

Chapter 1

Introduction

There is a basic philosophical question that involves metaphysics, physics and epistemology: Can we explain what the world is like through a fundamental physical theory? This question corresponds to the historic disagreement among scientists and philosophers concerning how to regard physical theories to which people commonly refer as the realist–antirealist debate. The position of the antirealist is the one according to which we should not believe that physics reveals to us something about reality but rather we should be content with physics to be, for example, just empirically adequate. In contrast, the realist is strongly inclined to say not only that physics tells us about reality, but also that it is our only way to actually do metaphysics. I am a realist insofar as I believe that physics actually informs us about the world. That is, I agree on what Tim Maudlin claims in his *Suggestions from Physics for Deep Metaphysics* (Maudlin m.):

[...] metaphysics, i.e. ontology, is the most generic account of what exists, and since our knowledge of what exists in the physical world rests on empirical evidence, metaphysics must be informed by empirical science.

The main problem I would like to investigate in my dissertation is therefore the following: Granted that physics provides information about the world, what does it mean to explain the world around us in terms of a fundamental physical theory? I believe that this question can be reformulated in this way: Is there a general structure that a fundamental physical theory should have in order to allow us to understand what the world is like?

There are some notions that I believe helped me in finding an answer to this question: the notion of primitive ontology and the one of typicality. The notion of primitive ontology is connected to but not exhausted by another notion, the one of “local beable”. This has been introduced by John Stuart Bell in his *Are there Quantum Jumps?*

(Bell 1987). A “beable” is a speculative piece of ontology, it represents what is real according to the theory. It is “local” in the sense that its value can be assigned to a given bounded region of space-time. In the words of Bell:

These are the mathematical counterparts in the theory to real events at definite places and times in the real world (as distinct from the many purely mathematical constructions that occur in the working out of physical theories, as distinct from things which may be real but not localized, and distinct from the “observables” of other formulations of quantum mechanics, for which we have no use here).
(Bell 1987)

It has been suggested that there is some sort of distinction among the objects that are commonly accepted as the ontology of a fundamental physical theory. This is clear in the work of Detlef Dürr, Shelly Goldstein and Nino Zanghì, (DGZ 1992) and (DGZ 1997), in which they talk about a primitive ontology of a theory as opposed to the ontology of the theory in general:

[...] primitive ontology – the basic kinds of entities that are to be the building blocks of everything else (Except, of course, the wave function)[the parenthetical remark was in a footnote in the original].(DGZ 1992)

I believe that this notion is crucial for understanding what it means for a theory to give an account of the behavior of physical objects in the world we live in. Before fleshing this out, though, it is necessary to characterize the notion of primitive ontology more carefully as a significant ingredient of a fundamental physical theory.

I therefore start from some practical examples in order to try to clarify what the notion of primitive ontology is supposed to be and what is an adequate primitive ontology for a fundamental physical theory. At first I discuss about quantum mechanics. In order to do this properly, I start in Chapter 2 discussing the measurement problem. In that chapter I also analyze the different solutions of the measurement problem: Bohmian mechanics, GRW theory and many worlds quantum mechanics, what are called “quantum theories without observer”. I discuss the different possible quantum metaphysics that one can infer from them in Chapter 3.

Chapter 2

Pandora's Cat and Quantum Theories without Observers

In this chapter I will discuss what is believed to be a most discussed problem of quantum mechanics: the problem of the Schrödinger cat, also called the measurement problem. After having done so, I will argue that it is just a symptom, rather than the cause, of all the mysteries and the paradoxes of quantum mechanics. To anticipate the conclusion, a way to express the moral of the Schrödinger cat is to say that quantum mechanics is not a complete theory, that is, the problem of the Schrödinger cat is *the problem of the completeness of quantum mechanics*. One could say that quantum mechanics is not complete in the sense that it is unable to account for the properties we believe macroscopic object should have. This is what I would call *the problem of indefinite properties*. But again, why is quantum mechanics not able to do so? What is the origin of this problem? When we talk about a property, we have in mind the idea of something having that property. But what is that “something” if quantum mechanics is true? I will argue that this problem is parasitic on what we could call *the problem of the lack of a clear ontology*: as clearly Shelly Goldstein (Goldstein 1998) pointed out first, it is not clear what quantum mechanics is about. In the process of trying to figure out what the theory is about we end up with various alternatives: The wave function? The observer? The results of measurement? Particles? Fields? Strings? Only after having answered this question, one can proceed to investigate whether the primitive ontology is an adequate one. The primitive ontology is not adequate if it is not able to represent physical objects and their properties. In that case, we say that the theory is not complete. Some of the primitive ontologies are almost straightforwardly inadequate like, for example, the proposal that quantum is about the observer, as we will see. What about the wave function itself? As we will discuss later, it is usually said that what the

problem of the Schrödinger cat is telling us is that the wave function alone, *if it evolves according to Schrödinger equation*, cannot completely describe physical objects. But what about a wave function that does not evolve according to Schrödinger's equation? This is, I believe, the real question that needs to be answered. I will argue that the moral of the problem of the cat is that the description provided by the wave function alone is *never* complete. That is, the wave function cannot be what quantum is fundamentally about, or, in other words, the wave function cannot be the *primitive ontology* of the theory. And the reason for this is that it is a too abstract mathematical object. This is what I will call *the problem of the adequacy of the primitive ontology*. We will see how, once one has a theory with an adequate primitive ontology, one can account for the properties and the behavior of macroscopic objects in three-dimensional space so that the problem of indefinite properties does not arise.

2.1 The Schrödinger Cat: The Measurement Problem

But what is the problem of the Schrödinger cat? It might seem I am being a little bit redundant to explain this problem again: is it not clear already what it is supposed to be, at least in the philosophical community? Nonetheless, I believe that the situation is still quite subtle. Indeed, we will see how it is not so obvious what the problem is and what the solutions are *really* telling us about what the world is like.

Let us start from the beginning. What is “standard” quantum mechanics, the one that is found in physics textbooks? Let us call “state” of a system all that needs to be specified in order to completely describe any physical system. Then the basic assumptions of this theory, that we could call, for reasons that will appear clear later, *bare quantum mechanics*, are the following:

- There is an object, called the wave function and usually written with the Greek letter ψ , that represents the state of the system of any physical system, and
- It evolves in time according to a given differential equation, called the Schrödinger equation.

The wave function $\psi = \psi(q)$ is a function of the configuration $q = (q_1, \dots, q_N)$, where

N is the number of “particles”. Note that in quantum mechanics so far there are no particles, just the wave function. So the use of the word “particle” in this context will be just for convenience. Each $q_i \in \mathbb{R}^3$ with $i = 1, \dots, N$ is a degree of freedom of the wave function such that q lives in a space of dimension $d = 3N$. This space, \mathbb{R}^{3N} , is called *configuration space*, since it could be identified with the space of the configurations of N particles in three-dimensional space, if there were any particles. Therefore, the wave function is a mathematical object defined on configuration space. In addition, it is complex-valued:

$$\psi : \mathbb{R}^{3N} \rightarrow \mathbb{C} \quad (2.1)$$

The space of all wave functions forms an Hilbert space, that is a space that generalizes the notion of Euclidean space. Roughly speaking, this space has a linear structure (it is a vector space) on which an inner product is defined (and therefore it is possible to talk about distance, angles, orthogonality), and such that it is complete (that is, there are not any pathological behavior in taking the limits). In addition, the two wave functions ψ and $c\psi$, where $c \in \mathbb{C}$ is such that $|c|^2 = 1$, are not physically distinct: that is, ψ and $c\psi$ represent the same physical object. Therefore, the wave functions actually form equivalence classes in a projective Hilbert space, defined by the relation: $v \sim w$ when $w = cv$, for any $c \neq 0, c \in \mathbb{C}$. This is what is meant when it is said that wave functions are projective objects. More precisely, they are *rays* in Hilbert space. As anticipated, the wave function ψ evolves in time according to Schrödinger’s equation:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi, \quad (2.2)$$

where $\hbar = h/2\pi$, h being the Planck constant $h = 6.63 \cdot 10^{-34} \text{ m}^2\text{kg/s}$, H is the usual nonrelativistic Hamiltonian, that, for spinless particles, is of the form

$$H = - \sum_{k=1}^N \frac{\hbar^2}{2m_k} \nabla_k^2 + V, \quad (2.3)$$

containing as parameters the masses m_k of the “particles” as well as the potential energy function V of the system. One can also write the wave function at a time t ψ_t

as evolved according to Schrödinger's equation from the wave function at time $t = 0$ ψ_0 through an operator U , $U = e^{-\frac{i}{\hbar}Ht}$, that describes the Schrödinger evolution:

$$\psi_t = U\psi_0. \quad (2.4)$$

It is important to emphasize that a crucial feature of Schrödinger's equation is that it is linear: if ψ_1 is a possible description of a physical system at a given time t , and so is ψ_2 , then also the sum of the two, namely $\psi_1 + \psi_2$ (up to a normalizing factor), provides a possible description of that physical system at time t . States of this form are called *superposition states*.

Let us now turn to the measurement problem, which has been formulated for the first time by Erwin Schrödinger in his seminal paper (Schrödinger 1983). The situation discussed in the experiment considered by Schrödinger, as shown in Figure 2.1, is as follows: there is a cat in the box, together with a bottle of poison. This bottle is connected to a device that is triggered by the decay of a radioactive nucleus in such a way that if the nucleus decays the poison will be diffused into the box killing the cat. If the nucleus does not decay, nothing happens to the cat. At a given time, the nucleus can or cannot decay. That is, the possible states of the nucleus are: $\psi_{decayed}$ and $\psi_{undecayed}$. Because of linearity, there is another possible quantum state, namely the one described by a wave function of the form $\psi = \psi_{decayed} + \psi_{undecayed}$. If the wave function provides the complete description of the world, and if it evolves according to Schrödinger's equation, then the microscopic superposition of the nucleus amplifies macroscopically to the state that describes cat: A cat alive and dead at the same time? This state is a superposition state, therefore it describes the system being at the same time into two (macroscopically disjoint) states of affair, in this case a dead and an alive cat. This does not correspond to a state that we find in the world, and therefore something wrong is going on. Note that this state is what bare quantum mechanics predicts to happen in any experimental situation. This is why this problem is also called *the measurement problem*. In fact, suppose we want to measure something, take for example the current in a wire. The standard rules of quantum mechanics say that

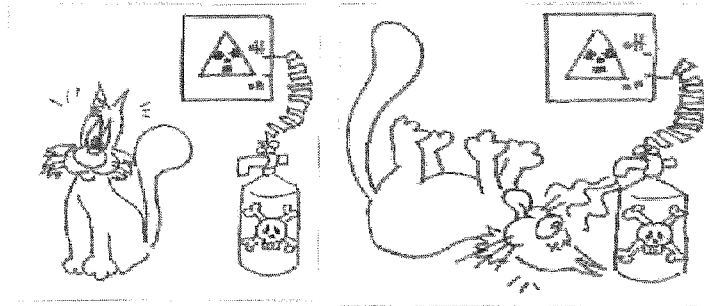


Figure 2.1: The Schrödinger cat experiment.

the results of the experiment will be given, mathematically, by the eigenvalues of some appropriate self-adjoint operator. A vector z is an eigenstate of a given operator A if and only if

$$Az = \xi z,$$

where $\xi \in \mathbb{C}$ is the corresponding eigenvalue. If z is an eigenstate, the operator A just transforms it into a multiple of itself. Because of this reason, it has been considered reasonable for eigenvalues to represent properties. Suppose that the state of the system is an eigenstate ψ_α of the “observable being measured” (in this case, the current, represented by an operator A). That is, $A\psi_\alpha = \alpha\psi_\alpha$. Let the apparatus, the pointer, have an initial wave function ϕ_0 so that the initial total state is $\psi_\alpha\phi_0$. Note that, since we want the measurement to be genuinely such, Schrödinger’s evolution should not change the state ψ_α of the system being measured, that is $U\psi_\alpha = \psi_\alpha$. The pointer state, since we want it to be a genuine measurement apparatus, will have to evolve into $\phi_t = U\phi_0 = \phi_\alpha$, such that the information we want to measure about the system will be displayed macroscopically by the position of the pointer. That is, ϕ_α represents the position of the pointer in the direction α corresponding to the state of the system

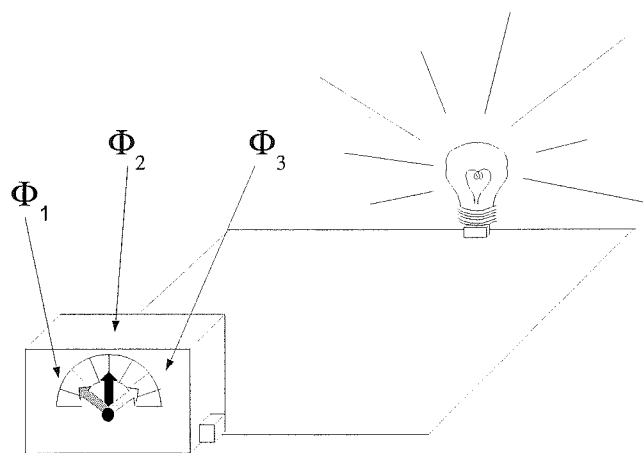


Figure 2.2: Pointers do not point in bare quantum mechanics.

being ψ_α . The final state of the comprehensive system (system+apparatus) is therefore $\Psi_t = \psi_\alpha \phi_\alpha$. Note that α represents a possible results for the experiment: 1 ampere, 2 ampere, 3 ampere, and so on. Now suppose the state of the system is not an eigenstate, but instead it is a superposition of the form $\psi_0 = \sum_\alpha c_\alpha \psi_\alpha$. The initial wave function of the total system is therefore $\Psi_0 = \psi_0 \phi_0 = \sum_\alpha c_\alpha \psi_\alpha \phi_0$. The final wave function will be $\Psi_t = U \Psi_0 = U \sum_\alpha c_\alpha \psi_\alpha \phi_0$ and because of the linearity of the evolution U we will have $\Psi_t = \sum_\alpha c_\alpha \psi_\alpha U \phi_0 = \sum_\alpha c_\alpha \psi_\alpha \phi_\alpha$. This final wave function is analogous to the one of the cat in macroscopic superposition of life and death and describes macroscopic superpositions of pointer positions pointing in different directions. We arrived to a really troublesome conclusion, since the final state corresponds to measurement not having results. In fact, each term of the superposition represents a pointer pointing somewhere (ψ_α corresponds to the pointer pointing to the value α), so that this state describes a pointer not pointing anywhere, as shown in Figure 2.2!

To sum up, then, the three claims:

1. The wave function provides the complete description of any physical system,
2. The wave function evolves according to Schrödinger's equation,

3. Measurements have results

are incompatible. This is, in a nutshell, the problem of the Schrödinger cat as it is usually presented.

2.2 Bell's Alternatives

If we assume that measurements have indeed results, then the lesson we should draw from the measurement problem is that either (1) or (2) must be false. This leads us straightforwardly to the following alternatives:

- Deny (1) and add something to complete the description provided by the wave function, or
- Deny (2) and allow the wave function to evolve according to an equation different from Schrödinger's evolution.

These are the famous alternatives proposed by the physicist John Stuart Bell:

either the wave function, as given by the Schrödinger equation, is not everything, or is not right (Bell 1987).

This means that (if we assume measurements to have results) there are only two possibilities to solve the measurement problem and to make quantum mechanics a precise (i.e. not ambiguous) fundamental physical theory. In the first possibility the complete description of a physical system is given by the wave function, which still evolves according to Schrödinger's equation, *and* by some additional ("hidden", because it is not "suggested by" the Schrödinger equation) variable. The other only possibility consists in assuming that the wave function provides the complete description of the system but its evolution in time is given by an equation that differs by the one of Schrödinger.

2.3 Some Problematical Attempts

Bell's alternatives leave open the possibility for more than two theories. In fact, depending on what we add to the description of the wave function or how we change

Schrödinger's evolution, we have different theories. A variety of these theories have been proposed, and some are more satisfactory than others. In the following section I will start analyzing some problematical attempts, while in Section 2.4 I will discuss more satisfactory solutions of the measurement problem.

2.3.1 Von Neumann and the Collapse

The first attempt to solve the measurement problem was provided by the famous physicist and mathematician John von Neumann (von Neumann 1932). The basic idea is to postulate that Schrödinger's equation ceases to be valid during a measurement situation. In that case, when a measurement occurs, the evolution is determined by a nonunitary transformation, often called "collapse" or "reduction" of the wave function. This evolution is, of course, incompatible with Schrödinger's evolution since it is random and irreversible: every time there is a measurement, the wave function is not in a superposition state anymore but collapses randomly into one of the terms of the superposition. This evolution is undeniably *ad hoc*, postulated just in order to eliminate all the other terms of the superposition but the one that happens to be the result of the measurement.

We could classify this way of solving the problem as following route number 2: the Schrödinger evolution is not valid all the time, but only in between measurements. Or we could also classify it as a way of completing quantum mechanics: the wave function does not provide the complete description, we need also to add the "observer": she is playing a crucial role in the theory because every time she makes an observation or a measurement, there is a change in the evolution of the wave function. When the measurement occurs, the object is not described by a superpositions state but rather by one of its terms, as shown in Figure 2.3.

Still, here is something to think about: in the defining terms of this theory there is the notion of measurement and observer. But what is an observer? What is a measurement? As Bell has emphasized, any fundamental physical theory should not use in its definition, among its fundamental entities, such vague concepts. Rather, they should be derived by something more fundamental:

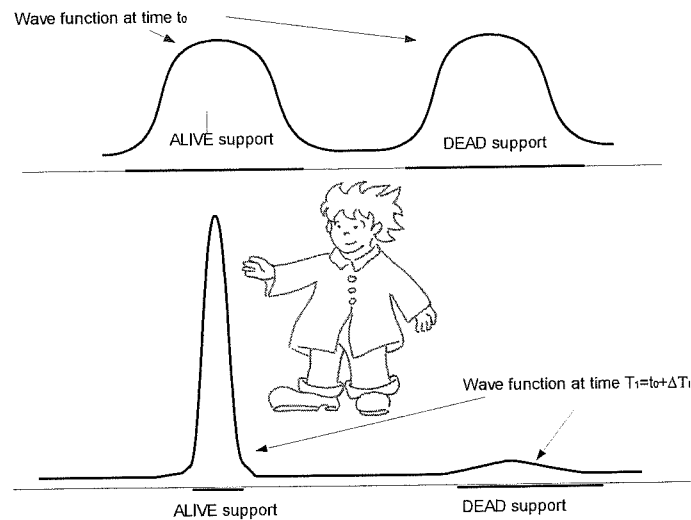


Figure 2.3: Von Neumann and the Schrödinger cat.

What exactly qualifies some physical systems to play the role of ‘measurer’? Was the wavefunction of the world waiting to jump for thousands of years until a single-celled living creature appeared? Or did it have to wait a little longer, for some better qualified system... with a PhD? (Bell 1987)

After all, is not a measurement a physical process? And is not the observer a physical object as well? In this theory they have a special status, since something distinctive happens when someone “observes” something. But what makes them special? If we want to take this theory seriously, we need to clarify what the observers are and what makes them so fundamentally different from the rest.

2.3.2 Wigner and Consciousness

Fritz London and Edmond Bauer (London and Bauer 1983) proposed that it is human consciousness which defines what an observer is. In the 1960s, the physicist Eugene Wigner followed up on this proposal. In his *Remarks on the Mind-Body Question* (Wigner 1967), Wigner argued that what characterizes the observer is consciousness and that the collapse of the wave function happens because of an interaction of the consciousness on the physical system. In order to make his point, Wigner elaborated

an extension of the experiment of the Schrödinger cat, what is called the experiment of Wigner's friend, as shown in Figure 2.4. In this experiment, in addition to the cat, there is also Wigner and one of his scientist friends. The latter performs the experiment of the Schrödinger cat while Wigner is outside the laboratory. According to Wigner, the state of the system (cat+box+scientist) is a superposition state of (the atom having decayed, the cat having died, the friend having seen the dead cat) and (the atom having not decayed, the cat being still alive, the friend having seen the alive cat). But at some point, Wigner comes back and learn the result of the experiment. The idea is to show how consciousness is necessary: if instead of a conscious observer we have some apparatus, as we saw, the linearity of the wave function implies that the wave function is in a linear sum of possible states. In contrast, a conscious observer must be in either one state or another, and this is what makes conscious observations different. Consciousness will then act on the physical state to make it collapse into one of the terms.

As one can see, this is straightforwardly a denial of the closeness of the physical world and, as such, is a very radical position. That implies, also, that in order to construct a quantum mechanics without the measurement problem following this route we need to have a theory of how consciousness works and of how it interacts with the physical world.

2.3.3 Bohr and the Copenhagen Interpretation

Another attempt to solve the measurement problem is the one provided by Niels Bohr, also called the Copenhagen interpretation of quantum mechanics. According to Bohr, quantum mechanics is not a complete theory: it needs classical mechanics for its own foundations. For a detailed account of this view, see (Landau and Lifschitz 1977). So, in this sense, it is a solution that follows route 1. Therefore, there are two fundamental theories: the classical and the quantum theory. But while the classical theory deals with a clear and distinct world, made of cats, planets, tables and chairs, the quantum world is so obscure and far away from our ordinary experiences that it is impossible *in principle* to have a clear picture of it. The idea is that we do not have, intrinsically, the

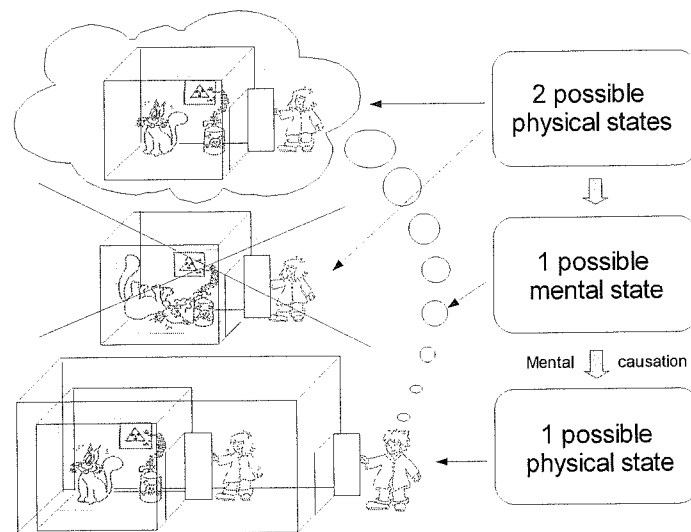


Figure 2.4: Wigner and the Schrödinger cat.

appropriate concepts to be used in the framework of quantum mechanics: the best we can do is to supplement the quantum description with a classical one.

This is where Bohr's (in)famous "wave-particle duality" came from: "wave" and "particle" are words that we use in ordinary language, but they do not really adequately account for what happens in the microscopic world. They provide at best a partial description of it. This idea came to Bohr probably reflecting on some experiments that at the time shown how what we would regard as particles sometimes would behave like waves and *vice versa*. One of the most famous examples of wave-like behaviors of particles is the two-slits experiments. In this experiment, as shown in Figure 2.5, particles are sent toward a screen with two small slits on it. A second screen is placed behind it to detect the particles. The particles, arriving one by one, hit the second screen, forming a spot. What one would expect is to find on the second screen, after a while, the image of the two slits, corresponding to the arrival of those particles which were not stopped by the first screen. What is found instead is an interference patterns, like a wave passing through the two slits would have produced. Suppose that previous experiments have identified the "entities" sent toward the screen as particles: for example, they showed a track in the bubble chamber. Then how can we explain

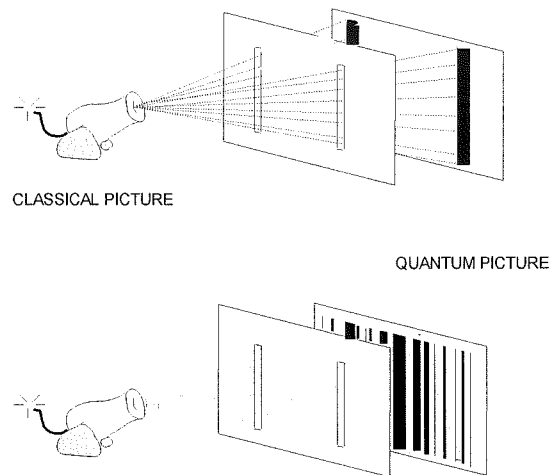


Figure 2.5: The two-slit experiments.

such a wave-like behavior? Bohr thought that we simply cannot: particles and waves are just our inadequate concepts, they do not reflect much about what the quantum world is like. It happens that we do not have, intrinsically, the appropriate concepts to describe quantum reality. Bohr actually went much farther than that. In fact, he concluded that not only in principle it is impossible for us to understand or coherently talk about the quantum world, but also there is no fact of the matter about it. Of course, one is not forced to reach such a radical conclusion from the discussion we have done so far: one thing is that we cannot describe something, another is that there is no reality to it! Nonetheless, nowadays this is what one commonly hears in physics departments all over the world.

Be that as it may, if we assume the quantum world to be real, according to Bohr we have two descriptions, one in which there are superpositions, that would be appropriate for the microscopic world, and one in which there are not, the classical one. Since the cat is a macroscopic object, by definition, she is never in a superposition state. A decaying nucleus instead obeys to the quantum laws, so it can stay in superpositions state. The nucleus, decaying or not, will have influence on what will happen to the cat but there is no paradox: the cat is either dead or alive because she is a classical object,

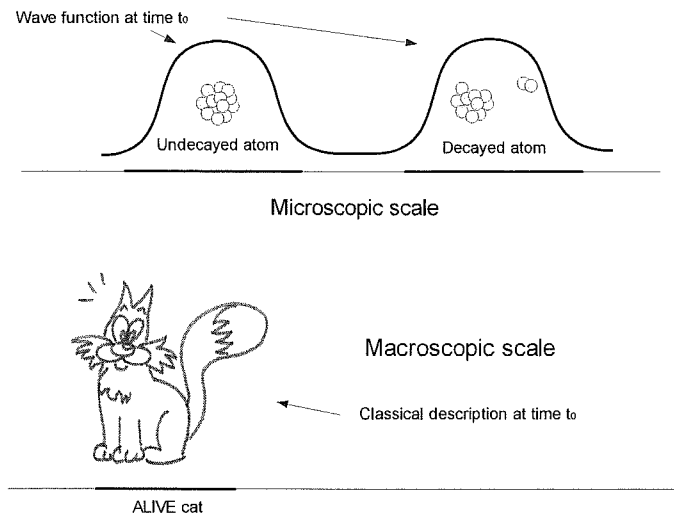


Figure 2.6: Bohr and the Schrödinger cat.

while it is not problematical for the nucleus to be in a superposition state because it is a quantum object, see Figure 2.6.

After a little thought it should be clear how this theory is intrinsically not very satisfactory: in fact, where is the cut between the “solid” classical world and the “wavy” quantum world? How many “particles” must an object have in order to be called macroscopic? This introduces a fundamental ambiguity into the theory, as Bell would put it, that should not be there in a fundamental physical theory:

Thus in contemporary quantum theory it seems that the world must be divided into a wavy quantum system, and a remainder that is in same sense classical... It introduces a fundamental ambiguity into fundamental physical theory (Bell 1987).

2.4 Quantum Mechanics without Observer

Albert Einstein was really troubled by quantum mechanics. The debate between Einstein and Bohr has been often considered the paradigmatic debate about the foundations of quantum theory. While Einstein insisted on the possibility and the necessity of having a formulation of quantum mechanics in which the observer did not play any fundamental role, and of describing the macroscopic world through such a theory, Bohr

solved the measurement problem in his own way, postulating that there are actually two fundamental descriptions, namely quantum and classical mechanics, that are complementary the one to the other. According to most physicists, Bohr was the winner of the debate: “We might not like it, but that is how the world is” is the often repeated slogan. After all, we are the product of evolution and there is nothing that *guarantees* that we will be able to properly describe reality. For example, Colin McGinn (McGinn 1989) has argued that consciousness is something exactly of that sort: we evolved in such a way that we do not actually have the right abilities to grasp what consciousness really is. Bohr is saying the same for the concept we use in physics. Therefore, we should embrace reality: we *thought* we had the appropriate concepts to understand the world but we *do not*. Wigner instead is insisting that consciousness has to play an active role in physics. Either way, the best we can do is to construct such unsatisfactory theories as these ones and we should learn how to live with them.

That might well be the case, but before accepting something like this one should better look at all the alternatives and see whether they work and whether we are really forced to give up any hope of understanding the microscopic world through physics as we know it. Why should we believe that we cannot comprehend the world if there are alternatives in which we actually can? I think that physicists have been a little too hasty in following Bohr or Wigner: before giving up so quickly to the strangenesses of the quantum world and adjusting to the idea that we have to give up the possibility of really understand what the world is like, we should at least analyze whether there are some other alternatives. So, what about them? Are really there theories in which the observer, or consciousness, does not play any crucial role and in which we do not need to postulate a quantum and a classical world? The answer is positive: There are more than one quantum theory that solve the measurement problem, provide a coherent representation of the microscopic world, that do not need the “observer”, the specification of the notion of “classical” or consciousness in its formulation. They can be labeled with the name of “quantum theories without observer”. The terminology is due to Karl Popper, who first used this expression in his article *Quantum Mechanics Without the Observer* (Popper 1967) and later used in particular by Bell (Bell 1987)

and Goldstein (Goldstein 1998). At the end of the day, since there are clear examples of quantum theories without observer, there is nothing that forces us to choose what von Neumann, Wigner and Bohr have proposed. Therefore, I think that we can claim that the real winner of the debate between the two scientists was Einstein, not Bohr: what Bohr believed to be impossible does actually exist, in more than one version. Let us discuss in more detail what these quantum theories without observer are.

2.4.1 Quantum Theory without Observer 1: Additional Variables

This quantum theory without observer solves the measurement problem taking the first of the two routes described above: the wave function does not provide the complete description of the system but something needs to be added to it. Historically, since these additional variables are not suggested by the Schrödinger equation (considered to be the core of quantum mechanics), this theory has also been called “theory of hidden variables”. It was first proposed, under the name of “pilot-wave theory”, by Louis de Broglie (de Broglie 1928) at the famous Solvay congress of 1927 for a one-particle system. Wolfgang Pauli had some objection to the theory and that discouraged de Broglie so much that he decided not to continue his investigations on the matter any further. The theory was proposed again, in a more general framework, by David Bohm in 1952 (Bohm 1952). In Bohm’s paper, the theory was presented in terms of the so-called “quantum potential” and this was very unfortunate, for a variety of reasons. A better formulation of Bohm’s theory is the one provided by Detlef Dürr, Shelly Goldstein and Nino Zanghí in their *Quantum Equilibrium and the Origin of Absolute Uncertainty* (DGZ 1992) under the name of Bohmian mechanics, in which the theory is presented by a system of two coupled first-order differential equations in which there is no mention of the mysterious quantum potential (for a review of the problems of the formulation in terms of the quantum potential, see (DGZ 1992)).

In this theory the complete description of the state of an N point-like particles system is given by the couple (Ψ_t, Q_t) , where $\Psi_t = \Psi_t(q)$ is the wave function of the system and $Q_t = (Q_1(t), \dots, Q_N(t))$ represents the configuration of the N particles

composing the system, q_k being the position of particle k in \mathbb{R}^3 ¹. Therefore, the variable Q belongs to the *configuration space* \mathbb{R}^{3N} , that is the space of the possible positions that a physical system composed of N particles may have.

The world “particle” in this theory should be taken seriously: there are *really* particles in the world if Bohmian mechanics is true, just like there would have been if classical mechanics were true. Each $Q_k(t)$ is the actual trajectory of the k -th particle in three-dimensional space \mathbb{R}^3 . This is a very big difference with bare quantum mechanics in which the state of the same system is given only by the wave function and there are no particles with no positions and no trajectories whatsoever.

In Bohmian mechanics, the wave function ψ evolves according to Schrödinger’s equation, as in bare quantum theory:

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi, \quad (2.5)$$

where H is the usual nonrelativistic Schrödinger Hamiltonian. The particles evolve according to the so-called “guide” or “guiding” equation, which is determined by the wave function:

$$\frac{dQ_k}{dt} = v_k^\psi(Q_1, \dots, Q_N) = \frac{\hbar}{m_k} \operatorname{Im} \frac{\psi^* \nabla_k \psi}{\psi^* \psi}(Q_1, \dots, Q_N), \quad (2.6)$$

where m_k , $k = 1, \dots, N$, are the masses of the particles. That is, the wave function defines a velocity field v^ψ for the particles. This is the sense in which the wave function “guides” the motion of the particles, and this is where the original name of the theory, “pilot-wave theory”, comes from. Note, though, that this is a velocity field in configuration space, not in three-dimensional space.

It should be noted that in Bohmian mechanics the wave function Ψ is the wave function of the entire universe. But the fact that we have configurations in Bohmian mechanics allows us to define the wave function of a given physical system (smaller than the universe): it is the *conditional* wave function, as described in (DGZ 1992), in which

¹A remark about notations: usually capital letters are used for the *actual* values. Otherwise variables are written in lower case. Roughly speaking, where there is a Q_k (and not a q_k), there is particle k .

we simply plug in the actual configurations of the particles that constitutes everything but the system. If Y represents the actual configuration of all the particles but those that compose the system of interest, which are described by the configuration x , then the wave function of the system is given by $\psi(x) = \Psi(x, Y)$.

Equations (2.5) and (2.6) form a complete specification of the theory. What we have in Bohmian mechanics is a dynamical system for the variables (Ψ, Q) . Without any other axiom (about properties assignment, for example), all the results obtained in the framework of non relativistic quantum mechanics follow from the analysis of this system.

As a consequence of Schrödinger's equation and of Bohm's law of motion, we have an important consequence: the distribution $|\psi(q)|^2$ is "equivariant" (see (DGZ 1992)). This means that if the configuration $Q_t = (Q_1(t), \dots, Q_N(t))$ of a system is random with distribution $|\psi_t|^2$ at some time t , then this will be true also for any other time t . Thus, we can consistently assume the *quantum equilibrium hypothesis*, which asserts that whenever a system has wave function ψ_t , its configuration Q_t is random with distribution $|\psi_t|^2$ ². This hypothesis is not as hypothetical as its name may suggest: it follows in fact from the law of large numbers under the assumption that the initial configuration of the universe is typical (i.e., not-too-special) for the $|\Psi|^2$ distribution, with Ψ the initial wave function of the universe. As a consequence of the quantum equilibrium hypothesis, a Bohmian universe, even if deterministic, appears random to its inhabitants. In fact, the probability distributions observed by the inhabitants agree exactly with those of the quantum formalism (see (DGZ 1992) for details). We will come back on this issue in Chapter 8.

Here is a rough idea of why in Bohmian mechanics the problem of the cat does not arise (see Figure 2.7). Even if the wave function is in a superposition, the complete description of the system is given by the wave function and by the positions of particles. Therefore, as far as they are concerned, they are either "here" or "not here". The particles composing the cat are initially in a certain configuration that corresponds to

²The word "equilibrium" here is used as Boltzmann would have used it: the equilibrium state is the state in phase space that has the biggest size.

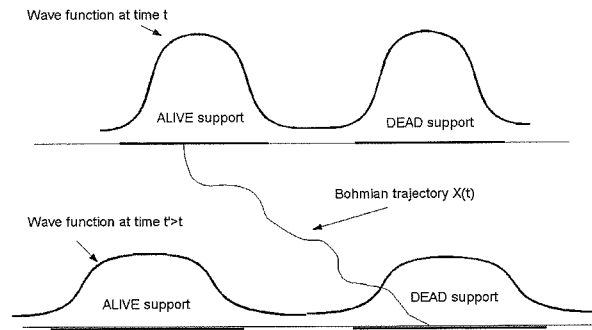


Figure 2.7: Bohmian mechanics and the Schrödinger cat.

an alive cat. With that we mean simply (and roughly) that the particles are so arranged that they form molecules that interact with each other as they do in living things. Call the set of configuration corresponding to an alive cat L and the one corresponding to a dead cat D . Note that the state of a dead cat is macroscopically distinct from that of an alive cat: it is possible to specify macroscopic quantities to differentiate the two states, like for example the temperature of the cat. Due to the fact that the wave function is what it is, the configuration of particles will evolve into a given final configuration. At the final time (which is enough to assume that the experiment is over) the particles will end up, again, either “here” or “not here”. In particular, they will end up either in that set L of configurations corresponding to an alive cat or in the set D corresponding to a dead cat. The conditional wave function of the cat will make it the case that the bump whose support does not contain the particles of the cat will not influence the other bump, so that we can consider it to evolve *as if* there was only just the bump containing the particle.

2.4.2 Quantum Theory without Observer 2: Nonlinear Evolution of the Wave Function

We have discussed already how von Neumann's approach can be regarded as a theory that takes route 2 (that is, the denial that the wave function evolves according to Schrödinger's equation) to solve the measurement problem. This theory had the problem that it needed a definition of "observer" in order to establish when we have the collapse of the wave function. A much more satisfactory realization of this second possibility that does not involve the notion of observer at any level is the theory called "spontaneous collapse" or "spontaneous localization" theory. The project was initiated by Philip Pearl (Pearl 1976) in the 70s and developed further by Gian Carlo Ghirardi, Alberto Rimini and Tulio Weber in the 80s (GRW 1986). Other names under which this theory is known are "dynamical reduction" theory and, more simply, "GRW" theory, from the initials of the names of the developers of the theory. In this theory the wave function does not evolve according to Schrödinger's equation. Rather, it evolves according to a different equation in which the superposition wave function spontaneously "collapses" to one of its terms.

We can imagine that the deterministic evolution is for some time undisturbed and then, at an entirely random moment, it is interrupted by the stochastic one, after which the deterministic evolution again prevails. These "jumps" happen at random times with an average frequency λ that in the original GRW model is of the order of $\lambda \sim 10^{-15}\text{s}^{-1}$ (that roughly means that the stochastic evolution prevails every 300 millions years). This parameter should be intended as a new constant of nature.

In von Neumann's theory, the collapse rule tells us that observation changes randomly the state of an object from the initial wave function ψ to one of the possible results states. In GRW instead such a rule is a fundamental law of nature. When nature acts on the wave function, it localizes it in a neighborhood of a given position $x \in \mathbb{R}^3$. But what is the wave function transformed to? One simple possibility is to assume that the initial wave function gets multiplied by a Gaussian with a given dispersion, σ . The parameter σ should to be considered an additional constant of nature: The empirical

predictions of the theory decide what is the most suitable value for it, which is $\sigma \sim 10^{-7}$ m. With these values for λ and σ , the predictions of GRW theory, for suitably short times, are indistinguishable from those of standard quantum mechanics.

More technically, the situation is the following. Consider a quantum system described by an N -“particle”.³ wave function $\psi = \psi(q_1, \dots, q_N)$, $q_k \in \mathbb{R}^3$, $k = 1, \dots, N$; for any point x in \mathbb{R}^3 (the “center” of the collapse that will be defined next), define the *collapse operator*

$$L_i(x) = \frac{1}{(2\pi\sigma^2)^{3/2}} e^{-\frac{(\widehat{Q}_i - x)^2}{2\sigma^2}}, \quad (2.7)$$

where \widehat{Q}_i is the position operator of “particle” i . Let ψ_{t_0} be the initial wave function. Then ψ evolves in the following way:

1. It evolves unitarily, according to Schrödinger’s equation, until a random time $T_1 = t_0 + \Delta T_1$, so that

$$\psi_{T_1} = U_{\Delta T_1} \psi_{t_0}, \quad (2.8)$$

where ΔT_1 is a random time distributed according to the exponential distribution with rate $N\lambda$.

2. At time T_1 it undergoes an instantaneous collapse with random center X_1 and random label I_1 according to

$$\psi_{T_1} \mapsto \psi_{T_1+} = \frac{L_{I_1}(X_1)^{1/2} \psi_{T_1}}{\|L_{I_1}(X_1)^{1/2} \psi_{T_1}\|}. \quad (2.9)$$

I_1 is chosen at random in the set $\{1, \dots, N\}$ with uniform distribution. Different labels identify different particles. The center of the collapse X_1 is chosen randomly with probability distribution.

$$\mathbb{P}(X_1 \in dx_1 | \psi_{T_1}, I_1 = i_1) = \frac{\int \psi_{T_1}^* L_{i_1}(x_1) \psi_{T_1} dx_1}{\int \psi_{T_1}^* L_{i_1}(x_1) \psi_{T_1} dx_1} \frac{L_{i_1}(x_1)^{1/2} \psi_{T_1}}{\|L_{i_1}(x_1)^{1/2} \psi_{T_1}\|} dx_1. \quad (2.10)$$

3. Then the algorithm is iterated: ψ_{T_1+} evolves unitarily until a random time $T_2 =$

³Note that there are no real particles in this theory: the word “particle” is used only for convenience in order to be able to use the standard notation and terminology.

$T_1 + \Delta T_2$, where ΔT_2 is a random time (independent of ΔT_1) distributed according to the exponential distribution with rate $N\lambda$, and so on.

In other words again, the evolution of the wave function is the Schrödinger evolution interrupted by collapses. When the wave function is ψ , a collapse with center x and label i occurs at rate

$$r(x, i|\psi) = \lambda \langle \psi | L_i(x) \psi \rangle, \quad (2.11)$$

and when this happens, the wave function changes to $L_i(x)^{1/2} \psi / \|L_i(x)^{1/2} \psi\|$.

Thus, if between time t_0 and any time $t > t_0$, n collapses have occurred at the times $t_0 < T_1 < T_2 < \dots < T_n < t$, with centers X_1, \dots, X_n and labels I_1, \dots, I_n , the wave function at time t will be

$$\psi_t = \frac{L_{t,t_0}^{F_n} \psi_{t_0}}{\|L_{t,t_0}^{F_n} \psi_{t_0}\|} \quad (2.12)$$

where $F_n = \{(X_1, T_1, I_1), \dots, (X_n, T_n, I_n)\}$ and

$$L_{t,t_0}^{F_n} = U_{t-T_n} L_{I_n}(X_n)^{1/2} U_{T_n-T_{n-1}} L_{I_{n-1}}(X_{n-1})^{1/2} U_{T_{n-1}-T_{n-2}} \dots L_{I_1}(X_1)^{1/2} U_{T_1-t_0}. \quad (2.13)$$

Since T_i , X_i , I_i and n are random, ψ_t is also random. It should be observed that (unless t_0 is the initial time of the universe) also ψ_{t_0} should be regarded as random, being determined by the collapses that occurred at times earlier than t_0 . However, *given* ψ_{t_0} , the statistics of the future evolution of the wave function is completely determined; for example, the joint distribution of the first n collapses after t_0 , with particle labels $I_1, \dots, I_n \in \{1, \dots, N\}$, is

$$\mathbb{P}(X_1 \in dx_1, T_1 \in dt_1, I_1 = i_1, \dots, X_n \in dx_n, T_n \in dt_n, I_n = i_n | \psi_{t_0}) = \lambda^n e^{-N\lambda(t_n-t_0)} \|L_{t_n,t_0}^{f_n} \psi_{t_0}\|^2 dx_1 dt_1 \dots dx_n dt_n, \quad (2.14)$$

with $f_n = \{(x_1, t_1, i_1), \dots, (x_n, t_n, i_n)\}$ and $L_{t_n,t_0}^{f_n}$ given, *mutatis mutandis*, by (2.13).

The rate of collapses is given by $N\lambda$, where N is the numbers of “particles” of the system, so that in the case of a macroscopic object in which $N \sim 10^{23}$, we have 10^8 collapses per seconds. That is, using the words of Bell (also, see Figure 2.4.2):

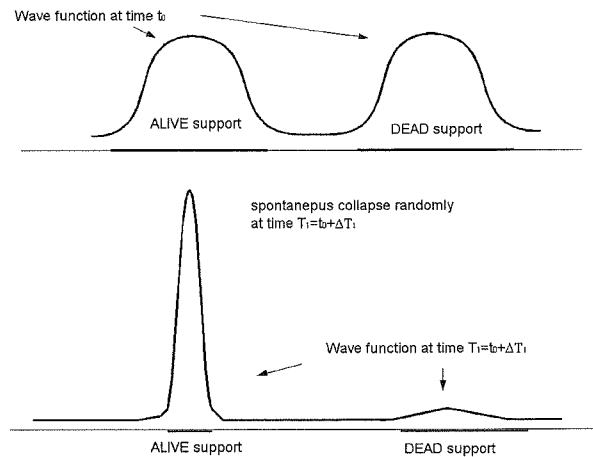


Figure 2.8: GRW theory and the Schrödinger cat.

any embarrassing macroscopic ambiguity in the usual theory is only momentary in the GRW theory. The cat is not both dead and alive for more than a split second (Bell 1987).

2.4.3 Quantum Theory without Observer 3: Measurement have no Results

Following Bell's alternative presented in Section 2.2, it seems that in Bohmian mechanics we add some entity to complete the description provided by the wave function, while in GRW theory we change the evolution of the wave function. One might be willing to keep both the completeness of the description of the wave function and the simplicity of the linearity of the Schrödinger evolution. Then, the only option is to reject the idea that measurement have results: in that case, option (3) listed in Section 2.2 would be false. This is the approach first proposed by Hugh Everett in his PhD thesis *Relative State Formulation of Quantum Mechanics* (Everett 1950), even if he framed his intent in slightly different terms.

The idea is that the wave function, even if it can stay in superposition, provides the complete description of the universe even if it does not seem to be the case. This is

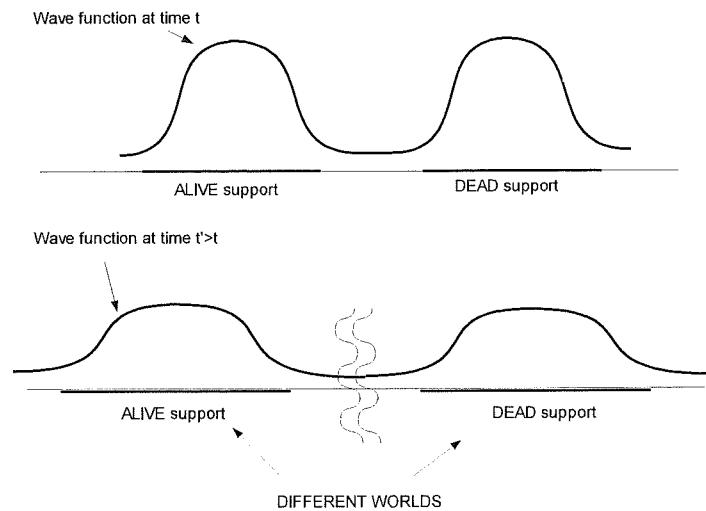


Figure 2.9: The many worlds theory and the Schrödinger cat.

due to the fact that the universe is very different from what we think it is. Each term of the superposition wave function represents different states of affair, corresponding to different measurement results. If the wave function is in a superposition state, we come to the conclusion that measurements do not have results if we think as the results being realized all in *this* space-time. If instead we have a more “liberal” view of what the universe is, we might interpret each term as living in a different space-time, in a different world. In this respect, it is not that measurement do not have results, they do not have results in this space-time. Rather, all the possible measurement results indeed are realized in the *multiverse*, the space of all space-times in which each result can be thought of realizing, see Figure 2.9. For this reason this theory is also called “many worlds” quantum mechanics.

This is a rather wild idea, and, as presented, it is also very vague. In fact, stated like this, it is not so obvious how it should be intended. First of all, when a measurement happens it seems we are supposed to think that there is a “splitting” of the world in a number of other words, one for any possible result of the measurement. But how exactly should we intend this splitting? One should specify what these “words” are: Are they really different space-times or are in the same space-time but they are superimposed

into the same space-time but “transparent” the one to the other? There is no right or wrong answer: the point is that depending on what the answer to this question is, we have a different theory.

Attempts have been made to make precise this theory, focusing on the specification of what the worlds are in terms of *subjective* experiences of the observer. In this way, we have solutions of the measurement problem which are in some sense very close to Wigner’s solution. This is the original proposal of Everett, that called his theory “relative state” formulation of quantum mechanics, since the states are relative to the observer. A possible solution, also based on the attempt to account for the observer’s experiences and proposed by David Albert and Barry Loewer in their *Interpreting the Many-Worlds Interpretation* (Albert and Loewer 1988), is the so called “many minds” theory: while physical states described by the wave function linearly evolve, mental states are not in superpositions but they collapse. Also in this theory, as in Wigner’s one, we explicitly mention the mind. But the two theories are different, since in Wigner’s picture consciousness acts on physical bodies to make them collapse, while here we just have collapse of the mental states without any mental causation.

Be that as it may, as soon as we talk about the observer’s subjective experiences, we need to invoke a theory of mind, as in Wigner’s theory. And therefore the same objections raised for that theory will apply also in this case. The challenge is therefore to figure out whether a many worlds theory in which consciousness or mental states do not play any role can be developed. And the conclusion is that there does not seem to be any. Let us in fact ask: What are the differences between bare quantum mechanics and the many worlds theory? After all, they both have the same ingredients: there is just the wave function ψ , and it evolves according to Schrödinger’s equation. We have seen that the bare theory has the measurement problem, so why this one does not? The strategy of a many world theory is to show that our perceptions of the macroscopic world are, somehow, not in superposition. And in order to succeed in this project they will have to provide a theory of perception (on the recognition of this, see David Wallace in his *Everett and Structure* (Wallace 2003)).

To conclude, I wish to notice that there is a huge amount of people that find the

many worlds approach rather fascinating. I think there are at least three reasons for this. The first is the need of preserving the mathematical beauty and the simplicity of the bare theory, which is accomplished directly in many worlds, since in its framework the wave function provides the complete description and it evolves linearly. It should be noted, though, that it is not so obvious what counts as simpler: completeness of the wave functions, linearity of Schrödinger's equation, and a theory of the mind or one of the alternatives listed above? Another reason that is usually provided in favor of the many worlds account is that the reconciliation of quantum mechanics with relativity seems to be easier in the case of many worlds theory than in the other alternatives. So many worlds really seems to be the most promising of all: no additional variables, no ugly dynamics, maybe a little of philosophy of mind, but a more or less natural relativistic invariance. In this respect, no surprise that most of the physicists or philosophers that declare their position in this matter claim their favorite theory to be Everett. When we will discuss what symmetry properties are in Chapter 6, it will be evident that exactly the opposite is the case. Another, not so noble, reason for being fascinated by many worlds quantum mechanics is that, somehow, this theory seems to solve all the problems but also keep some of the paradoxical flavor of quantum mechanics. People are fascinated by strange and "mysterious" consequences and this was one of the appeal of quantum mechanics at some point. But maybe it went too far, as the Schrödinger's cat paradox has shown. Now, a theory like the many worlds theory, which is less extreme but not so much seems just perfect: it allows for parallel universes (in some respect), to time travel (in some respect, see (Deutsch and Lockwood 1994)), and some weird theories of personal identity (see (Barrett 2003)), so we can write a lot of articles about it! Bell (Bell 1987) talked about "romantic" pictures of the world provided by quantum mechanics as opposed to "unromantic" ones. To the former group belong Bohmian mechanics and GRW theory (at least in some version of it, as we will see), and many worlds theory seems to fit very naturally among the romantic ones (see also (Tumulka 2007) for a nice comparison between romantic and unromantic views of quantum mechanics).

2.6 The Problem of the Lack of a Clear Ontology

In all this chapter we have talked about wave functions, operators, measurement results, particles... but what is “out there” in the world? The question that should be asked now is the following: *What is quantum mechanics about?* This question is a tricky one: when I say “about”, what do I really mean? One might think that the question above could be paraphrased as follows: *What is there in the world if quantum mechanics is true?* That is, what is the ontology of the theory? If we do not specify which one, among the mathematical variables that appear into the specification of the theory, correspond to what is “out there”, the theory remains just empty mathematics.

But the question above could also mean the following: *What are tables and chairs made of?* In fact, there seems to be no reason to restrict what exists to what exists *in the physical world*, since, for example, numbers and laws could exist without being physical. Since quantum mechanics is a fundamental *physical* theory, the question: “What is it about?”, should be intended in my opinion as: “What are tables and chairs made of?”, and not as the more general question: “What is there if the theory is true?”

There are a lot of variables in a fundamental physical theory. If we want to be realist about it, we need to decide which of the different variables represent physical objects and therefore we should distinguish them from the other variables that appear into the theory as well but do not represent physical objects. Some of the variables of the theory are mathematical constructs so they exist in the same sense as numbers exist: the operators in quantum mechanics, the Hamiltonian, the potential energy, and so on. Some of the variables may correspond to something physically real: positions of particles, field values, strings, the wave function and so on. I will call (the mathematical representation of) what constitutes physical objects at the fundamental level *primitive ontology* of the theory, as suggested by Dürr, Goldstein and Zanghí, (DGZ 1992).

2.7 The Problem of the Adequacy of the Primitive Ontology

Are all mathematical variables adequate primitive ontology? To answer this question one should determine, among the variables in the theories, which ones are suitable to represent physical objects. We have already seen how some of the primitive ontology proposed for quantum theories without observer are, more or less obviously, inadequate: von Neumann's theory, and the Copenhagen interpretation include in their very formulation some notions that are intrinsically vague, and this is unacceptable for a fundamental physical theory. And when we try to make precise what an observer, say, means, we end up with the need of consciousness to explain the physical world. For this reason, people have considered the quantum theories without observer, Bohmian mechanics, the GRW theory, and many worlds. Let us now see in more detail what we have learned so far about them. Suppose we start from a realistic attitude toward quantum mechanics and we want to know what tables and chairs are. Our first guess would be that they are made of wave functions, what is called *monism* about the wave function: after all, isn't it *the* object of quantum mechanics? The entity that evolves in time according to Schrödinger's equation? The possibility of regarding quantum mechanics as a theory about the wave function is something that also Schrödinger would have liked very much, of course: he was one of the first who tried to allow for the wave function to be the complete description of a system. But, as we have seen in the cat experiment, the wave function of the entire system, evolving to the Schrödinger equation, had in it

the living and the dead cat (pardon the expression) mixed or smeared out in equal parts (Schrödinger 1983).

Schrödinger therefore dismissed the idea of the wave function representing reality in a complete way because of the presence of these macroscopic superpositions. Now, it seems that in GRW we can avoid this problem allowing for quantum jumps. That is, it seems that we can describe the world completely just with the wave function, if we allow it to have a nonlinear evolution. As Bell pointed out, Schrödinger would have probably liked those quantum jumps: at the same time they do not allow for

macroscopic superpositions and it seems we can still think about the wave function as providing the complete description, without the need of adding anything to the theory. This is not exactly right: also in GRW we need to add something to the description provided by the wave function. As we saw in the problem of the tails, we need to add some rule to make a connection between our macroscopic talk and the talk of fundamental physics. I will argue that in GRW we have to add much more than that, since these practical rules are needed to define the primitive ontology of the theory.

So, the question is the following: Can the wave function be an adequate primitive ontology? Can the wave function represent physical objects? That is, is it possible for the wave function to be what the theory is about? I will argue in Chapter 5 that it does not: it is intrinsically an inadequate primitive ontology. Not so much because it can be in superposition (otherwise with fields we will be in trouble all the time), but rather because it lives in \mathbb{R}^{3N} , configuration space, which is a space of very large number of dimensions: if we think that the world is made of, just to play it safe, an Avogadro number of “particles”, then the dimension of configuration is $M = 3N = 3 \times 10^{23}$! And, as such, it does not have any possibility of representing, by itself, an object in three-dimensional space, unless we already assume there are particles in \mathbb{R}^3 . If that is right, then, not even in GRW tables and chairs are made of wave functions.

Before dealing with this problem, in the next chapter we will analyze the different interpretation of the different solutions of the measurement problem in more detail. This will allow me to divide them into two groups and to compare them in Chapter 4 and 5, arguing that the one in which the wave function is taken as the primitive ontology of the theory is very problematical, at best, and should be abandoned.

The terminology “local beables” has been introduced by Bell in the framework of GRW theory:

These are the mathematical counterparts in the theory to real events at definite places and times in the real world (as distinct from the many purely mathematical constructions that occur in the working out of physical theories, as distinct from things which may be real but not localized, and distinct from the “observables” of other formulations of quantum mechanics, for which we have no use here). A piece of matter then is a galaxy of such events (Bell 1987).

“Beable” is short for “maybeable”, things that might exist if the theory is true. Among the variables in the theory there are certain entities that can mathematically represent localizable objects and some that cannot. If with localized we mean that they can be attached to a region of space-time, then, for example, positions, fields and strings are local beables, while the wave function is not. Therefore, if a local beable is what can represent physical objects, then the wave function cannot do that. The notion of local beable seems similar to the notion of primitive ontology. But, for reasons that will be clear later, not all local beables (such as the electric and magnetic fields in classical electrodynamics) need to be regarded as part of the primitive ontology. If so, then in classical electrodynamics fields do not constitute physical objects. We will come back on the differences between local beables and primitive ontology in Chapter 6.