Decoherence allows quantum theory to describe the use of itself

Armando Relaño

1Departamento de Estructura de la Materia, Física Térmica y Electrónica, and GISC, Universidad Complutense de Madrid, Av. Complutense s/n, 28040 Madrid, Spain

We show that the quantum description of measurement based on decoherence fixes the bug in quantum theory discussed in [D. Frauchiger and R. Renner, Quantum theory cannot consistently describe the use of itself, Nat. Comm. 9, 3711 (2018)]. Assuming that the outcome of a measurement is determined by environment-induced superselection rules, we prove that different agents acting on a particular system always reach the same conclusions about its actual state.

I. INTRODUCTION

In [1] Frauchiger and Renner propose a Gedankenexperiment to show that quantum theory is not fully consistent. The setup consists in an entangled system and a set of fully compatible measurements, from which four different agents infer contradictory conclusions. The key point of their argument is that all these conclusions are obtained from certain results, free of any quantum ambiguity:

‘In the argument present here, the agents’ conclusions are all restricted to supposedly unproblematic “classical” cases.’ [1]

The goal of this paper is to show that this statement is not true, at least if ‘classical’ states arise from quantum mechanics as a consequence of environment-induced superselection rules. These rules are the trademark of the decoherence interpretation of the quantum origins of the classical world. As it is discussed in [2], a quantum measurement understood as a perfect correlation between a system and a measuring apparatus suffers from a number of ambiguities, which only dissipate after a further interaction with a large environment —if this interaction does not occur, the agent performing the measurement cannot be certain about the real state of the measured system.

The main conclusion of this paper is that the contradictory conclusions discussed in [1] disappear when the role of the environment in a quantum measurement is properly taken into account. In particular, we show that the considered cases only become ‘classical’ after the action of the environment, and that this action removes all the contradictory conclusions.

The paper is organized as follows. In Sec. II we review the Gedankenexperiment proposed in [1]. In Sec. IIIA, we review the consequences of understanding a quantum measurement just as a perfect correlation between a system and a measuring apparatus; this discussion is based on [2, 3]. In Sec. IIIB, we re-interpret the Gedankenexperiment taking into account the conclusions of Sec. IIIA. In particular, we show that the contradictory conclusions obtained by the four measuring agents disappear due to the role of environment. In Sec. IV we summarize our main conclusions.

II. THE GEDANKENEXPERIMENT

This section consists in a review of the Gedankenexperiment proposed in [1]. The reader familiarized with it can jump to section III.

A. Description of the setup

The Gedankenexperiment [1] starts with an initial state in which a quantum coin $R$ is entangled with a $1/2$-spin $S$. A spanning set of the quantum coin Hilbert space is $\{|\text{head}\rangle_R, |\text{tail}\rangle_R\}$; for the $1/2$ spin we can use the usual basis, $\{|↑\rangle_S, |↓\rangle_S\}$.

The experiment starts from the following state:

$$|\text{init}\rangle = \sqrt{\frac{1}{3}} |\text{head}\rangle_R |↑\rangle_S + \sqrt{\frac{2}{3}} |\text{tail}\rangle_R |→\rangle_S.$$ (1)

From this state, four different agents, $W$, $F$, $\overline{W}$, and $\overline{F}$, perform different measurements. All these measurements are represented by unitary operators that correlate different parts of the system with their apparatus. Relying on the Born rule, they infer conclusions only from certain results —results with probability $p = 1$. These conclusions appear to be contradictory.
To interpret the results of measurements on the initial state, Eq. (1), it is useful to rewrite it as different superpositions of linearly independent vectors, that is, by means of different orthonormal basis. As is pointed in [2, 3], this procedure suffers from what is called basis ambiguity — due to the superposition principle, different basis entail different correlations between the different parts of the system. This problem is specially important when all the coefficients of the linear combination are equal [4]; however, it is not restricted to this case. We give here four different possibilities for the initial state given by Eq. (1):

\[ |\text{init}\rangle_{(1)} = \sqrt{\frac{1}{3}} |\text{head}\rangle_{R} |\downarrow\rangle_{S} + \sqrt{\frac{1}{3}} |\text{tail}\rangle_{R} |\downarrow\rangle_{S} + \sqrt{\frac{1}{3}} |\text{tail}\rangle_{R} |\uparrow\rangle_{S}. \]  

(The term in \(|\text{head}\rangle_{R} |\uparrow\rangle_{S}\) does not show up, because its probability in this state is zero).

\[ |\text{init}\rangle_{(2)} = \frac{\sqrt{5}}{6} |\text{head}\rangle_{R} |\rightarrow\rangle_{S} - \frac{\sqrt{1}}{6} |\text{tail}\rangle_{R} |\leftarrow\rangle_{S} + \frac{\sqrt{2}}{3} |\text{tail}\rangle_{R} |\rightarrow\rangle_{S}. \]  

(Again, the probability of the term in \(|\text{tail}\rangle_{R} |\leftarrow\rangle_{S}\) is zero).

\[ |\text{init}\rangle_{(3)} = \frac{\sqrt{2}}{3} |h+t\rangle_{R} |\downarrow\rangle_{S} + \frac{1}{6} |h+t\rangle_{R} |\uparrow\rangle_{S} - \frac{1}{6} |h-t\rangle_{R} |\rightarrow\rangle_{S}. \]  

(And again, the probability of the term \(|h-t\rangle_{R} |\rightarrow\rangle_{S}\) is zero).

\[ |\text{init}\rangle_{(4)} = \frac{\sqrt{3}}{4} |h+t\rangle_{R} |\rightarrow\rangle_{S} - \frac{1}{12} |h+t\rangle_{R} |\leftarrow\rangle_{S} - \frac{1}{12} |h-t\rangle_{R} |\rightarrow\rangle_{S} - \frac{1}{12} |h-t\rangle_{R} |\leftarrow\rangle_{S}. \]

It is worth to remark that all \(|\text{init}\rangle_{(1)}\), \(|\text{init}\rangle_{(2)}\), \(|\text{init}\rangle_{(3)}\) and \(|\text{init}\rangle_{(4)}\) are just different decompositions of the very same state, \(|\text{init}\rangle\). In the equations above we have used the following notation:

\[ |\rightarrow\rangle_{S} = \sqrt{\frac{1}{2}} |\uparrow\rangle_{S} + \sqrt{\frac{1}{2}} |\downarrow\rangle_{S}, \]  

\[ |\leftarrow\rangle_{S} = \sqrt{\frac{1}{2}} |\uparrow\rangle_{S} - \sqrt{\frac{1}{2}} |\downarrow\rangle_{S}, \]  

\[ |h+t\rangle_{R} = \sqrt{\frac{1}{2}} |\text{head}\rangle_{R} + \sqrt{\frac{1}{2}} |\text{tail}\rangle_{R}, \]  

\[ |h-t\rangle_{R} = \sqrt{\frac{1}{2}} |\text{head}\rangle_{R} - \sqrt{\frac{1}{2}} |\text{tail}\rangle_{R}. \]

All the statements that the four agents make in this Gedankenexperiment are based on different measurements performed on the initial state, given by Eq. (1): their results are easily interpreted relying on Eqs. (2)-(5). The procedure is designed to not perform two incompatible measurements. That is, each agent works on a different part of the setup, so the wave-function collapse after each measurement does not interfere with the next one. As a consequence of this, each agent can infer the conclusions obtained by the others, just by reasoning from their own measurements.

To structure the interpretation of the Gedankenexperiment, we consider the following hypothesis for the measuring protocol:

**Hypothesis 1 (Measurement procedure)** To perform a measurement, an initial state in which the system, \(S\), and the apparatus, \(A\), are uncorrelated, \(|\psi\rangle = |s\rangle \otimes |a\rangle\), is transformed into a correlated state, \(|\psi'\rangle = \sum c_{i} |s_{i}\rangle \otimes |a_{i}\rangle\), by the action of a unitary operator \(U^{M}\). We assume that both \(|s_{i}\rangle\) and \(|a_{i}\rangle\) are linearly independent. Therefore, if the outcome of a measurement is \(|a_{j}\rangle\), then the agent can safely conclude that the system is in the state \(|s_{j}\rangle\).

B. Development of the experiment

Equations (2)-(5) provide four different possibilities to establish a correlation between the system and the apparatus. Each of the four agents involved in the Gedankenexperiment works with one of them. Follows a summary of the main results; more details are given in [1].
Measurement 1.- Agent $\mathcal{F}$ measures the state of the quantum coin $R$ in the basis $\{ |\text{head}_R \rangle , |\text{tail}_R \rangle \}$. According to hypothesis 1 above, this statement is based on the following facts. Agent $\mathcal{F}$ starts from Eq. (4). Then, she performs a measurement by means of a unitary operator that correlates the quantum coin and the apparatus in the following way

$$(c_1 |\text{head}_R \rangle + c_2 |\text{tail}_R \rangle) \otimes |\mathcal{F}_0 \rangle \rightarrow c_1 |\text{head}_R \rangle |\mathcal{F}_1 \rangle + c_2 |\text{tail}_R \rangle |\mathcal{F}_2 \rangle.$$  

That is, for any initial state of the coin, $|R \rangle = c_1 |\text{head}_R \rangle + c_2 |\text{tail}_R \rangle$, the state $|\mathcal{F}_1 \rangle$ of the apparatus becomes perfectly correlated with $|\text{head}_R \rangle$, and the state $|\mathcal{F}_2 \rangle$ becomes perfectly correlated with $|\text{tail}_R \rangle$. This procedure is perfect if $\langle \mathcal{F}_1 |\mathcal{F}_2 \rangle = 0$, but this condition is not necessary to distinguish between the two possible outcomes. Since the same protocol must be valid for any initial state, the only constraint for coefficients $c_1$ and $c_2$ is $|c_1|^2 + |c_2|^2 = 1$.

As a consequence of this, the measurement performed by agent $\mathcal{F}$ consists in

$$(\sqrt{\frac{2}{3}} |\text{head}_R \rangle |→\rangle_S - \sqrt{\frac{1}{6}} |\text{head}_R \rangle |←\rangle_S + \sqrt{\frac{1}{6}} |\text{tail}_R \rangle |→\rangle_S) \otimes |\mathcal{F}_0 \rangle \rightarrow$$

$$\rightarrow \sqrt{\frac{2}{3}} |\text{head}_R \rangle |\mathcal{F}_1 \rangle |→\rangle_S - \sqrt{\frac{1}{6}} |\text{head}_R \rangle |\mathcal{F}_1 \rangle |←\rangle_S + \sqrt{\frac{1}{6}} |\text{tail}_R \rangle |\mathcal{F}_2 \rangle |→\rangle_S.$$  

Furthermore, the quantum coin together with the agent $\mathcal{F}$ become the laboratory $\mathcal{L}$:

$$|\text{head}_R \rangle \otimes |\mathcal{F}_1 \rangle \equiv |h\rangle_\mathcal{L},$$  

$$|\text{tail}_R \rangle \otimes |\mathcal{F}_2 \rangle \equiv |t\rangle_\mathcal{L},$$  

and therefore the state of the whole system becomes

$$|\text{init}\rangle_2 = \sqrt{\frac{2}{3}} |h\rangle_\mathcal{L}|→\rangle_S - \sqrt{\frac{1}{6}} |h\rangle_\mathcal{L}|←\rangle_S + \sqrt{\frac{1}{6}} |t\rangle_\mathcal{L}|→\rangle_S.$$  

The main conclusion obtain from this procedure can be written as follows:

**Statement 1.**- If agent $\mathcal{F}$ finds her apparatus in the state $|\mathcal{F}_2 \rangle$, then she can safely conclude that the quantum coin $R$ is in the state $|\text{tail}_R \rangle$. Then, as a consequence of Eq. (13), she can also conclude that the spin is in state $|→\rangle_S$, and therefore that agent $W$ is going to obtain $|\text{fail}_L \rangle$ in his measurement (see below for details).

Measurement 2.- Agent $\mathcal{F}$ measures the state of the spin $S$ in the basis $\{ |↑\rangle_S , |↓\rangle_S \}$. Again, according to hypothesis 1, this statement is based on a perfect correlation between the apparatus and the spin states. In this case, agent $\mathcal{F}$ starts from Eq. (2). Taking into account the previous measurement, hers gives rise to the following correlation:

$$\left( \sqrt{\frac{1}{3}} |h\rangle_\mathcal{L}|↓\rangle_S + \sqrt{\frac{1}{3}} |t\rangle_\mathcal{L}|↓\rangle_S + \sqrt{\frac{1}{3}} |t\rangle_\mathcal{L}|↑\rangle_S \right) \otimes |F_0 \rangle \rightarrow$$

$$\rightarrow \sqrt{\frac{1}{3}} |h\rangle_\mathcal{L}|F_1 \rangle |↓\rangle_S + \sqrt{\frac{1}{3}} |t\rangle_\mathcal{L}|F_1 \rangle |↓\rangle_S + \sqrt{\frac{1}{3}} |t\rangle_\mathcal{L}|F_2 \rangle |↑\rangle_S.$$  

It is worth to note that this measurement is totally independent from the previous one.

As it happened with agent $\mathcal{F}$, agent $\mathcal{F}$ becomes entangled with her apparatus, and both together conform the laboratory $\mathcal{L}$:

$$|↓\rangle_S \otimes |F_1 \rangle \equiv |→1/2\rangle \mathcal{L},$$  

$$|↑\rangle_S \otimes |F_2 \rangle \equiv |→1/2\rangle \mathcal{L}.$$  

Then, the whole system becomes

$$|\text{init}\rangle_1 = \sqrt{\frac{1}{3}} |h\rangle_\mathcal{L}|→1/2\rangle \mathcal{L} + \sqrt{\frac{1}{3}} |t\rangle_\mathcal{L}|→1/2\rangle \mathcal{L} + \sqrt{\frac{1}{3}} |t\rangle_\mathcal{L}|→1/2\rangle \mathcal{L}. $$  

The main conclusion obtained by agent $\mathcal{F}$ can be written as follows:

**Statement 2.**- If agent $\mathcal{F}$ finds her apparatus in state $|F_2 \rangle$, then she can safely conclude that the spin $S$ is in state $|↑\rangle_S$. Then, as it is shown in Eq. (18), she also is certain that laboratory $\mathcal{L}$ is in state $|→\rangle_\mathcal{L}$, and therefore she can
safely conclude that agent \( F \) has obtained \(|\text{tail}\rangle_R \) in her measurement. Finally, according to Statement 1, she can be sure that agent \( W \) is going to obtain \(|\text{fail}\rangle_L \) in his measurement.

**Measurement 3.**- Agent \( W \) measures the laboratory \( L \) in the basis \( \{ |\text{fail}\rangle_L, |\text{ok}\rangle_L \} \), where \(|\text{fail}\rangle_L = (|\uparrow\rangle_L + |\downarrow\rangle_L) / \sqrt{2} \), and \(|\text{ok}\rangle_L = (|\uparrow\rangle_L - |\downarrow\rangle_L) / \sqrt{2} \).

Starting from (11), and taking into account all the previous results, this measurement implies:

\[
\left( \frac{\sqrt{2}}{3} |\text{fail}\rangle_L |\text{fail}\rangle_L - \frac{1}{2} |\text{fail}\rangle_L |\text{ok}\rangle_L - \frac{1}{2} |\text{ok}\rangle_L |\text{fail}\rangle_L + \frac{1}{2} |\text{ok}\rangle_L |\text{ok}\rangle_L \right) \otimes |W\rangle \rightarrow
\]

\[
\rightarrow \frac{\sqrt{2}}{3} |\text{fail}\rangle_L |\text{fail}\rangle_L \otimes |W\rangle_1 + \frac{1}{2} |\text{fail}\rangle_L |\text{ok}\rangle_L \otimes |W\rangle_1 - \frac{1}{2} |\text{ok}\rangle_L |\text{fail}\rangle_L \otimes |W\rangle_1 + \frac{1}{2} |\text{ok}\rangle_L |\text{ok}\rangle_L \otimes |W\rangle_2 .
\]  

Again, as the measurement is not on either the spin \( S \) or the quantum coin \( R \) or both, it is fully compatible with the previous ones. And again, agent \( W \) becomes entangled with his apparatus, in the same way that agents \( F \) and \( F \) did. However, since no measurements are done over this new composite system, we do not introduce a new notation: state \(|\text{fail}\rangle_L \) can be understood as \(|\text{fail}\rangle_L \otimes |W\rangle_1 \), and \(|\text{ok}\rangle_L \) as \(|\text{ok}\rangle_L \otimes |W\rangle_2 \).

The main conclusion that agent \( W \) obtains can be written as follows:

**Statement 3.**- If agent \( W \) finds his apparatus in state \(|W\rangle_2 \), then he can safely conclude that laboratory \( L \) is in state \(|\text{ok}\rangle_L \). Hence, as a consequence of Eq. (19), he can also conclude that laboratory \( L \) is in state \(|+\rangle_L \).

Therefore, from statement 2, agent \( W \) knows that agent \( F \) has obtained \(|\uparrow\rangle_S \) in her measurement, and from statement 1, he also knows that agent \( F \) has obtained \(|\text{tail}\rangle_R \). Consequently, agent \( W \) can be certain that agent \( W \) is going to obtain \(|\text{fail}\rangle_L \) in his measurement on laboratory \( L \).

The key point in (11) lays here. As all the agents use the same theory, and as all the measurements they perform are fully compatible, they must reach the same conclusion. This conclusion is:

*Every time laboratory \( L \) is in state \(|\text{ok}\rangle_L \), then laboratory \( L \) is in state \(|\text{fail}\rangle_L \). Hence, it is not possible to find both laboratories in states \(|\text{ok}\rangle_L \) and \(|\text{fail}\rangle_L \), respectively.*

It is worth to remark that agent \( W \) must also obtain the same conclusion from statements 1, 2 and 3.

**Measurement 4.**- As the final step of the process, agent \( W \) measures the laboratory \( L \) in the basis \( \{ |\text{fail}\rangle_L, |\text{ok}\rangle_L \} \), where \(|\text{ok}\rangle_L = (|\text{tail}\rangle_L - |\text{ok}\rangle_L) / \sqrt{2} \), and \(|\text{fail}\rangle_L = (|\text{fail}\rangle_L + |\text{ok}\rangle_L) / \sqrt{2} \).

Starting from Eq. (13), the result of this final measurement is

\[
\left( \frac{\sqrt{2}}{3} |\text{fail}\rangle_L |\text{fail}\rangle_L + \frac{1}{12} |\text{fail}\rangle_L |\text{ok}\rangle_L - \frac{1}{12} |\text{ok}\rangle_L |\text{fail}\rangle_L + \frac{1}{12} |\text{ok}\rangle_L |\text{ok}\rangle_L \right) \otimes |W\rangle \rightarrow
\]

\[
\rightarrow \frac{\sqrt{2}}{3} |\text{fail}\rangle_L |\text{fail}\rangle_L \otimes |W\rangle_1 + \frac{1}{12} |\text{fail}\rangle_L |\text{ok}\rangle_L \otimes |W\rangle_2 - \frac{1}{12} |\text{ok}\rangle_L |\text{fail}\rangle_L \otimes |W\rangle_1 + \frac{1}{12} |\text{ok}\rangle_L |\text{ok}\rangle_L \otimes |W\rangle_2 .
\]  

Therefore, and despite the previous conclusion that agent \( W \) has obtained from statements 1, 2, and 3, after this measurement he can conclude that the probability of \( L \) being in state \(|\text{ok}\rangle_L \) and \( L \) in state \(|\text{ok}\rangle_L \) is not zero, but 1/12.

This is the contradiction discussed in (11), from which the authors of this paper conclude that quantum theory cannot consistently describe the use of itself.

As Eq. (20) establishes that the probability of obtaining \(|\text{ok}\rangle_L |\text{ok}\rangle_L \) after a proper measurement is \( p = 1/12 \), and as the same theory, used to describe itself, allows us to conclude that this very same probability should be \( p = 0 \), the conclusion is that quantum theory cannot be used as it is used in statements 1, 2 and 3. That is, quantum theory cannot consistently describe the use of itself.

In the next section, we will prove that this is a consequence of hypothesis 1, that is, a consequence of understanding a measurement just as a perfect correlation between a system and a measuring apparatus. If we consider that a proper measurement requires the action of an external environment, as it is discussed in (2), quantum theory recovers its ability to speak about itself. Environmental-induced super-selection rules determining the real state of the system after a measurement removes all the contradictions coming from statements 1, 2 and 3.
III. ENVIRONMENT-INDUCED SUPERSELECTION RULES

A. The problem of basis ambiguity

In [2, 3], W. H. Zurek shows that a perfect correlation, like the one summarized in hypothesis 1, is not enough to determine the result of a quantum measurement. The reason is the basis ambiguity due to the superposition principle. To understand this statement, let us consider a simple measurement in which the state of the quantum coin determines the result of a quantum measurement. The reason is the basis ambiguity due to the superposition principle.

Note that this is the very same state as the one written in Eq. (23) — it is obtained from the system can either be in state |A⟩, and between |tail⟩ and |A⟩. Furthermore, if such apparatus states verify ⟨A|A⟩ = 0, the measurement is perfect. Starting from an initial state

\[ |\Psi_0⟩ = \left( \sqrt{\frac{1}{3}}|\text{head}\rangle_R + \sqrt{\frac{2}{3}}|\text{tail}\rangle_R \right) \otimes |A⟩ \]

the final state of the composite system, quantum coin plus apparatus, is

\[ |\Psi⟩ = \sqrt{\frac{1}{3}}|\text{head}\rangle_R \otimes |A⟩ + \sqrt{\frac{2}{3}}|\text{tail}\rangle_R \otimes |A⟩. \]

This measurement fulfills the conditions for Hypothesis 1; indeed, it is equivalent to the one that the agent F performs in measurement 1. However, the basis ambiguity allows us to rewrite (24) in the following way:

\[ |\Psi⟩ = \sqrt{\frac{1}{2}}\left( \sqrt{\frac{1}{3}}|\text{head}\rangle_R + \sqrt{\frac{2}{3}}|\text{tail}\rangle_R \right) \otimes |A⟩ + \sqrt{\frac{1}{2}}\left( \sqrt{\frac{2}{3}}|\text{head}\rangle_R - \sqrt{\frac{2}{3}}|\text{tail}\rangle_R \right) \otimes |A⟩. \]

Note that this is the very same state as the one written in Eq. (23)—it is obtained from |Ψ0⟩ as a consequence of the action of \( U^{RA} \). The new states of the apparatus,

\[ |A⟩ ≃ \sqrt{\frac{1}{2}}(|A'_1⟩ + |A'_2⟩), \]

\[ |A⟩ ≃ \sqrt{\frac{1}{2}}(|A'_1⟩ - |A'_2⟩), \]

also fulfill ⟨A|A⟩ = 0, so they also give rise to a perfect measurement.

Let us reinterpret measurement 1, as described in the previous section, taking into account this result. Hypothesis 1 establishes that a measurement is performed when a perfect correlation between a system and an apparatus has been settled. But, as both Eqs. (23) and (24) fulfill this requirement, and both represent the very same state, |Ψ⟩, the action of the operator \( U^{RA} \) is not enough to be sure about the final state of both the system and the measuring apparatus. Indeed, the only possible conclusion we can reach is:

Measurement \( U^{RA} \) cannot determine the final state of the system: if the outcome of the apparatus is “ONE”, the system can either be in state |head⟩R or state \( \sqrt{\frac{2}{3}}|\text{head}\rangle_R + \sqrt{\frac{2}{3}}|\text{tail}\rangle_R \); and if the outcome of the apparatus is “TWO”, the system can either be in state |tail⟩R or state \( \sqrt{\frac{2}{3}}|\text{head}\rangle_R - \sqrt{\frac{2}{3}}|\text{tail}\rangle_R \).

Hence, measurement 1, understood as the (only) consequence of Eq. (10) seems not enough to support the conclusion summarized in statement 1. Both |tail⟩R and \( \sqrt{\frac{2}{3}}|\text{head}\rangle_R - \sqrt{\frac{2}{3}}|\text{tail}\rangle_R \) are fully compatible with the output “TWO” of the measuring apparatus.

But what has really happened? What is the real state of the quantum coin after the measurement is completed? To which state does the wave function collapse? We know that experiments provide precise results — Schrödinger cats are always found dead or alive, not in a weird superposition like \( \sqrt{\frac{1}{2}}|\text{alive}\rangle - \sqrt{\frac{1}{2}}|\text{dead}\rangle \), so it is not possible that both possibilities are true. To answer this question, we introduce the following assumption:

**Assumption 1 (“Classical” reality)** An event has certainly happened (at a certain time in the past) if and only if it is the only explanation for the current state of the universe.
This assumption just reinforces our previous conclusion — from the measurement $U^{RA}$, that is, from Eq. \(10\), we cannot make a certain statement about the state of the system. Both $|\text{tail}\rangle_R$ and $\sqrt{\frac{2}{3}} |\text{head}\rangle_R - \sqrt{\frac{1}{3}} |\text{tail}\rangle_R$ are compatible with the real state of the universe, given by $|\Psi\rangle$ and the measurement outcome “TWO”.

This is why W. H. Zurek establishes that something else has to happen before we can make a safe statement about the real state of the system. The procedure described in Hypothesis 1 constitutes just a pre-measurement. The measurement itself requires another action, performed by another unitary operator, to determine the real state of the system. This action is done by an external (and large) environment, which becomes correlated with the system and the apparatus. As is described in \([2, 3]\), after the pre-measurement is completed, the system plus the apparatus interacts with a large environment by means of $U^E$. Let us suppose that the result of this interaction is

$$
|\Psi\rangle_E = \sqrt{\frac{1}{3}} |\text{head}\rangle_R \otimes |A_1\rangle \otimes |\mathcal{E}_1\rangle + \sqrt{\frac{2}{3}} |\text{tail}\rangle_R \otimes |A_2\rangle \otimes |\mathcal{E}_2\rangle,
$$

(27)

with $\langle \mathcal{E}_1 | \mathcal{E}_2 \rangle = 0$. Then, this interaction establishes a perfect correlation between environmental and apparatus states, in a similar way that the pre-measurement correlates the system and the apparatus. The main difference between these two processes is given by the following theorem:

**Theorem 1 (Triorthogonal uniqueness theorem \([4]\))** Suppose $|\psi\rangle = \sum_i c_i |A_i\rangle \otimes |B_i\rangle \otimes |C_i\rangle$, where $\{|A_i\rangle\}$ and $\{|C_i\rangle\}$ are linearly independent sets of vectors, while $\{|B_i\rangle\}$ is merely noncollinear. Then there exist no alternative linearly independent sets of vectors $\{|A'_i\rangle\}$ and $\{|C'_i\rangle\}$, and no alternative noncollinear set $\{|B'_i\rangle\}$, such that $|\psi\rangle = \sum_i d_i |A'_i\rangle \otimes |B'_i\rangle \otimes |C'_i\rangle$. (Unless each alternative set of vectors differs only trivially from the set it replaces.)

In other words, this theorem establishes that the state $|\Psi\rangle_E$ is unique, that is, we cannot find another decomposition for the very same state

$$
|\Psi\rangle_E = \sqrt{\frac{1}{2}} \left( \sqrt{\frac{1}{3}} |\text{head}\rangle_R \otimes |A'_1\rangle \otimes |\mathcal{E}'_1\rangle + \sqrt{\frac{2}{3}} \left( \frac{1}{3} |\text{head}\rangle_R - \frac{2}{3} |\text{tail}\rangle_R \right) \otimes |A'_2\rangle \otimes |\mathcal{E}'_2\rangle \right),
$$

(28)

with $\langle \mathcal{E}'_1 | \mathcal{E}'_2 \rangle = 0$. Hence, the interaction with the environment determines the real state of the system plus the apparatus. The action of $U^E$ gives rise to Eq. \(27\). To obtain a state like the one written in Eq. \(28\), a different interaction with the environment is mandatory, $U^{E'}$. Thus, we can formulate an alternative hypothesis:

**Hypothesis 2 (Real measurement procedure (adapted from \([2]\))** To perform a measurement, an initial state in which the system, $S$, the apparatus, $A$, and an external environment $E$ are uncorrelated, $|\psi\rangle = |s\rangle \otimes |a\rangle \otimes |\varepsilon\rangle$, is transformed: i) first, into a state, $|\psi'\rangle = \left( \sum_i c_i |s_i\rangle \otimes |a_i\rangle \otimes |\varepsilon_i\rangle \right)$, by means of a procedure called pre-measurement; and ii) second, into a final state, $|\psi''\rangle = \left( \sum_i c_i |s_i\rangle \otimes |a_i\rangle \otimes |\varepsilon_i\rangle \right)$. This final state determines the real correlations between the system and the apparatus. If $\langle \varepsilon_i | \varepsilon_j \rangle = 0$, then, after tracing out the environmental degrees of freedom, the state becomes

$$
\rho = \sum_i |c_i|^2 |s_i\rangle \langle s_i| |a_i\rangle \langle a_i|.
$$

(29)

Therefore, the measuring agent can safely conclude that the result of the measurement certainly is one of the previous possibilities, $\{|a_i\rangle |s_i\rangle\}$, each one with a probability given by $p_i = |c_i|^2$. The states $\{|a_i\rangle |s_i\rangle\}$ are called “pointer states”. They are selected by the environment, by means of environmental-induced superselection rules; they constitute the “classical” reality.

This hypothesis establishes that only after the real correlations between the system, the apparatus and the environment are settled, the observation of the agent becomes certain. Tracing out the environmental degrees of freedom, which are not the object of the measurement, the state given by Eq. \(27\) becomes:

$$
\rho_E = \frac{1}{3} |\text{head}\rangle_R \langle A_1| + 2 \frac{2}{3} |\text{tail}\rangle_R \langle A_2|.
$$

(30)

In other words, the agent observes a mixture between the system being in state $|\text{head}\rangle_R$ with the apparatus in state $|A_1\rangle$ (with probability $p_1 = 1/3$), and the system being in state $|\text{tail}\rangle_R$ with the apparatus in state $|A_2\rangle$ (with probability
interaction with the environment determines that Schrödinger cats are always found either dead or alive because the interaction with the environment determines that \(|\text{dead}\rangle\) and \(|\text{alive}\rangle\) are the pointer “classical” states.

B. Re-interpretation of the Gedankenexperiment

Let us re-interpret the first statement of the Gedankenexperiment, in the terms discussed above. Agent \(\overline{F}\) cannot reach any conclusion about the real state of the quantum coin before the pointer states are obtained by means of the interaction with a large environment \(\mathcal{E}\). The key point is that the environment \(\mathcal{E}\) interacts with the whole system, that is, with the quantum coin \(R\), the apparatus, and the spin \(S\), because the three of them are entangled. So, let us assume that a correlation like Eq. (23) has happened as a consequence of the pre-measurement. In such a case, taking into account that the quantum coin \(R\) is entangled with the spin \(S\), the state after the pre-measurement is

\[
|\Psi\rangle = \sqrt{\frac{1}{3}} |\text{head}\rangle_R |\downarrow\rangle_S |A_1\rangle + \sqrt{\frac{2}{3}} |\text{tail}\rangle_R |\rightarrow\rangle_S |A_2\rangle .
\]  (31)

The next step in the process is the interaction with the environment, which determines the pointer states of the system composed by the quantum coin and the spin. There are several possibilities for such an interaction. Let us consider, for example,

\[
U^\mathcal{E} = |\varepsilon_1\rangle |\text{head}\rangle_R |\downarrow\rangle_S |A_1\rangle \langle \text{head}\rangle_R |\downarrow\rangle_S (A_1 + |\varepsilon_2\rangle |\text{tail}\rangle_R |\rightarrow\rangle_S |A_2\rangle ,
\]  (32)

and

\[
U^\mathcal{E}' = |\varepsilon'_1\rangle |\text{head}\rangle_R |\downarrow\rangle_S |A_1\rangle + |\varepsilon'_2\rangle |\text{tail}\rangle_R |\downarrow\rangle_S |A_2\rangle + \frac{1}{3} |\text{tail}\rangle_R |\rightarrow\rangle_S |A_2\rangle .
\]  (33)

If the real interaction with the environment is given by Eq. (32), the final state of the system, after tracing out the environmental degrees of freedom, is

\[
\rho^\mathcal{E} = \text{Tr}_E [U^\mathcal{E} |\Psi\rangle] = \frac{1}{3} |\text{head}\rangle_R |\downarrow\rangle_S |A_1\rangle \langle \text{head}\rangle_R |\downarrow\rangle_S (A_1 + \frac{2}{3} |\text{tail}\rangle_R |\rightarrow\rangle_S |A_2\rangle + \frac{1}{3} |\text{tail}\rangle_R |\rightarrow\rangle_S |A_2\rangle .
\]  (34)

On, the contrary, if the real interaction with the environment is given by Eq. (33), the final state is

\[
\rho^\mathcal{E}' = \text{Tr}_E [U^\mathcal{E}' |\Psi\rangle] = \frac{1}{3} |\text{head}\rangle_R |\downarrow\rangle_S |A_1\rangle \langle \text{head}\rangle_R |\downarrow\rangle_S (A_1 + \frac{1}{3} |\text{tail}\rangle_R |\downarrow\rangle_S |A_2\rangle + \frac{1}{3} |\text{tail}\rangle_R |\rightarrow\rangle_S |A_2\rangle ,
\]  (35)

At this point, the question is: what is the real state of the system after the measurement is completed?

- Eq. (32) establishes that it is a mixture in which the agent can find the system either in \(|\text{head}\rangle_R \) and \(|\downarrow\rangle_S\), with probability \(p = 1/3\), or in \(|\text{tail}\rangle_R \) and \(|\rightarrow\rangle_S\), with probability \(p = 2/3\). It is worth to remark that this is not a quantum superposition, but a classical mixture. That is, due to the interaction with the environment, \(U^\mathcal{E}\), the state of the system is compatible with either a collapse to \(|\text{head}\rangle_R |\downarrow\rangle_S\), with \(p = 1/3\), or a collapse to \(|\text{tail}\rangle_R |\rightarrow\rangle_S\), with \(p = 2/3\). This is what agent \(\overline{F}\) concludes in statement 1.

- But the other possible interaction with the environment, Eq. (33), establishes that the real state of the system is a mixture in which the agent can find the system in \(|\text{head}\rangle_R \) and \(|\downarrow\rangle_S\), with probability \(p = 1/3\), \(|\text{tail}\rangle_R \) and \(|\downarrow\rangle_S\), with probability \(p = 1/3\), and \(|\text{tail}\rangle_R \) and \(|\rightarrow\rangle_S\), with probability \(p = 1/3\).

At this stage, the key point is the following. As the measurement performed by agent \(\overline{F}\) only involves the quantum coin \(R\), her apparatus only reads \(|\text{head}\rangle_R \) with probability \(p = 1/3\), and \(|\text{tail}\rangle_R \) with probability \(p = 2/3\). But both (32) and (33) are compatible with this result. Hence, following assumption 1, agent \(\overline{F}\) cannot be certain about the state of the spin \(S\), and thus, she can neither be certain about what agent \(\overline{W}\) is going to find when he measures the state of the laboratory \(L\). The only way to distinguish between (34) and (35) is to perform a further measurement on
the spin $S$. Such a procedure would provide the pointer “classical” states of the system composed by the quantum coin and the spin —its outcome would determine whether the interaction with the environment is given by Eq. (32) or by Eq. (33). But such a procedure would be incompatible with the measurement performed by agent $F$. Hence, agent $F$ has to choose between: i) not being certain about the real state of the quantum spin $S$, and therefore not being able to reach any conclusion about the measurement that agent $W$ will do in the future; or ii) performing a further measurement which would invalidate the conclusions of this Gedankenexperiment.

This conclusion is enough to rule out the contradictions discussed in [1]. As agent $F$ cannot be certain about the outcome what agent $W$ will obtain in his measurement, none of the four agents can conclude that it is not possible to find laboratory $L$ in state $\ket{\text{ok}}_L$, and laboratory $T$ in state $\ket{\text{ok}}_T$, at the same time. Hence, the outcome of measurement 4, whatever it is, becomes fully compatible with all the conclusions obtained by all the agents.

It is worth to note that the same analysis can also be done over measurements 2 and 3. The conclusions are pretty the same.

### IV. CONCLUSIONS

The main result of this paper is to show that assumption 1 and hypothesis 2 allow quantum theory to consistently describe the use of itself. This conclusion is based on the decoherence interpretation about quantum measurements [2]. Hence, a further statement can be set down:

*To make quantum theory fully consistent, in order it can be used to describe itself, the decoherence interpretation of measurements (and origins of the classical world) is mandatory.*

In any case, the main conclusion of this paper is applicable to other interpretations of quantum mechanics. Decoherence interpretation of the measurement process establishes that the wave-function collapse is not real —the measuring agent sees the system as if its wave-function had collapsed onto one of the pointer states selected by the environment, even though the whole wave function remains in a quantum superposition. However, this interpretation is not really important for experimental results; from this point of view, it is compatible with the Copenhagen interpretation, because it assigns the same probabilities to all of the possible outcomes. Furthermore, it is also compatible with Everett many-worlds interpretation [5]: the branches onto which the universe splits after a measurement are determined by the environmental-induced super-selection rules. The key point is that real “classical” states are not ambiguous, but they are the (unique) result of the interaction between the measured system, the measuring apparatus, and a large environment.

### ACKNOWLEDGEMENTS

The author is supported by Spain’s MINECO/FEDER Grant No. FIS2015-70856-P, and acknowledges C. M. Lóbez and A. L. Corps for their valuable comments.

---

* armando.relano@fis.ucm.es

[1] D. Frauchiger and R. Renner, *Quantum theory cannot consistently describe the use of itself*, Nat. Comm. 9, 3711 (2018).
[2] W. H. Zurek, *Decoherence, einselection, and the quantum origins of the classical*, Rev. Mod. Phys. 75, 715 (2003).
[3] W. H. Zurek, *Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse?*, Phys. Rev. D 24, 1516 (1981).
[4] A. Elby and J. Bub, *Triorthogonal uniqueness theorem and its relevance to the interpretation of quantum mechanics*, Phys. Rev. A 49, 4213 (1994).
[5] H. Everett, "*Relative State*” Formulation of Quantum Mechanics, Rev. Mod. Phys. 29, 454 (1957).