Why the Many-Worlds Interpretation of quantum mechanics needs more than Hilbert space structure

Meir Hemmo†, Orly Shenker‡
In: Rik Peels, Jeroen de Ridder, René van Woudenberg (eds.),
Scientific Challenges to Common Sense Philosophy.
Oxford: Routledge. Pages 61-70.

Abstract

McQueen and Vaidman argue that the Many Worlds Interpretation (MWI) of quantum mechanics provides local causal explanations of the outcomes of experiments in our experience that is due to the total effect of all the worlds together. We show that although the explanation is local in one world, it requires a causal influence that travels *across* different worlds. We further argue that in the MWI the local nature of our experience is not derivable from the Hilbert space structure, but has to be added to it as an independent postulate. This is due to what we call the factorisation-symmetry and basis-symmetry of Hilbert space.

Keywords

basis-symmetry of Hilbert space; branching structure; causation; causation across worlds; decoherence; experience; locality; locality of interactions; many-worlds; factorisation-symmetry of Hilbert space; preferred basis problem; quantum state; Schrödinger's equation; stability (of memories); survival advantage

Introduction

In their contribution to this volume, McQueen and Vaidman argue that common sense requires that explanations in physics be not only causal,[1] but also *local* (they give some necessary conditions for what counts as 'local'; see end of section 3 in their paper). Their main claim is that the Many-Worlds Interpretation (MWI) of quantum mechanics (originally due to Everett ([1957]) provides local explanations of the outcomes of experiments that in other interpretations of quantum mechanics seem to require (some sort of) non-locality. In this sense, they argue, the MWI restores common sense to quantum mechanics.

The causal role of the worlds

We accept here the necessary conditions assumed by McQueen and Vaidman on what counts as local; and we grant (for the sake of the argument) their position that fundamental physics describes causal processes. To see what is at stake here, consider, for example, what happens in the experiment of the *nested* Mach-Zender Interferometer (MZI) (Figure 5 in McQueen and

† Philosophy Department, University of Haifa; meir@research.haifa.ac.il
‡ The Hebrew University of Jerusalem; orly.shenker@mail.huji.ac.il
[1] McQueen and Vaidman's argument is meant to be independent of any specific view about causation; see (Ben-Menahem [2018]) for the linkage between causation and locality (and other related concepts, e.g. determinism).

Vaidman). They admit that the locality of the explanation is restored *not* by looking at what happens in one or another 'parallel' world, but rather in *all* parallel worlds taken together. But what does this exactly mean? In the MWI it turns out that physical facts in *one* world depend not only on whatever happens in *that* world, but rather on interactions *literally* occurring in *other* parallel worlds, so that causation might be spatio-temporally local, but only in virtue of these other-worldly interactions. Let us see how this idea plays out.

Take, for example, the world in which the particle is detected by D2 (in the nested MZI, see Figures 5 and 10 in McQueen and Vaidman). Call this world, *our* world. The explanation McQueen and Vaidman suggest of what causes D2 to click in our world is that there is a continuous trace of the motion of the particle in *our* world from the origin of the experiment along path C up to D2. However, in our world there is also a trace in the inner box, which cannot be explained locally by the motion of our particle through the inner box, since no trace in our world leads to our particle passing through the inner box (see McQueen and Vaidman, Figure 5). Here the intuition is that a *local* explanation of the *trace* in the inner box (which appears as strong as the trace along path C) based on physical matters of fact in our world requires that there are traces of the particle in *our* world that lead to the inner box. But there are no such traces in our world. According to McQueen and Vaidman the physical facts that explain this inner trace occur in *another* world. In that other world, there is a continuous trace of the motion of a copy of our particle from the origin through the inner box up to detector D3 (see Figure 10c). And their point about local causation is that this motion of the copy particle in the *other* world leaves a trace in the inner box *also* in our world.

So the situation is this: (i) a *copy* of our particle is detected by (a copy of detector) D3 in some *other* world, but not in ours, and therefore this copy particle might be said to exist in that world; (ii) this copy particle leaves a continuous trace in the *other* world from the origin of the experiment via the (copy) inner box up to D3 (we shall omit from now on the term 'copy' unless it is needed); (iii) our particle does not leave a trace that leads to the inner box, so if *its* motion were the cause of the trace we see in the inner box, this trace would be created either by some sort of a nonlocal influence in our world (from path C to the inner box), or else by some sort of a nonlocal motion in which our particle travels along *both* path C and the inner box (without leaving a trace that leads to or from the inner box; (iv) to avoid this sort of nonlocality, the claim (on this proposal) is that it is the *copy* particle in the *other* world that causes (or creates, or what have you) the trace in the inner box also in *our* world.

We grant that in the MWI, one might take the trace in the inner box in *our* world to be some sort of a 'photograph' of the *other* 'parallel' world, where in the 'photograph' we have a 'record' of a segment of the trace left locally by the copy particle in the (copy) inner box of the other world. But we don't see that this local behavior in the other world leads to some sort of a *local* picture of how things play out causally *in our world*, even on this way of looking at the trace (that is, as a 'photograph' in our world of segments of the motion of the copy particle in the other world), since also on this way of thinking we have a cross-world causation by which the 'record' of the trace is formed. Same points (*mutatis mutandis*) arise with respect to the other experiments described by McQueen and Vaidman.

The upshot is that particles in the MWI leave traces *both* in the world where they exist, but also in *other* worlds, in which *they don't exist*. So: the explanation is said to be local, since spacetime splits together with the particles, so that there is no influence at space-like separation *within* a world, but the causal influence literally *travels across 'parallel' worlds*. This seems to us to stretch the concept of locality beyond common sense, if not beyond breaking point, even if one understands causation weakly in terms of counterfactual dependence rather than by straightforward physical interaction.

This leads to the more general question of how precisely one should understand the concept of 'worlds' in the MWI, which is *indispensable*, but quite tricky. In this context it seems to us that an old problem that has been mounted against the MWI[2] (called the 'preferred basis problem') which is believed by many to have been solved by decoherence[3] is still open for reasons that have escaped the literature up to now. We shall sketch the argument here, which is based on (Hemmo and Shenker [2019a]).

Consequences of the preferred basis problem in the MWI

The problem of the preferred basis stems from the mathematical fact that the quantum state is symmetric (or invariant) under the (infinitely many) choices of basis of Hilbert space in which it can be written. By this we mean that given the Hilbert space structure, a choice of basis in which the quantum state is described makes no difference with respect to the physical state and the facts that obtain in the universe when it is in this state, it makes no difference with respect to the time evolution of the quantum state, and it makes no difference with respect to the predictions of future facts. Moreover, the standard description of local interactions (as well as the decoherence interaction) presupposes a *factorisation* of the set of all degrees of freedom of the universe into subsets (which are the subsystems), for example, a measured system (say the spin + position of an electron in a Stern-Gerlach device), a photographic screen, an observer, and environment. This standard facorisation is intuitive and justified by our experience. However, there are theoretically other factorisations: for example, into: the electron; the left hemisphere of the observer's brain + one cubic meter of air molecules in the laboratory; and the rest of the degrees of freedom of the universe. We call this the *factorisation-symmetry* of Hilbert space by which we mean the following: There are many (possibly an infinite number of) factorisations of the universal Hilbert space into sets of degrees of freedom (or subsystems), such that given the quantum state of the universe, all the factorisations are on equal footing; in other words, there are no facts determining a preferred factorisation. But our experience corresponds (by and large) to the standard factorisation, and in addition also to certain local states of macroscopic systems given the standard factorisation. In this sense the standard factorisation and the local basis of states are preferred, but there is no deeper account of why they are preferred. In particular, the structure of the interactions does not explain this preference, because it presupposes it. When one appeals to the structure of the interactions in the universe, say the decoherence interaction, or the fact that the interactions between macroscopic systems are local, one presupposes the factorisation that features in our experience of the total set of degrees of freedom. This is acceptable, but we should note already at this stage, that it does not explain our experience. In other factorisations, the structure of the interactions between the subsystems, induced by the same total Hamiltonian, is different.

Let us illustrate this idea by the following figure:

[2] There are many versions of the MWI: see Everett's ([1957]) 'relative-state' formulation; and later versions, for example: (DeWitt [1970]; Zeh [1973], [2001]; Deutsch [1985]; Zurek [1993]; Saunders [1995]; Vaidman [1998], [2014]; Wallace [2012]). Our argument applies to all the versions.
[3] For decoherence, see (Zurek [1993]; Joos *et al*. [2003]).

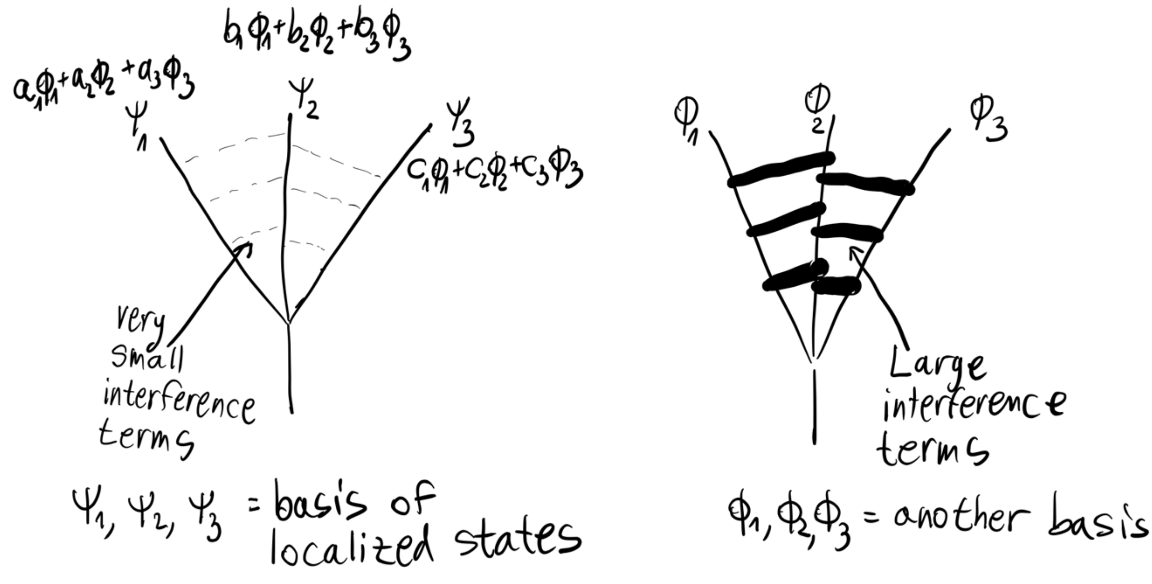

Figure 1: Branching and other structures

On the left side of Figure 1 we depict (very roughly) the branching structure which matches our classical-like experience is described by the psi-basis, which is the basis of localised states, and the *same* structure as it is described by another basis, we call the phi-basis. In these two bases, the interference terms (denoted by the thin black lines) between the branches are small and the interaction Hamiltonian picks out the psi-basis as *dynamically* preferred, in the sense that the interference terms in this basis are small and the states match our experience. Formally, in the phi-description of the branching on the left side we do not add up the similar terms in the different branches. It is a fact that under the time evolution of the universal state the interference between the branches in the psi-basis are very small; and this fact is common to the two descriptions of the branching structure on the left side: in terms of the psi-basis and the phi-basis. By contrast, on the right side of the figure, we depicted a *different* structure that (by the symmetry of bases in Hilbert space) equally exists in the *same* quantum state when it is written in the phi-basis. In this structure (on the left side of the figure) the interference terms (denoted by the thick black lines) between the branches are large.

Vaidman ([2014], [2019]) and McQueen (private correspondence) acknowledge the factorisation-symmetry and the symmetry of bases in Hilbert space as well as the preferred basis problem that follows from these symmetries for the MWI:

> Mathematically, one can decompose the wave function of the universe into a superposition of orthogonal components, not just as in [state (5)], but in many other ways that will not provide a familiar world's picture in every branch. So, critics might say that the proposal is circular: I define by fiat what I want to explain. First, a simple definition that is confirmed by observation sounds to me like a legitimate strategy. But there is also a more specific answer. The basis of the decomposition is indeed preferred. (Vaidman [2019], p. 100)

But the question is: which *facts* make the decomposition corresponding to our experience *preferred*? Or: what makes it the case that our experience is described by *components* of the state in the preferred basis (and factorisation)? According to Vaidman and McQueen the *local* structure of the interactions singles out the decomposition of states such as (5) in terms of the psi-states in which macroscopic systems are in localised states (see the left side of Figure 1). But as we argued above this claim already presupposes our experience, it does not follow from the structure of the universal Hilbert space alone. In measurements, for example, the interaction

Hamiltonian depends on the position of a macroscopic pointer, or the position of ink marks on a piece of paper, or the position of neurons in our brains, etc., and even if one disregards the decoherence interaction with the environment, the position basis, or more generally the expansion of the state in terms of narrowly peaked Gaussians in position are preferred. In our example of Figure 1, the psi-states (corresponding to the branch structure on the left side of the figure) are the localised states which match our experience, whereas the phi-states (depicted on the right side of the figure) are delocalised superpositions of the psi-states and do not match our experience. But what in the Hilbert space structure accounts for this asymmetry between the psi-states and the phi-states? Vaidman ([2019]) argues, in a way that might seem to undermine this point, that the localised states are preferred, because they are stable over time:

> Until now I have not mentioned time evolution. Everything was considered at a particular moment. But we cannot experience anything at zero time. We need an order of 0.1 seconds to identify our experience. Thus, the world needs some finite time to be defined. The world has to be stable, at least on the scale of seconds. Locality of interactions in nature ensures that only the decomposition of wave functions corresponding to well-localized macroscopic objects can be stable. A quantum state describing the superposition of a macroscopic object in separate locations with a particular phase evolves almost immediately into a mixture that has a large component with a different phase. This obvious fact is analyzed in numerous papers using the buzzword 'decoherence.' (Vaidman [2019], p. 100)

This goes along the tradition of Everett's original argument form 1957. But why, for example the world or our experience *has* to be stable, as Vaidman and McQueen require? Of course, as a matter of empirical fact, the world as we experience it, *is* stable. But the MWI should derive this fact from its fundamental postulates and laws, not assume it. Here, as we mentioned earlier, evolutionary arguments to the effect that stability of the preferred states is essential for survival come in.[4] The idea is that since our experience is associated with components of the universal state, the components need be stable over time in order for biological systems to evolve and survive along the branch structure defined by these components. It is true that in biology the standard description of evolutionary survival is in terms of adaptive systems that are immersed in some environment that survive stably over time. But how does this condition become a constraint on fundamental physics which is compatible also with universes in which there are no biological systems at all? After all, whether or not there is experience of our kind that is stable over time, or for that matter, whether or not certain biological kinds survive, need not be a factor that determines whether or not something is real. If it is true, this fact should be added to the Hilbert space structure.

Let us suppose (for the sake of the argument) that stability of *components* of the universal state in some basis is indeed a condition for survival. But given the MWI the universal state *now*, in the present moment, is a state like (5): why do we not experience now the unstable phi-basis in which we are in superpositions of the localised states? In this case, if we grant the evolutionary argument, we would presumably cease to evolve as experiencing agents, and we would not be around to ask questions about our experience. But this is just bad luck for us. Why should the laws of physics care about our luck, or our evolution in the first place? Perhaps it is true, but if so, some additional structure backing this up and breaking the basis-symmetry of Hilbert space is needed, in much the same way that in the standard view about classical statistical mechanics (for example) the past hypothesis is added to break the time-symmetry of the equations of motion and account for the increase of entropy towards the future (see e.g., Feynman [1965]).

Our conclusion applies also to the versions of the MWI which rely on decoherence to define the preferred basis (see Zurek [1993]; Zeh [2001]). One might say that the most obvious justification of choosing the decoherence basis (or the localizes states basis) as preferred (as

[4] Vaidman, McQueen, private correspondence; it seems to us that this is also Zurek's view (1993).

well as the corresponding factorisation) to define the worlds or the branching is that in this basis the on-diagonal elements of the reduced states of the macroscopic systems that appear in our experience can be *directly interpreted as the relative frequencies of the states of these systems in our experience.* Of course, this can be done (and this is what is usually done)! But our point is that this 'obvious' addition requires adding structure to Hilbert space beyond the structure given by the quantum state. That is, one must add structure to Hilbert space that will underlie the facts that make it the case that we have this experience. In particular, Wallace ([2012]) introduces a high-level law (he calls Dennett's criterion) for the emergence of the worlds (or our experience), the role of which is to make some patterns in the quantum state *real*. This high-level criterion is based on functionalist ideas in the foundations of the special sciences. However, in his influential paper on the functional-state hypothesis, Putnam ([1975], p. 436) has already noted that: "the functional-state hypothesis is not incompatible with dualism!" Moreover, it is provable (regardless of quantum mechanics) that if a functionally-defined property is not identical to a micro-physical property (that is, in quantum mechanics, of the quantum state), then the functional-state hypothesis implies that any token of the quantum state from which the functional property emerges must itself have some non-physical property (see Hemmo and Shenker [2019a], [2019b]). So: functionalism is not only compatible with dualism, it *entails* additional *non*-physical structure.

Vaidman and McQueen are aware of the crucial difference with respect to the preferred basis problem and the account of our experience that holds between, on the one hand, the MWI, and on the other hand collapse and hidden variable theories (like, respectively, the GRW theory and Bohm's theory) in which the additional laws that are added to the Hilbert space structure (respectively, GRW collapses or flashes, Bohmian trajectories) account for our familiar macroscopic experience. In the GRW theory the states that we experience are singled out by the flashes or the collapses of the wavefunction, in Bohm's theory by the trajectories in 3D space (or perhaps in 3N space; this is debated in the literature; see Albert and Ney 2013). But in the MWI if one only pre-supposes the Hilbert space structure, there is no account for why in the first place we experience the components of the universal state in the preferred basis rather than in some other (stable or not) basis, or why we do not experience the entire superposition of states like (5) despite the fact that we are in such states and our brain states are superposed in the way depicted in Figure 1. One has to accept that our experience corresponds to the localised psi-states familiar from classical mechanics as a brute fact. It seems to us that Vaidman (2019, p. 98) acknowledges this point when he says:

> In quantum mechanics without collapse we must add a postulate to connect to our experience, because mathematics does not provide a (unique) picture corresponding to what we see around us.

Perhaps for Vaidman and McQueen adding a postulate such as the locality of the interactions in our universe, which as we argued presupposes our experience, is nevertheless more justified than the GRW collapses or Bohm's trajectories. However, it follows from our argument that, contrary to the received wisdom, the MWI is not more parsimonious and therefore it has no advantage over other theories that solve the measurement problem (such as Bohm's ([1952]) theory, or the collapse theory by Ghirardi, Rimini and Weber ([1986]; see Bell [1987]; or the many-minds theory of Albert and Loewer [1988]). All these theories introduce *additional* laws or structure and additional elements of reality over and above the Schrödinger equation for the quantum state, and, as we argued in this paper, the MWI is *no exception* in this regard. They all solve the measurement problem by changing drastically quantum mechanics, for good or for worse. Here, obviously, different questions may come up, such as the compatibility of the extra laws with relativity theory. But Ockham's razor does not cut in favour of the MWI.

Many often reject the MWI on the grounds that the multiplicity of the worlds is extravagant. This does not strike us as a good argument; it seems to us that none of the interpretations of quantum mechanics is common-sensical. While the set of common sense beliefs is not uniquely and sharply delineated and is often given by examples that appear to be psychologically irresistible and intuitively true, each and every interpretation of quantum mechanics is strongly incompatible with some of the most central common sense beliefs. In this sense quantum mechanics in all its interpretations shutters our common sense, if one takes it to be true. The result is that: Naïve realism leads to physics, and physics, if true, shows that naïve realism is false. Therefore, naïve realism, if true, is false; therefore, it is false (Russell, 1940, p. 15). The question arises: how do the common sense beliefs come about in a quantum-mechanical world and what justifies relying on empirical evidence which we understand common-sensically as confirming quantum mechanics to begin with? This question is addressed and answered by Shenker (2020).

Acknowledgement

We thank Lev Vaidman, Kelvin McQueen, Christoph Lehner, and Guy Hetzroni for discussions of the MWI. We also thank an anonymous reviewer for helpful comments. The research on which this paper is based is supported by the Israel Science Foundation (ISF), grant number 1148/2018.

References

Albert, D. and Loewer, B. [1988]: 'Interpreting the Many Worlds interpretation', *Synthese* **77**, pp. 195–213.
Ben-Menahem, Y. [2018]: *Causation in Science*, Princeton University Press.
Deutsch, D. [1985]: 'Quantum Theory as a Universal Physical Theory', *International Journal of Theoretical Physics*, **24**, pp. 1–41.
DeWitt, B. S. [1970]: 'Quantum Mechanics and Reality', *Physics Today*, **23** (9), pp. 30–35.
Everett, H., III [1957]: '`Relative State' Formulation of Quantum Mechanics', *Reviews of Modern Physics*, **29**, pp. 454-462.
Feynman, R. [1965]: *The Character of Physical Law*, Cambridge: MIT Press.
Hemmo, M. and Shenker, O. [2019a]: 'Two Kinds of High-Level Probability," *The Monist* **102**, pp. 458–477.
Hemmo, M. and Shenker, O. [2019b]: 'Why Quantum Mechanics is Not Enough to Set the Framework for Its Interpretation," forthcoming.
Joos, E., H.D. Zeh, D. Giulini, C. Kiefer, J. Kupsch, I. O. Stamatescu [2003]: *Decoherence and the Appearance of a Classical World in Quantum Theory*, Heidelberg: Springer.
McQueen, K. and Vaidman, L. [2019]: 'How the Many Worlds Interpretation Brings Common Sense to Paradoxical Quantum Experiments," this volume.
Putnam, H. [1975]: 'The Nature of Mental States', in Putnam, H., *Mind, Language and Reality*, pp. 429–40, Cambridge: Cambridge University Press [1975]; Originally published as 'Psychological Predicates', in: William H. Capitan and Daniel D. Merrill (Eds.), *Art, Mind and Religion*, pp. 37–48, Pittsburgh, PA: University of Pittsburgh Press (1967).
Russell, B. [1940]: *An Inquiry into Meaning and Truth*, London: Allen and Unwin
Saunders, S. [1995]: 'Time, Decoherence and Quantum Mechanics', *Synthese*, **102**, pp. 235–266.

Shenker, O. [2020] ‘A Scientific Realist Account of Common Sense’, Jeroen de Ridder, Rene van Woudenberg, and Rik Peels (Eds.) *Cambridge Companion to Common Philosophy*, Cambridge University Press.

Vaidman, L. [2014]: ‘Many-Worlds Interpretation of Quantum Mechanics’, *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (Ed.), URL = <https://plato.stanford.edu/archives/fall2018/entries/qm-manyworlds/>.

Vaidman, L. [2019]: ‘Ontology of the Wave Function and the Many-Worlds Interpretation’, in Lombardi, O., Fortin, S., López, C., Holik, F., (Eds.) *Quantum Worlds: Perspectives on the Ontology of Quantum Mechanics*, pp. 93-106, Cambridge: Cambridge University Press.

Wallace, D. M. [2012]: *The Emergent Multiverse*, Oxford: Oxford University Press.

Zeh, D. H. [1973]: ‘Towards a Quantum Theory of Observation’, *Foundations of Physics*, **3**, pp. 109-116.

Zeh, D. [2001]: *The Physical Basis of the Direction of Time* (4th ed.). Berlin: Springer.

Zurek, W.H. [1993]: ‘Preferred States, Predictability, Classicality and the Environment-induced Decoherence’, *Progress in Theoretical Physics*, **89**, pp. 281-312.