Introductory lecture on Bohmian mechanics

Florian Hoffmann, Nicola Vona

April 15, 2014

This lecture is aimed at students, who are familiar with some basic quantum mechanics. Its aim is to convey the basic ideas and show some applications of Bohmian mechanics in a rather accessible way. Consequently, it sometimes ignores subtleties and demonstrates things in a simplified fashion. For a more thorough account we refer the reader to the literature mentioned in the bibliography.

Introduction

What Bohmian mechanics is

Bohmian mechanics is a theory of quantum phenomena based on a particle picture. On a fundamental level, it explains the world around us in terms of point particles that move on deterministic paths. When used to describe experimental situations, Bohmian mechanics yields the textbook formalism of quantum mechanics, and can therefore be considered as an explanation of the axioms of standard quantum mechanics. Using Bohmian mechanics, one can easily understand all of the quantum puzzles.

Motivation

So, why exactly is Bohmian mechanics needed?

When one studies quantum mechanics one learns how to formally describe the state of a system and how to find its evolution in time; on top of that, one also learns a set of rules that are used to get from the state of the system predictions about the statistics of the outcomes of experiments. Without any doubt, this framework is very effective considering its empirical predictions. But how do these rules come about? What is their content? What is their meaning? What do they tell us about the physical world?

To date, there have been four main attempts to answer these questions. Two of them, namely many worlds and operationalism, do not change the basic equations of the theory. A third possibility is represented by the spontaneous localization models. They are modifications of the Schrödinger equation that embed the collapse of the wave function in the free dynamics.

We will not elaborate further on any of these three attempts here, but we will only concentrate on the fourth one, namely Bohmian mechanics, that also changes the theoretical structure of ordinary quantum mechanics, not by modifying the Schrödinger equation, but rather by complementing it with a new equation. Bohmian mechanics does directly spring from the quantum formalism if one just dares to ask "what if there really are particles, really moving on definite paths?".

There are two main merits of Bohmian mechanics. At first, it proves that quantum randomness can be explained within a deterministic theory, in a way similar to how Newtonian mechanics explains thermodynamics. At second, in a Bohmian framework all of the famous quantum puzzles can be easily understood. This is a consequence of the fact that the rules for the empirical predictions that one uses in quantum mechanics rest within Bohmian mechanics on a mathematically rigorous and conceptionally clear basis. Bohmian mechanics is a fundamental theory that can be applied to any physical system. If one specializes it to measurement situations, one gets an effective theory of measurement outcomes that is nothing else than ordinary quantum mechanics. In this sense there is no friction between ordinary quantum mechanics and Bohmian mechanics: the latter does not alter the former, rather it provides a way to explain it.

But there are also some words of warning in order here. Even though Bohmian mechanics is a deterministic theory about the motion of actual particles, this motion is in no way classical. The Bohmian dynamics drastically differ from what one would expect classically; reasonably, one should not have expected otherwise. After all Bohmian mechanics does not originate from the will to carry over classical mechanics to the quantum realm, but springs from the question whether there is a way to understand the quantum formalism in terms of particles having an objective and definite position at any time.

1 The state of a system and its dynamics

Consider at first a system composed of a single particle. Ordinary quantum mechanics describes the state of such a system by means of a function on the space of possible positions of the particle , $\psi(\mathbf{q})$ with $\mathbf{q} \in \mathbb{R}^3$; this function is called *wave function*. Additionally, in Bohmian mechanics the particle has an actual position $\mathbf{Q} \in \mathbb{R}^3$. In other words, in Bohmian mechanics, a particle really is a point in physical space, characterized by its position and by a wave function.

It is important not to confuse the actual particle position $\mathbf{Q}$ with the variable $\mathbf{q}$ of the wave function. The latter is just the variable of a function, that can be evaluated at any point, whereas $\mathbf{Q}$ represents the actual position of the particle, and has a definite value.

As time passes, the position and the wave function change, therefore we have to write them as functions of time: $\mathbf{Q}(t)$, and $\psi(\mathbf{q}, t)$. If the wave function at some time t_0 is known, it is possible to calculate it for every time t by using the Schrödinger equation, as in ordinary quantum mechanics:

$$i\hbar \frac{\partial}{\partial t}\psi(\mathbf{q}, t) = H\psi(\mathbf{q}, t). \tag{1}$$

Here H is the usual Hamilton operator, e.g. for a particle of mass m in a potential V we

have

$$H = -\frac{\hbar^2}{2m}\Delta + V. \tag{2}$$

Conversely, to determine the actual position at time t, it is not enough to know the position at time t_0, but also the wave function is needed. Indeed, the evolution of the actual position is dictated by the wave function through the guiding equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{Q}(t) = \frac{\hbar}{m}\mathrm{Im}\frac{\psi^*\nabla\psi}{|\psi|^2}(\mathbf{Q}(t), t), \tag{3}$$

where Im denotes the imaginary part.

There are several ways to motivate this equation. One is asking for the simplest Galilei invariant vector field that can be derived from a function on the space of possible positions [1, sec. 8.1]. Another motivation, easliy accessible to those already familiar with QM, is the following. It is easy to prove, that from the Schrödinger equation we get the relation

$$\frac{\partial}{\partial t}|\psi|^2 = -\nabla\cdot\mathbf{j}, \tag{4}$$

with

$$\mathbf{j} := \frac{\hbar}{m}\mathrm{Im}(\psi^*\nabla\psi) \tag{5}$$

called the *quantum current*. Equation (4) is a continuity equation and it states that the quantity $|\psi|^2$ is "never created nor destructed", but rather it just moves around in space along the trajectories that have velocity $\mathbf{j}/|\psi|^2$. Therefore, the wave function and the Schrödinger equation naturally define the trajectories which are solution of equation (3).

The quantity $|\psi(\mathbf{q}, t)|^2$ is interpreted in ordinary quantum mechanics as the probability density of finding the particle at the position $\mathbf{q}$ if one were to look for it at time t. We will see in the next section that similarly in Bohmian mechanics $|\psi(\mathbf{q}, t)|^2$ is interpreted as the probability density for the actual position $\mathbf{Q}(t)$ to *be* at the location $\mathbf{q}$ at time t. This implies of course also that you find it there if you look for it!

Summing up, the wave function $\psi(\mathbf{q}, t_0)$ and the actual position $\mathbf{Q}(t_0)$ at some time t_0 uniquely determine the motion of the particle for all later and earlier times. Consequently, Bohmian mechanics is an entirely deterministic theory in which the particle moves on a continuous path through space and has a well defined position at any time.

If the system consists of more than one particle, say there are N particles, then according to Bohmian mechanics each particle has an actual position $\mathbf{Q}_k \in \mathbb{R}^3$, $k = 1, \ldots, N$; the collection of these positions $Q := (\mathbf{Q}_1, \ldots, \mathbf{Q}_N) \in \mathbb{R}^{3N}$ is called *configuration* of the system. The wave function is in this case a function on the space of possible positions of the N particles, $\psi(\mathbf{q}_1, \ldots, \mathbf{q}_N)$, with $\mathbf{q}_k \in \mathbb{R}^3$, and $k = 1, \ldots, N$, that we can also write as $\psi(q)$ with $q \in \mathbb{R}^{3N}$. Indicating by Δ_k and ∇_k the Laplacian and the gradient with respect to the coordinates of the k-th particle, respectively, the Schrödinger equation is still equation (1), with

$$H = -\sum_k \frac{\hbar^2}{2m_k}\Delta_k + V, \tag{6}$$

and the guidance equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{Q}_k(t) = \frac{\hbar}{m_k}\mathrm{Im}\frac{\psi^*\nabla_k\psi}{|\psi|^2}(Q(t), t). \tag{7}$$

Note that on the left hand side there is $\mathbf{Q}_k$, while on the right hand side there is Q: the velocity of the kth particle depends on the position of all of the N particles of the system.

If the particles have spin, the only difference is that the wave function is spinor-valued rather than a scalar, and ψ^* in the guidance equation should be replaced by $\psi^\dagger$, the Hermitian conjugate. Spin is a property of the wave function, not of the particles, and its treatment in Bohmian mechanics and in ordinary quantum mechanics are exactly the same.

Summing up, the axioms of Bohmian mechanics are:

1. There are particles. That means point particles, always having definite position. The physical world described by Bohmian mechanics is made out of these particles. Objects like atoms in an experiment just as the experimental setup, the apparatus, the furniture of the lab etc. are made out of these particles.

2. The particles move. To describe the motion of the particles, a new entity is needed, a wave function ψ, defined on the space of the possible configurations of all particles. The wave function ψ evolves in time according to the Schrödinger equation (1). Then, the motion of the particles is controlled by ψ through the Bohmian velocity law (3).

To use Bohmian mechanics to derive statistical prediction, we need in addition a statistical axiom:

3. $|\psi|^2$ gives the right notion of typicality for the whole universe.

The meaning of this last axiom will be the topic of the next section.

From these ingredients the entire formalism of textbook quantum mechanics follows as theorems.

This concludes the presentation of the theory. In the next sections we will present some analysis of it. But before that, let us make a short remark about the possible extension of the theory to the relativistic regime. The biggest problem is that to calculate the velocity of the ith particle at time t from the Bohmian law (3) one needs to substitute into the argument of the wave function the position of all other particles at the same time t. This operation is clearly problematic in a relativistic setting, where simultaneity depends on the reference frame. The easiest possibility would be to introduce some preferred reference frame, that specifies the needed simultaneity to use the Bohmian velocity law (3). This does not contradict relativity theory because one finds that the predictions for experiments are not frame dependent; anyhow, this idea is of course against the spirit of relativity. A more satisfactory solution is for instance to introduce a distinguished foliation of spacetime derivable from something already present in the theory, like for example the wave function. There are several possibilities to do that, all making the same empirical predictions. Which one is the correct one is an open research question.

2 Equivariance and Born Rule

As we have seen Bohmian mechanics is a deterministic theory of quantum phenomena: specification of the positions and of the wave function at some time t_0, uniquely determine the entire future evolution, and therefore also the positions at any future time.

Now one might be confused, because it is often thought that the most fundamental features of quantum mechanics are randomness and indeterminism: do we not have evidence of randomness in quantum systems? Do we not have experiments on identically prepared systems yielding empirical distributions? How can this be if the foundation of quantum behavior is entirely deterministic, as Bohmian mechanics states?

The question is actually the much more general one of how randomness enters a deterministic theory, which was answered conclusively for the first time in the work of Boltzmann, Smoluchowski, and Einstein at the beginning of the 20th century. This question is intimately related to what randomness and probability are, which is a hard question. When one speaks about randomness in physics one speaks about the typical behavior of empirical distributions – the relative frequencies of outcomes of measurements. In the following we will explain this statement.

The situation in Bohmian mechanics is very similar to the way randomness enters in classical statistical mechanics. Let us explain the Newtonian situation with an easy example. Consider a machine in which one can load one thousand coins and that then tosses them one after the other. The outcome of all the coin tosses is determined by the initial positions and momenta of all the constituents of the coins, the machine, the table the coins fall on, the air surrounding them and what else is relevant. For every initial condition of this combined system, one has a unique final state, which gives a unique series of heads and tails. For practical purposes to which series of heads and tails a certain initial state leads is of little relevance; what is really interesting is the question what typical initial states do, that means what the overwhelming majority of initial states do. To judge what the overwhelming majority is, one needs a way of counting states. As the space of initial conditions is a continuum, we cannot simply count states, but rather need a measure to evaluate the size of a subset. In a moment we will see how to get a handle on finding the correct typicality measure. Summarizing we can state that analyzing the situation for coin tossing, one finds that typical initial data yield sequences of outcomes which contain approximately as many heads as tails.

Now, let us have a look at Bohmian mechanics. Let us consider a machine which shoots thousand electrons all with the same wave function φ one after the other through a double slit. As Bohmian mechanics is a deterministic theory, the positions at which the particles hit the screen are uniquely determined by the initial conditions of the biggest relevant system, i.e. the actual configuration Q and the wave function Ψ of that system. From now on we will call the biggest relevant system simply the "universe". The randomness we want to explain here is the randomness that is present in standard quantum mechanics, i.e. given a wave function the observed positions are still random. So let us take a fixed Ψ and ask: what do typical initial configurations Q do? This is analogous to what we did in the coin tossing example, where we asked what the overwhelming majority of initial configurations do. As above, we need a measure of typicality on the space of possible configurations Q.

How do we know what the correct measure of typicality is? This measure should be singled out by the dynamics, otherwise its relevance would be dubious. Let us first look at the familiar example of Newtonian physics to explain this. Let ρ be the density corresponding to the measure of typicality, which is a measure on phase space, and v the vector field generating the dynamics, that is the Hamiltonian vector-field. The notion of

what is typical or atypical should be transported by the dynamics, i.e. if a subset of phase space is typical today, the same set evolved until tomorrow will be typical tomorrow, with respect to tomorrow's notion of typicality. This is expressed by the continuity equation

$$\frac{\partial}{\partial t}\rho + \nabla(\rho \mathbf{v}) = 0. \tag{8}$$

Moreover, the procedure to judge which sets are typical or atypical should be the same for all times, it should be invariant under the dynamics. The easiest possibility is to have a density that simply does not change in time, i.e. such that $\partial_t \rho = 0$; this is called a stationary one. For Newtonian mechanics we can, in this case, simply assign equal weight to all states permitted by energy conservation, getting the microcanonical density.

In Bohmian mechanics the velocity field v guiding the configuration is determined by the wave function that in general depends on time. As a consequence, we cannot ask for a stationary density for the measure of typicality. From the guidance equation we know, however, that $\rho = |\Psi|^2$ satisfies the continuity equation. We used exactly this observation as one possible motivation for the guiding equation in the last section. Moreover, the prescription "take the modulus squared of Ψ" is a method to judge what is typical or atypical built out of Ψ in a way independent of time. This property is called equivariance and replaces the stationarity of the measure we had in the Newtonian case. Therefore the dynamics naturally indicates that the density $\rho = |\Psi|^2$ yields the correct measure of typicality. This makes plausible what we stated as Axiom 3 in the last section.

Now that we have a measure of typicality for Bohmian mechanics, let us go back to the question what typical initial configurations do. We will learn in one of the next sections how to derive the wave function for a subsystem from the wave function Ψ of the universe. Let us for the moment just assume that there is a way of finding the wave function of a subsystem given the wave function Ψ. It turns out that in Bohmian mechanics one can prove the following theorem ([1, ch. 11], [2, ch. 2]):

> For every wave function Ψ and for typical initial configurations Q of the universe, subsystems with given wave function φ have configurations distributed according to $|\varphi|^2$.

For our example with the double slit this means that the distribution of the spots where the particles hit the screen is approximately given by $|\varphi|^2$, where φ is the wave function of each single electron. The density $|\varphi|^2$ was already considered by Max Born in the early days of quantum mechanics, and this is why this distribution is called Born rule.

At this point one might be urged to ask the following: If the particles always have definite positions, is it not possible to prepare a system not only with a definite wave function but also, at the same time, with a definite configuration by very elaborate experimental setups? Is it conceivable, that, with very clever means, we can actually prepare a system with a configuration specified to a greater accuracy than the familiar uncertainty principle allows? The answer is *no*. The quantum mechanical interaction between a system and an apparatus is always an interaction of wave functions. From that one can show that if one prepares a system with prescribed wave function φ, no matter what kind of apparatus one uses, it is impossible to prepare the system with configuration specified more accurately than $|\varphi|^2$.

A further consequence of the $|\varphi|^2$-distribution is that all empirical predictions of Bohmian mechanics agree with the ones of standard quantum mechanics. In Bohmian mechanics as well as in standard quantum mechanics the outcomes of position measurements are $|\varphi|^2$-distributed, therefore both theories make the same predictions for position measurements. Now, consider for example a spin measurement. One sends an electron though an inhomogeneous magnetic field, produced by a Stern-Gerlach magnet, and records whether it gets deflected up or down. Thus, measuring the spin actually amounts to measuring the final position of the electron. This is true in general: every measurement in the end is a measurement of position; the outcome of the measurement must necessarily result in something moving in space, e g. a pointer showing the outcome on a scale, electrons moving in a circuit, and so on. Thus from the fact that Bohmian mechanics and standard quantum mechanics agree on the predictions for position measurements one gets that they agree on every measurement. One should keep in mind that the $|\varphi|^2$-distribution in Bohmian mechanics is not artificially tailored in order to reproduce the predictions of ordinary quantum mechanics, but it comes about naturally if one just asks: *what if there really are particles with definite positions at all times?*

3 Collapse, measurement problem and decoherence

In Bohmian Mechnics the description of the measurement process naturally follows from the formulation of the theory. A measurement is the interaction, for a finite amount of time, between two quantum mechanical systems, the system to be measured and the measuring apparatus, which will from now on be called system and apparatus, respectively. In Bohmian mechanics, both systems can be treated as ordinary quantum systems, and their interaction is described by the usual Schrödinger evolution. We will denote by x the configuration of the system, by y that of the apparatus, and by T the duration of the process.

The characterizing property of a measuring device is that at the end of the interaction it is in one of several macroscopically distinguishable configurations, like for instance pointers pointing in different directions. So there will be different apparatus wave functions corresponding to different outcomes of the measurement e.g. pointer positions, and these wave functions will have (practically) disjoint support in configuration space. Let us call the outcome wave functions Φ_i and Φ_0 the ready state of the apparatus, i.e. the state prior to the measurement.

The generic apparatus will behave in the following way. If the system starts out in some initial state φ_i at time $t = 0$, then interacts with the apparatus, the latter should evolve to a certain final state Φ_i depending on the initial state of the system to be measured. In other words, the time evolution during the measurement will be as follows

$$\varphi_i(x)\Phi_0(y) \longrightarrow \varphi_i^T(x)\Phi_i(y). \tag{9}$$

This is the justification why one can say the apparatus performs a measurement: it reacts to certain system states in a definite way, e.g. by a pointer pointing to definite values on a scale. Now, as we mentioned above, it is important that the Φ_i correspond to macroscopically distinguishable states. On the level of the wave function this means that Φ_i and Φ_j have

basically no overlap on configuration space for $i \neq j$, i.e. Φ_i is unequal to zero almost only in regions of configuration space where Φ_j is almost zero and vice versa.

Now if we take a sum of the φ_i as an initial state of the system, the time evolution yields for the state after the measurement

$$\sum_i c_i \varphi_i(x)\Phi_0(y) \longrightarrow \sum_i c_i \varphi_i^T(x)\Phi_i(y). \tag{10}$$

This is simply a consequence of the linearity of the Schrödinger equation. The result (10) is very puzzling. The state after the measurement is a superposition of macroscopically distinguishable states, so e.g. a superposition of a pointer pointing to the right and a pointer pointing to the left. This is not what we see in the experiment, we do not see a superposition of pointer positions, but we see just one of the summands. This is the famous measurement problem of standard quantum mechanics. Standard quantum mechanics attempts to solve this problem by postulating that in a measurement situation the Schrödinger dynamics breaks down and the wave function follows for a while a different dynamics that makes it collapse to just one of the summands. This solution, although very common, is not really viable. If one assumes that the collapse happens only in connection with measurement set-ups, then strictly speaking one can not anymore understand macroscopic bodies as made out of atoms: something more must be present that is responsible for the collapse, otherwise the atoms would always follow Schrödinger's dynamics, no matter how many they are. Moreover, it is unclear for which systems the collapse is supposed to happen and for which not.

Sometimes it is claimed that the measurement problem is not a true problem in standard quantum mechanics because of *decoherence*. Let's briefly consider this reasoning. For any pair of summands in (10), $\varphi_i^T\Phi_i$ and $\varphi_j^T\Phi_j$, with $j \neq i$, the overlap in configuration space is negligible, they practically have disjoint support on configuration space. Then, for the wave function of system and apparatus after the measurement we have

$$|\psi(T)|^2 = \Big|\sum_i c_i \varphi_i^T\Phi_i\Big|^2 = \sum_i |c_i\varphi_i^T\Phi_i|^2 + \sum_{i\neq j} c_i^*\varphi_i^{T\,*}\Phi_i^*\; c_j\varphi_j^T\Phi_j \approx \sum_i |c_i\varphi_i^T\Phi_i|^2. \tag{11}$$

This is the mathematically precise formulation of the statement "macroscopically distinguishable". The process that brings the initial state to a state with such a structure is a continuous one. The different summands in the wave function do at first overlap, but the overlap gets smaller and smaller as times goes on. This phenomenon is what is called decoherence. In a state like the final one no quantum interference can be preset because the summands have no overlap, therefore the righthand side of (11) is indistinguishable from a *statistical mixture*, in which the apparatus has a definite state among the Φ_is and we simply do not know which one. Nevertheless, this does not solve the problem, indeed it does not explain how in a single run of the experiment the system passes from a sum to just one of the summands. In doing that one is replacing a logical AND by a logical OR. In a statistical mixture we are ignorant about which state was actually prepared and therefore we say that the system is in one of several possible states, linked by a logical OR: this state, or that, or that or... In (11) the sum corresponds to a logical AND between the summands. Just because we can neglect some of the terms occurring in taking the square, which is meant by our $\approx$ in (11) above, does of course not change a logical AND to a logical OR!

Thus decoherence alone does not explain how a sum is transformed into just one of its summands. A step is still missing!

In Bohmian mechanics, however, the expression (10) is not a problem, because in addition to the wave function we still have the particle positions at all times.
Let us carefully analyze the situation from a Bohmian point of view. The apparatus is of course also made out of particles described by Bohmian mechanics. Let us say it is made out of M particles whereas the system is made out of N particles. Let $X \in \mathbb{R}^{3N}$ be the actual configuration of the system, $Y \in \mathbb{R}^{3M}$ that of the apparatus, and $Q(t) \coloneqq (X(t), Y(t))$.

The relevant wave function is the combined wave function of system and apparatus, e.g. initially $\sum_i c_i \varphi_i(x)\Phi_0(y)$, that is a function on $\mathbb{R}^{3(N+M)}$. As we know the wave function and the actual configuration at time t_0 uniquely determine the actual configuration $Q(t)$ for all other times t. The actual configuration $Q(t) \coloneqq (X(t), Y(t))$ is guided by the wave function and therefore moves just in that part of configuration space where the wave function is not equal to zero. This can easily be seen as at all times the probability to find $Q(t)$ in some region of configuration space is proportional to the modulus squared of the wave function. In other words there is zero probability to find $Q(t)$ where the wave function is zero. So after the end of the interaction of system and apparatus the configuration $Q(T)$ of system and apparatus will be in the support of just one of the summands on the right hand side of (10), and the way it moved there in configuration space is uniquely determined by its value at t_0. In Bohmian mechanics there is clearly just one pointer position at the end of the experiment, namely the one the actual configuration assumes.

Because the configuration $Q(t)$ just moves where the wave function is not zero, it will never move from one branch of the wave function to another, disjoint one. If after the experiment the different summands in (10) will never overlap again, which for a measurement situation is the case, the configuration $Q(t)$ will stay in its brach, e.g. $\varphi_i^T \Phi_i$ for all times $t \geq T$. The practical implication is that we can ignore the other branches for the future, and we get an effective collapse of the wave function. We will examine this in more detail in the next section.

4 Conditional/Effective wave function

A measurement situation is a special case of the more general situation of a system made of several subsystems. The treatment of subsystems is of crucial importance in physics, and we will now discuss it in a Bohmian framework.

In physics books and presentations, many discussions start with the words "consider a system composed of…": have you ever thought that this does not make sense if taken literally? It does not, because the system meant here is a certain set of physical objects interacting among themselves, but completely independent from any other physical object in the universe, i.e. these objects form a universe on their own. But we have only one universe! Of course nobody making the above statement means that. What is meant is *consider a subsystem of the actual universe, composed of…, and such that it is negligibly influenced by the state of the rest of the universe.*

If one speaks about classical mechanics, then for a given subsystem it is trivial to check if this is the case or not: it is enough that everything else is very far away. For example,

think of the magnetic force. Two magnets always exert a force on each other, and if you want to check with a compass where North is, you have to avoid to have other magnets around, but you do not have to worry about the influence of the magnetic field of Mars: it is simply too far away to matter.

In quantum mechanics, on the contrary, it is not that easy. For example, think of a single electron in outer space, very far away from everything else. Still, it could be entangled with another particle on a distant planet, and its state would then depend on what happens to the distant particle, no matter how far away it is. Clearly, for quantum mechanics spatial distance is not the main feature needed to have a subsystem independent of any object not belonging to it. For quantum systems a deeper analysis is required, and Bohmian mechanics neatly provides us with the needed tools for such an analysis.

Let us look at this analysis in some detail. We will use the word *universe* to designate the biggest system that can be considered relevant to the discussion. Let us denote by Q the Bohmian configuration of the universe, and let us split it as $Q = (X, Y)$, where X denotes the configuration of the subsystem of interest, that we will call x-system, and Y that of the rest of the universe, or y-system. The problem is to understand when the x-system has a Bohmian description on its own, that means that we have to build a wave function for it alone, satisfying a Schrödinger equation on its own, independently of the rest of the universe and guiding the X configuration with the usual Bohmian equation. What we know, is that the whole universe has a certain wave function $\Psi(x, y)$, and that the X configuration is guided by it according to the law[1]

$$\dot{X}(t) = \mathrm{Im}\frac{\Psi^* \nabla_x \Psi}{\Psi^* \Psi}(X(t), Y(t), t). \tag{12}$$

It is then obvious to define

$$\varphi(x, t) = \frac{\Psi(x, Y(t), t)}{\sqrt{\int |\Psi(x, Y(t), t)|^2 \, \mathrm{d}x}}, \tag{13}$$

indeed the function φ guides the X configuration:

$$\dot{X}(t) = \mathrm{Im}\frac{\varphi^* \nabla_x \varphi}{\varphi^* \varphi}(X(t), t). \tag{14}$$

The wave function $\varphi(x, t)$ is called *conditional wave function* of the x-system, given that the rest of the universe follows the Bohmian trajectory $Y(t)$. This wave function depends only on x and guides the x-configuration, but it is not yet what we were looking for. Indeed, the conditional wave function can always be defined even if the x- and y-systems are strongly dependent on each other. The question is now when the conditional wave function $\varphi(x, t)$ satisfies a Schrödinger equation on its own, independent of the y-system. In this case the conditional wave function is called the *effective wave function* for the x-subsystem.

To understand when this is the case, it is convenient to start from an obvious example: suppose that the wave function of the universe has a product structure, i.e.

$$\Psi(x, y) = \varphi(x)\Phi(y). \tag{15}$$

[1]For ease of notation in this chapter we set $\hbar$ and the mass of each particle to be equal to one, which amounts to a change of variables. The masses can easily be restored by writing the equations (12) and (14) component wise or by writing a mass matrix in front of the right hand sides.

Applying eq. (13) we find that the conditional wave function of the x-subsystem is always $\varphi(x)$, regardless of the Y coordinate. In this case, the conditional wave function $\varphi(x)$ is an effective wave function, indeed we can describe the x-system using just this wave function, and forget about the rest of the universe: the function $\varphi(x)$ satisfies a Schrödinger equation on its own. Of course, the product structure is a condition much too strong for the wave function of the universe. Even assuming that it holds at some time t_0, it is immediately destroyed by any interaction of the x- and y-systems, even by the weakest interaction that one can imagine.

Let us now look at a second example. Consider that the wave function of the universe is the sum of two functions

$$\Psi(x,y) = \Psi_1(x,y) + \Psi_2(x,y) \tag{16}$$

with disjoint y-support, i.e. if Y is such that $\Psi_1(x,Y)$ is different from zero, then $\Psi_2(x,Y)$ is equal to zero, and the other way around. Then, when for example $Y \in \operatorname{supp}\Psi_1(x,Y)$, i.e. when $\Psi_1(x,Y)$ is different from zero, the conditional wave function of the x-system is independent of $\Psi_2(x,Y)$. This splitting of the support of the wave function is not a rare case: the y-coordinate is the collection of the coordinates of any other particle in the universe

$$y = (y_1, y_2, \ldots) \tag{17}$$

and in order to have the splitting for the y-coordinate it is sufficient to have it for one single particle y_i. Moreover, for a y_i, such a splitting occurs as soon as one has two clearly different alternative states, like for example the positions of the pointer of a measuring apparatus.

The more general structure needed to have an effective wave function for the x-system comes from the combination of these two simple cases. Consider that, during a certain interval of time I_t, the wave function of the universe is of the form

$$\Psi(x,y) = \varphi(x)\Phi(y) + \Psi^{\perp}(x,y), \tag{18}$$

where $\Phi(y)$ and $\Psi^{\perp}(x,y)$ have disjoint y-supports. Then, the conditional wave function of the x-system is $\varphi(x)$ whenever Y is in the support of $\Phi(y)$, and $\varphi(x)$ satisfies a Schrödinger equation on its own. This means that, if we know that during the time interval I_t the configuration Y of the rest of the universe is in the support of $\Phi(y)$, then we can study the x-system as an independent system, characterized by the wave function $\varphi(x)$, that evolves according to the Schrödinger equation. Of course, in this case $X \in \operatorname{supp}\varphi$.

This solves our problem. Note that to derive from the description of the whole universe an evolution for the x-system as independent of the y-system, the concept of the conditional wave function is needed, and this is a concept that makes explicit use of the Bohmian configuration.

We can use the considerations described above also for the problem of the collapse of the wave function. Indeed, in a measurement situation the wave function after the measurement is of the form (10), i.e.

$$\Psi(x,y,T) = \sum_i c_i \varphi_i^T(x)\,\Phi_i(y), \tag{19}$$

that is a special case of (18). Therefore, after the experiment the x-sytem can be described independently of the y-system, because it admits an effective wave function. The effective wave function for the x-system is given by eq. (13), and is $\varphi_i^T(x)$ whenever the actual pointer position $Y(T)$ lies in the support of $\Phi_i(y)$. Hence, the usual collapse rule of quantum mechanics is recovered, with the effective wave function for the x-system corresponding to the collapsed wave function. Which φ_i one gets in the end is determined by the initial positions X and Y.

This procedure for getting the wave function of a subsystem from the wave function of the universe is what one uses to derive the Born rule for subsystems from the measure of typicality of the universe [1, ch. 11].

5 Operators

According to Bohmian mechanics every object is composed of physical point particles moving around and interacting with one another. No special role is reserved to conscious beings nor to measurement processes. An observer is, from the Bohmian point of view, just a bunch of particles, and an observation is just an interaction of this bunch of particles with some other particles not belonging to that bunch. During such a process, that has to be described by the Schrödinger evolution as any other physical process, nothing extraordinary can happen.

How comes then, that *observables* and the related *self-adjoint operators* are so important in quantum mechanics? It is so because quantum mechanics is a theory *about measurements*, so these things are fundamental by definition. On the contrary, Bohmian mechanics is a theory *about particles that move*, and the usual formalism describing the measurements must be derived as a special case. Self-adjoint operators are in this view just mathematically convenient ways to summarize all the relevant information concerning the process: the measurement formalism is nothing else than a very sophisticated bookkeeping technique.

Let's have a look at the derivation of the measurement formalism according to Bohmian mechanics. An experiment is a physical process in which a system of interest, characterized by the configuration x, comes into interaction with another physical system, called apparatus and described by the configuration y. This interaction has to fulfill some requirements in order to allow us to understand it as a measurement. We say that the y-system is measuring the x-system if there exists a set of y-states $\{\Phi_i\}_{i=0,1,\ldots}$ with the following properties:

- they have disjoint supports, i.e.

$$\forall\, y \in \operatorname{supp}\Phi_i \quad \Rightarrow \quad \Phi_j(y) = 0, \;\; j \neq i; \tag{20}$$

- they are normalized, i.e. $\langle \Phi_i | \Phi_i \rangle = 1$;

- starting from the ready state $\Phi_0(y)$ and from a certain state $\varphi_i(x)$, the Schrödinger evolution brings the composite system to the state $\varphi_i^T(x)\Phi_i(y)$, i.e. the states before and after the interaction are related by

$$\varphi_i(x)\,\Phi_0(y) \longrightarrow \varphi_i^T(x)\,\Phi_i(y); \tag{21}$$

By linearity, we have that starting from the x-state $\varphi(x) = \sum_i c_i \varphi_i(x)$, the evolution leads to

$$\varphi(x)\,\Phi_0(y) = \sum_i c_i \varphi_i(x)\,\Phi_0(y) \longrightarrow \sum_i c_i \varphi_i^T(x)\,\Phi_i(y). \tag{22}$$

The probability that after the interaction the Bohmian configuration of the apparatus lies in the support of Φ_i is

$$\mathbb{P}_i = \int \mathrm{d}x \int_{\mathrm{supp}\,\Phi_i} \mathrm{d}y \left| \sum_i c_i \varphi_i^T(x)\Phi_i(y) \right|^2 = \int \mathrm{d}x \int_{\mathrm{supp}\,\Phi_i} \mathrm{d}y \left| c_i \varphi_i^T(x)\Phi_i(y) \right|^2 = |c_i|^2. \tag{23}$$

Note that the x-system admits an effective wave function both before and after the interaction process, and these wave functions are $\varphi(x) = \sum_i c_i \varphi_i(x)$ and $\varphi_i^T(x)$, respectively. The probability that the final effective wave function of the x-system is φ_i^T is equal to $\mathbb{P}_i$.

Given such a measurement, the x-states φ_i that are associated to it must satisfy some properties. At first, they must be orthogonal. Indeed, consider the norm of the initial state

$$\|\varphi\Phi_0\|^2 = \int \mathrm{d}x \int \mathrm{d}y \left| \sum_i c_i \varphi_i(x)\Phi_0(y) \right|^2 = \int \mathrm{d}x \left| \sum_i c_i \varphi_i(x) \right|^2 \tag{24}$$

$$= \sum_i |c_i|^2 + \sum_{i \neq j} c_i^*\, c_j \int \mathrm{d}x \varphi_i^*(x)\, \varphi_j(x) \tag{25}$$

and that of the final one

$$\left\| \sum_i c_i \varphi_i^T \Phi_i \right\|^2 = \int \mathrm{d}x \int \mathrm{d}y \left| \sum_i c_i \varphi_i^T(x)\Phi_i(y) \right|^2 = \sum_i |c_i|^2. \tag{26}$$

By unitarity of the evolution these two norms must be equal, therefore

$$\sum_{i \neq j} c_i^*\, c_j \int \mathrm{d}x\, \varphi_i^*(x)\, \varphi_j(x) = 0. \tag{27}$$

The coefficients c_i are arbitrary, hence it must be

$$\int \mathrm{d}x\, \varphi_i^*(x)\, \varphi_j(x) = \langle \varphi_i | \varphi_j \rangle = 0. \tag{28}$$

One more property derives from the fact that one can start with an arbitrary x-state and the apparatus will always show a result after the measurement. It must be possible to calculate the probability of each final position of the pointer of the apparatus. Therefore, it must be possible to decompose any x-state in the form $\sum_i c_i \varphi_i$, i.e. the states $\{\varphi_i\}_i$ must constitute a basis. To be more precise, we should consider that several initial states $\varphi_{i,\alpha}$ can trigger the same answer i of the apparatus. The set of states $\{\varphi_{i,\alpha}\}_{i,\alpha}$ are characteristic of the experiment, and the experiment is not able to distinguish the states with the same i and different αs. The mathematical object that encodes all the information regarding the set $\{\varphi_{i,\alpha}\}_i$ relative to the outcome i is the projector

$$P_i = \sum_\alpha |\varphi_{i,\alpha}\rangle \langle \varphi_{i,\alpha}| = P_i^2 = P_i^\dagger. \tag{29}$$

The properties of the states $\varphi_{i,\alpha}$ translate to the following properties of the projectors P_i:

$$P_i P_j = 0, \quad i \neq j, \tag{30}$$

$$\sum_i P_i = 1. \tag{31}$$

Moreover, given the initial x-state φ, we can calculate the probability of the outcome i by

$$\mathbb{P}_i = |c_i|^2 = \langle \varphi | P_i | \varphi \rangle . \tag{32}$$

One last feature of a typical experiment remains to be discussed. Until now, we have identified the different answers of the apparatus by an index i. This is not usual for an experiment. Normally, different answers of an experiment correspond to different *values*, i.e. to different numbers printed on the scale of the pointer, or the like. Let's denote the value corresponding to the answer i by λ_i. It is then useful to build the operator

$$\hat{A} = \sum_i \lambda_i P_i \tag{33}$$

with which we can calculate all the relevant statistical predictions for the experiment. For instance, the mean of the measured values for the initial state φ is

$$\mathbb{E}_\varphi(\lambda) = \sum_i \lambda_i \mathbb{P}_i = \sum_i \lambda_i \langle \varphi | P_i | \varphi \rangle = \langle \varphi | \sum_i \lambda_i P_i | \varphi \rangle = \langle \varphi | \hat{A} | \varphi \rangle . \tag{34}$$

Moreover, their variance is

$$\begin{aligned} \mathbb{E}_\varphi(\lambda^2) - \mathbb{E}^2_\varphi(\lambda) = \\ = \sum_i \lambda_i^2 \langle \varphi | P_i | \varphi \rangle - \langle \varphi | \hat{A} | \varphi \rangle^2 = \sum_i \lambda_i^2 \langle \varphi | P_i^2 | \varphi \rangle - \langle \varphi | \hat{A} | \varphi \rangle^2 = \\ = \sum_{i,j} \lambda_i \lambda_j \langle \varphi | P_i P_j | \varphi \rangle - \langle \varphi | \hat{A} | \varphi \rangle^2 = \langle \varphi | \hat{A}^2 | \varphi \rangle - \langle \varphi | \hat{A} | \varphi \rangle^2 . \end{aligned} \tag{35}$$

It is easy to verify that the operator $\hat{A}$ is selfadjoint by construction, i.e. $\hat{A} = \hat{A}^\dagger$, indeed $\lambda_i \in \mathbb{R}$ and $P_i^\dagger = P_i$.

Summarizing, given an experiment, we can compute all the relevant probabilities and statistics using the projectors P_i and the values λ_i, both encoded in the selfadjoint operator $\hat{A}$, i.e. restricting Bohmian mechanics to measurement processes one recovers the usual formalism of ordinary quantum mechanics. Because the description of any measurement process is the same in Bohmian mechanics and in quantum mechanics, there is then no experiment for which the two theories can possibly give different predictions.

This derivation shows how the measurement formalism follows from the Schrödinger evolution and proves that the promotion of measurements from special cases of the Schrödinger evolution to fundamental processes described by *ad hoc axioms* is not a necessity, but a *deliberate theoretical choice*.

It must be noted that this derivation is not exclusively Bohmian, indeed it has been developed by Günter Ludwig in the framework of ordinary quantum mechanics. Nevertheless, in this case one still needs the collapse postulate to get a definite answer for each run of the experiment, while the Bohmian version uses the concept of the effective wave function at that point. Therefore, the derivation of the measurement formalism gets very natural and clear in the context of Bohmian mechanics, where the measurement problem is solved without reference to additional axioms.

References

[1] Detlef Dürr, Stefan Teufel, *Bohmian Mechanics*. Springer, 2010

[2] Detlef Dürr, Sheldon Goldstein, Nino Zanghi *Quantum Physics without Quantum Philosophy*. Springer, 2013