The Classical Stance: Dennett's Criterion in Wallacian quantum mechanics

Ruward Mulder*

1st July, 2024

Abstract

David Wallace's 'Dennett's Criterion' plays a key part in establishing realist claims about the existence of a multiverse emerging from the mathematical formalism of quantum physics, even after decoherence is fully appreciated. Although the philosophical preconditions of this criterion are not neutral, they are rarely explicitly addressed conceptually. I tease apart three: (I) a rejection of conceptual bridge laws even in cases of inhomogeneous reduction; (II) a reliance on the pragmatic notion of usefulness to highlight quasi-classical patterns, as seen in a decoherence basis, over others; and (III) a structural realist or 'functional realist' point of view that leads to individuating those patterns as real macroscopic objects at the coarse-grained level, as they are seen from the Classical Stance (analogous to Dennett's Intentional Stance). I conclude that the justification of Dennett's Criterion will be intimately tied up with the fate of strong forms of naturalism, and in particular that Wallacian quantum mechanics is a key case study for concretely evaluating his 'math-first' structural realism (Wallace 2022).

Contents

* *University of Cambridge, Trinity College, CB2 1TQ, Cambridge, UK.* *ram202@cam.ac.uk*.

1. **Introduction**

The pioneering book *The Emergent Multiverse* is arguably the most thoroughly worked-out solution to the measurement problem of quantum mechanics that is presented as a single-authored complete book. In it, David Wallace (2012, p. 2) argues, along Everettian lines, to interpret the wavefunction "[...] literally, straightforwardly – naively, if you will – as a direct description of the physical world, just like any other microphysical theory" and that this by itself leads to a multiplicity of actually existing universes, all equally 'real'. However, a mathematical formalism by itself is silent about ontological commitment: active interpretation of the quantum formalism is required, in the form of some criterion or other, to identify what part of the wavefunction represents reality and what part does not. This paper is meant to make explicit how that active interpretation proceeds in-between the lines in the book.

Indeed, Wallace provides an interpretative criterion which he calls *Dennett's Criterion*, in honour of Daniel Dennett's (1991) "Real Patterns" view of identifying mental states with the part of the lower-level neurophysiology that plays the functional role of psychological mental states—this is mimicked in the current context by identifying classical worlds with patterns in the underlying quantum-mechanical wavefunction. Wallace's Dennett's Criterion is a way to highlight some patterns admitted by a higher-level theory (such as classical mechanics) over all possible patterns in a lower-level framework (such as the quantum formalism); and it does so on the basis of pragmatic virtues[1]:

> **Dennett's Criterion**. A macro-object is a pattern, and the existence of a pattern as a real thing depends on the usefulness – in particular, the explanatory power and predictive reliability – of theories which admit that pattern in their ontology.

The application of Wallace's Dennett's Criterion to Everettian quantum mechanics is then to identify certain patterns within the universal wavefunction with patterns that we know from classical mechanics, and to consider such patterns as more 'real' than other patterns. In short, this pattern-matching occurs at the 'coarse-grained level', a higher-level domain where we consider a large set of composite quantum states, rather than a 'fine-grained' level such as a single particle system. At the coarse-grained level, the preferred patterns are the quasi-classical patterns, which arise in the wavefunction when it is expressed in a decoherence basis. These quasi-classical patterns behave approximately as classical systems (in real space) would behave: they fulfil a 'Classical Role'.

This functionalist emphasis on behaviour, drawing our attention towards dynamics and away from kinematics, even without first specifying what the wavefunction is about (or represents), is refreshing and powerful. However, there remain ontological questions to be

[1] Dennett's Criterion is found, always the same to the letter, in several places (cf. Wallace 2003, p. 93; 2010, p. 58; and finally in the book 2012, p. 50).

addressed: while it is reasonable to consider these quasi-classical patterns as real – at least from the vantage point of pure wave mechanics – should not any other pattern in the universal wavefunction be considered just as real? For *in lieu* of a 'differentiating criterion' there is a Hilbert space democracy of bases: expressing the wavefunction with one or the other set of basis vectors is to be regarded as ontologically equivalent to each other. This goes for a coarse-grained level just as much as for a fine-grained level.

Decoherence, by itself, does not solve the preferred basis problem (see also Hemmo and Shenker 2022). The salient point is that dynamical preference and ontological preference are not automatically the same thing: there is a missing link. It is the philosophical addition of Dennett's Criterion that functions as the missing link and as the above-mentioned differentiating criterion.

Hence my overall thesis: Wallace's underlying view takes non-trivial positions across several philosophical debates, which do necessary work for his proposed solution of the measurement problem, and in particular in establishing the realist claims that there is a multiplicity of real classical worlds, forming a multiverse. I will highlight three preconditions of the application of Dennett's Criterion to quantum mechanics:

(I) **Reduction through instantiation**. Patterns at different levels of grain are approximately matched through instantiation, a relation that goes through purely on a quantitative level, without qualitative bridge laws linking classical and quantum-mechanical terms;

(II) **Pragmatic individuation**. The quasi-classical patterns in a decoherence basis are made salient with respect to other possible patterns in the wavefunction through the pragmatic virtues of explanatory power and predictability;

(III) **The Classical Stance**. Akin in spirit to Daniel Dennett's Intentional Stance, we need to adopt a certain structural realist or 'functional realist' point of view at the coarse-grained level, which implies a privileged macroscopic ontological status for the patterns pragmatically preferred by clause (II) above, if they are structurally equivalent or if they play the Classical Role (which includes being deterministic and definite).

Thus I maintain that this additional philosophical baggage should be made explicit and more precise, so that one can make up one's mind whether the resolution of the measurement problem in the emergent multiverse way is worth this philosophical price.

It would of course be too ambitious to try and *complete* this cost-benefit analysis here. But I stress that such an analysis is important for two recent debates, about 'math-first structural realism' (Wallace 2022, cf. Jacobs 2024) and 'Dennettian functionalism' (Knox and Wallace 2024), and it seems Wallacian quantum mechanics is to be used as a concrete case study for both. I should also add that I intend the preconditions (I)–(III) as necessary philosophical assumptions needed to justify Dennett's Criterion, but *not* as jointly sufficient to deduce Dennett's Criterion.[2] Furthermore, to emphasise different philosophical themes, preconditions (I)–(III) are presented as independent of each other, but conditions (II) and (III) can be collapsed by those of a more pragmatic stripe, by deflating 'ontology' as much as inflating 'pragmatics'.

In §2 some crucial background is reviewed and illustrated, namely the preferred basis problem, the ontology problem and decoherence. In §3 Wallace's Dennett's Criterion is further discussed in this context. The three themes of (I) intertheoretic reduction, (II) pragmatic highlighting, and (III) functionalist stance realism, will occupy §4–§6 respectively. In the concluding §7, I outline a worry about philosophical monism as resulting from a restrictive form of naturalism.

2. Wallacian quantum mechanics

Hugh Everett III (1956) proposed an unexpected non-collapse strategy to solve the measurement problem: erasing the non-unitary dynamics from the formalism and denying our experience that the macroscopic world is indeed free of superpositions. The pioneering step was to apply the quantum formalism to the entire universe, including macroscopic objects. Then our experience only *appears* definite, as our perceptions are contained in one 'branch' of a superposed universal wavefunction: such a branch should be understood as we understand our classical universe, including stars and galaxies and all the other macroscopic objects like measurement devices, mountains, cities, the Earth, and ourselves.[3]

Next to this infamous measurement problem, a related but less often emphasised problem is the so-called *ontology problem*. This is a problem of scientific representation, primarily concerned with the extraction from (or imbuing onto) the formalism some ontology. In our context, that would concern the question what the wavefunction represents, whether as a ray in Hilbert space or a state in the $3N$-dimensional configuration space.[4]

[2] I thank an anonymous reviewer for pressing this clarification.

[3] 'The' quantum measurement problem is notoriously difficult to pin down. As Hugh Everett III saw it in his unedited "long" thesis (Everett III 1956; cf. Barrett and Byrne 2012), it is the postulation of two rules of evolution in von Neumann textbook quantum mechanics: the continuous, unitary and deterministic Schrödinger dynamics and the discontinuous, non-unitary and indeterministic projection that takes over when measuring. Without specification of what is meant by 'measurement' and disentangling it from physical interactions which are themselves described by unitary evolution (the Measurement Meaning Problem, Muller 2023, p. 27), these two dynamical rules are seemingly in conflict (the Measurement Explanation Problem, *ibid.*, p. 25). Note that by removing the non-unitary process and modelling the apparatus as a quantum state itself, we generically obtain a superposition of outcomes, conflicting with the definite outcomes of experience (the Reality Problem of the Classical World, *ibid.*, p. 24), which is in turn claimed to be solved by Everettians by saying that all of these outcomes *actually* obtain as distinct worlds.

[4] The ontology problem is often less emphasised than that third interpretative problem of quantum me-

Everett's theory also suffers from the ontology problem, but his concern was mainly with solving the measurement problem—it is plausible that Everett, contra Wallace, was not aiming to solve the ontology problem.[5] To keep branches stable for a long time, one can formulate a *no-interference criterion*, making sure there is little to no interference between the branches. One way to satisfy the no-interference criterion is to specify a preferred basis. The quantum formulation – read straightforwardly – does not come with such a criterion, as it adheres to *Hilbert space democracy*: one can represent the same physical system in a rotated basis. If in some basis representation the interference does not show up, the same system in a rotated basis generically will. A preferred basis would break Hilbert space democracy. Absence of justification for such a preference is often called the *preferred basis problem*, and a solution to it goes a long way towards solving the ontology problem because it specifies a basis in which the ontology is to be represented best.

In the modern approach, the idea is that the dynamical mechanism of *decoherence* can pick out the preferred basis. Formulated by Heinz-Dieter Zeh (1970), decoherence is the phenomenon where quantum systems lose coherence (become less entangled) through interaction with their environment. Using the decoherence mechanism as a way to underpin Everett's approach became popular from the 1970s onwards—a dynamically preferred *decoherence basis* is seen as a preferred basis, not on a par with other bases (cf. Crull 2021). Regarding the ontology problem, the hope here is this: to equate dynamical preference with ontological preference. Wallace's (2011) claim is that, when the wavefunction is expressed in a decoherence basis, we see autonomous substructures that *behave* similarly to classical objects even though they are not (or not obviously) ontologically equivalent to them, so-called 'quasi-classicality'.

The central aim is obtaining macroscopic definiteness. Take a Schrödinger's cat set-

chanics: the locality problem. Yet, in order for a locality criterion to get off the ground, one needs to know what the ontology is that is behaving non-locally. The most-studied example of this is that Bell locality is in need of the specification of local beables in order to be evaluated (Bell 1987). In Travis Norsen's words: "[i]f, according to a theory, there *are* no physically real objects in ordinary 3-dimensional space, then concepts like "local" and "non-local" are simply, radically, fatally, inapplicable" (Norsen 2017, p. 292). The same problem is encountered in the context of electrodynamics (cf. Maudlin 2018, emphasising the Coulomb gauge; cf. Mulder 2021, pp. 17–21, emphasising the Lorenz gauge): to even begin to give a local explanation of the Aharonov-Bohm effect, one first needs to assess what part of the electromagnetic potentials counts as local beables.

[5] Note that Everett himself was *not* a scientific realist about quantum mechanics. For in his long thesis (Barrett and Byrne 2012), he presented his theory not as claims about reality but as restoring logical consistency, under the constraint of saving the phenomena by means of securing a correct model in the empirical substructure of pure wave mechanics. Jeffrey Barrett (2011) has coined this position *empirical faithfulness*. Everett likely takes such a view despite him dubbing all branches as 'equally real'; he also consistently wrote 'real' in scare quotes. Wallace (2012, p. 2) states that "Everett's unpublished work, however, has recently made it quite clear that he understood the many-worlds implications of his view, and that he refrained from making them clearer essentially for political reasons", citing Peter Byrne. However, Byrne (2012, Chs. 35, 38, 39) makes clear that in Everett's relative state interpretation all the branches are "equally real" independent of the choice of basis; furthermore, for Everett branching is time-symmetric, a "tree both ways" (*ibid.*, p. 315).

up. We have a closed system consisting of an initial (macroscopic) state of a cat, where for times $t < t_M$ the entire state is in the state $|\text{🐈}\rangle$; we also have a two-level (microscopic) system superposed in the spin-basis, with 'up'-state $|\uparrow\rangle$ and 'down'-state $|\downarrow\rangle$. A mechanism is set in motion that interacts with the two-level system, thereby becoming entangled with it, and, via intermediate interactions, eventually becomes entangled with a measuring apparatus, represented by quantum state $|M_i\rangle$ (each i representing a distinguishable pointer state). This triggers the breaking of a flask filled with sevoflurane, so that – due to the spreading of this anaesthetic gas – the cat state will quickly evolve to a 'sleeping cat state' $|\text{💤}\rangle$ after a time t_M. Hence we start with a macroscopically definite state where the cat is awake,

$$(c_\uparrow |\uparrow\rangle + c_\downarrow |\downarrow\rangle) \otimes |M_i\rangle \otimes |\text{🐈}\rangle , \tag{2.1}$$

which after t_M has evolved to

$$c_\uparrow |\uparrow\rangle \otimes |M_\uparrow\rangle \otimes |\text{🐈}\rangle + c_\downarrow |\downarrow\rangle \otimes |M_\downarrow\rangle \otimes |\text{💤}\rangle . \tag{2.2}$$

This state is macroscopically *in*definite, because there are two terms, each representing a different state of macroscopic affairs.

Where Schrödinger used this to tease out a problem with (the completeness of) the quantum formalism, Wallace makes the following turn: what we have here are two bits of formalism that 'give rise to' or 'instantiate' (see §4) the *structures* of both an awake cat and a sleeping cat. So, whereas before there was one superposed state, there are now two macroscopic states, distinct from each other. It seems that "[s]uperposition has become multiplicity at the level of structure" (Wallace 2012, p. 61).

Since in practice the system cannot be closed completely, the cat will also interact with its immediate environment, including us human observers, forming a von Neumann chain. Thus it will eventually become entangled with the entire universe (dropping $\otimes$ for brevity):

$$c_\uparrow |\uparrow\rangle |M_\uparrow\rangle |\text{🐈}\rangle |\text{rest of universe}\rangle + c_\downarrow |\downarrow\rangle |M_\downarrow\rangle |\text{💤}\rangle |\text{rest of universe}\rangle , \tag{2.3}$$

so that after a time t_F sufficiently long for the dynamics to unfold, a 'quasi-duoverse' arises,

$$|\text{quasi-duoverse}\rangle = c_\uparrow |\text{universe with } \uparrow \text{ and } \text{🐈}\rangle + c_\downarrow |\text{universe with } \downarrow \text{ and } \text{💤}\rangle . \tag{2.4}$$

Wallace's point then is that even though the concept of Duoverse does not exist in the 'fundamental' ontology of quantum physics, a real Duoverse can be seen as *emerging from* the quantum formalism.

This interpretation should at first sight raise a suspicion concerning the preferred basis problem, namely that in Hilbert space one has the freedom to rotate the basis vectors as one pleases, creating other branching structures. In one such rotated basis, we have

$$|A\rangle = |\uparrow\rangle\,|M_\uparrow\rangle\,|\text{awake cat}\rangle + |\downarrow\rangle\,|M_\downarrow\rangle\,|\text{sleeping cat}\rangle \tag{2.5}$$

and

$$|B\rangle = |\uparrow\rangle\,|M_\uparrow\rangle\,|\text{awake cat}\rangle - |\downarrow\rangle\,|M_\downarrow\rangle\,|\text{sleeping cat}\rangle\,, \tag{2.6}$$

so that the situation represented by Eq. (2.4) is written as

$$c_{\uparrow+\downarrow}\,|A\rangle + c_{\uparrow-\downarrow}\,|B\rangle\,. \tag{2.7}$$

According to exactly the same analysis as above, one can argue that a multiplicity of structures will also be instantiated from this. These structures will not resemble classical cats, but there are parts of the cat structure that were hitherto associated with a (sleeping or awake) classical cat.

Why is there a multiplicity in the sleeping/awake basis of Eq. (2.4) and not in another, such as Eq. (2.7)? Wallace's argument is that a particular basis is picked out by the *dynamics*: at a coarse-grained level, the dynamical process of decoherence suppresses interference when the wavefunction is represented in a decoherence basis. This 'preference' of a decoherence basis needs to be justified, as it is a critical assumption in interpreting Eq. (2.4) as a multiplicity arising from superposition.

Hence, we must deny that mere superposition is the cause of the multiplicity, even if there are objectively special bases in which the branches of the superposition are not interfering. We need something extra. That extra is to regard a decoherence basis not just as a dynamically special basis, but as ontologically preferred, at least at the macroscopic level: patterns in this basis are more real (i.e., 'real_2', see §5) than others in the sense that they are useful patterns amidst all possible patterns.This is established through Dennett's Criterion, namely that the pattern instantiated through Eq. (2.7) does not map onto the mathematical structure of classical mechanics (§4), will not be useful (§5) and does not fit the Classical Role (§6).

Before discussing Wallace's Dennett's Criterion (§3), a note on the decoherence being *approximate* as opposed to exact. In physical situations, decoherence is a gradual process in time: there is a regime in time where there is large interference and a regime in which the interference is sufficiently suppressed (approximately zero).[6] For the gray area in-between these regimes it may therefore not be possible to speak about whether a world has already emerged or not, even though in most real-life cases these periods are very short due to

[6]This is the case for real-life situations, but theoretically this need not be: Blackshaw, Huggett, and Ladyman (2024) construct a toy model consisting of a one-dimensional line of spin degrees of freedom, for which interference is suppressed everywhere instantaneously (although this dynamical suppression itself spreads out locally).

decoherence of a system in a large environment being very rapid.[7] Per world, this is an approximate matter, whereas from the point of view of the universal wavefunction, worlds split-off continuously. It is therefore somewhat misleading to speak of 'the' decoherence basis, since a basis rotated slightly away from a basis in which decoherence is prevalent will also be basis in which there is a high degree of decoherence (cf. Butterfield 2001). A decoherence basis should further be thought of as a *coarse-grained basis*: the direct product of the many eigenstates corresponding to many degrees of freedom. It should not be confused with a 'fine-grained basis', such as the position basis of a single particle. 'The' decoherence basis should thus be understood not as a unique basis but as a *range* of possible bases in which decoherence is present to a high degree.

3. Dennett's Criterion

Everettian quantum mechanics automatically includes decoherence. Applying Dennett's Criterion to it will, I believe, lead to real worlds as patterns that approximately obey dynamical equations of the same mathematical form as those describing classical systems, e.g, describing the trajectories of classical particles in three-dimensional space. In my own words, Dennett's Criterion in the context of quantum mechanics would be rendered as:

> A quasi-classical world is a pattern, and the existence of a worldly pattern as a real classical world depends on the usefulness – in particular, definite outcomes and deterministic prediction, recovered by approximate robustness under decoherence – of classical physics which admits the classical pattern of a world in its ontology.

Although Wallace presents Dennett's Criterion as one that follows naturally from careful reflection on how we pursue science, he does not defend or justify it in detail. Indeed, he claims (Wallace 2012, p. 63) that it is based on "the same principles we apply right across science". Without further justification, this is a radical claim: certainly any philosopher carefully reading Dennett's Criterion will readily acknowledge it is not philosophically neutral and will search for the underlying view that justifies it.

The rest of this paper concerns that underlying view. This is warranted since the justification of Dennett's Criterion is much less discussed than other aspects of Wallace's

[7] Zeh, Erich Joos, and others (Joos and Zeh 1985, Joos et al. 2003, p. 67) calculated how fast diagonalization in the position basis occurs for the localization of macroscopic objects. They found that after scattering with environmental particles (in this case polarized photons) the diagonal entries of the reduced density matrix ($x \neq x'$) in the position basis acquire an exponential decay function in time,

$$\rho(x,x') \rightarrow \rho(x,x') \exp\left[-\Lambda t (x-x')^2\right], \quad (2.8)$$

for Λ the 'localization rate', depending on the scattering cross-section and the particle flux. This localization rate is often larger than the rate at which systems reach thermal equilibrium! A dust particle of $10^{-5} cm$, for example, will decohere at a characteristic timescale of 10^{-13} *sec*, due to interaction with air molecules only.

approach – such as the Deutsch-Wallace decision-theoretic approach to deriving the Born rule – although there are notable exceptions in (Lewis 2016, p. 66; Ney 2013; McQueen 2015, p. 13; Møller-Nielsen 2016, pp. 79–80) and a more thorough analysis in (Hemmo and Shenker 2022).[8] The next three sections (§4–§6) attempt to fill this gap.

Wallace himself gives two explicit hints about the underlying view. First, he mentions in Section 8.8 (2012, pp. 314–315) that his treatment in Chapter 2 (where he outlines Dennett's Criterion) is a 'structuralist' approach (without mentioning this in Chapter 2 itself), but hastens to add that one need not be a structural realist for his approach to work (which I believe to be questionable, see §6).

Second, in the Introduction (2012, p. 3), there is an explicit appeal to 'naturalism' as one of the core assumptions: "There is, however, a strong, but largely tacit, philosophical premise running throughout the books [*sic*]: *naturalism*, of the kind advocated by Quine (1969) and more recently by Ladyman and Ross (2007)." It may well be that Wallace has Dennett's Criterion in mind as a corollary of a particular kind of naturalism, but I would argue this is a particularly strong kind. For any form of naturalism that goes beyond the injunction to 'take science seriously in your philosophical methodology' is bound to be contentious, and a naturalism from which Dennett's Criterion can be deduced would indeed go far beyond this injunction. But I am getting ahead of myself.

Dennett's 'mild realism', brought out most explicitly in his "Real Patterns" (1991), which was a more general follow-up to the *Intentional Stance* (1971, 1981) created much turbulence in the philosophical lake (cf. Hill 1994; Haugeland 1998; Ross, Brook, and Thompson 2000; Millhouse 2022), and it would take us too far to wade through all the remaining eddies here. Rather, I will bring out two elements of importance for our discussion: the idea of a stance, and the concept of non-useful patterns.

Dennett's realism involves the idea that a pattern can correspond to something real if there are underlying principles implying that the pattern fulfils some functional role. Which patterns are recognised among a multitude of recognisable patterns depends on the specific functional role that is relative to a point of view, or *stance*, that we adopt towards it. Two individuals may discern quite different patterns from the data in front of them, if those individuals have different goals or standards of accuracy. That is not to say that the pattern is merely subjective or is absent when there are no agents in existence. It is only to say that the pattern is not more important than another: *importance is assigned*. In other words, the pattern is not principally observer-dependent: Dennett (1991, p. 34) insists that a pattern that "exists in some data – is real – if there is a description of the data that is more efficient

[8]The topic is also discussed well in high-quality graduate work, such as (Janssen 2008, p. 135; Newey 2019, p. 28). Further, it appears in the *Stanford Encyclopedia of Philosophy* entry on decoherence (Bacciagaluppi 2016) and a book on nanotechnology (Cooley and Lynn 2020, pp. 255–260); the criterion (in amended form) is also used in (Dürr 2019) but in the different context of gravitational energy.

than the bit map, whether or not anyone can concoct it.” Nevertheless, human interests exert *some* influence on the salience of a pattern, namely in the choice of highlighting it at the expense of other patterns.

Dennett gives us three different stances as strategies towards the behaviours of things around us—be it objects, animals, or machines. First, the Physical Stance is adopted by laboratory scientists who attempt to explain the behaviour of their target system – such as the discrete spectral lines of atomic phenomena – in terms of natural laws, its physical components and their interactions. Second, the Design Stance, on the contrary, is adopted when such physical language is not helpful and a more general design language will do instead. The fact that a washing machine adequately cleans cotton clothes is more adequately explained (assuming the system is working correctly) by appealing to running the ‘cotton programme’ which is designed for that purpose, and in engineering terms like ‘water temperature’ and ‘centrifuge RPMs’, rather than hydrodynamics and chemical equations representing how lipids dissolve in a soapy solution. Finally, the Intentional Stance is a way of understanding and predicting the behaviour of things in functional terms by treating those things as rational agents, by ascribing to them mental states that drive that behaviour. In one way, then, the Intentional Stance is a predictive strategy that allows us to make sense of complex systems, even without fully understanding the internal workings of those systems. This is to regard mental beliefs and desires as patterns or structures in natural systems that play the role that we usually ascribe to those mental states as such, in the context of folk psychology.

Thus the idea is that the Intentional Stance is a point of view from which recognizing beliefs becomes the same as recognizing patterns that play the roles of those beliefs. This does not go as far as to *reduce* the mental states to physical states, but instead *individuates* them as patterns in a background of physical things, so that we from there on treat them as autonomous (given that the patterns are sufficiently stable under noise). Hence the patterns are functionally specified in terms of (agglomerates of) the underlying ontology, such as neuron activity, even if not directly identified with that ontology. Moreover, when we distinguish a pattern against a background of constituents, we recognize that such a pattern need not be exact: the pattern can deviate a bit from the underlying stable thing that it is supposed to represent. Can certain patterns be ‘unreal’? Yes, according to Dennett:

> If we go so far as to distinguish them [the abstract entities] as *real* (contrasting them, perhaps, with those abstract objects which are *bogus*), that is because we think they serve in perspicuous representations of real forces, “natural” properties, and the like.” (Dennett 1991, p. 29)

Hence, there is underlying objective behaviour that can be appealed to, so that the specific pattern is to be discerned: it is in that sense that the patterns are ‘real’.

Assuming an underlying reality of atoms, we can easily imagine bogus patterns as the behaviour of mereological sums of, for example, the atoms that make up the water of the Bosphorus together with the atoms in your right toe. There seems to be no useful behaviour exhibited by this combination of atoms: it is not 'cutting nature at the joints' and can be considered bogus. At the same time, it would be unproblematic for Dennett were the assumption of an underlying swirl of atoms to be false, as long as there is *something* real that underlies the patterns. If the recognised patterns can be better understood in terms of a continuous matter field, or in terms of the field excitations of a field theory, this would not invalidate the recognition of the patterns. In this way, Dennett can be rather indifferent about the underlying ontology, as long as there is one.

Unsurprisingly, we see there is indeed considerable overlap between Wallace's position and Dennett's. Wallace, too, through his formulation of Dennett's Criterion, emphasises the role of behaviour in the recognition of patterns and adopting a realist stance towards them as long as that is useful (further discussed in §5). Both, also, are rather indifferent about the underlying ontology: just as it does not matter for Dennett whether the underpinning ontology of mental states consists of minuscule billiard balls or matter fields, likewise Wallace does not give a clear underlying ontology that the wavefunction represents at a fine-grained level.

Dennett is rather safe to assume there is a more fundamental level to appeal to, which represents an ontology that can behave in a patterned way. He appeals to an ontologically reductive picture where the material we care about (such as brains) is made out of smaller material stuff (neurons and neural activity), the useful patterns in which – when seen from the Intentional Stance – amount to mental states. Although similar in spirit, the coarse-grained behaviour that Wallace is focusing on is not automatically the behaviour of smaller physical constituents, but that of a smaller number of composite basis states, strung along by direct products. Quantum mechanics by itself applies to the macroscopic as well as to the microscopic, and as far as the quantum formalism is concerned, there is no principled distinction between the coarse-grained and the fine-grained, except for the mere number of factor Hilbert spaces one is working with. In this sense, both coarse- and fine-grained levels do not come with a preferred basis, even when decoherent evolution is present. As such, before *applying* Dennett's Criterion to the decoherent patterns at the coarse-grained level, those patterns are still fully quantum-mechanical, and not different (in particular: not classical) as far as the formalism is concerned. To regard them as classical macroscopic objects, one must look at patterns only when expressed in a decoherence basis, and then view these patterns from the vantage point of a definite, non-quantum mechanical world, and one does this by adopting the Classical Stance (§6).

Also, Dennett's patterns can be safely assumed to occur in real space, whereas Wallace's patterns occur in an abstract mathematical space.[9] More generally, the conceptual language of classical mechanics does not match that of quantum theory. As such, there is a conceptual (or qualitative) gap between the patterns that are being matched—an issue we turn to next (§4).

4. Reduction through instantiation: quantitative pattern-matching without qualitative grip

We now come to the three philosophical preconditions of Dennett's Criterion that I would like to highlight: (I) reduction through instantiation, (II) pragmatic individuation, and (III) the Classical Stance. In the coming three sections I explore these themes in turn.

The part of Dennett's Criterion that is concerned with the vertical relation between theories is easily distilled: patterns of a lower-level theory are matched to patterns admitted into the ontology of a higher-level theory. That is, the patterns in the universal wavefunction of the lower-level quantum theory are matched with patterns of classical physics. These are the patterns arising in the universal wavefunction in a decoherence basis matched to the patterns of classical objects obeying classical dynamics in real space. Hence, the pattern of the macro-object is supposed to represent some derivative ontology (or perhaps agglomerate of lower-level constituents) of the more fundamental ontology of the lower-level theory. That derivative ontology then corresponds to the ontology of some higher-level theory.

To flesh out more carefully *how* to match the patterns at different levels, Wallace speaks of 'instantiation' (to be understood as 'identifying an instance' of something) as a relationship between two theories inside a certain domain. Classical mechanics, for example, reigns in the domain of the Solar system; and molecular physics instantiates a theory of classical point particles subject to Newton's laws. In zoology, within the domain of typical behavioural patterns of animals, we find for many cases an instantiation of game theory: stable evolutionary strategies arise from selfish and rational actors. Thus Wallace's pattern-

[9] That is, this is the case for the Emergent Multiverse interpretation. One can also approach the issue starting from *Spacetime State Realism* (Wallace and Timpson 2010, cf. Wallace 2012, Ch. 8), according to which the quantum ontology is a state-valued field in spacetime, making it more suitable to mesh with quantum field theory. In this theory the state is more than a codification of expectation values, but constitutive of a real field with non-local relations between spacetime regions that do not supervene on the intrinsic properties of these regions (i.e., there is non-separability). The observables are the expectation values of Heisenberg operators, invariant under spacetime-dependent phase transformations. A full evaluation of Dennett's Criterion in the context of Spacetime State Realism is beyond the scope of the current paper – especially in light of having to address various criticisms (Arntzenius 2012; Lewis 2013; Baker 2014, 2015; Ismael and Schaffer 2020; Swanson 2020) – but it is clear that assuming this interpretation sufficiently bridges the gap between the wavefunction's Hilbert space and classical space (or spacetime) and therefore re-establishes the possibility of seeking for Dennettian patterns in real space (I thank an anonymous reviewer for pointing this out). This does not mean there will be a fully homogeneous reduction, with no qualitative gap, since the state-valued fields in spacetime are not given in classical terms, but in terms of Heisenberg operators and a dynamics that is not invariant under spacetime-dependent phase transformations (cf. Arntzenius 2012, §3.13).

matching is more elastic than traditional accounts of reduction. Although it may include fully-fledged intertheoretic reductions of the type where thermodynamics is said to be reduced to statistical mechanics, it can also be much smaller in scope, allowing for context-dependence:

> Crucially: this 'reduction', on the instantiation model, is a local affair: it is not that one theory is a limiting case of another *per se*, but that, *in a particular situation,* the 'reducing' [lower-level] theory instantiates the 'reduced' [higher-level] one. (2012, p. 55, original emphasis)

The working definition of instantiation is:

> Given two theories A and B, and some subset D of the histories of A, we say that A instantiates B over domain D iff there is some (relatively simple) map ρ from the possible histories of A to those of B such that if some history h in D satisfies the constraints of A, then $\rho(h)$ (approximately speaking) satisfies the constraints of B. (Wallace 2012, p. 34)

The idea is clear: it is an attempt to weakly reduce theory B to theory A, or parts of those theories, within a domain D, allowing also for approximations. Hence, there exists a map that approximately maps histories constrained by theory A in domain D to histories constrained by theory B,

$$\exists \rho : \rho(h^{A_D}) \to h^{B_D}. \tag{4.1}$$

Although Wallace admits there is work to be done to make the definition of instantiation more precise,[10] especially to apply it to examples such as the one on zoology and game theory above, this is not delivered upon in later work. For example, in his recent work on what he calls the 'math-first' approach to scientific theories, he refers *back* to the 2012 book:

> On the math-first view, reduction is something like *instantiation*: the realizing by some substructure of the low-level theory's models of the structure of the higher-level theory's models. In the important case of state-space instantiation, for instance (discussed in more detail in (Wallace 2012, ch.2) [i.e. *The Emergent Multiverse*]), the lower-level theory instantiates the higher-level one if (roughly) there is a map from the lower-level state space to the higher-level state space

[10] See also (Franklin 2023; Franklin and Robertson 2023) for a more precise way to flesh out emergence in terms of the screening-off of the underlying micro-physics in the context of rainforest realism (also a strong form of naturalism). I disagree, however, with taking the criterion of instantiation as the criterion of emergence: it seems more philosophically explicit to see Dennett's Criterion as establishing the emergence claim, because it is here where one outlines the ontological commitment to higher-level patterns. The instantiation criterion is merely a means of fleshing out the way in which patterns are matched. It may of course be that the philosophical preconditions of Dennett's Criterion are shared by (Franklin 2023), making it seem like it is the instantiation criterion that is doing the work.

> that commutes with the dynamics and leaves invariant any commonly-interpreted structures (for instance, spacetime structure) in the two theories. (Wallace 2022, p. 16, *original emphasis*)

Back to quantum mechanics: how, then, are the classical worlds instantiated by quantum mechanics? Armed with this working definition of instantiation, we want to interpret Eq. (2.4) as representing two structures that instantiate the structure of a universe containing the sleeping cat *and* another one containing the waking cat. In this way, a relationship between a substructure of quantum mechanics and classical mechanics is established, at least within the domain in which the mapping approximately holds. Given this pattern-matching between mathematical structures at different levels, it is wise to investigate parallels with the philosophical debate over reduction. In particular the lack of conceptual, or qualitative, bridge laws, which are so pertinent to more traditional debates.

One should distinguish between the concepts of reduction and emergence. Regarding *emergence*, what is at stake is the autonomy and robustness of higher-level concepts or structures relative to the lower-level domain (cf. Bedau 1997; Chalmers 2006).[11] Of course, Wallace's titular stated aim precisely is emergence, not reduction: the multiverse is supposed to be *emergent* from pure wave mechanics. But the emergence of a multiplicity of real classical worlds requires the whole package of Dennett's Criterion, where the pattern-matching is fleshed out by the (reductive) instantiation relation. This emergence relation itself is not completely fleshed out, because Wallace wishes to solve the measurement problem without getting bogged down in the details of emergence:

> These details [of 'explanatory usefulness'], however, are not crucial for our purposes. This is not a book about the philosophy of emergence: it is a book about the measurement problem. (Wallace 2012, p. 58)

Regarding *reduction*, the most prominent account is the Nagel-Schaffner model, where Nagel (1961; see also Hempel 1966, p. 77; Ager, Aronson, and Weingard 1974) holds that a successful reduction has two requirements, namely (A) the derivation of the higher-level theory from the lower-level theory, with the help of (B) conceptual bridge laws (for inhomogeneous cases, as explained below). Derivability here is somewhat notorious, as it should be understood at least in spirit as a fully-fledged *deduction*, which in practice can prove quite unworkable due to the complexities of actual scientific theories. But whereas Nagel was operating under a rather strict deductive-nomological framework, Nagelian reduction has often been reinvented with suitable relaxing modifications. Schaffner's (2006) extension is to loosen up requirement

[11] There is, of course, a vast literature on the topics of emergence and reduction interrelate, which cannot be summarised here, but for the readability of the rest of this section it is sufficient to have *rejected* two common obstacles, namely that (1) emergence is incompatible with reduction and that (2) emergence is nothing else but supervenience when reduction fails. See (Butterfield 2011b; 2011a; cf. Dewar 2019).

(A) to include partial, or 'patchy', derivations, amounting to (A$'$) the derivation of higher-level structures from the lower-level structures, allowing for approximations and occasional exceptions or failure. Wallace's instantiation has a similar caveat in that it allows for approximations and depends on the case at hand, i.e. it is not a "sweeping reduction" (Schaffner 2006, p. 379). Indeed, the salient difference between this traditional account and Wallace's account lies not in (A) but in (B).

Bridge laws, or 'connectability assumptions', can be understood as rules or definitions that bridge the qualitative gap between a concept of one theory with that of another. The terms 'light ray' and 'refraction index' of linear optics are not part of electrodynamics, and should thus be brought into the qualitative scope of that theory by *defining* them in terms of 'electric field' and 'magnetic field', and 'permittivity' and 'permeability', so as to establish not just a quantitative matching of maths, but also a qualitative matching of concepts; and likewise for the more famous (and more contentious) reduction of 'temperature' of thermodynamics to 'kinetic energy' of the atoms of molecular kinetics.

Instantiation, then, says nothing about bridge laws in any *qualitative* way. Instead, by dealing only with the (approximate) mapping $\rho(h)$ between structures of different theories, this criterion is entirely mathematical.[12] As such, the mapping $\rho(h)$ is playing the role of a bridge law in the sense that the mapping 'bridges the gap' between the higher and lower levels quantitatively, leaving a qualitative gap.

Is a qualitative gap always a problem? In the philosophy of reduction, there are exceptional cases where bridge laws (B) are not required, which are called the homogeneous cases, where the ontology of the reducing theory and the ontology of the reduced theory are already formulated in the same terms *before* the reduction is carried out. An example of such a homogeneous reduction is Galileo's law of falling bodies being reduced by Newton's laws of motion and law of gravitation. In cases of inhomogeneous reduction they are always required.[13] One would be hard-pressed, however, to see the current case as a homogeneous one, i.e. regarding the relation between wavefunctions in configuration space as coinciding with the ontology of classical mechanics.

If one follows The Emergent Multiverse interpretation, one is nevertheless likely to reject (B), arguing that a quantitative mapping – just the map $\rho(h)$ – is all that is really needed for a successful reduction. This would amount to a (controversial) stance in the reductionism debate that conceptual bridge laws are not necessary even in cases of inhomogeneous reduc-

[12] Joshua Rosaler (2015) makes a similar distinction between 'formal' and 'empirical' reduction in physics. The former is a two-place relation between higher- and lower-level theories that can proceed *a priori*, purely through mathematical analysis; the latter is a three-place relation between higher- and lower-level theories and the domain of real-world cases, bringing in synthetic facts, so that empirical reduction does not require the higher-level mathematical framework to be wholly subsumed by the lower-level theory and instead focuses only on the physical parts.

[13] I thank an anonymous reviewer for pressing me on this clarification.

tion; or that the distinction between homogeneous and inhomogeneous reduction should be collapsed altogether. This should be justified. Indeed this is (part of) the math-first structural realist project (see below), where the role of language (and thus the mismatch in vocabulary) is deemed unimportant or at best secondary.

Thus a qualitative leap is taken when interpreting the superposition of autonomous quasi-worlds as a multitude of overlapping classical worlds set on a classical space. These quasi-classical worlds behave classically in the sense that they approximately obey the classical equations of motion.[14] Mathematically, going from the Hilbert space of the quasi-duoverse of Eq. (2.4) to individual identifications of two quasi-universes like

$|$universe with $\uparrow$ and  $\rangle$ & $|$universe with $\downarrow$ and  $\rangle$ (4.2)

may be convincing on the basis of a suitable autonomy criterion satisfied by robustness under decoherence. But to move from here to the duo of real macroscopic classical worlds

$$\text{‘classical universe with sleeping cat’ \& ‘classical universe with awake cat’,} \tag{4.3}$$

is underpinned by Dennett's Criterion, without which the language used cannot be justified.

By emphasising that there is a qualitative mismatch between the theories I do not intend to raise any concerns about approximation. Even if the autonomy of the quasi-classical worlds *were* to be exact, not approximate, one still needs to say something qualitative to establish an ontological identification, rather than merely a quantitative match.

Nor do I intend the word ‘qualitative’ to indicate anything vague or wishy-washy, such as an appeal to a conceptual metalanguage that transcends both theories, or merely an added ‘qualitative note’ that guides us psychologically towards understanding the old concepts in terms of the new. On the contrary, without a qualitative link, the reductive relation between the theories is more wishy-washy than with it.

This additional understanding derives from bridging the predicates of the higher-level theory with the predicates of the lower-level one, such that the former is explicitly shown to be a part of the latter. A common technique to achieve this is *definitional extension*, and is precise: take both theories to have distinct non-logical vocabularies even if the same or similar terms occur in both theories that should be differently interpreted (such as ‘light-ray’

[14] It is important to note that although the quantum formalism interpreted via Dennett's Criterion includes the quasi-classical worlds that behave classically, the argument does not work the other way around. That is, classical mechanics is not capable of making all the predictions that the quantum formalism can also account for, such as those quintessential *quantum* effects as the stability of matter, double-slit interference patterns, or radioactive decay. This can be recognised also in the asymmetry of instantiation, which allows for theory A to be effective in a larger domain that theory B, the instantiation only taking place within D. That is, theory A weakly reduces theory B, but not *vice versa*. I thank an anonymous reviewer for pressing me to clarify this point.

in linear optics and Maxwell electrodynamics, or 'classical' and 'quasi-classical' in the current context); and then proceed by adding to the lower-level theory some definitions (the bridge laws) that each capture a term of the higher-level theory, in such a way that the higher-level theory is fully encapsulated by the extended lower-level theory.[15]

Indeed, such "predicate precisification" (Wallace 2022, p. 14) is precisely what Wallace would see as an artifact of the language-first view of scientific theories. For him, it is the mathematical formalism that is *given*, and only subsequently marred by straitjacketing it into terms of objects, properties and relations.[16] Indeed, it is on a language-based account (roughly coinciding with the syntactic conception of theories) where one obtains a notion that objects and relations of the reduced theory are *part of* the lower-level theory. The former's theorems can be derived from the latter's. For Wallace, this is not necessary, and sufficiently explained by a quantitative mapping of the patterns if a (relatively simple) map $\rho(h)$ exists.

This brings out a theme that may be framed as differences between *Lewisian reduction* and *Dennettian reduction* (Lorenzetti 2024; Knox and Wallace 2024).[17] These frameworks broadly emphasise different aspects of reduction: where the Lewisian kind is primarily concerned with ontological questions, the Dennettian kind seems to be concerned with epistemological ones. I say primarily, because the Dennettian, apart from epistemology, relegates ontological questions to a hypothesised fundamental level; and Lewisian reduction is fallible: a particular attempt may fail if no unique realizer is found, for reasons of *synthetic* fact. This latter point is likely to make the two views irreconcilable, since the Dennettian appears to make *analytic* statements about the realizer: Dennettian functionalism cannot fail. This is not the place to attempt a full resolution of this debate, and so it remains to be seen in future work to what extent these different accounts are incompatible or not. The point here is that Dennetian reduction (or something similar) is needed for the multiverse project to go through in terms of instantiation.

This requires the rejection that qualitative bridge laws in inhomogeneous cases of reduction are needed, or the rejection of the distinction between homogeneous and inhomogeneous reduction itself. Thus there appears to be a seed of the above-mentioned recent work

[15]See (Butterfield and Gomes 2023) for a full overview as well as didactic account of the details of this account of Lewis-style functionalist reduction (Lewis 1970, 1994), and see (Gomes and Butterfield 2022) for an application to geometrodynamics. In this case, functional identification provides the way to bridge the qualitative gap by simultaneously specifying several concepts through the roles they play in the theories. In some lucky cases – lucky meaning a successful but fallible empirical search – a unique realizer can be found that plays this role, allowing them to be identified.

[16]On the math-first approach, the mathematical structure of a theory is directly compared – without the use of conceptual language – to empirical data structures. The upshot is that structural realists will then not get bogged down in metaphysical debates about the precise nature of 'structure'. Much can be said about this approach (and doubtless, much will be), but here it is sufficient to look at the role of the bridge laws: in the math-first approach these conceptual connections will not come up.

[17]As far as I know, this opposition was coined by Eleanor Knox at a conference talk on June 16^{th}, 2021 (MetaScience project's ""Going Up?" Realisation and Composition across the Sciences").

on the math-first approach to structural realism (Wallace 2022), since in this approach the centrality of language itself is rejected. As such, the inhomogeneity of terms does not arise. Curiously, this recent work does not mention the emergent multiverse as a case study to support the math-first approach, nor does it mention Dennett's Criterion.

For now, then, the ontological question remains: a pattern in a wavefunction is an entirely different animal than a classical world. At least at first sight, one needs to specify a qualitative match between patterns in the universal wavefunction and the patterns in the real space of the formalism of classical mechanics. Mathematical similarity between theories is no guarantee for ontological similarity—they may have very different philosophical consequences.

5. Pragmatic individuation: a parallel world near Van Fraassen

What do we mean by 'preference' when we speak of the preferred basis, or of the decoherence basis being preferred? From §2, we see that there is a kind of *dynamical preference* that nature has for evolving in a particular way rather than another, based on synthetic facts that inform the specific forms of interaction Hamiltonian operators. But surely it is not the dynamics itself that does any of the preferring: a dynamically preferred basis is just another way of saying that there is some basis in which some special dynamical behaviour occurs which does not occur in other bases. Thus, a decoherence basis is preferred in the sense that the patterns of the wavefunction that can be identified in that basis is robust under decoherence, whereas another basis will generally show large and rapidly-changing interference terms. Does such a particularly unique kind of dynamical behaviour make the patterns seen in a decoherence basis *more real*? Appealing to pragmatic considerations, Wallace appears to argue just that. In the formulation of Dennett's Criterion, this takes the shape of an appeal to "usefulness – in particular, the explanatory power and predictive reliability – of theories".

In everyday uses of the word 'pragmatic', one often takes it as the ignoring of any details or truths that distract one from reaching a particular goal in the fastest way. Different goals specify different contexts in which different details becomes salient. Given that people have divergent goals, different sets of salient facts will obtain for them. Wallace, rightfully, warns us against this subjective or observer-dependent understanding of 'pragmatic' in his quantum mechanics:

> I want to stress that talk of 'explanation' here is not meant to imply that it is just a *pragmatic* matter whether or not, for example, we take seriously tigers and other such macro-objects. (2012, p. 57)

Indeed, decoherence is an objective dynamical feature of the quantum formalism, regardless of what you or I may want—and the same goes for the existence of tigers. And indeed, when you are in a position to choose between using one pattern or another, you want to pick the pattern that is most useful to meet your goals.[18] In this sense, reasoning from our practical human needs, it often seems reasonable to choose the quasi-classical patterns of Eq. (2.4) over those in a rotated basis of Eq. (2.7). Through Dennett's Criterion, however, the context is specified not by individual or subjective goals but by the explanatory power and predictive reliability of theories in a more objective sense. Let us discuss these in turn.

Explanatory power is often hailed as one of the goals of science, and is often (but not always) taken as the main motivation for scientific realism: as an inference to the best explanation, the true theory is that one which provides the best explanation of the available evidence. Decoherence is real in that it can serve as a dynamical explanation for why quantum computers only work at exceedingly low temperatures (namely to sustain the superposition by preventing the entanglement from spreading over the environment). However, decoherence by itself cannot be sufficient to explain the existence of real classical worlds. That gap is supposedly bridged by the philosophy of emergence, here in the form of Dennett's Criterion.

Predictability is the quality or characteristic of being able to be foreseen with a good degree of accuracy. It implies the ability to anticipate a future outcome or future behaviour based on available information. Indeed, the quasi-classical pattern *is* useful, precisely because it re-establishes (within a given world) deterministic prediction of definite outcomes of experiments. Surely, definite and (approximately) deterministically behaving things seem more tractable (disregarding chaotic behaviour) than dealing with indefinite descriptions and probabilities. In the words of Żurek (2002, p. 21): "classical reality can be regarded as nearly synonymous with predictability."

Now we can answer a question that should have been bothering the reader while reading §2: why is Eq. (2.4) more useful than Eq. (2.7)? As the former is written in a decoherence basis, decoherence ensures that the sleeping cat state and waking cat state of Eq. (2.4) will *not* interfere with each other. Thus these three-dimensional subsystems can be regarded as detached from each other. This is useful in the sense of predictability, as Żurek would have it: individual structures approximately mimicking classical behaviour. I am less sure about how explanatorily useful it is (this surely depends on the details of one's favourite theory of explanation), but presumably the relative stability of the quasi-classical patterns singles out such patterns as describable by classical macroscopic variables, which can in turn feature in an explanation.

[18] Of course there is also the traditional pragmatist conception of truth, as laid out by Peirce, James, Dewey and others, where the correspondence view of truth is rejected (cf. Misak 2013). A statement is then not true in light of its correspondence with any propositions or states of affairs, but because of its usefulness to the inquiry at hand. There appear to be themes in common with *The Emergent Multiverse*, especially as regarding only the quasi-classical pattern as real due to its usefulness (§5–§6), and Wallace regularly (e.g. 2010) uses the words 'pragmatic' and 'practical', but none of the authors of the traditional or neo-pragmatist schools are cited, and the concept of truth itself is left untouched, and so I will leave this project to the reader. Certainly, the math-first approach has an *anti-representationalism* in common with neo-pragmatism.

If classical reality is synonymous with predictability, what about quantum-mechanical reality? If Hilbert space democracy of bases holds true, and if the quasi-classical patterns are to be considered real, then surely any other pattern should be considered as just as real.

Let us briefly explore an analogy with the relativity postulate in special relativity: just as all sets of Hilbert-space basis vectors are treated equally in quantum mechanics, so are all inertial frames in special relativity. The relativity postulate ensures that dynamical laws take the same form in each inertial frame. In Hilbert space, the superposition principle ensures that for any arbitrary choice of basis the evolution from an initial state to a final state will lead to the same outcomes for observables. At the level of formalism, this is as much the case for the rotated basis of Eq. (2.7) as it is for the Duoverse in Eq. (2.4).

Desiring to break the tie in relativity theory, one could appeal to important robust dynamical features: all inertial observers agree on the intensity of the cosmic microwave background. This introduces a preferred frame of reference when it comes to physical facts, the so-called cosmic frame. Yet, this does not conflict with the relativity postulate, which entails that different choices of inertial frames will lead to agreement of Lorentz-invariant quantities. Indeed, the mistake that the cosmic microwave background breaks the relativity of inertial frames has been made before, e.g., by Hermann Bondi and Peter Bergmann (cited and refuted in Muller 1992).

In this respect, a decoherence basis in quantum mechanics is similar to the cosmic frame in cosmology. There is an objective *physical fact* about a decoherence basis, namely that it is the basis in which interference terms vanish. It would be a mistake, however, to take this to entail that patterns in this particular basis are more real than patterns in another basis. Indeed, this is not exactly the claim that is being made in the Emergent Multiverse interpretation.

For those who follow Wallace's approach, the follow-up is going to be as follows. Indeed, we could speak of something such as a Hilbert space democracy of bases, but this really only makes sense *from a fine-grained point of view*. At a coarse-grained level, certain patterns arise that fit classical mechanics, and we know that classical mechanics is a very useful theory, particularly regarding its domain-specific predictability and explanatory power.[19] Any other 'patterns' that arise – if you insist on calling them that[20] – do not fit classical mechanics and are hence not to be taken macroscopically seriously. One could say that this democracy of *bases* at the fine-grained level does not lead to the equal treatment of *patterns* at a coarse-

[19]Note, here, that the claim is certainly not that expressing the wavefunction in just any basis at the fine-grained level will lead to the same patterns at the coarse-grained level, as if there is a convergence whatever one chooses as a fine-grained basis: this is not the case.

[20]I speak of all kinds of patterns within the universal wavefunction as if they are on a par, i.e. whether they are useful (i.e. real_2) or useless (merely real_1), admitted by higher-level theories or not. In opposition to Dennett's 'bogus patterns', Wallace does not seem to use the word 'pattern' for useless patterns. But this is just semantics.

grained level. But this is not as radical as it sounds: to make this consistent one should make a distinction between two different conceptions of reality. On the one hand, we have $real_1$ as that which is being quantified over in formulating the theory, providing the ontology at the fine-grained level; in this case that could be the wavefunction, to the extent that it is a good candidate for ontology (although in our case there is some agnosticism or deflationism about the fine-grained ontology). On the other hand, we have $real_2$ as those coarse-grained patterns – consisting of, or recognised in, the real_1 fine-grained ontology – which are pragmatically individuated and (as seen from the Classical Stance, §6) serve as macroscopic objects. As such, real_2 implies real_1, and is thereby logically stronger.[21]

However, there is that nothing in the formalism that forces this interpretational move of identifying something as real_2 instead of merely useful. An interpretation needs to be added, namely an implicit but specific pragmatic individuation which – through Dennett's Criterion – justifies the individuation of real_2 patterns.

Let us to that end investigate another analogy, namely between Wallace's contextual highlighting of the quasi-classical pattern and Bas van Fraassen's (1977; 1980) pragmatic theory of explanation:

> As in all explanations, the correct answer consists in the exhibition of a single salient factor in the causal net, which is made salient in that context by factors not overtly appearing in the words of the question. (Van Fraassen 1980, p. 132)

According to Van Fraassen, scientific explanations, like all explanations, do not stand alone; they require certain facts about the context in which we ask for explanation and in which we offer explanation. Without a specification of these contextual factors, *the* explanation simply does not exist, akin to clarifying what 'yesterday' refers to without specifying the context of the time when that word is uttered. Thus, according to Van Fraassen, explanations are not two-place relations between theories and facts, but three-place relations between theories, facts, and contexts. In a nutshell: an explanation is always an answer to a 'why-question', which is determined by (a) a *topic*, which is the proposition about the fact which we seek to explain; (b) a *contrast class*, which is a set of propositions that contrasts alternatives to the topic; and (c) a *relevance relation*, which signals a requested reason for that question to be worthy of asking.

For example, one why-question I am trying to answer in this paper is 'Why did Wallace write chapter 2?'. The topic would be 'Wallace wrote chapter 2'. The contrast-class

[21] I thank an anonymous reviewer for pressing me to give the different senses of 'real' a label so as to keep them apart. Note also that indeed it is true that the quasi-classical worlds are "structures instantiated within the quantum state, but they are no less real for all that" (Wallace 2012, p. 63), but that real_2 is in fact a stronger sense of existence, as being a macroscopic object as a real pattern rather than a 'bogus' pattern, whereas all the structure in the wavefunction (presumably) counts as real_1.

includes varying alternatives to which the topic can be contrasted, like 'Wallace *composed* chapter 2', 'Wallace wrote *chapter 3*', or '*Daniel Dennett* wrote chapter 2'. The relevance relation appeals to a reason which is relevant for the topic. For example, the answer 'because Wallace thinks chapter 2 is important' is irrelevant in most contexts, since it is contextually understood implicitly: such an answer would probably not satisfy the questioner.

Van Fraassen's broader project (of constructive empiricism and the rejection of Putnam's (1979) no-miracles argument) is to divorce explanation and truth, instead recognising it as a thoroughly human need and rejecting explanation as a goal of science (cf. Gopnik 2000). Wallace, as we have seen, intends nothing of the kind: the quasi-classical patterns are objectively and truly there, independent of the observer, and they are objectively there *because* they are explanatory.

But despite differences, Van Fraassen and Wallace both appeal to pragmatic virtues in order to give a special role to some patterns over others. Van Fraassen highlights patterns in a web of causal relations by their relevance; Wallace highlights patterns in the universal wavefunction by their usefulness. Both ways of highlighting are informed by explanatory power. Just like the objective threads highlighted in the causal web of science, the quasi-classical patterns are also objective. According to Van Fraassen in the quotation above, what makes a particular thread in the causal web a *salient* answer are "factors not overtly appearing in the words of the question," and because explanation is not part of science proper, he is here appealing to cognitive facts of the human mind. Thus the context that accounts for a certain causal thread as explanatorily satisfying is subjective. And even though explanatory power and predictive reliability seem like concepts relative to us humans, Wallace insists that existence$_2$ of the quasi-classical patterns via pragmatic virtues are not subjective or observer-dependent (or at least not 'just' so).

Looking at Dennett's Criterion, this objectivity is achieved through an existence claim about macroscopic objects, since it says that the existence of a macro-object depends on its being admitted into the ontology of useful theories. Via what precise arguments this real$_2$ ontology comes about is the topic of §6; here we consider how – given that one accepts the Classical Stance described there – one way of establishing the existence claim is through an anthropic argument.

Weak anthropic reasoning is a way to explain some aspects of nature conditional on human existence. It can be used as a selection criterion for what *we* can or cannot encounter (cf. Davies 1984). For example, via weak anthropic reasoning one concludes that it is not a coincidence that the Earth finds itself cosily within the Goldilocks zone—that is, a theoretical region so fine-tuned as to make life possible, so that, for example, moving a little more towards or away from the sun would be catastrophic to human life, and so on. The reason this is not surprising is that, given the uncountably many planets in the universe, at least one of them had to have the right conditions.

A similar story can be told for picking a decoherence basis among all available bases. Although Wallace never mentions the word 'anthropic' in the context of the preferred basis, he – with reference to the idea of information gathering and utilization systems (IGUSes) in (Saunders 1993) – might be thinking much along the same lines:

> In fact there will be a subset of history spaces which are much more convenient: we are information-processing systems, and it can be shown that any such system picks out a consistent history space (Saunders 1993). (Reverting to the subsystem description given earlier, the point is (in part) that such a system needs to store memories and if it chooses an encoding of memories into states which are not diagonal in a decoherence basis, they will not last long.) (Wallace 2001, p. 649)

Thus, here is a statement of, and solution to, Everett's 'determinate records problem' (cf. Barrett 2011): it is because we store memories and because we last for long periods of time that we are bound to find classical patterns to be convenient. For evolutionary reasons, perhaps, we only observe the observables as defined in a decoherence basis.[22] In that case, the philosophical preconditions of Dennett's Criterion could be formulated as:

> (II') amongst the objectively existing patterns in any basis, the patterns in a decoherence basis are subjectively salient because of our human condition, rooted in our evolutionary background.

In this sense, one can establish the subjective importance of the quasi-classical patterns, namely as important to *us*, and these are real patterns because all the patterns in the wavefunction were already real to begin with (i.e. real$_1$).

Hemmo & Shenker (2022, pp. 11-12) have reservations about such an evolutionary solution, as they envisage Kochen-Specker type objections when regarding the quasi-classical observables that our brains are evolved to observe to clash with observables associated with operators that our brains do not recognise. That is, assuming that all patterns in the wavefunction are real, the contextuality of quantum mechanics might spoil the claim about the existence of real classical worlds. I further follow Hemmo & Shenker in their observation that this evolutionary argument ought to be a matter of synthetic fact, instead of posited as a necessary explanation:

[22] This appears to be quite close to Zeh's own interpretation of quantum mechanics which amounts to a functionalist version of the many-minds interpretation: there *really* is a superposition but this ends at the level of the brain. The multiplicity of Eq. (2.4) is not one of worlds but of minds. It is then a matter of synthetic fact that our mental states correspond to eigenstates of a decoherence basis of some observables of human brains. Such a position would be a serious contender to the emergent multiverse interpretation, even more so because it sits very naturally with the functionalist tenets of Dennett's Criterion. Thus it seems that Dennett's Criterion applied to the quantum formalism does not automatically deliver a unique macroscopic ontology.

> Perhaps one could argue that as a matter of empirical fact our mental states are associated with the eigenstates of some observables of our brain, but not with others, and this is why we do not have an experience of these other observables, although they are equally real. It is the job of brain science to discover which observables of our brain are the mental ones, and it may as well be discovered that, for example, that the mental observables are the decohering ones. (Hemmo and Shenker 2022, p. 12)

Regardless of the success of the anthropic argument, however, already conceptually the precondition (II') is not sufficient to justify the existence claim spelled out in Dennett's Criterion, or at least not without adding an ontological component. For Dennett's Criterion is a selection criterion for patterns, picking out the useful ones, and then identifying those selected patterns with the ontological status of being a (real_2) macroscopic object. For Wallace, then, it seems that all we mean by 'existence' is pragmatic salience, but that identification is only possible with a certain kind of pragmatist approach to ontology—that is, real_2 ontology emerges from a specific pragmatic view, and we could call it 'pragmatist ontology', or 'functionalist ontology', or 'stance ontology', and to which we will turn now. Once this 'Classical Stance' is occupied, the anthropic claim is (if it works) one way of getting to that pragmatist ontology.

6. The Classical Stance

In Section 3 we saw that, for Dennett, it is not necessary to minutely specify the underlying ontology in which we recognise patterns: atoms, matter fields, or real-space quantum fields, will do. Though similar in spirit, Wallace's recognition of the quasi-classical patterns is at a more abstract, mathematical level. Looking at just the universal wavefunction and its behaviour, it is not just that there could be multiple different ontologies in which patterns could form, it is not clear which ontology qualifies in the first place. Where Dennett recognises patterns in real space, Wallace recognises them in an abstract and high-dimensional mathematical space. It is wise to remember the ontology problem: without knowing what quantum mechanics is *about* on a fine-grained level, we also do not know what the quasi-classical patterns on a coarse-grained level are *in*.

Is this a problem for Wallacian quantum mechanics? Not necessarily. On the one hand, if we had a clear ontology of quantum mechanics, then one should certainly specify how this ontology connects to the coarse-grained ontology of classical worlds that Wallace identifies. But this connection is supplied by the part of Dennett's Criterion that deals with reduction, through instantiation—already discussed in §4. On the other hand, one can remain agnostic or deflationist about the ontology of quantum mechanics at a fine-grained level: even without specifying an underlying ontology – or even in its absence altogether – we may be able to speak of patterns in a more abstract way. Indeed, this is the sense one gets when reading Wallace's other work, in which he is "against wavefunction realism" (2020) in that he does not consider the 'fundamental' ontology of the world to be the high-dimensional space of quantum mechanics, such as configuration space[23]; and he frames quantum mechanics as a "framework theory" (2017) that fits a large range of ontologically divergent concrete theories.

Although the specification of a clear ontology is a worthy goal (if not pursued by the physicist, then all the more reason for a philosopher to try), this paper is not about solving the ontology problem at a fine-grained level: for current purposes it is sufficient to remember the more abstract sense of pattern-recognition at play.

To address the ontological status of the quasi-classical patterns at the coarse-grained level, we need to be more specific about what ontological status means. Forget for now the wavefunction, or the Everett interpretation, or even Dennett's Criterion, and allow me to make a general (if not pedantic) point about existence: surely we all agree that there *are* countless objects (or structures, or properties, or patterns, or variables) in the world that we never refer to—not even in thought, let alone in speech. Recall the mereological sum of the water of the Bosphorus together with your right toe (§3). Or, more seriously, think about two macro-variables expressed in terms of patterns in the behaviour of micro-variables of the kinetic theory: one being $T = \sum_i \left\langle m_i v_i^2 \right\rangle$ and the other $A = \sum_i \left\langle m_i^2 v_i^3 \right\rangle$. These objects (or structures, or properties, or patterns, or variables) are *real*, which is, after all, what 'are' means. But of course this kind of existence does not conflict with us judging things like A as cognitively useless.

Whatever the real_1 ontology at a fine-grained level, then, how to pick out the real_2 coarse-grained ontology? One step towards the answer was already outlined in §5, where the patterns in a decoherence basis are made salient through usefulness. I see three options for a second step to proceed. First, the patterns in the decoherence basis, even though they are special, are not more real than other patterns, so that there is a myriad of other patterns available to be quantified over and define (less stable) macroscopic objects. Second, we adopt the language (introduced in §5) of real_1 and real_2. The former sees every pattern in the wavefunction as ontologically on a par, while the latter identifies macroscopically real objects with that what is useful at that level. This is what Dennett (1991, p. 29) intends to say by calling his realism "mild and intermediate".[24] It is not just instrumentalism about the patterns, some of them useful and others not, but a *realism*: whereas patterns exist_1 (as

[23] Again, one could alternative evaluate the theory of Spacetime State Realism, which, if successful, does provide a real-space ontology, see fn. 9.

[24] Dennett (1991, p. 27): "Philosophers generally regard such ontological questions [whether there are really beliefs] as admitting just two possible answers: either beliefs exist or they do not. There is no such state as quasi existence; there are no stable doctrines of semirealism. Beliefs must either be vindicated along with the viruses or banished along with the banshees."

consisting in real_1 ontology), the useful patterns exist_2 (as real_2 macroscopic objects).[25]

To recognise the stable patterns in the neurons and neuronic activity in the brain as mental states, Dennett (see §3) adopts the Intentional Stance as a point of view from where these patterns can be discerned. As seen from the Intentional Stance, the mental states exist_2 as stance ontology. Similarly, to recognise the stable patterns in the universal wavefunction as classical worlds, the application of Wallace's Dennett's Criterion requires adopting the *Classical Stance* as a point of view from where the quasi-classical patterns can be discerned.

As such, the Classical Stance serves as an existence claim for macroscopic objects: the patterns in a decoherence basis are not only useful but *ontologically revealing* at the macroscopic level. As seen from the coarse-grained level (that is, taking a perspective, or stance), only the decoherence bases represent macroscopic reality at that level. Akin to Dennett's Physical Stance, we can think of a 'Quantum Stance' at the fine-grained level, which is purely wavefunction and enjoys a Hilbert space democracy of bases. As an example, consider the neutral kaons. K^0 can be in a superposition in terms of K_{long} and K_{short} (and the same for $\bar{K}^0$); alternatively, K_{long} can be in a superposition of K^0 and $\bar{K}^0$ (and the same for K_{short}). Neither is preferred, but their ontological difference is not without consequences: decay rates and decay products diverge radically.[26] As seen from the Classical Stance, however, only the classical patterns deserve an ontological status as macroscopic objects, rotating the basis sufficiently far away from a decoherence basis leads one to a representation in which there are no macroscopic objects, and thus Hilbert space democracy is swept aside for an *ancien régime classique.*

How does one flesh out the Classical Stance? Wallace presents it as something like this: a useful piece of structure derives its ontological significance through being the correct piece of structure or through fulfilling the correct role. It is not entirely clear to me whether Wallace appeals to functional role playing, or structure, or both. Yet, since functional and structural ways of speaking are often meshed together, it make sense to consider both routes.

Structural realism is the claim (in the current context) that there is a structural continuity between classical mechanics and those structures in the wavefunction that exist_2. In §4 I pointed out that the kind of pattern-matching that is implied by Wallace's instantiation criterion disregards the need for bridge laws or a conceptual matching, taking as sufficient a merely quantitative mapping between structures at the quantum level and the classical level. In this sense, one takes one's realism about classical structures and uses that as a constraint

[25] This is not a semantic dispute: for Everett, there is only the universal wavefunction and any state discerned within it is a relative state, anchored in the point of view of a part of a composite system, which the rest of the universal wavefunction is relative to. From any anchor, a branching structure emerges, but different branches do *not* have a special ontological status: Everett's branching wave functions are all on the same footing, whether in a decoherence basis or not, or at a fine-grained or a coarse-grained level.

[26] See Gerard 't Hooft (2003, Ch. 7) for the implication of the ontological differences, calling them the 'crazy kaons'.

to ontologically commit to only certain structures of quantum mechanics as real$_2$ (at the coarse-grained level).

Thus, a selective realism takes shape that reasons from the higher-level theory back to the lower-level: if 'classical' is not merely to be synonymous with 'macroscopic scale', there can be quantum structures additional to classical structural reality, but at the macroscopic level we should dismiss those that conflict with classical structure. In that case Wallace's position would have a theme in common with a position often attributed to Niels Bohr – or otherwise associated with the Copenhagen theories – which holds that we cannot speak about quantum-mechanical reality in the absence of a higher-level measuring apparatus. In both cases, classicality is regarded as a constraint on ontological commitment. For Wallace this would be a mathematical constraint, whereas for Bohr it is a constraint on language.

Functionalism is (generally) the thesis that uses a functional role to specify or individuate simultaneously many items by their patterns of relations (both to each other and to other non-problematic items). Consider the standard case of the relation between mental states such as pain and an underlying ontology of neurons in the brain playing the functional role of pain. Typically, human beings suffer pain when undergoing a particular kind of neural activity (*not* C-fibers, cf. Puccetti 1977; Aranyosi 2013) which is (through contingent scientific research) recognised as playing the causal role of pain, namely to cause signs of discomfort and to try to stop the cause of the pain. From the Intentional Stance, these role-playing patterns are real$_2$ mental states. For the quasi-classical structures to be real$_2$ classical worlds seen from the Classical Stance, one needs a specific functional role, which can be fleshed out as

> **The Classical Role**. Behaving in such a way that we obtain definite values for measurement outcomes through the deterministic laws of classical physics.

Wallace's claims about real$_2$ macroscopic objects can thus be couched in functionalist terms: the quasi-classical patterns form the ontology at a coarse-grained level precisely because they approximately fulfill the Classical Role. As explained in the previous section, this role is indeed (approximately) fulfilled by the quasi-classical patterns recognised in a decoherence basis. That is, definiteness is safeguarded by the suppression of interference, and deterministic behaviour according to classical laws is satisfied when recognising that the individual three-dimensional quasi-classical patterns by themselves sweep out trajectories that functionally mimic the trajectories of a classical system.

Note that as there is a difference between *structuralism* and *structural realism*, it may be helpful to distinguish between *functionalism* and what I would call *functional realism*. Where structuralism and functionalism are purely semantic tools for describing or defining a certain piece of structure or set of items, the realisms go beyond this semantic individuation and come with an commitment to the existence of that structure or set of items.

7. Is the price right?

Without Dennett's Criterion, the measurement problem of quantum mechanics is not solved. I have given three main philosophical preconditions needed to justify this criterion in the context of quantum mechanics. First, the rejection of conceptual/qualitative bridge laws even in the case of inhomogeneous reduction, relying on merely quantitative pattern-matching. Whether this is a problem or not lies in further evaluation of the reduction literature, and in particular the fate of Wallace's proposal of math-first structural realism (in which there is no conceptual mismatch because concepts are linguistic). Second, the reliance on usefulness to make salient a quasi-classical pattern over others, as represented in a decoherence basis. Third, the need for a Classical Stance, analogous to Dennett's Intentional Stance, which requires the patterns to fulfill the Classical Role and as such to imbue the pragmatically individuated patterns with (macroscopic, real_2) existence on top of merely being quantified over (i.e. are real_1) at a fine-grained level. I do not claim to have been exhaustive and more can (and should) be made philosophically more precise—although even further refinement is perhaps better to leave to the originators of the interpretation.

Although Dennett's Criterion is an exciting attempt to pragmatically mesh the ontic and the epistemic to solve the measurement problem, its use carries with it sufficient metaphysical baggage that makes it hard to believe that this is "just quantum mechanics itself, read literally, straightforwardly – naively, if you will – as a direct description of the physical world, just like any other microphysical theory" (Wallace 2012, p. 2). The fate of the emergent multiverse interpretation will thus be thoroughly tied up with these central and extensive debates in the philosophy of science more generally. Given that the mathematical intricacies raised in the aftermath of Everett's work are now well understood, future work on Wallace's version of the Everett interpretation should take up from this philosophical terrain.

That being said, I believe that once the background philosophy (I)-(III) is accepted, this indeed delivers the claimed ontology: real worlds as autonomous classical degrees of freedom exhibiting emergent classical behaviour. But some argue that even when Dennett's Criterion is accepted, the ontology it delivers is somehow bad. One line of argument to this end is the alleged circularity or incoherence of the decision-theoretic derivation of the Born rule (Baker 2006; Kent 2010; Dawid and Thébault 2014; Thébault and Dawid 2015). Although these authors do not argue directly against the background philosophy of Dennett's Criterion, they say that the resulting ontology of real classical worlds through decoherence is *unable* to produce a Born rule, at least not without presupposing it. And indeed: without the functional identification of quasi-classical structures with real classical worlds, there seems to be no way – *as far as the formalism goes* – to use pure wave mechanics to make contact with the empirical results, which all rely on probability distributions. However, were the philosophical preconditions of Dennett's Criterion accepted, then also everyday classically deterministic and definite observations would count as empirical support for quantum mechanics, as utilised by Franklin (2023). This functional identification is especially vivid in Fig. 1 of that work (*ibid.*, p. 9), where the middle figure depicts the orbit of Hyperion modelled by quantum mechanics, with decoherence suppressing interference effects, which is subsequently quantitatively matched to the bottom figure, where Hyperion is described by a classical chaotic model.

Certainly, the proponent of the view would argue that Dennett's Criterion is applied in science across the board, for example to establish the emergence of thermodynamic temperature from statistical physics, and a good naturalist should therefore readily accept it. Yet, in the case of temperature, no philosopher ever saw the need to explicitly write down Dennett's Criterion. This, I think, gets to the heart of the matter: although Wallace is wise to emphasise the long-neglected role of dynamics in interpretative matters, quantum theory is not – or not yet sufficiently understood to be – 'just like any other microphysical theory'. Namely, it suffers from the ontology problem and the preferred basis problem. Perhaps we can deflate these problems, and Dennett's Criterion could be a promising way to do so. As a result, the notion how exactly the quantum state represents the physical world (or indeed a world with continuously many 'worlds' within it) is not as clear-cut as elsewhere in physics. But then again: this specific stance about scientific representation may be a philosophical price worth paying.

How does one determine whether a philosophical price is worth paying? The judgement whether a philosophical addition is desirable, or even justified, warrants some discussion about our metametaphysical method. It is reasonable to assume that Wallace (2012, pp. 3–4) intends Dennett's Criterion to be justified by a version of naturalism, which, properly construed, would entail the philosophical preconditions pointed out in this paper. Indeed, in the introduction he tells us that his only (!) philosophical assumption is naturalism, understood as "[...] the doctrine that in studying science and its philosophical implications, we have no tool better than the successful practices of the sciences themselves" and then later on:

> As I noted in the Introduction, a basic premise of this book is naturalism: the thesis that we have no better guide to metaphysics than the successful practice of science. From a naturalistic perspective, we should regard the conceptual puzzles of emergence as worthy of serious study. But it would be a mistake to eschew the use of emergence in solving the quantum measurement problem until those puzzles are fully solved. (Wallace 2012, p. 54)

Metaphorically understood, surely – and rightfully – it is characteristic of our field to take scientists and their work seriously. In this perhaps rather weak sense, I endorse a form of naturalism understood as taking science as the starting point for philosophical inquiry (as long as 'science' is understood as a sufficiently broad amalgam of epistemic activities). Where possible, intuitions about how nature 'ought to be' should be rejected in favour of successful scientific practice. In this sense, philosophy is the handmaiden of science, not theology: it would be foolish for a contemporary thinker to philosophise about time travel without taking into account relativity theory; and it may be foolish even more so to study conscious experience without taking neuroscience on board; it even seems true that most philosophical mistakes of the past were made when metaphysics leapt too far ahead of epistemology. But this is not the end of the story.

Literally understood, the methods of science do not neatly map onto the methods of philosophy, nor can philosophers follow scientists in all respects. Scientists are primarily incentivised to pursue discovery, not understanding. But understanding is one of the primary goals of philosophy. There exist strong forms of naturalism that claim a strict continuity of philosophy with science, the strongest instantiations of which are (arguably) the language-first approach to ontic structural realism (Ladyman and Ross 2007) and the mathematics-first approach by Wallace himself (2022). This latter mathematics-first approach – as it stands – remains on a high level of generality and does not engage with quantum mechanics or the multiverse approach directly, but it is unsurprising that it has much in common with the philosophy behind *The Emergent Multiverse*. What is surprising is that although this later work relies on instantiation, it does not appeal to Dennett's Criterion. Even so, it is straightforward to presuppose that, partly or wholly, the implicit philosophical baggage of (Wallace 2012) is made explicit in (Wallace 2022). However, how exactly the tenets (I)-(III) are derived from the assumption of naturalism ("a basic premise of this book") remains elusive, and should perhaps be reserved for future work.

A strong form of naturalism such as the kind alluded to in the previous paragraph is bound to be contentious. There is even the question whether a fully naturalized metaphysics is even *possible* to begin with (cf. Jaksland 2016 for an extensive overview). Importantly, it is also *restrictive*, in the sense of excluding much philosophy by fiat, leading to a philosophical monism that may not be constructive in the current context. Peter Godfrey-Smith likewise dismisses the restrictive part of naturalism:

> A particular kind of freedom is important in philosophy; you never know where it is going to go next, and where the next interesting idea might come from. When naturalism falls into a constraining mode (do things *this* way and *don't* do that...), I step away from it. (Godfrey-Smith 2003, p. 326)

The title of (Wallace 2022) makes clear that his position is now *stated*, i.e., clarified. The resolution of the measurement problem along Wallacian lines may very well depend on the future *defence* of this approach in the context of a strong form of naturalism. Although an intellectual *chef-d'oeuvre*, I believe that the philosophy behind Dennett's Criterion is vulnerable—each of its moving parts is philosophically contentious, whereas each component is necessary to establish a solution to the measurement problem. Contrary to the dynamical robustness resulting from decoherence, the philosophy behind the application of Dennett's Criterion to the wavefunction is prone to wild philosophical interference, the stabilisation of which would have intense consequences for how we speak of reality.

Acknowledgements

This work is based on the results of a Master's thesis, completed at Utrecht University in August 2018, and recently updated according to recent progress in the field. I wholeheartedly thank Guido Bacciagaluppi and Ronnie Hermens for extensive supervision; Neil Dewar and Jeremy Butterfield for extensive supervision and for spurring me on to rework the thesis; Henrik Røed Sherling, Isaac Wilkins, and the rest of the Cambridge Foundations of Physics 2022-23 reading group for reading earlier versions; Helene Scott-Fordsmand for talking me through the argument; Dennis Dieks, Alexander Franklin and David Wallace for kind and illuminating correspondence, and two anonymous reviewers for their time and their sharp counterarguments. I am grateful for the support from the Tarner Scholarship in Philosophy of Science and History of Ideas, Trinity College, Cambridge.

References

Ager, Tryg A., Aronson, Jerrold L., and Weingard, Robert (1974). "Are Bridge Laws Really Necessary?" *Noûs* 8.2, pp. 119–134 (cit. on p. 14).

Aranyosi, István (June 2013). "Chapter 3: Return of the C Fibers; or, Philosophers' Lack of Nerve". *The Peripheral Mind: Philosophy of Mind and the Peripheral Nervous System*. Oxford University Press (cit. on p. 27).

Arntzenius, Frank (2012). *Space, Time, & Stuff*. Edited by Cian Sean Dorr. New York: Oxford Univ. Press (cit. on p. 12).

Bacciagaluppi, Guido (2016). "The Role of Decoherence in Quantum Mechanics". *Stanford Encyclopedia of Philosophy*. Edited by E.N. Zalta (cit. on p. 9).

Baker, David J. (2006). "Measurement Outcomes and Probability in Everettian Quantum Mechanics". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 38.1, pp. 153–169 (cit. on p. 28).

— (2014). "Review: Frank Arntzenius: Space, Time, and Stuff". *Philosophy of Science* 81.1, pp. 171–174 (cit. on p. 12).

— (2015). "The Philosophy of Quantum Field Theory" (cit. on p. 12).

Barrett, Jeffrey A. (2011). "On the Faithful Interpretation of Pure Wave Mechanics". *British Journal for the Philosophy of Science* 62.4, pp. 693–709 (cit. on pp. 5, 23).

Barrett, Jeffrey A. and Byrne, Peter (2012). *The Everett Interpretation of Quantum Mechanics*. Princeton, NJ: Princeton University Press (cit. on pp. 4, 5).

Bedau, Mark A. (1997). “Weak Emergence”. *Philosophical Perspectives* 11, pp. 375–399 (cit. on p. 14).

Bell, J. S. (1987). “La Nouvelle Cuisine”. *Speakable and unspeakable in quantum mechanics: collected papers on quantum philosophy*. Edited by John S. Bell. Cambridge University Press, pp. 232–248 (cit. on p. 5).

Blackshaw, Nadia, Huggett, Nick, and Ladyman, James (2024). “Everettian Branching in the World and of the World”. Preprint (cit. on p. 7).

Butterfield, Jeremy (2001). “Some Worlds of Quantum Theory”. *Quantum Physics and Divine Action*. Edited by R. J. Russell, N. Murphy, and C. J. Isham. Vatican Observatory Publications, pp. 111–140 (cit. on p. 8).

— (2011a). “Emergence, Reduction and Supervenience: A Varied Landscape”. *Foundations of Physics* 41.6, pp. 920–959 (cit. on p. 14).

— (2011b). “Less is Different: Emergence and Reduction Reconciled”. *Foundations of Physics* 41.6, pp. 1065–1135 (cit. on p. 14).

Butterfield, Jeremy and Gomes, Henrique (2023). “Functionalism as a Species of Reduction”. *Current Debates in Philosophy of Science: In Honor of Roberto Torretti*. Edited by Cristián Soto. Springer Verlag, pp. 123–200 (cit. on p. 17).

Byrne, Peter (2012). *The Many Worlds of Hugh Everett III: Multiple Universes, Mutual Assured Destruction, and the Meltdown of a Nuclear Family*. Oxford University Press (cit. on p. 5).

Chalmers, David J. (2006). “Strong and Weak Emergence”. *The re-emergence of emergence: the emergentist hypothesis from science to religion*. Edited by Philip Clayton and Paul Davies. Oxford University Press (cit. on p. 14).

Cooley, Kit and Lynn, Vic (2020). *Nanotechnology: Principles and Practices*. Ed-Tech Press (cit. on p. 9).

Crull, Elise M. (2021). “Chapter 13: Quantum Decoherence”. *The Routledge Companion to Philosophy of Physics*. Edited by Eleanor Knox and Alastair Wilson. Routledge (cit. on p. 5).

Davies, Paul (1984). “Inflation in the universe and time asymmetry”. 312, pp. 524–527 (cit. on p. 22).

Dawid, Richard and Thébault, Karim P. Y. (2014). “Against the Empirical Viability of the Deutsch-Wallace-Everett Approach to Quantum Mechanics”. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 47, pp. 55–61 (cit. on p. 28).

Dennett, Daniel C. (1971). “Intentional Systems”. *Journal of Philosophy* 68.4, pp. 87–106 (cit. on p. 9).

— (1981). *The Intentional Stance*. Cambridge, Massachusetts: MIT Press (cit. on p. 9).

Dennett, Daniel C. (1991). "Real Patterns". *Journal of Philosophy* 88.1, pp. 27–51 (cit. on pp. 2, 9, 10, 25).

Dewar, Neil (2019). "Supervenience, Reduction, and Translation". *Philosophy of Science* 86.5, pp. 942–954 (cit. on p. 14).

Dürr, Patrick M. (2019). "Against 'Functional Gravitational Energy': A Critical Note on Functionalism, Selective Realism, and Geometric Objects and Gravitational Energy". *Synthese* 199.S2, pp. 299–333 (cit. on p. 9).

Everett III, Hugh (1956). "The Theory of the Universal Wavefunction". *The Everett Interpretation of Quantum Mechanics*. Edited by Jeffrey A. Barrett and Peter Byrne. Princeton, NJ: Princeton University Press, pp. 72–172 (cit. on p. 4).

Franklin, Alexander (2023). "Incoherent? No, Just Decoherent: How Quantum Many Worlds Emerge". *Philosophy of Science* (cit. on pp. 13, 29).

Franklin, Alexander and Robertson, Katie (2023). "Emerging Into the Rainforest: Emergence and Special Science Ontology". Preprint (cit. on p. 13).

Godfrey-Smith, Peter (2003). *Theory and Reality: An Introduction to the Philosophy of Science*. Chicago: University of Chicago Press (cit. on p. 30).

Gomes, Henrique and Butterfield, Jeremy (2022). "Geometrodynamics as Functionalism About Time". *From Quantum to Classical. Fundamental Theories of Physics*. Edited by C. Kiefer. Springer (cit. on p. 17).

Gopnik, Alison (2000). ""Explanation as orgasm and the drive for causal understanding: The evolution, function and phenomenology of the theory-formation system."". *Cognition and explanation*. Vol. 8. Cambridge, MA: MIT Press, pp. 299–323 (cit. on p. 22).

Haugeland, John (1998). "Chapter 11: Pattern and Being". *Having Thought: Essays in the Metaphysics of Mind*. Cambridge, Mass.: Harvard University Press, pp. 267–291 (cit. on p. 9).

Hemmo, Meir and Shenker, Orly (2022). "The Preferred Basis Problem in the Many-Worlds Interpretation of Quantum Mechanics: Why Decoherence Does Not Solve It". *Synthese* 200.3, pp. 1–25 (cit. on pp. 3, 9, 23, 24).

Hempel, Carl Gustav (1966). *Philosophy of Natural Science*. Englewood Cliffs, N.J., Prentice-Hall (cit. on p. 14).

Hill, Christopher, ed. (1994). *The Philosophy of Daniel Dennett*. University of Arkansas Press (cit. on p. 9).

Hooft, Gerard 't (2003). *In Search of the Ultimate Building Blocks*. Cambridge: Cambridge University Press (cit. on p. 26).

Ismael, Jenann and Schaffer, Jonathan (2020). "Quantum Holism: Nonseparability as Common Ground". *Synthese* 197.10, pp. 4131–4160 (cit. on p. 12).

Jacobs, Caspar (2024). "Stating Maths-First Realism, or: How to Say Things with Models". *Preprint* (cit. on p. 3).

Jaksland, Rasmus (2016). *The Possibility of Naturalized Metaphysics*. Ph.D. Dissertation, University of Copenhagen (cit. on p. 30).

Janssen, Hanneke (2008). *Reconstructing Reality: Environment-Induced Decoherence, the Measurement Problem, and the Emergence of Definiteness in Quantum Mechanics*. Master Thesis, Radboud University Nijmegen, supervised by Jos Uffink and Klaas Landsman (cit. on p. 9).

Joos, Erich and Zeh, Heinz-Dieter (1985). "The Emergence of classical properties through interaction with the environment". *Zeitschrift für Physik B, Condensed Matter* 59, pp. 223–243 (cit. on p. 8).

Joos, Erich and Zeh, Heinz-Dieter; Kiefer Claus; Domenico Giulini; Joachim Kupsch; Ion-Olimpiu Stamatescu (2003). *Decoherence and the Appearance of a Classical World in Quantum Theory*. Heidelberg: Springer (cit. on p. 8).

Kent, Adrian (2010). "One World Versus Many: The Inadequacy of Everettian Accounts of Evolution, Probability, and Scientific Confirmation". *Many Worlds?: Everett, Quantum Theory, & Reality*. Edited by Simon Saunders et al. Oxford University Press (cit. on p. 28).

Knox, Eleanor and Wallace, David (2024). "Functionalism Fit for Physics". Preprint (cit. on pp. 3, 17).

Ladyman, James and Ross, Don; Don Spurrett; John Collier (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford University Press (cit. on pp. 9, 30).

Lewis, David K. (1970). "How to Define Theoretical Terms". *Journal of Philosophy* 67.13, pp. 427–446 (cit. on p. 17).

— (1994). "Reduction of Mind". *Companion to the Philosophy of Mind*. Edited by Samuel Guttenplan. Blackwell, pp. 412–431 (cit. on p. 17).

Lewis, Peter J. (2013). "Review of David Wallace, The Emergent Multiverse: Quantum Theory according to the Everett Interpretation". *Notre Dame Philosophical Reviews* (cit. on p. 12).

— (2016). *Quantum Ontology: A Guide to the Metaphysics of Quantum Mechanics*. New York, NY: Oxford University Press USA (cit. on p. 9).

Lorenzetti, Lorenzo (2024). "Two Forms of Functional Reductionism in Physics". *Synthese* 203.2 (cit. on p. 17).

Maudlin, Tim (2018). "Ontological clarity via canonical presentation: electromagnetism and the Aharonov- Bohm effect". *Entropy* 20, p. 465 (cit. on p. 5).

McQueen, Kelvin J. (2015). "Four Tails Problems for Dynamical Collapse Theories". *Studies in the History and Philosophy of Modern Physics* 49, pp. 10–18 (cit. on p. 9).

Millhouse, Tyler (2022). "Really Real Patterns". *Australasian Journal of Philosophy* 100.4, pp. 664–678 (cit. on p. 9).

Misak, Cheryl (2013). *The American Pragmatists*. Oxford, England: Oxford University Press (cit. on p. 19).

Møller-Nielsen, Thomas (2016). "Weak Discernibility, Again". *Ergo: An Open Access Journal of Philosophy* 3 (cit. on p. 9).

Mulder, Ruward A. (2021). "Gauge-Underdetermination and Shades of Locality in the Aharonov-Bohm effect". *Foundations of Physics* 51.2, pp. 1–26 (cit. on p. 5).

Muller, F.A. (1992). "On the Principle of Relativity". *Plenum Publishing Corporation* 5.6 (cit. on p. 20).

— (2023). "Six Measurement Problems of Quantum Mechanics". *Non-Reflexive Logics, Non-Individuals, and the Philosophy of Quantum Mechanics: Essays in Honour of the Philosophy of Décio Krause*. Edited by Jonas R. B. Arenhart and Raoni W. Arroyo. Springer Verlag, pp. 225–259 (cit. on p. 4).

Nagel, Ernest (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York, NY, USA: Harcourt, Brace & World (cit. on p. 14).

Newey, Simon M.T. (2019). *Decoherence and its Role in Interpretations of Quantum Physics*. Ph.D. dissertation, University of Leeds (cit. on p. 9).

Ney, Alyssa (2013). "Ontological Reduction and the Wave Function Ontology". *The Wave Function: Essays on the Metaphysics of Quantum Mechanics*. Edited by Alyssa Ney and David Z. Albert. Oxford University Press, pp. 168–183 (cit. on p. 9).

Norsen, Travis (2017). *Foundations of Quantum Mechanics: An Exploration of the Physical Meaning of Quantum Theory*. Cham: Imprint: Springer (cit. on p. 5).

Puccetti, Roland (1977). "The Great C-Fiber Myth: A Critical Note". *Philosophy of Science* 44.2, pp. 303–305 (cit. on p. 27).

Putnam, Hilary, ed. (1979). *Philosophical Papers: Volume 1, Mathematics, Matter and Method*. New York: Cambridge University Press (cit. on p. 22).

Quine, Willard van Orman (1969). "Epistemology Naturalized". *Ontological Relativity and Other Essays*. Edited by Willard van Orman Quine. Columbia University Press (cit. on p. 9).

Rosaler, Joshua (2015). "'Formal' vs. 'Empirical' Approaches to Quantum-Classical Reduction". *Topoi* 34.2, pp. 325–338 (cit. on p. 15).

Ross, Don, Brook, Andrew, and Thompson, David, eds. (2000). *Dennett's Philosophy: A Comprehensive Assessment*. MIT Press (cit. on p. 9).

Saunders, Simon (1993). "Decoherence, Relative States, and Evolutionary Adaptation". *Foundations of Physics* 23.12, pp. 1553–1585 (cit. on p. 23).

Schaffner, Kenneth F. (2006). "Reduction: The Cheshire Cat Problem and a Return to Roots". *Synthese* 151.3, pp. 377–402 (cit. on pp. 14, 15).

Swanson, Noel (2020). "How to Be a Relativistic Spacetime State Realist". *British Journal for the Philosophy of Science* 71.3, pp. 933–957 (cit. on p. 12).

Thébault, Karim P.Y. and Dawid, Richard (2015). "Many Worlds: Decoherent or Incoherent?" *Synthese* 192.5, pp. 1559–1580 (cit. on p. 28).

Van Fraassen, Bas (1977). "VII. The Pragmatics of Explanation". *Philosophy of Science: Contemporary Readings*. Edited by Yuri Balashov and Alexander Rosenberg. Routledge, p. 56 (cit. on p. 21).

Van Fraassen, Bas C. (1980). *The Scientific Image*. New York: Oxford University Press (cit. on p. 21).

Wallace, David M. (2001). "Worlds in the Everett Interpretation". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 33.4, pp. 637–661 (cit. on p. 23).

— (2003). "Everett and Structure". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 34.1, pp. 87–105 (cit. on p. 2).

— (2010). "Decoherence and Ontology, Or: How I Learned to Stop Worrying and Love Fapp". *Many Worlds? Everett, Quantum Theory, and Reality*. Edited by Simon Saunders et al. Oxford University Press (cit. on pp. 2, 19).

— (2011). "Decoherence and its Role in the Modern Measurement Problem". *Philosophical Transactions of the Royal Society A* 370, pp. 4576–4593 (cit. on p. 5).

— (2012). *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation*. Oxford: Oxford University Press (cit. on pp. 2, 5, 6, 8, 9, 12–14, 18, 21, 28–30).

— (2017). "Against Wavefunction Realism". *Current Controversies in the Philosophy of Science*. Edited by B. Weslake and S. Dasgupta. Routledge (cit. on p. 25).

— (2020). "On the Plurality of Quantum Theories: Quantum Theory as a Framework and its Implications for the Quantum Measurement Problem". *Scientific Realism and the Quantum*. Edited by Steven French and Juha Saatsi. Oxford University Press (cit. on p. 25).

— (2022). "Stating Structural Realism: Mathematics-First Approaches to Physics and Metaphysics". *Philosophical Perspectives* 36.1, pp. 345–378 (cit. on pp. 3, 14, 17, 18, 30).

Wallace, David M. and Timpson, Christopher G. (2010). "Quantum Mechanics on Spacetime I: Spacetime State Realism". *British Journal for the Philosophy of Science* 61.4, pp. 697–727 (cit. on p. 12).

Zeh, Heinz-Dieter (1970). "On the interpretation of measurement in quantum theory". *Foundations of Physics*, pp. 69–76 (cit. on p. 5).

Żurek, Wojciech H. (2002). "Decoherence and the Transition from Quantum to Classical-*Revisited*". *Los Alamos Science* 27 (cit. on p. 19).