Eds.) (2010). *Many Worlds? Everett, Quantum Theory, and Reality*, Oxford. Oxford University Press.
- Schlosshauer, M. (2007). *Decoherence and the Quantum-to-Classical Transition*. Berlin: Springer.
- Struyve, W. and H. Westman (2007). A minimalist pilot-wave model for quantum electrodynamics. *Proceedings of the Royal Society of London A* 463, 3115–3129.
- Tumulka, R. (2006). Collapse and relativity. In A. Bassi, T. Weber, and N. Zanghi (Eds.), *Quantum Mechanics: Are There Quantum Jumps? and on the Present Status of Quantum Mechanics*, pp. 340. American Institute of Physics Conference Proceedings. Available online at <http://arxiv.org/abs/quant-ph/0602208>.
- Wallace, D. (2003). Everett and Structure. *Studies in the History and Philosophy of Modern Physics* 34, 87–105.
- Wallace, D. (2008). The interpretation of quantum mechanics. In D. Rickles (Ed.), *The Ashgate Companion to Contemporary Philosophy of Physics*, pp. 197–261. Burlington, VT: Ashgate.
- Wallace, D. (2010a). Decoherence and ontology: Or: How I learned to stop worrying and love FAPP. In S. Saunders, J. Barrett, A. Kent, and D. Wallace (Eds.), *Many Worlds? Everett, Quantum Theory, and Reality*, pp. 53–72. Oxford: Oxford University Press.

- Wallace, D. (2010b). Taking particle physics seriously: a critique of the algebraic approach to quantum field theory. Forthcoming.
- Wallace, D. (2011). *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation*. Oxford: Oxford University Press.
- Wallace, D. and C. Timpson (2007). Non-locality and gauge freedom in Deutsch and Haydens formulation of quantum mechanics. *Foundations of Physics* 37(6), 951–955.
- Wallace, D. and C. Timpson (2010). Quantum mechanics on spacetime I: Spacetime state realism. *British Journal for the Philosophy of Science* 61, 697–727.
- Zurek, W. H. (1991). Decoherence and the transition from quantum to classical. *Physics Today* 43, 36–44. Revised version available online at <http://arxiv.org/abs/quant-ph/0306072>.