The Unreasonable Effectiveness of Decoherence

Davide Romano (dromano@letras.ulisboa.pt)

Centre of Philosophy, University of Lisbon

Paper draft for the volume: *Quantum Mechanics and Fundamentality* (V. Allori, editor), Synthese Library.

Draft: May 30th, 2021

Abstract

This paper aims to clarify some conceptual aspects of decoherence that seem largely overlooked in the recent literature. In particular, I want to stress that decoherence theory, in the standard framework, is rather silent with respect to the description of (sub)systems and associated dynamics. Also, the selection of position basis for classical objects is more problematic than usually thought: while, on the one hand, decoherence offers a pragmatic-oriented solution to this problem, on the other hand, this can hardly be seen as a genuine ontological explanation of why the classical world is position-based. This is not to say that decoherence is not useful to the foundations of quantum mechanics; on the contrary, it is a formidable weapon, as it accounts for a realistic description of quantum systems. That powerful description, however, becomes manifest when decoherence theory itself is interpreted in a realist framework of quantum mechanics.

1. Introduction: decoherence theory

Decoherence theory is the best answer to the classical limit problem, i.e. the problem of deriving the classical world (classical systems obeying Newtonian mechanics) at the macroscopic regime from the microscopic quantum world (quantum systems obeying the Schrödinger equation). However, there is no consensus in the literature on what has been really achieved by the results of decoherence. On the one hand, Schlosshauer (2007) and Crull (2015, 2019) argue that decoherence explains the emergence of classical objects by providing a robust description of emergent well-localized systems in the position basis. On the other hand, Ballentine (2008) and Okon & Sudarsky (2016) argue (even if for different reasons) that the program of decoherence is essentially flawed and does not help in the job to connect quantum with classical mechanics. In this paper I will not take position for one of the two sides; instead, I will focus on three conceptual problems that have been largely overlooked in the debate so far and that emerge every time we attempt to interpret (or un-interpret) decoherence in standard quantum mechanics: (i) the lack of a precise characterization for the status of subsystems (**sect. 2**); (ii) the lack of an objective dynamics for classical objects from decoherence (**sect. 3**) and (iii) the selection of position basis as the privileged basis of classical systems (**sect. 4**).

In the following, I will briefly introduce the formalism of decoherence theory.[1]
Decoherence theory starts when a system $|S\rangle$ interacts with a second system $|E\rangle$, generally called "the environment".[2] Assuming, for simplicity, that S is a superposition of two states: $|S\rangle = c_1|S_1\rangle + c_2|S_2\rangle$ and that S- relative states get correlated with E-relative states in the S-E interaction:

[1] For a comprehensive presentation of the formalism of decoherence theory, see Schlosshauer (2007, 2019).

[2] The "environment" can be generally thought of as external or internal degrees of freedom with respect to the degrees of freedom representing our system of interest. Spatial degrees of freedom (position coordinates) may be, for example, "the environment" for spinor degrees of freedom of a spin ½ particle (spin-up, spin-down).

$|S_1\rangle|E\rangle \xrightarrow{\hat{H}_{int}} |S_1\rangle|E_1\rangle$
$|S_2\rangle|E\rangle \xrightarrow{\hat{H}_{int}} |S_2\rangle|E_2\rangle$

then the interaction eventually leads to the system-environment entangled state:[3]
$|S\rangle|E\rangle = (c_1|S_1\rangle + c_2|S_2\rangle)|E\rangle \xrightarrow{\hat{H}_{int}} c_1|S_1\rangle|E_1\rangle + c_2|S_2\rangle|E_2\rangle = |S, E\rangle.$

However, what we are really interested in decoherence theory is the description and behavior of the (initial) systems *S*, now subsystem of the larger entangled state $|S, E\rangle$, under the action of the "external" environment. We can think of a table (*S*) scattered by surrounding air molecules (*E*) or the famous Schrödinger's cat (*S*) interacting with an unfortunate chain of events (radioactive atom, Geiger counter, relais, hammer, poison vial: all collectively described by *E*). Since in quantum mechanics we cannot assign a state vector (or a wave function) to subsystems, we represent the subsystem *S* by the reduced density matrix:[4]

$\hat{\rho}_S = Tr_E(\hat{\rho}_{SE}),$

which is connected to the usual quantum distribution of eigenvalues by the trace rule:

$Tr(\hat{\rho}_S \hat{O}_S) = \langle \hat{O}_S \rangle.$

where $\hat{\rho}_S$ is the *S*-subsystem density matrix, $\hat{O}_S$ an observable acting "locally" on the Hilbert space on the subsystem *S* (we can think of a property belonging only to *S*, such as the position of electrons on the final screen of a double-slit experiment interacting with light photons between the silts and the screen).The dynamics of $\hat{\rho}_S(x, x', t)$ will make the reduced density matrix diagonal in a very short time (this process is quasi-instantaneous for decoherence models at macroscopic scale), which, in turn, will make quantum interference between different components impossible to detect in a measurement. For example, when decoherence is induced by scattering of environmental particles on a system *S*, the evolution of the *S*-subsystem reduced density matrix (in the long wavelength limit) will be:

$\rho_S(x, x', t) = \rho_S(x, x', 0)e^{-\Lambda(x-x\prime)^2}$

where $\rho_S(x, x', 0)$ is the reduced density matrix at the initial time ($t = 0$), Λ a constant of the model (the *scattering constant*) and $(x - x')$ the spatial separation between two points of the subsystem *S*. The term $\Lambda(x - x')^2$ plays thus the role of an exponential rate, and $\tau = \frac{1}{\Lambda(x-x\prime)^2}$ is the characteristic decoherence time of the model. As a result, the S-subsystem reduced density matrix is progressively diagonalized at a quadratic exponential rate. After very short time, ρ_S will become (approximately) diagonal, and only relative components that describe well-defined states will "survive" the dynamical process (for example: only a well-defined alive cat *and* a well-defined dead cat):[5]

[3] Technical note: the coefficients of the entangled state superposition will generally be different from those of the initial *S* state superposition. However, this difference will not be relevant for the present discussion.

[4] Strictly speaking, this is a *density operator*, while the *density matrix* is the density operator expressed in a particular basis (generally in the position basis). However, as this difference will not be relevant, I will just use the term density matrix in both cases.

[5] The term ε in the off-diagonal components of the matrix stands for "negligible quantity": as the diagonalization process is mathematically described by a decreasing (quadratic) exponential, it will reach the zero value only asymptotically.

$$\hat{\rho}_S = \begin{vmatrix} |c_1|^2 & \varepsilon \\ \varepsilon & |c_2|^2 \end{vmatrix} \qquad [5]$$

where ε on the off-diagonal components represents a "negligible quantity" and $|c_1|^2$ and $|c_2|^2$ are the probabilities to obtain, respectively, the observables' eigenvalues associated to the eigenstates $|S_1\rangle$ and $|S_2\rangle$ when a measurement is performed on the subsystem S of the total entangled state $|S, E\rangle$.[6] We finally arrived at the central result of decoherence: open systems, i.e. systems in interaction with the environment, will be effectively described by *diagonal reduced density matrices*, generally interpreted as *improper mixtures of well-localized states*. I will analyze and (critically) discuss this important conclusion in the next section.

2. **The (lack of) description for subsystems**

2.1 **Decoherence does not transform "pure states into mixtures of well-localized states"**

Eq. (5) represents a diagonal reduced density matrix and it is generally called an "improper mixture": even though the resulting matrix looks like a mixed-state density matrix (a density matrix computed for a mixed state, i.e. a classical sum of pure states), it does not actually represent a genuine mixed state, as in this case all the diagonal components of the matrix are equally real[7]. instead, in the case of a mixed-state density matrix, the diagonal components represent classical epistemic probabilities over the real state of the system, and therefore only one component represents the "real state" (a pure state) of the system. One of the principal results of decoherence is thus expressed as follows:

Claim: *decoherence transforms pure states into improper mixtures of well-localized states*

However: while the definition of an improper mixture is pretty clear, i.e. it is a shorthand to say that a given superposition state looks like a real mixture (a mixed-state) when represented in a density matrix, we may ask: why does the subsystem *S* should be described by (an improper mixture of) *well-localized states*? What is it in the formalism of decoherence that justifies this claim?
Usually, this claim follows from the consideration that, as the reduced density matrix becomes diagonal, the coherence length (the distance over which quantum interference between different components of the superposition can be detected) shrinks, approaching negligible values in a very short time. As we cannot detect quantum interference between appreciably distant regions in space, it follows that the reduced density matrix represents the system being a (improper) mixture of well-localized states. However, the problem with this argument is that, in the standard context, it is not possible to assign wave functions to subsystems of an entangled state. That is: in decoherence theory, it is not possible to assign a wave function to the system interacting with the environment, as this is a subsystem of the total system-environment entangled state. This is indeed the very first reason why we describe subsystems via reduced density matrix in decoherence. So, the question is: how can we claim that the subsystem *S* is a well-localized state (or an improper mixture thereof) if we cannot even assign to *S* a state vector or a wave function? This seems to be physically illegitimate at best, and conceptually wrong.

Consider, for example, the Brownian model discussed by Zurek et al. (2003): a system represented by a quantum harmonic oscillator in interaction with a thermal bath of harmonic oscillators at a constant temperature T. After having analyzed the dynamics of ρ_S (in particular, the evolution of the linear entropy), the authors argue that the pointer state of the model (the specific state that

[6] While this is a natural interpretation of the reduced density matrix, it could be interesting to ask whether a reduced density matrix may have a more general significance independently from measurement interactions.
[7] As the dynamics of $\rho_S(x, x', t)$ is linear, it cannot eliminate any state of the superposition (see, e.g., Adler (2003)).

remains stable in the interaction with the environment) is characterized by position and momentum uncertainties identical to the ground state of a quantum harmonic oscillator. From this mathematical result, it is suggested that the subsystem pointer state *is t*he ground state of the quantum harmonic oscillator, i.e. a Gaussian wave packet. However, while the results concerning the subsystem linear entropy and the position-momentum uncertainties are physically relevant, the conclusion that the subsystem is described by a Gaussian wave packet is definitely an illegitimate step, as it involves a physical representation (assigning a state vector to the subsystem) that goes beyond standard quantum mechanics.

To sum up:

- the so-called improper mixtures are not mixtures at all: they represent *coherent superpositions of non-interfering relative components* of the subsystem of a larger entangled state;
- the well-localized wave packets resulting from the decoherence are not "wave packets" at all, as in quantum mechanics we cannot assign wave functions (so, wave packets) to subsystems.

Concretely: we cannot think of the decoherence process in standard quantum mechanics as a mechanism that separates the initial entangled state into different non-overlapping components and that re-assigns a wave function to these components. This operation cannot be done in the standard context until a measurement is performed: only a measurement interaction with a macroscopic device[8] makes the initial (system-environment-measurement device) entangled state collapse into one of the decohered components (eigenstates of the measured observable), with probability given by the usual Born rule. The unique eigenstate of the measured observable selected in the measurement will be eventually assigned a wave function. We are thus led to the following:

Counterclaim: *Decoherence in standard quantum mechanics does not transform pure states into mixtures (of any sort) of well-localized states but makes it impossible to detect quantum interference between distant spatial components of the subsystem if we perform a local measurement on the subsystem*. *The subsystem is a superposition of non-overlapping components in configuration space.*

2.2 Undesired consequences: the status of classical objects in decoherence theory

The fact that subsystems cannot be generally described by a state vector/wave function is an ontological problem often overlooked in the philosophy of decoherence. This poses a problem not only for the ontology of any realistic microscopic system but also (and, maybe, most importantly, insofar as we are looking for a clear account of the quantum to classical transition) for the *status of classical objects*. It is standardly accepted that classical objects emerge in decoherence theory as macroscopic systems (systems composed of many degrees of freedom) in a strong interaction with an external (or internal) environment, i.e. when $H_{int} \gg H_S + H_E$. By definition, classical objects are subsystems maximally entangled with the environment and mathematically described by perfect diagonalized reduced density matrix. Furthermore, they must provide classical numerical values for observables such as position, energy and momentum when "locally" measured via $\langle O_S \rangle = Tr_E(O_S \rho_S)$ [more on this in sect. 3]. The problem is that, from the ontological point of view, a classical object is difficult to characterize if on the one hand, *the subsystem reduced density matrix is the only tool we have at our disposal* and, on the other hand, the ontology of a quantum system

[8] Note that the system-environment interactions in decoherence theory are measurement-like interactions, as there is a coupling between the system's and environment's degrees of freedom and the environment relative states get correlated with the system relative states. However, these interactions do not produce collapse of the wave function, as the master equations are linear and do not select one particular component. In order to have collapse of the wave function, the system has to interact—by definition--with a macroscopic measurement device.

usually refers to the wave function and not to the density matrix or, even worse, to the reduced density matrix. It is worth noting indeed that, in this context, the usual debate on the nature of the wave function in quantum mechanics becomes completely irrelevant for the metaphysical characterization of realistic quantum systems *and* emergent classical objects, as the latter are, by definition, macroscopic quantum systems entangled with the environment.

Does it mean that classical objects cannot be clearly defined within quantum mechanics? Certainly not. To my view, this is only a sign that the standard interpretation of quantum mechanics cannot conceptually explain what decoherence achieves from the pragmatic/empirical point of view. The empirical predictions of decoherence clearly indicate that subsystems after having interacted with the environment do become independent/autonomous systems. But quantum mechanics, in the standard framework, is unable to provide such an account. And this is not a specific problem taking place at the macroscopic regime, examples can be given also at the microscopic one. Think, for example, at the description of atomic orbitals: the explanation of why we do not see a superposition of orbitals but only definite orbitals is again provided by decoherence theory:[9] interaction between different orbitals leads to decoherence in the energy basis (in the microscopic regime $H_S \gg H_E + H_{int}$). This process selects well-defined separate components in the energy basis: atomic orbitals. Strictly speaking, atomic orbitals should not be described by wave functions, but only by reduced density matrices. However, in practice we do assign a wave function to each atomic orbital and the fact that we are able to assign such a wave function to a state resulting from decoherence is not a sign that decoherence does not work (it works perfectly); instead *it is a sign that the standard interpretation cannot do all the job in the interpretation of decoherence*. Subsystems selected by decoherence are, practically speaking, systems that become independent from the total entangled state. And so are classical objects. But this fact is simply not accounted for in standard quantum mechanics, where subsystems are described by density matrices even after the decoherence process. Indeed, the idea to assign wave functions to subsystems of entangled states goes beyond standard quantum mechanics, and can be described more precisely by non-standard interpretations of quantum mechanics: the Everett/Many Worlds Interpretation (MWI), the de Broglie—Bohm theory (dBB) and the spontaneous collapse models (e.g. GRW, CSL). This point will be briefly analyzed in the next subsection (while a complete analysis would require a separate work).

2.3 Interpreting decoherence

The problem of the characterization of subsystems, which is not solved and in principle unsolvable in the standard interpretation of quantum mechanics,[10] is actually solved when decoherence theory is embedded in a realist interpretation of quantum mechanics: MWI, GRW or Bohm's theory.

In **GRW theory,** the spontaneous collapse process is represented through the multiplication of the ordinary wave function (solution of the Schrödinger equation) by a Gaussian wave packet whose center is random and $|\psi|^2$- distributed. An important feature of this theory is that the spontaneous collapse rate is amplified for an N-particle system. That is: while the standard collapse frequency for a 1-particle system is $\lambda = 10^{-16}s$, the collapse frequency for an N-particle system will be $\Gamma = N\lambda$. For a macroscopic system, i.e. a system composed of a number of particles of the order of the

[9] See e.g. Crull (2015).

[10] It is worth noting, however, that in some recent (more refined) interpretations of the quantum formalism this problem does not arise or could not arise, for example: Rovelli's account of decoherence in relational quantum mechanics (Di Biagio & Rovelli 2021); Myrvold's ontology of quantum states in terms of distribution of values of dynamical variables (Myrvold 2018); Chen's density matrix realism (Chen 2018). In all of these interpretations, the problem discussed above requires a careful and distinct analysis, which will be developed in an extended version of the present paper.

Avogadro's number (10^{23}), the collapse rate will be very effective and the system will be "always" well-localized (as far as we can reasonably detect at the macroscopic scale and for classical relevant timescales). The significance of decoherence is thus clear in this context: the interaction with the environment is the physical mechanism which largely increases the number of degrees of freedom of a GRW system, making the spontaneous collapse very effective via the amplification mechanism. A 1-particle microscopic system $S(x)$ in interaction with the environment $E(y)$ will be part of an entangled state: $\Psi_{SE}(x, y)$. As the number of interactions with the surrounding environment will increase (at least) linearly in time (for collisional models and quantum Brownian motion models), the system will undergo a spontaneous collapse after short time (the precise time depending on the interaction and environment model, but it is reasonable to expect the GRW spontaneous collapse rate to approach the standard decoherence time very quickly). This eventually explains why, in the GRW theory, microscopic "decohered" subsystems (such as atomic orbitals) and ordinary classical systems are well-defined states to which a wave function can be unproblematically assigned. The significance of decoherence theory in GRW is therefore not related to the decoherence effect, i.e. the separation of components in configuration space. Strictly speaking, there is no decoherence process in GRW theory, as the dynamics of the theory is non-linear. Nevertheless, the interaction with the environment supplements the GRW system with additional degrees of freedom, which trigger the amplification mechanism of the collapse rate and make open quantum systems (so, by definition, classical systems) (quasi-)instantaneously and (quasi-)perfectly well-localized.[11]

In **MWI and dBB theory**, decoherence effect is the result of the separation of the initial system-environment entangled state into effectively separate components[12] in configuration space. As, in the case of GRW theory, the collapse is more likely to happen for systems with a high number of degrees of freedom, analogously, in MWI e DBB theory, the separation between components in configuration space is more likely to happen when the configuration space becomes "large enough". That is: when the number of degrees of freedom of the entangled state increases, also increases the possibility that different relative environmental states have negligible overlap with each other.[13]
The separation of different components in configuration space is the well-known *branching process* in MWI (e.g., Wallace 2012) and the less well-known (but equally important) *effective factorization* process in dBB theory (Bohm & Hiley (1987), DGZ (1992); Holland (1993)), i.e. the physical process that permits to assign an *effective wave function*[14] to subsystems.

3. No dynamics for the subsystem S – Ehrenfest theorem

A second aspect which is generally overlooked in the presentation of the classical limit via decoherence is the relation between quantum and classical dynamics. A classical object is not only a system with a well-defined position in three-dimensional space; it also moves in space according

[11] But instantaneously well-localized for relevant timescales at the macroscopic regime and perfectly well-localized with respect to macroscopic localization.
[12] Separate components stand here for "components whose overlap is negligible in configuration space", since the condition of no-overlap is not realistic: tails of Gaussian environmental particles will overlap in any region of space.
[13] Even though there is no rigorous formulation, this is not so different from the standard decoherence condition of "orthogonality" of states, that in most cases is reached approximately and asymptotically. For an analysis of the decoherence condition in dBB theory, see Romano (2016a). For a comparative analysis of that condition in dBB theory and MWI, see Rosaler (2015, 2016).
[14] The notion of *effective wave function* is originally introduced and defined in Bohm & Hiley (1987).

to Newtonian dynamics.[15] That is: when a classical system with position *x* is affected by a classical (gravitational[16]) potential $V(x)$, it will accelerate according to Newton's second law of dynamics:

$F = m\ddot{x}$ | Newton's 2nd law: an object with position x accelerates due to the force F

where $F = -\nabla_x V(x)$ is a classical gravitational force generated by the potential $V(x)$, generated (in turn) by the existence of a physical gravitational field filling up three-dimensional space (at least, the finite region of space over which the gravitational field is non-negligible and well-defined). Newtonian dynamics, i.e. the fact that a classical object evolves in time according to a Newtonian trajectory when affected by classical potentials, is a not a secondary feature but a central one for the correct representation of a classical object. So, the question is: how does decoherence theory account for this important feature of classical mechanics? Stating the obvious: without a precise account of Newtonian dynamics in the macroscopic regime, all we have achieved from the quantum to classical transition via decoherence is a (quantum) object well-localized in space mathematically represented in the position basis. This is an excellent starting point, of course, but not the whole story: for an object to be classical, it has to evolve in time according to (approximate)[17] Newtonian trajectories. The connection between decoherence theory and the subsystems' dynamics is given by a generalized form of the Ehrenfest theorem. This can be easily shown in two steps:

1. **Trace rule**: we can compute the mean value of any observable acting on the subsystem of a larger entangled state using the reduced density matrix ρ_S through the trace rule:

$$Tr\left(\hat{\rho}_S \hat{O}_S\right) = \langle \hat{O}_S \rangle$$

where $\hat{\rho}_S$ is the subsystem reduced density matrix and $\hat{O}_S$ a Hermitian operator acting "locally" on the subsystem's Hilbert space.

2. **Ehrenfest theorem for open systems**: we can use the trace rule defined in step (1) to compute the mean value of the momentum, the acceleration and the "force" operators:

Ehrenfest theorem:

$$m\frac{d}{dt}\langle \hat{x} \rangle = \langle \hat{p} \rangle$$
$$\frac{d}{dt}\langle \hat{p} \rangle = -\langle \nabla V(\hat{x}) \rangle$$

[15] Or Lagrangian mechanics, or Hamiltonian mechanics. However, as the ontology of classical systems (system dynamics and interaction between systems) is generally built on Newtonian mechanics, I will consider the quantum to classical dynamics transition as the transition from quantum dynamics to Newtonian mechanics.

[16] I consider here only the gravitational potential since, strictly speaking, there is no electromagnetic field in non-relativistic quantum mechanics. Classical electrodynamics should emerge from quantum electrodynamics, i.e. a different mathematical and physical framework. It is true that electromagnetic interactions are described also in non-relativistic quantum mechanics, e.g. the textbook presentation of the proton-electron interaction in a Hydrogen atom. However, this kind of analysis is rather phenomenological and relies on semi-classical assumption (particles described by wave functions interacting through classical electromagnetic forces). On the other hand, there is no gravitational interaction either in non-relativistic quantum mechanics, as a real quantization of the gravitational field is only done in quantum gravity. Yes, any realistic description in non-relativistic quantum mechanics is trickier than usually thought.

[17] We want to derive Newtonian trajectories approximately and not exactly, as any deviation of the order e.g. of the atomic scale cannot be detected at the macroscopic scale.

where $\langle \hat{x} \rangle = Tr(\hat{\rho}_S \hat{x})$ and $\langle \hat{p} \rangle = Tr(\hat{\rho}_S \hat{p})$, respectively. When the mean value of the potential can be reasonably approximated with the potential of the mean value of the position operator,[18] the second equation above becomes:

"Newtonian" dynamics

$$\frac{d}{dt}\langle \hat{p} \rangle = -\langle \nabla V(\hat{x}) \rangle \cong -\nabla V \langle \hat{x} \rangle$$

which is generally taken to describe a quantum system moving along a Newtonian trajectory.
From an instrumentalist/pragmatic point of view, this is correct: the mean value of the acceleration operator will be Newtonian-like distributed. However, *this formula is not a description of an object moving in space according to a specific Newtonian trajectory*. The mean value of an operator is a statistical quantity (e.g. Ballentine (1998, Ch. 2); Bowman (2008, Ch. 5))--a quantity computed after a (reasonably) long sequence of separate measurements on identically prepared systems[19]--which is ontologically meaningless for the description of the dynamics of an individual system.
The fact that the subsystem of interest does not have a state vector does not help in this situation either, as we can represent the classical system only by means of the reduced density matrix. The problem of finding or recovering Newtonian dynamics in the macroscopic regime combines together two conceptual problems already present in standard quantum mechanics:

- the definition of the physical state for the subsystem of a larger entangled state;
- the definition of the subsystem's dynamics independently from observables (which implicitly involve the use of ensembles, statistical distributions and measurement operations).

To conclude: the Ehrenfest theorem (in the standard or generalized form) does not describe individual systems moving on Newtonian trajectories: it provides, instead, a statistical description which can be consistently applied only to ensembles. No account is provided for trajectories of individual systems. What we obtain from the Ehrenfest theorem is not a classical object moving according to Newtonian mechanics, but a discrete sequence in time of position and momentum eigenvalues which are formally compatible with a Newtonian trajectory. Ontologically speaking, though, Newtonian trajectories are not there: they simply cannot be described (neither exactly nor approximately) in the standard framework of quantum mechanics, even introducing decoherence.

4. **The basis problem: does decoherence really help?**

A further important problem of the quantum to classical transition is to explain why the familiar classical world is *position-based*, i.e. why position is a privileged variable for classical objects, starting from the relative freedom we have in the basis representation of quantum systems. The fact that position is a privileged variable in the classical world comes from the evidence that all classical objects have a well-defined position in space and they are described by continuous trajectories. On the contrary, a quantum system can be generally be described in different bases (e.g. position, momentum, energy), the specific choice being usually taken for pragmatic reasons, such as the eigenbasis of the observable we want to measure in a designed experiment or the energy basis for the particle in a box. Thus, we can frame the problem of the basis as follows:

[18] This approximation turns out to be exact for linear and quadratic potentials (Shankar 1994) insofar the wave function is well-localized (Wallace 2012).
[19] As these are notions we are already familiar from the basic course in quantum mechanics, I will not enter here in the detail of the difference between a statistical quantity and a quantity describing an individual system. For the interested reader, however, I have analyzed this issue more carefully in another paper (Sakurai, sect. 3).

Basis problem: *why is position basis a privileged basis in the classical world, starting from the physical equivalence of the basis representation in quantum mechanics*?

The standard answer given by decoherence is that, when the interaction between the system and the environment is very strong, the only states that remain stable under the action of decoherence are position eigenstates.[20] In turn, this leads to diagonalization of the reduced density matrix in that particular basis, and, therefore, decoherence will select subsystem's states *well-localized in position.*

Decoherence claim: *The classical world is position-based because of decoherence. In particular, the continuous interaction between a system with its surrounding environment selects position as the privileged basis for the representation of that system in the macroscopic regime*.

This explanation, however, does not consider one particular fact: all decoherence models of the quantum to classical transition describe the interaction between the system and the environment as a function of the position coordinates. That is, if we consider a system $S(x)$ and an external environment $E(y)$, the interaction Hamiltonian (which describes the physical coupling between the two) will be of the form: $\widehat{H}_{int}(x, y)$. Now:

- as the basis "selected by decoherence" is the basis of eigenstates of those operators ($\hat{O}$) that commute with the total Hamiltonian: $\left[\widehat{H}, \hat{O}\right] = \widehat{H}\hat{O} - \hat{O}\widehat{H} = 0$,
- given that in the macroscopic regime the total Hamiltonian can be approximated with the interaction Hamiltonian $\widehat{H}_{tot} \cong \widehat{H}_{int}$, since $\widehat{H}_{int} \gg \widehat{H}_S + \widehat{H}_E$
- and since $\widehat{H}_{int}(x, y)$ already describes interactions between the system and its surrounding environment in position basis,
- *then* position basis is selected: $\left[\widehat{H}_{int}, \hat{X}\right] = 0$

This looks like a consistent physical description, but not really as an ontological argument that *explains* why the classical world is position-based. In this procedure, we are implicitly assuming that in the classical world "things happen" in position, as we describe the interactions between system and environment already in that particular basis. There is no cogent or specific reason to describe those interactions in position basis if not a very important one: we *know* how to describe interactions in position basis, we are *familiar* with this kind of interaction from classical mechanics and we *expect* classical objects to emerge in position basis.

4.1 Example: quantum Brownian motion

Consider, for example, the quantum Brownian motion model.[21] In this model, the system is linearly coupled in position with a collection of quantum harmonic oscillators at constant temperature T (thermal bath). Each environmental particle (harmonic oscillator) interacts independently and only with the system and is mathematically represented by the position-basis Hamiltonian:

$$\widehat{H}_E = \frac{\hat{p}}{2m} + \frac{1}{2} m\omega^2 \hat{q}^2$$

The interaction between the system and the generic (i-th) harmonic oscillator is represented by a bilinear coupling in the position variable:

[20] Following Zurek's terminology, they are generally called *pointer states*.

[21] I take this specific example as the Brownian model is one of the principal models for the quantum to classical transition. Other important models, such as the collisional model, would be perfectly equivalent for the example.

$$\widehat{H}_{int} = \hat{x} \otimes \sum_i c_i \, \hat{q}_i$$

where $\hat{x}$ is the system position operator, c_i the coupling strength between the system and the environmental particle and $\hat{q}_i$the environmental particle position operator. As the interaction Hamiltonian $\widehat{H}_{int}(x, q)$ is a function of the system and "environment" position, in the approximation $\widehat{H}_{tot} \cong \widehat{H}_{int}$ the master equation will lead to decoherence in position. Furthermore, as states with different momenta will form macroscopic superpositions in position in a short time, these will be also decohered by the model (though at a slower rate with respect to macroscopic superpositions in the position basis). As a consequence, the pointer states of this model will be states well-localized in position and momentum. In particular: minimum-uncertainty Gaussian wave packets.[22]
In this example, we see in what sense position is already a privileged variable in the model: the environment is described in the position basis, and the interaction with the system is described by a coupling of the position coordinates. One could equivalently describe the same model in energy or momentum eigenstates, probably obtaining different results concerning the selected pointer states. Nevertheless, we do want to express the system-environment interaction in the position basis because what we are looking for in this model is an emergent description of classical objects.

Position in-position out: if we introduce position in the model from the start via $\widehat{H}_{int}(x, y)$, then it is no surprise that we get position out for the subsystem preferred basis. As it is no surprise that in different regimes (when the condition $\widehat{H}_{int} \gg \widehat{H}_S + \widehat{H}_E$ does not hold), other bases will be selected. For example, in the microscopic regime, where the self-Hamiltonian of the system dominates the dynamics: $\widehat{H}_S \gg \widehat{H}_{int} + \widehat{H}_E$, the energy basis will be generally selected, leading to decoherence in that particular basis and the formation of separate energy eigenstates, the atomic orbitals.
We thus reach the following:

Counterclaim: *decoherence theory consistently describes/accounts for the selection of position basis in the macroscopic regime, but it does not provide a genuine explanation of why the classical world is position-based. Decoherence selects position basis in the macroscopic regime since the system-environment interactions are already expressed in position basis. This introduces a position-basis representation, which becomes particularly relevant in all cases in which:* $\widehat{H}_{int} \gg \widehat{H}_S + \widehat{H}_E$.

While this scheme/procedure is perfectly fine for pragmatic purposes, this cannot be taken as a genuine ontological explanation of the selection of position basis at the macroscopic regime. In doing this, we would mix up the (sometimes hidden) phenomenological assumptions that physicists put in concrete models to make them work with the genuine ontological explanations coming from the theory. Decoherence theory does not explain why position basis emerges in the classical world, as this is only "explained" by introducing position as a privileged variable for the system-environment interaction.

5. Conclusion

The way in which decoherence theory describes the emergence of the classical world from quantum mechanics is trickier than usually thought. In particular, I have shown that a precise characterization of the emergent classical objects is lacking in the standard context, as well as the description of Newtonian trajectories for individual systems. The basis problem—the problem to understand why the classical world is position-based—is solved pragmatically, but not ontologically: this leaves the

[22] See e.g. Schlosshauer (2019, Sect. 4.2).

debate open for future research. Finally, I have suggested that decoherence theory itself proves to be philosophically very helpful when interpreted in a realist framework of quantum mechanics.

References

- S. L. Adler (2003), Why decoherence has not solved the measurement problem: a reply to P. W. Anderson, *Studies in History and Philosophy of Modern Physics*, 34(1): 135-142.
- G. Bacciagaluppi (2020), The role of decoherence in quantum mechanics, *Stanford Encyclopedia of Philosophy, substantive revision* 2020: https://plato.stanford.edu/entries/qm-decoherence/
- L. Ballentine (1998), *Quantum Mechanics: A Modern Development*, World Scientific.
- D. Bohm & B. Hiley (1987), An ontological basis for the quantum theory: non-relativistic particle systems, *Physics Report*, 144 (6): 321-375.
- D. Bohm & B. Hiley (1993), *The Undivided Universe: An Ontological Interpretation of Quantum Theory*, Routledge, London.
- G. E. Bowman (2008), *Essential Quantum Mechanics*, Oxford University Press, Oxford.
- E. Chen (2018), Quantum mechanics in a time-asymmetric universe: on the nature of the initial quantum state, *The British Journal for the Philosophy of Science*.
- E. Crull (2015), Less interpretation and more decoherence in quantum gravity and inflationary cosmology, *Foundations of Physics* 45: 1019-1045.
- E. Crull (2019), Quantum decoherence, in E. Knox & A. Wilson (eds.), *The Routledge Companion to Philosophy of Physics*, 1st edition, 2021.
- A. Di Biagio & C. Rovelli (2021), Stable facts, relative facts, *Foundations of Physics* 51 (30).
- D. Dürr, S. Goldstein & N. Zanghì (1992), Quantum equilibrium and the origin of absolute uncertainty, *Journal of Statistical Physics* 67: 843-90.
- D. Dürr & S. Teufel (2009), *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory*, Springer, Berlin.
- P. R. Holland (1993), *The Quantum Theory of Motion: An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics*, Cambridge University Press, Cambridge.
- W. Myrvold (2018), Ontology for collapse theories, in Shan Gao (ed.): *The Collapse of the Wave Function*, Cambridge University Press, 97-123.
- E. Okon & D. Sudarsky (2016), Less decoherence and more coherence in quantum gravity, inflationary cosmology and elsewhere, *Foundations of Physics* 46 (7): 852-879.
- D. Romano (2016a), Bohmian classical limit in bounded regions, in L. Felline, A. Ledda, F. Paoli & E. Rossanese (eds.), *New Directions in Logic and the Philosophy of Science*, SILFS series: 303-317, College Publications, London, 2016.
- D. Romano (2016), *The Emergence of the Classical World from a Bohmian Universe*, PhD thesis, University of Lausanne.
- J. Rosaler (2015), Is de Broglie--Bohm theory specially equipped to recover classical behavior?, *Philosophy of Science* 82 (5): 1175-1187.
- J. Rosaler (2016), Interpretation neutrality in the classical domain of quantum theory, *Studies in History and Philosophy of Modern Physics*, 53: 54-72.
- R. Shankar (1994), *Principles of Quantum Mechanics*, Springer, Berlin.
- M. Schlosshauer (2007), *Decoherence and the Quantum-to-Classical Transition*, Springer, Berlin.
- M. Schlosshauer (2019), Quantum decoherence, *Physics report*, 831: 1-57.
- D. Wallace (2012), *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation*, Oxford University Press, Oxford.