

# THE STORY OF THE CAT THAT WAS BOTH DEAD AND ALIVE

VALIA ALLORI

Department of Philosophy Northern Illinois University

Wednesday January 30th @ noon, Illinois Room (Holmes Student Center)

This is the tale, maybe a bit romanticized, of the most quoted (and surely the most abused) cat in the history of physics and philosophy- the Schrödinger cat. In 1935 Erwin Schrödinger, one of the founders of quantum mechanics, developed a thought experiment (involving a cat) devised to show that there had to be something wrong with quantum theory: if quantum mechanics is correct, then we would end up observing, for instance, cats who are dead and alive at the same time, which is absurd. A lot of ink has been spent to discuss this experiment, commonly called the Schrödinger's cat experiment, and the problem that it underlies, the measurement problem. Nonetheless, there is still something to be said after all these years, and this is what will be discussed in this lecture, after having summarized what has happened in this regard in the last 80 years.

## *The Birth: The Measurement Problem*

In his very famous 1935 article<sup>1</sup>, the Austrian physicist Erwin Schrödinger, perplexed by some consequences of Quantum Mechanics, discussed a mental experiment to make these consequences evident to everyone. He started from the presupposition that Quantum Mechanics is a fundamental physical theory. In other words, it should describe every single physical system without exception using an object called “the wave function”  $\Psi$ , that evolves in time according to an equation proposed by Schrödinger himself, and that in fact carries his name. An important mathematical feature of this equation is that it is linear: that means that if  $\Psi_1$  and  $\Psi_2$  represent two possible states for the system at a given time  $t$ , then also their sum,  $\Psi_1 + \Psi_2$ , also describes a possible state of the same system at that time. These states are called “superposition states” because if  $\Psi_1$  represents, say an object in a given region of space that we can call  $R_1$ , and  $\Psi_2$  represents that object in a given region of space, then most natural representation of the superposition state seems to be that it represents the same object both in  $R_1$  and in  $R_2$ . That this state might represent a real physical state seems obviously problematical: how is it possible that a particle can be both here and there?

The scientists of the time seemed to find a way out: after all, we do not know what *really* happens in the microscopic world, we do not have any direct experience of it, we simply make inferences about how that might be starting from observations we make in the macroscopic world. Why we want necessarily impose our macroscopic categories also to a world of which we have no direct experience? It might well be that the microscopic world is indescribable with such categories, and to insist that it is will just lead to plain contradictions. With tables and chairs we have no choice: wither they are here or there, but why are we so sure that the same thing is true also for invisible to the naked eye particles? This was the reaction that scientists had with respect to superposition states that describe microscopic objects: they are simply counterintuitive, but there is nothing problematical in that, they maintained.

But what was worrying Schrödinger though was that superposition states could describe also macroscopic states. In this case, the problem cannot be solved as easily. To make his point clear, Schrödinger developed the mental experiment that we are going to describe now. Let us take a cat, and let us put her in a box. This box is equipped with a mechanism that is activated by the decay of a radioactive nucleus. If the nucleus does not decay, nothing happens, but if the nucleus does decay

---

<sup>1</sup> Schrödinger Erwin (1935): “Die Gegenwärtige Situation in der Quantenmechanik”, *Naturwissenschaften*, 1935, 23: 807-812; English translation: *The Present Situation in Quantum Mechanics: a Translation of Schrödinger's “Cat Paradox” Paper*, in: Wheeler; John Archibald; Zurek, Wojciech H.: *Quantum Theory and Measurement* (Princeton University Press, Princeton, 1983).

then the mechanism is activated, and some poison is diffused into the room, killing the cat. We assumed by hypothesis that Quantum Mechanics is a complete theory, that is every single physical system is describable by the wave function. In particular, the nucleus at the initial instant can be described by the wave function  $\Psi_{\text{not decayed}}$ , that represents the not decayed nucleus, or by the wave function  $\Psi_{\text{decayed}}$ , that represents the decayed nucleus. Furthermore, the wave function evolves in time according to the Schrödinger equation, which is linear. That means that the state  $\Psi_{\text{not decayed}} + \Psi_{\text{decayed}}$  represents a possible description of the system. But for how the experiment is designed, after a given amount of time -the one necessary to be sure that the mechanism has been activated-, this microscopic superposition state is “amplified” macroscopically to the cat: she is described now by a superposition state of life and death, a superposition of two states that are macroscopically distinct, but that happens at the same time.

The natural question is the following: what is the significance of this state? We cannot, at least not in an obvious way, make the same move we did above: after all, we do have experience of the cat, and if we imagine to do an experiment such the one described we do not expect to find, when we open the box, as a result of the experiment a cat both dead *and* alive! What we will find is instead always a cat who is dead *or* one that is alive: superpositions states like the ones described by Quantum Mechanics do not seem to be found in the macroscopic world. In general, the situation described in the experiment if the Schrödinger cat is in its entirety similar to every other experimental situation in which there are N possible final states, or, like sometimes it is (improperly) said, N possible experiment “results” described by the wave functions  $\Psi_1, \dots, \Psi_N$ . Because of the linearity of the Schrödinger equation, also the superposition state  $\Psi_1 + \dots + \Psi_N$  is a possible state of the system. Suppose for instance that we want to measure a current on a wire and suppose that we know that it is between 3 and 5 Ampere (included), and that it comes in discrete values. Thus, the possible experiment results will be the following: the pointer of the amperometer points “3 A”, represented by the wave function  $\Psi_3$ , the pointer of the amperometer points “4 A”, represented by the wave function  $\Psi_4$ , the pointer of the amperometer points “5 A”, represented by the wave function  $\Psi_5$ . And the superposition state  $\Psi_3 + \Psi_4 + \Psi_5$  represents a possible state of the system that no one ever observes: experiments *always* have results, while in the case of a superposition state it describes a situation in which the pointer does not point to any number on the scale! So, what do we do with that?

What we described so far is called “the problem of the Schrödinger cat”, or “the measurement problem”, or even again “the problem of completeness of Quantum Mechanics”, and the reason for the last name will be clarified shortly. Schrödinger was very upset by this situation, and he did not know what position to take in that regard. As we will see, there are possible alternatives that one can entertain. To understand where they come from, one can summarize the problem as follows. The subsequent three claims are incompatible:

- 1-the wave function provides the complete description of the system
- 2- the wave function evolves according to the Schrödinger equation,
- 3-Experiments have results

If we exclude the possibility of denying the last claim, one could say that, as the Irish physicist John Stuart Bell put it:

*Either the wavefunction, as given by the Schrödinger equation, is not everything, or is not right.*<sup>2</sup>

Let us now turn to analyze the different possibilities and their consequences.

### *The Childhood: Quantum Theories with the Observer*

A first possibility, corresponding to deny the first claim above, is the one proposed (maybe in

---

<sup>2</sup> Bell, John Stuart (1987). *Speakable and Unsayable in Quantum Mechanics* (Cambridge: Cambridge University Press,1987).

an indirect way) by the famous Hungarian mathematician John von Neumann<sup>3</sup>. The idea is that the “observer” plays an important role in the theory. The wave function evolves according to the Schrödinger equation, but only until someone observes what happens. At this point, what is called the “collapse” or “reduction” of the wave function happens: the wave function transforms, instantaneously, from the sum of more than one terms into one term only of the superposition state. Therefore, we do not ever observe superpositions because the act of observing them makes them “disappear”.

It is clear, as soon one thinks about it, that this solution of the measurement problem is not really very satisfactory at all: what is *exactly* an observer? Some, including the great Hungarian mathematician Eugene Wigner<sup>4</sup>, arrived to the conclusion that the presence of consciousness is what will determine that something is an observer. As a consequence, in our project of constructing a fundamental physical theory it seems important to explain what consciousness is, and how it influences the behavior of physical objects. As a methodological criterion, before accepting such a radical view one should consider whether there are more sensible alternatives. Is it really the only possibility to solve the measurement problem? If the answer is negative, that is if there are actually other solutions to the problem in which it is not necessary to invoke consciousness to construct the theory, I do not see what reason one might have *-ceteris paribus-* to choose the radical position. In addition, after all, what is so special in the act of observation? Isn't it a physical process like any other? Should it not be explained by the theory as well (together with all the other physical processes)? Here the situation seems to be the contrary: it is not the measurement to be explained, but rather it is used to provide all explanations!

These few lines should be enough to make you appreciate how this solution, Quantum Mechanics with the observer, is really problematical, and that is not appropriate for the “observer” to appear in the formulation of a fundamental physical theory. At this point the natural question one should ask is the following: “Are there formulation of Quantum Mechanics that do not suffer from the problem of the Schrödinger cat and that do not use the “observer” to solve it? E' heartening to know that the answer is affirmative. In the next sections we will dedicate our attention to these theories, albeit briefly.

### *The Adolescence: The Copenhagen Interpretation*

A solution of the measurement problem *a la* von Neunamm that still denies the completeness of the description provided by Quantum Mechanics but that goes in the direction of excluding the “observer” from the formulation of the theory is the one proposed by one who is regarded one of the greatest physicists of the time, the Danish Niels Bohr<sup>5</sup> and then thus takes the name of Copenhagen Interpretation. The basic idea is that physics is not able to describe the world because it is one of our creations, and we as human beings are intrinsically limited. There are two worlds, the classical and the quantum one, and both of them are offering us a foreshortening on reality, neither of which is complete and exhaustive. Quantum Mechanics alone does not describe a complete description: to its description one needs to add the one provided by Classical Mechanics. According to Bohr though, this is not a symptom that the theory being inadequate, but rather of the fact that a coherent view is intrinsically impossible for us as human beings.

This seems to be suggested by the fact that, in many recent (for the time) experiments particles seemed to behave like waves, and *vice versa*. An example of particles that behaves like

---

3 von Neumann, John (1932). *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer University Press, 1932); English Translation: *Mathematical Foundations of Quantum Mechanics* (Princeton: Princeton University Press, 1955).

4 Wigner, Eugene (1967). *Symmetries and Reflection* (Bloomington: Indiana University Press, 1967).

5 Bohr, Niels (1949). *Discussion with Einstein on Epistemological Problems in Atomic Physics*, in: P.A. Schilpp (ed.), *Albert Einstein, Philosopher-Scientist* (Evanston: Library of Living Philosophers, 1949).

wave is illustrated by the famous two slits experiment: let us consider a screen with two slits, and let us send particles -one by one- against such screen. A second screen is positioned behind the slits to record “what happens”. As the experiment proceeds, small spots form on the screen, presumably indicating that a particle hit that point, and it happens that the spots more and more concentrate in certain locations rather than in others to form a classic interference pattern. This result is incredible: how come particles should form such a pattern on a screen? The natural thing that should have happened is to see two luminous strips on the second screen corresponding to the two slits, where the particles have not been blocked by the first screen and could reach the second one. Instead in this case it is like as if a wave, not a particle, passed through the slits! But how come a particle can behave like a wave?

Bohr, among the others, thought that to solve this mystery one should recognize that the concepts of particle and wave are not adequate, and introduced his “wave-particle duality”: at best “wave “ and “particle” are two concepts that provide only partial descriptions of the microscopic world. This led Bohr to formulate afterwards the Copenhagen Interpretation in which, as we have seen, one abandons the idea of describing the microscopic world and limits oneself to a description at the level of measurements.

One can see how not the measurement problem does not exist anymore: by definition, macroscopic objects are “classical” objects, that behave in a perfectly “normal” way. Therefore, they are never described by a superposition state, by definition. In this theory, the complete description of the system is given by the wave function and by the macroscopic classical variables, the ones we would call the experiment results. The wave function evolves according to the Schrödinger equation and the classical variables evolve according to the classical laws, but when there is interaction the dynamics change: the wave function collapses, and the classical variables undergo random changes that are classically unpredictable, statistically governed by the wave function.

Let us note how this version of Quantum Mechanics is very bizarre: the macroscopic objects, like tables and chairs, are not constituted by particles, but on the contrary they provide the foundation of all the theory, and it would be more proper to consider tables and chairs to “constitute” what we call (improperly) call the “microscopic particles”. But being bizarre, in itself, is not an objection. Let us notice instead the following thing: how should we intend the term “macroscopic”? Where the separation line is between the quantum and the classical world? How many “particles” must an object have in order to be considered “classical”? This fundamental ambiguity of the formulation of the theory is by itself deeply problematical, especially if it is not a forced choice.

### *The Maturity: Quantum Theory without Observers*

Why should we build a theory since the foundation on the concept of macroscopic object, which is intrinsically vague? Bohr would say that we cannot help it, that we are intrinsically limited. But how much truth is there in this claim? Maybe there are other ways of interpreting the “craziness” of Quantum Mechanics without abandoning the idea of being able to provide a coherent description of the world through our physical theories. There are actually at least two theories that are taking up this challenge: one is called theory of spontaneous localization or Ghirardi-Rimini-Weber (GRW) Theory; another is the so-called theory of hidden variables, also called (more suitably) Bohmian Mechanics.

Bohmian Mechanics was initially developed by the French physicist Louis de Broglie<sup>6</sup> with the name of pilot-wave theory as an explanation of quantum phenomena alternative to the one of

---

6 de Broglie Louis (1928): in Solvay Congress (1927), *Electrons et Photons: Rapports et Discussions du Cinquième Conseil de Physique tenu à Bruxelles du 24 au 29 Octobre 1927 sous les Auspices de l'Institut International de Physique Solvay* (Paris: Gauthier-Villars, 1928).

Bohr, and was later generalized by the American physicist David Bohm<sup>7</sup>. Unfortunately, this theory was initially called “theory of hidden variables” because in it the description of a physical system provided by the wave function is completed adding the specification of the position of particles, the so called “hidden variables”, but this name does not make justice of the theory. According to Bohmian Mechanics, every physical object, either macroscopic or microscopic, is composed of particles: they have a well-defined position in three-dimensional space, and evolve in time according to an equation determined by the wave function, which in turn evolves in time according to the Schrödinger equation. Since physical objects, cats included, are made of particles, and since it is the wave function that is in superposition, the measurement problem simply vanishes: the cat is either dead or alive, depending on what her particles are doing. The “shape” of the wave function, in particular the fact that it is in superposition, does not matter: the wave function should not be considered as constitutive of physical objects as a “concrete” field in three-dimensional physical space, first of all because mathematically it is not: it “lives” in a space of extremely high dimensionality, of the order of the Avogadro number.

Once this is understood, one easily sees how there is no measurement problem in particular, and no problem in general<sup>8</sup>. In addition, one can see how the particles in this theory are not what the term “hidden variables” might suggest they are, but rather they are what really constitutes physical objects.

The GRW theory, presented for the first time by the Italian physicists Giancarlo Ghirardi, Alberto Rimini and Tulio Weber<sup>9</sup>, in contrast to Bohmian Mechanics denies that the wave function evolves according to the Schrödinger equation<sup>10</sup>: instead of invoke the “operator” as the responsible for the “collapses” of the wave function, this theory insert the collapse directly into the evolution equation. In GRW, the wave function evolves for most of the time according to the Schrödinger equation, but every once in a while it collapses into one of the terms of the superpositions. The rate at which this reduction happens is determined by a factor proportional to the number of “particles” that compose the object: the bigger the object and the less time it takes for the wave function to collapse. In this sense, macroscopic superpositions are never observed: if the object is macroscopic, it remains in superposition for a very short time.

But, having said so, something remains that still does not work: in what sense saying that the macroscopic object remains in a superposition a very short time helps in solving the problem of the cat? Surely, this explains why we do not see such superposition objects, but independently on how long the object remains in superposition, if the time is not exactly zero the problem of explaining what it means to have a state of a dead and alive cat remains! We need then think about what is the real problem of quantum mechanics, which is only suggested by the problem of the Schrödinger cat: *What mathematical object is able to suitably represent what exist in Nature*<sup>11</sup>? More precisely. Can the wave function represent physical objects? Reflections about this<sup>12</sup> lead to the conclusion that this

---

7 Bohm, David (1952). “A Suggested Interpretation in Terms of “Hidden Variables”: Part I and Part II,” *Physical Review*, 1952, 85: 166-179 and 180-193.

8

For a detailed description of this theory and to see how it can describe the other quantum phenomena, see: Allori, Valia; Zanghì, Nino (2004). “What is Bohmian Mechanics?”, *International Journal of Theoretical Physics*, 2004, 43: 1743-1755 ; Goldstein, Sheldon (2001). *Bohmian Mechanics*, in: the Stanford Encyclopedia of Philosophy (<http://plato.stanford.edu/entries/qm-bohm>) and references therein.

9 Ghirardi, Giancarlo; Rimini, Alberto; Weber, Tulio (1986). “Unified Dynamics for Microscopic and Macroscopic Systems”. *Physical Review D*, 1986, 34: 470–491.

10 See: Allori, Valia; Goldstein, Sheldon; Tumulka Roderich., Zanghì Nino (in sottomissione). “On the Common Structure of Bohmian Mechanics and the Ghirardi-Rimini-Weber Theory”, [quant-ph/0603027](http://arxiv.org/abs/quant-ph/0603027). For a clear discussion of how this claim should be intended.

11 This point has been emphasized for the first time by Goldstein in Goldstein, Sheldon (1998). “Quantum Theory Without Observers”, *Physics Today*, 1998, Part One: March 1998, 42–46; Part Two: April 1998, 38–42.

12 In addition to the article by Goldstein mentioned in the reference above, see also: Allori, Valia; Goldstein, Sheldon; Tumulka Roderich., Zanghì Nino (in submission). “On the Common Structure of Bohmian Mechanics and

is impossible, or at best very problematical: the aim of a fundamental physical theory should be to describe objects in three-dimensional space and the wave function, living in a much bigger space, cannot do it, at least not alone. As a consequence, in the GRW theory, tables and chair cannot be “made of” wave functions, not so much because it is “in superposition” but because it is the wave function. In contrast to what originally suggested, the description provided by the wave function needs *always* to be completed, independently by its temporal evolution. The proposal made by Bell - either the wave function is incomplete or it does not evolve according to the Schrödinger equation - is, in this sense, a false dichotomy. Realizing this, Ghirardi<sup>13</sup> proposed a formulation of GRW (that we will call GRWm) in which every physical object is described by a field,  $m(x)$ , in three-dimensional space, that one can identify with the density of mass of the object, that is determined by the wave function. Therefore, the description of objects is not, in this theory, of the corpuscular kind (there are no particles in GRWm), but rather continuous, provided by the  $m$  field. Let us not that the choice made by Ghirardi is not the only possible choice: *a priori*, there seem not to be any limitation (a part of being in three-dimensional space or in space-time) to the choice of what mathematical object can represent the building blocks of the whole universe, the so called *primitive ontology* of the theory. In particular, Bell<sup>14</sup> proposed that physical objects are not constituted by the  $m$  field, but by what we could call “flashes”: events in space-time. According to this theory, named GRWf, these flashes are what exists at the fundamental level, and physical objects are nothing but a “collection” of such flashes<sup>15</sup>.

#### *The Middle-Age Crisis: The Many-Worlds Interpretation*

Up to now we have not considered what would happen denying the last of the three claims in the first section: the fact that experiments have results. This is the road that Hugh Everett III, promising American physicist, thought of undergoing<sup>16</sup>. The idea is that all that *can* happen, *will* actually happen. The problem is to establish exactly “where”. In these theories one talks of “multiverse” : every time one performs an observation, all possible results are realized in one of the different universes that constitute the multiverse. For this reason, the theory is often called Many-World Theory. This approach, even if in the opinion of the author it is not very promising (first of all because it is based on the wave function as a fundamental physical object), has become very popular among physicists and philosophers, for different reasons: for the physicist, the appeal is that the theory leaves the mathematics of Quantum Mechanics completely unaltered and relegating all the problems to what is called “interpretation”; for the philosophers instead it leaves sufficient room for investigating metaphysical problems like time travel or the problem of personal identity, contrarily to what happens in the case of Bohmian Mechanics and GRW theory.

#### *A New Life? Quantum Mechanics and Relativity*

The choice of Bell that the world is made of flashes is strange, but not without motivation: in fact Bell noticed that, as recently the young German mathematician Roderich Tumulka<sup>17</sup> has

the Ghirardi-Rimini-Weber Theory”, [quant-ph/0603027](http://quant-ph/0603027).

<sup>13</sup> Benatti, Fabio; Ghirardi, Giancarlo; Grassi, Renata (1995). “Describing the Macroscopic World: Closing the Circle within the Dynamical Reduction Program”, *Foundations of Physics*, 1995, 25: 5–38.

<sup>14</sup> Bell, John Stuart (1987). *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press, 1987).

<sup>15</sup> For a detailed description of the various possibilities, and in particular GRWf, see: Allori, Valia; Goldstein, Sheldon; Tumulka Roderich., Zanghì Nino (in sottomissione). “On the Common Structure of Bohmian Mechanics and the Ghirardi-Rimini-Weber Theory”, [quant-ph/0603027](http://quant-ph/0603027).

<sup>16</sup> Everett, Hugh (1957). “Relative State Formulation of Quantum Mechanics”, *Review of Modern Physics*, 1957, 29: 454–462.

<sup>17</sup> Tumulka, Roderich (2006). “A Relativistic Version of the Ghirardi–Rimini–Weber Model”, *Journal of Statistical Physics*, 2006, 125: 821–840.

proven, with a formulation in terms of flashes Quantum Mechanics and relativity seems much closer. But in what respect? What is known is that Quantum Mechanics and relativity do not get along that well. What is not so known is what the problem is. The previous reflections hopefully were useful to understand that it is not even clear what the Quantum Mechanics we find in books is talking about: the observer? The measurement results? Particles? What, exactly? Without clarifying this, how can we even ask what is compatible with what? One needs to make a choice – and some are more satisfactory than others- and only then it will be possible to ask other questions. Therefore, in the realm of the so called Quantum Theories without observers, one can ask what are the symmetry properties of the theory, including the one of Lorentz invariance. In a sense, they are “properties” of the building blocks that constitute physical objects according to the theory. At this point, the problem becomes a mathematical problem: can we construct an equation for the time evolution of the primitive ontology with the desired property? If the answer is positive, as Tumulka has shown for GRWf, we are on board.... but this is another story!