

The Hilbert-space operator formalism within dynamical reduction models

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2007 J. Phys. A: Math. Theor. 40 13755

(<http://iopscience.iop.org/1751-8121/40/45/015>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.146

The article was downloaded on 03/06/2010 at 06:25

Please note that [terms and conditions apply](#).

The Hilbert-space operator formalism within dynamical reduction models

Angelo Bassi^{1,2,3}, GianCarlo Ghirardi^{1,3,4} and Davide G M Salvetti^{1,3}

¹ Dipartimento di Fisica Teorica, Università di Trieste, Strada Costiera 11, 34014 Trieste, Italy

² Mathematisches Institut der LMU, Theresienstr 39, 80333 München, Germany

³ Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, Strada Costiera 11, 34014 Trieste, Italy

⁴ The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, 34014 Trieste, Italy

E-mail: bassi@ts.infn.it, ghirardi@ts.infn.it and salvetti@ts.infn.it

Received 24 July 2007, in final form 26 September 2007

Published 23 October 2007

Online at stacks.iop.org/JPhysA/40/13755

Abstract

Unlike standard quantum mechanics, dynamical reduction models assign no particular *a priori* status to ‘measurement processes’, ‘apparata’ and ‘observables’, nor self-adjoint operators and positive-operator-valued measures enter the postulates defining these models. In this paper, we show why and how the Hilbert-space operator formalism, which standard quantum mechanics *postulates*, can be *derived* from the fundamental evolution equation of dynamical reduction models. Far from having any special ontological meaning, we show that within the dynamical reduction context the operator formalism is just a compact and convenient way to express the statistical properties of the outcomes of experiments.

PACS numbers: 03.65.–w, 03.65.Ca

1. Introduction

Dynamical reduction models (DRMs) provide, at the non-relativistic level at least, a coherent unified description of both microscopic quantum and macroscopic classical phenomena, and in particular give a consistent solution to the macro-objectification problem of quantum mechanics [1, 2]. They are defined by the following set of axioms:

AXIOM A: STATES. A Hilbert space \mathcal{H} is associated with any physical system and the state of the system is represented by a (normalized) vector $|\psi\rangle$ in \mathcal{H} .

AXIOM B: EVOLUTION (GRW MODEL). At random times, distributed like a Poissonian process with mean frequency λ , each particle of a system of N particles is subjected to a spontaneous localization process of the form

$$|\psi_t\rangle \longrightarrow \frac{L_n(\mathbf{x})|\psi_t\rangle}{\|L_n(\mathbf{x})|\psi_t\rangle\|}, \quad L_n(\mathbf{x}) = \sqrt[4]{\left(\frac{\alpha}{\pi}\right)^3} \exp\left[-\frac{\alpha}{2}(\mathbf{q}_n - \mathbf{x})^2\right], \quad (1)$$

where \mathbf{q}_n is the position operator associated with the n th particle and $|\psi_t\rangle$ is the wavefunction of the global system immediately prior to the collapse; the collapse processes for different particles are independent. Between two collapses, the wavefunction evolves according to the standard Schrödinger equation. The probability density for a collapse for the n th particle to occur around the point \mathbf{x} of space is

$$p(\mathbf{x}) = \|L_n(\mathbf{x})|\psi_t\rangle\|^2. \quad (2)$$

The standard numerical values [1] for the two parameters λ and α are $\lambda \simeq 10^{-16} \text{ sec}^{-1}$ and $\alpha \simeq 10^{10} \text{ cm}^{-2}$, respectively. A continuous formulation in terms of stochastic differential equations is also commonly used [2–6].

AXIOM C: ONTOLOGY. Let $\psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \equiv \langle \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N | \psi \rangle$ the wavefunction for a system of N particles (which for simplicity we take to be scalar) in configuration space. Then

$$m(\mathbf{x}, t) \equiv \sum_{n=1}^N m_n \int d^3x_1 \dots d^3x_N \delta^{(3)}(\mathbf{x}_n - \mathbf{x}) |\psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 \quad (3)$$

is assumed to describe the *density of mass*⁹ distribution of the system¹⁰ in three-dimensional space, as a function of time [2, 7].

As we see, within DRMs concepts like *measurement*, *apparata*, *observables* play no particular privileged role; like in classical mechanics, they merely refer to particular physical situations where a macroscopic physical system, which we call apparatus, interacts in a specific way with another physical system, e.g. a microscopic quantum system. However, such a macroscopic system, the apparatus, is ultimately described in terms of its fundamental constituents, and its interaction with other systems is ultimately described in terms of the fundamental interactions of nature. The basic idea should be clear: all physical processes are governed by the universal dynamics embodied in the precise axioms we have just presented. What is usually denoted as a ‘measurement’ of an observable by an apparatus is simply a precise physical process which is purposely caused by a human being under controlled conditions. In what follows we will use, to denote such a situation, the term ‘experiment’ to conform to the clear-cut position of JS Bell, summarized in the following lucid sentence [8]:

I am convinced that the word ‘measurement’ has now been so abused that the field would be significantly advanced by banning its use altogether, in favor, for example, of the word ‘experiment’.

Given this, the following question arises: why are experiments on microscopic quantum systems so efficiently described in terms of average values of self-adjoint operators, and more generally in terms of POVMs? Why is the Hilbert-space formalism so powerful in accounting for the observable properties of microscopic systems? The aim of this paper is to provide an answer to these questions, from the point of view of DRMs. We will show that, within DRMs, one can *derive* a well-defined role for self-adjoint operators and POVMs as useful (but not compelling) mathematical tools which allow to compactly express the statistical properties of microscopic systems subject to experiments. Accordingly, within DRMs, recovering the formal aspects of standard quantum mechanics is simply a matter of practical convenience

⁹ In the subsequent sections, for simplicity’s sake, we will not make reference to the mass density function anymore, and we will only keep track of the evolution of the wavefunction; however it should be clear that, in order to be fully rigorous, all statements about the properties of physical systems should be phrased in terms of their mass-density distribution.

¹⁰ The mass density in principle refers to the whole universe; however, as standard practice in Physics, one can make the approximation of considering only a part of it, which is sufficiently well isolated, and of ignoring the state of the rest of the universe.

and mathematical elegance. Stated in different terms, while experiments on quantum systems are nothing more than a particular type of interaction between a macroscopic (thus always well localized) system and a microscopic one, such that different macroscopic configurations of the macro-object correspond to different outcomes of the experiment, it is nevertheless simpler to refer to the statistical properties of the outcomes in terms of self-adjoint operators averaged over the initial state of the microscopic quantum system. Just a matter of practical convenience, nothing more.

The paper is organized as follows. In section 2, we show with a simple example how DRMs recover the operator formalism for the description of the statistical properties of the outcomes of quantum experiments. Section 3 is the core of the paper: we will prove in full generality how DRMs allow to associate a POVM with an experiment, i.e. how the statistical properties of the experiment can be represented as the average values of the effects of the POVM over the state before the ‘measurement’ of the microscopic quantum system. In section 4, we will show that when an experiment is reproducible, the POVM reduces to a PVM and the experiment can be represented by a unique self-adjoint operator, as it is typically assumed in standard text books on quantum mechanics. In section 5, we reconsider the so-called tail problem and show that it does not affect the dynamical reduction program. In section 6, we show with an explicit example that the Hilbert-space formalism can be used also to describe certain *classical* experiments. Section 7 contains some concluding remarks.

This paper takes most inspiration from [9], where the emergence of the operator formalism within the framework of Bohmian mechanics has been thoroughly analyzed.

2. Emergence of the operator formalism: a simple example

We begin our analysis by discussing a simple physical situation which should make clear how, and in which sense, the operator formalism of standard quantum mechanics ‘emerges’ in a natural way from the physical and mathematical properties of DRMs, even if it does not appear explicitly in the axioms defining these models.

2.1. A measurement situation

For simplicity’s sake, let us consider a spin-1 particle (its Hilbert space being $\mathcal{H} = \mathbb{C}^3$) and let us assume that it has been initially prepared in the normalized state

$$|\psi\rangle = a |S_z = +1\rangle + b |S_z = 0\rangle + c |S_z = -1\rangle, \quad (4)$$

where $|S_z = +1\rangle$, $|S_z = 0\rangle$ and $|S_z = -1\rangle$ are the three eigenstates¹¹ of S_z , and a , b and c are three complex parameters which can be varied according to the preparation procedure. Let us now perform a Stern–Gerlach type of experiment which measures the spin of the particle along the x direction. According to the rules of DRMs¹² (no other assumption is used, other than axioms A–C, in the simplified form suitable for this example), we can state that

- Throughout the entire process, the measuring device has always a well-defined macroscopic configuration. In the particular example we intend to discuss, there are three possible outcomes for the experiment, i.e. three possible final macroscopically different configurations of the measuring device, which we call, e.g. ‘outcome +1’, ‘outcome 0’ and ‘outcome –1’.

¹¹ There is of course nothing special in the choice of the eigenstates of S_z : we could have very well chosen another basis of \mathbb{C}^3 .

¹² See [10] for a detail analysis of this topic.

- The outcome is random and the probability distribution depends only on the initial state of the micro-system, according to a law which, with great accuracy, is equal to

$$\mathbb{P}^\psi(+1) = \frac{1}{4}|a + \sqrt{2}b + c|^2, \quad (5)$$

$$\mathbb{P}^\psi(0) = \frac{1}{2}|a - c|^2, \quad (6)$$

$$\mathbb{P}^\psi(-1) = \frac{1}{4}|a - \sqrt{2}b + c|^2. \quad (7)$$

It is worthwhile stressing that the above probabilities, which ultimately coincide with the standard quantum probabilities associated with the outcomes of a measurement of the spin along the x direction, are *not* postulated, but derive from the dynamics of DRMs, when applied to the specific measurement-like situation which has been chosen. Given these premises, we now show how one can associate an operator, namely the spin operator S_x , to this specific experiment.

2.2. Observables as operators

Let $\mathcal{O} = \{-1, 0, +1\}$ be the set of the possible outcomes of the experiment, and let \mathcal{G} be the power set of \mathcal{O} ($\mathcal{G} = \mathcal{P}(\mathcal{O})$), which is an algebra on \mathcal{O} . We can then define a probability measure on \mathcal{G} in the following obvious way:

$$\mathbb{P}^\psi(V) \equiv \sum_{v \in V} \mathbb{P}^\psi(v) \quad \forall V \in \mathcal{G}, \quad (8)$$

with $\mathbb{P}^\psi(v)$ given by equations (5)–(7). The above definition assumes that $|\psi\rangle$ is a fixed unit vector, while V can vary. Let us now reverse the roles of $|\psi\rangle$ and V : we fix an element $V \in \mathcal{G}$ and consider $\mathbb{P}^\psi(V)$ as a function of $|\psi\rangle$. It is easy to recognize that the probability distribution depends *quadratically* on the initial state $|\psi\rangle$ (i.e. on the coefficients a, b, c) of the micro-system: as a matter of fact, this is the very reason why one can attach an operator to the experiment. Such a property is even more evident if we introduce the following normalized and orthogonal states:

$$|v = +1\rangle = \frac{1}{2}[|S_z = +1\rangle + \sqrt{2}|S_z = 0\rangle + |S_z = -1\rangle], \quad (9)$$

$$|v = 0\rangle = \frac{1}{\sqrt{2}}[|S_z = +1\rangle - |S_z = -1\rangle], \quad (10)$$

$$|v = -1\rangle = \frac{1}{2}[|S_z = +1\rangle - \sqrt{2}|S_z = 0\rangle + |S_z = -1\rangle], \quad (11)$$

resorting to which we can express (8) in the following compact way:

$$\mathbb{P}^\psi(V) = \sum_{v \in V} |\langle v | \psi \rangle|^2. \quad (12)$$

If we allow $|\psi\rangle$ to run over the entire Hilbert space \mathcal{H} , not just over the unit sphere¹³ S^1 , then $\mathbb{P}^\psi(V)$ becomes a *quadratic* function from \mathcal{H} to \mathbb{R} , being the diagonal part of the sesquilinear form

$$\mathbb{P}^{\psi, \phi}(V) = \sum_{v \in V} \langle \psi | v \rangle \langle v | \phi \rangle \quad \psi, \phi \in \mathcal{H}, \quad V \text{ fixed}. \quad (13)$$

Given the bounded sesquilinear form $\mathbb{P}^{\psi, \phi}(V)$, the Riesz representation theorem allows us to express it in terms of a bounded linear operator O_V , in the following way:

$$\mathbb{P}^{\psi, \phi}(V) = \langle \psi | O_V | \phi \rangle, \quad (\text{obviously, in this simple case: } O_V \equiv \sum_{v \in V} |v\rangle \langle v|). \quad (14)$$

¹³ Of course, $\mathbb{P}^\psi(V)$ can be consistently interpreted as a probability only when $|\psi\rangle \in S^1$.

In this particular example we know also that the operator O_V is self-adjoint.

Going back to the original \mathbb{P}^ψ , we can then write

$$\mathbb{P}^\psi(V) = \langle \psi | O_V | \psi \rangle, \quad \forall V \in \mathcal{G}, \quad \forall |\psi\rangle \in \mathcal{H}, \quad (15)$$

and the set of operators $\{O_V\}_{V \in \mathcal{G}}$ forms a POVM, as one can easily prove. This is the result we wanted to arrive at: thanks to the particular dependence of the probability measure \mathbb{P}^ψ on ψ and to the Riesz representation theorem we have been able to express the statistical properties of the outcomes of the experiment in terms of average values of the effects of a POVM over the initial state of the microscopic system.

For our particular experiment, the set $\{O_V\}_{V \in \mathcal{G}}$ is more than a POVM: each of the eight operators O_V is in fact a projection operator, and therefore $\{O_V\}_{V \in \mathcal{G}}$ turns out to be a projection valued measure (PVM), which is the one associated with the spectral resolution of a self-adjoint operator

$$O = \sum_{v \in \mathcal{O}} v |v\rangle \langle v| \quad \Rightarrow \quad O = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \quad (16)$$

For this reason, we can rightly associate the operator O with our experiment, in the precise sense given here. By inspecting the components of the three vectors $|v\rangle$ given by equations (9)–(11) and the explicit form of the operator O displayed in equation (16), we can now finally recognize that O is indeed the component S_x of the spin along the x direction for a spin 1 particle, written in the basis $|S_z = +1\rangle, |S_z = 0\rangle, |S_z = -1\rangle$ of the eigenstates of S_z .

3. The operator formalism within dynamical reduction models

In this section, we prove in full generality what we have shown with the previous example, namely how the Hilbert-space operator formalism derives from DRMs as a tool to express the statistical properties of the outcomes of the experiments.

3.1. The link between experimental outcomes and macroscopic positions

When describing physical experiments, one usually identifies experimental outcomes with *real numbers*; however, what one actually sees (his empirical experience) as the outcome of a measurement is not a real number, but a specific configuration of a macroscopic object, namely a pointer being located in a well-defined region of space¹⁴. It is then necessary to set a link between real numbers interpreted as the outcome of an experiment and the position of the pointer of the measuring apparatus. We now wish to make precise the conditions which the state of the pointer satisfies whenever we experience a *perception* which we interpret as: ‘the outcome of the experiment is v ’, v being a real number.

3.1.1. The position of a macroscopic object. For simplicity’s sake, we will analyze the pointer by considering only the spatial degrees of freedom of its center of mass, ignoring its spatial extension and orientation, as well as all its microscopic degrees of freedom. According to the ontology of DRMs, a pointer is a distribution of mass which—being macroscopic—is localized within a small region of space coinciding with the spatial extension of the pointer

¹⁴ One might also consider different situations, like, e.g. the firing of a counter; however, what matters is that in all measurement situations the final states of the apparatus differs for a macroscopic mass density distribution. The reader will have no difficulty in transcribing the following analysis in such a way that it applies also to the just-mentioned situations in which there is not any pointer.

itself, but nevertheless it has small ‘tails’ spreading out to infinity; because of this, it is not possible to adopt the point of view according to which the pointer (and, in general, any macroscopic object) is located in some given region of space if and only if its mass density is entirely contained within that region; we will come back to this point in section 5. We then say that a macro-object A is located within a given region of space when almost all its mass density distribution is contained within that region.

As a consequence of the dynamical laws of DRMs, in all standard physical situations¹⁵ such as measurement processes, the region around which the wavefunction of the center of mass of a macro-object is localized is extremely small; accordingly, one can also take as the position of a macro-object the center of that region, which can be mathematically expressed by the formula

$$\mathbf{q}_t = \langle \psi_t | \mathbf{q} | \psi_t \rangle \in \mathbb{R}^3, \quad (17)$$

where \mathbf{q} is the position operator of the center of mass of A . We stress that \mathbf{q}_t should *not* be interpreted as the quantum average of the position operator \mathbf{q} , as done in standard quantum mechanics, but as the coordinates of the point in space around which, at time t , the mass density is appreciably different from zero. In the following, we will use (17) to denote the position of a macro-object.

Coming back to our pointer, let us assume that the pointer moves only along the graduated scale, so that we can treat it as a one-dimensional system: we will call $q_t \in \mathbb{R}$ its position along the scale. Since its dynamical evolution is intrinsically stochastic, q_t is a random variable $q_t : \Omega \rightarrow \mathbb{R}$ for any time t , where Ω is the sample space of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which the stochastic dynamics is defined; as a matter of fact, at the end of a measurement process we know that the pointer is located somewhere along the scale, but we do not know exactly where. Accordingly, the physically relevant information is embodied in the *probability distribution* of q_t :

$$\mathbb{P}_{q_t}[B] \equiv \mathbb{P}[q_t^{-1}(B)] \quad B \in \mathcal{B}(\mathbb{R}), \quad (18)$$

which gives the probability that q_t lies within the (measurable) subset B of the Borel σ -algebra $\mathcal{B}(\mathbb{R})$ of \mathbb{R} .

3.1.2. The calibration of an experiment. As we mentioned at the beginning of this section, we usually do not speak of the outcome of an experiment in terms of ‘the pointer being in a particular position in space’, but rather in terms of ‘the pointer signaling a particular outcome’, typically a real number; this sort of association between the spatial position of the pointer and the (numerical) outcome of the experiment is of course entirely a matter of convention, e.g., in a Stern–Gerlach type of experiment we usually associate the upper part of the screen with the outcome $+1$ and the lower part with the outcome -1 , but we do not directly observe the number $+1$ or -1 : what we observe are spots either in the upper or in the lower part of the plate or, if we want to keep referring to a movable pointer, what we observe is the pointer sitting in two macroscopically different positions along the scale, corresponding to the two possible outcomes. The set of outcomes $\{-1, +1\}$ is conventional, and we are free to use whatever set suits best the interpretation of the experiment. In accordance with [9], in what follows we will refer to the association between directly observed positions of the pointer and conventionally decided outcomes as to the *calibration* of the experiment. In all generality, we define the calibration of an experiment as follows.

¹⁵ Here, we do not take into account those pathological situations (whose probability of occurring is vanishingly small) in which a macroscopic object can be, for a very short time, in a superpositions of macroscopically different states.

Definition 1 (Calibration function). *Let \mathcal{O} be the set of possible outcomes of an experiment. The function*

$$f : \begin{cases} \mathbb{R} \rightarrow \mathcal{O} \\ q \rightarrow f(q) \end{cases} \quad (19)$$

from the space \mathbb{R} of the positions of (the center of mass of) the pointer to the space \mathcal{O} of the possible outcomes of the experiment, is called the calibration function of the experiment. For mathematical convenience, we assume f to be $\mathcal{B}(\mathbb{R})/\mathcal{G}$ -measurable, where \mathcal{G} is a chosen σ -algebra on \mathcal{O} .

By means of the calibration function, we can replace the probability distribution \mathbb{P}_{q_t} of equation (18) which gives the probability that the pointer lies within $B \in \mathcal{B}\mathbb{R}$, with the probability distribution

$$\mathbb{P}_{f(q_t)}[V] \equiv \mathbb{P}[q_t^{-1}(f^{-1}(V))] \quad V \in \mathcal{G}, \quad (20)$$

which represents the probability that the outcome of the experiment belongs to the measurable subset V of \mathcal{O} , the set of all possible outcomes.

3.2. The link between experimental outcomes and microscopic states

In this section we establish the main result of the paper, namely the link between the probabilities associated with the possible outcomes of an experiment and the pre-measurement states of the microscopic system whose properties are measured, showing how such probabilities can be represented by the mean value of the effects of a POVM over the state of the microscopic system right before the measurement.

3.2.1. The characteristic traits of measurement-like situations. We start our analysis by summarizing the most relevant properties of measurement processes, as they are described within DRMs; for a detailed analysis of these properties, we refer the reader to [10].

Let x_n be the n th position of the pointer along the graduate scale corresponding to the outcome $v_n \in \mathcal{O}$, where \mathcal{O} is the set of all possible outcomes, to which a σ -algebra \mathcal{G} is associated. Let t_F be the time at which the experiment ends. According to the dynamical laws typical of all DRMs, we can state that

Property 1. Throughout the whole measurement process, and in particular at its end, the center of mass of the pointer is always very well localized in space; making reference to the model analyzed in [10], the spread in position of the center-of-mass wavefunction of a pointer having the mass of 1 g is about $\sigma_q \simeq 10^{-14}$ m.

Property 2. Let us consider an interval δ_n , centered around the position x_n , whose extension d we take of the order of 10^{-5} cm; then, the probability η that, at the end of the measurement process, q_{t_F} does not lie inside *any* of the sub-intervals δ_n is very small, e.g. $\eta \simeq 10^{-17}$. This means that, with probability extremely close to 1, the pointer ends up in one of the positions along the graduate scale corresponding to one of the possible outcomes.

Since the wavefunction of the center of mass of the pointer has a spatial extension (though very small), the event $q_{t_F} \in \delta_n$ does not imply that most of the center-of-mass wavefunction lies almost entirely within δ_n : this happens, e.g. when q_{t_F} lies at the border of δ_n . In order to take this possibility properly into account, let us consider a new set of intervals Δ_n containing the intervals δ_n , each of which is centered around x_n , whose extension is equal to $d + \ell$ where ℓ is large compared to the typical spread σ_q of the center of mass of the pointer, e.g. $\ell \simeq 10^8 \times \sigma_q \simeq 10^{-6}$ m (see figure 1). We can then say that when q_{t_F} lies within δ_n ,

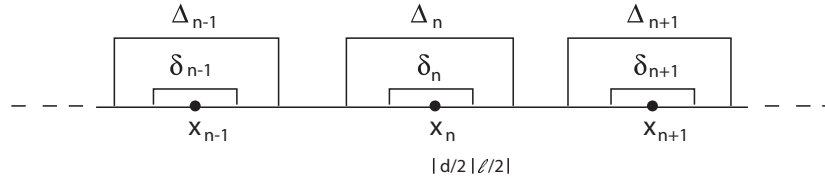


Figure 1. The picture shows the points x_n associated with the different outcomes of an experiment, together with the intervals δ_n and Δ_n used for the proof of theorem 1.

practically all the center-of-mass wavefunction lies within Δ_n , while the probability for q_{t_F} lying in $\Delta_n - \delta_n$, i.e. outside δ_n , is, according to property 2, vanishingly small.

Up to now we have spoken only of the position of the pointer, but as we have said it is custom to refer the probabilities to the (numerical) outcomes of the measurement. To this end we introduce the following calibration function:

$$f(q_{t_F}) = v_n \in \mathcal{O} \quad \text{iff: } q_{t_F} \in \Delta_n, \tag{21}$$

which simply means that the outcome of the experiment is v_n whenever the pointer, at the end of the measurement process, lies around the position x_n of the graduate scale. In the above definition, we have chosen Δ_n as the relevant intervals in order to make the proof of the following theorem simpler: however, on a macroscopic scale the two sets of intervals ($\{\delta_n\}$ and $\{\Delta_n\}$) are practically identical. Note that, since we want different outcomes to correspond to different macroscopic configurations of the pointer, the intervals Δ_n should not overlap, which means that the distance between two consecutive points x_n and x_{n+1} along the graduate scale should be bigger than $d + \ell \simeq 10^{-6}$ m; this is of course a perfectly reasonable assumption.

We are now in the position to prove the following theorem, which is the cornerstone for the subsequent derivation of the Hilbert-space formalism from the framework of DRMs.

Theorem 1. *Let us consider an experiment globally described by the state vector $|\psi_t\rangle$, let q_t be the position of the (center of mass of the) pointer at time t right before the measurement begins, and let t_F be the time at which the experiment ends. Then*

$$\forall V \in \mathcal{G} : \quad |\mathbb{P}_{f(q_{t_F})}[V] - \mathbb{Q}_{t_F}[V]| \leq 2 \left[\left(\frac{\sigma_q}{\ell} \right)^2 + \eta \right], \tag{22}$$

where the probability measure \mathbb{Q}_t is defined as follows:

$$\mathbb{Q}_t[V] \equiv \mathbb{E}_{\mathbb{P}}[\langle \psi_t | P_{\Delta_V} | \psi_t \rangle], \tag{23}$$

and P_{Δ_V} is the projection operator on the (measurable) subset $\Delta_V \equiv f^{-1}(V)$, which is the subset of \mathbb{R} of all positions of the pointer corresponding to a precise outcome v among the possible outcomes belonging to V .

The above theorem is very powerful: according to the very general structures of DRMs, the probability distribution $\mathbb{P}_{f(q_{t_F})}$ of the outcomes of an experiment can, for all practical purposes, be replaced by the probability measure \mathbb{Q}_{t_F} which has a the very simple mathematical expression given by equation (23). This is due to equation (22) and to the fact that, according to the typical numerical values for λ , η , σ_q and ℓ given before, one has

$$2 \left[\left(\frac{\sigma_q}{\ell} \right)^2 + \eta \right] \simeq 2 \times 10^{-16}, \tag{24}$$

a vanishingly small value. We point out that equation (23) should not be confused with the quantum average value of the projection operator P_B relative to the position of the pointer, as given by standard quantum mechanics. As a matter of fact, since an experiment always involves a macroscopic system, the one described by $|\psi_t\rangle$, a state vector which includes both the measured system and the measuring apparatus entails a dynamical evolution which is entirely different from that predicted by the Schrödinger equation, in particular it always keeps the measuring device in a state well localized in space. We now prove the theorem.

$$\begin{aligned} |\mathbb{P}_{f(q_{t_F})}[V] - \mathbb{Q}_{t_F}[V]| &\equiv |\mathbb{P}[q_{t_F}^{-1}(\Delta_V)] - \mathbb{E}_{\mathbb{P}}[\langle \psi_{t_F} | P_{\Delta_V} | \psi_{t_F} \rangle]| \\ &= \left| \int_{q_{t_F}^{-1}(\Delta_V)} d\mathbb{P} - \int_{\Omega} \langle \psi_{t_F} | P_{\Delta_V} | \psi_{t_F} \rangle d\mathbb{P} \right| \\ &\leq \int_{q_{t_F}^{-1}(\Delta_V)} (1 - \langle \psi_{t_F} | P_{\Delta_V} | \psi_{t_F} \rangle) d\mathbb{P} + \int_{\Omega - q_{t_F}^{-1}(\Delta_V)} \langle \psi_{t_F} | P_{\Delta_V} | \psi_{t_F} \rangle d\mathbb{P} \\ &= \int_{q_{t_F}^{-1}(\Delta_V)} \langle \psi_{t_F} | P_{\mathbb{R} - \Delta_V} | \psi_{t_F} \rangle d\mathbb{P} + \int_{q_{t_F}^{-1}(\mathbb{R} - \Delta_V)} \langle \psi_{t_F} | P_{\Delta_V} | \psi_{t_F} \rangle d\mathbb{P}, \end{aligned} \tag{25}$$

where Ω is the sample space on which the stochastic dynamics is defined. Let us define $\delta_V \equiv \cup_n \delta_n$ with $\delta_n \in \Delta_V$; the first of the two terms in the last line can be re-written as follows:

$$\int_{q_{t_F}^{-1}(\Delta_V)} \langle \psi_{t_F} | P_{\mathbb{R} - \Delta_V} | \psi_{t_F} \rangle d\mathbb{P} = \int_{q_{t_F}^{-1}(\delta_V)} \langle \psi_{t_F} | P_{\mathbb{R} - \Delta_V} | \psi_{t_F} \rangle d\mathbb{P} + \int_{q_{t_F}^{-1}(\Delta_V - \delta_V)} \langle \psi_{t_F} | P_{\mathbb{R} - \Delta_V} | \psi_{t_F} \rangle d\mathbb{P}; \tag{26}$$

according to the second property of DRMs listed above, the probability measure of the set $\Delta_V - \delta_V$ is smaller than η ; thus, taking also into account that $\langle \psi_{t_F} | P_{\mathbb{R} - \Delta_V} | \psi_{t_F} \rangle \leq 1$, we can write

$$\int_{q_{t_F}^{-1}(\Delta_V - \delta_V)} \langle \psi_{t_F} | P_{\mathbb{R} - \Delta_V} | \psi_{t_F} \rangle d\mathbb{P} \leq \eta. \tag{27}$$

Regarding the first term on the right-hand side of equation (26), we now show that the integrand $\langle \psi_{t_F} | P_{\mathbb{R} - \Delta_V} | \psi_{t_F} \rangle$ is extremely small for any $\omega \in \Omega$ such that q_{t_F} lies within δ_V . Defining $\psi(x) = \langle x | \psi_{t_F} \rangle$, we have of course

$$\langle \psi_{t_F} | P_{\mathbb{R} - \Delta_V} | \psi_{t_F} \rangle = \int d\{q\} \int_{\mathbb{R} - \Delta_V} dx |\psi_{t_F}(\{q\}, x)|^2, \tag{28}$$

where $\{q\}$ denote all degrees of freedom involved in the measurement, except the one for the center of mass of the pointer. Whenever q_{t_F} belongs to δ_V , we have that the parameter ℓ defining essentially the extension of the intervals Δ_n , satisfies $\ell^2 \leq (x - q_{t_F})^2$ for all $x \in \mathbb{R} - \Delta_V$; we get then the following inequality:

$$\begin{aligned} \ell^2 \int d\{q\} \int_{\mathbb{R} - \Delta_V} dx |\psi_{t_F}(\{q\}, x)|^2 &\leq \int d\{q\} \int_{\mathbb{R} - \Delta_V} dx (x - q_{t_F})^2 |\psi_{t_F}(\{q\}, x)|^2 \\ &\leq \int d\{q\} \int_{\mathbb{R}} dx (x - q_{t_F})^2 |\psi_{t_F}(\{q\}, x)|^2 \equiv \sigma_q^2. \end{aligned} \tag{29}$$

Equations (26), (27) and (29) lead to the following result:

$$\int_{q_{t_F}^{-1}(\Delta_V)} \langle \psi_{t_F} | P_{\mathbb{R} - \Delta_V} | \psi_{t_F} \rangle d\mathbb{P} \leq \left(\frac{\sigma_q}{\ell}\right)^2 + \eta \tag{30}$$

A completely symmetric argument holds for the second and last terms at the right-hand side of equation (25) as well, hence the theorem is proven.

To summarize, we have shown that in measurement-like situations it is fully legitimate, as a consequence of the reducing dynamics of DRM, to use the probability measure \mathbb{Q}_t in place of $\mathbb{P}_{f(q_t)}$ to compute the probabilities of the possible outcomes of an experiment.

3.2.2. *The emergence of the Hilbert-space operator formalism.* Let us now focus our attention on the probability measure \mathbb{Q}_t which, according to the definition (23), can be re-written also as

$$\mathbb{Q}_t[V] \equiv \text{Tr}[\rho_t(\mathbb{I}_{\{q\}} \otimes P_{\Delta_V})] \quad V \in \mathcal{G}, \tag{31}$$

where $\mathbb{I}_{\{q\}}$ is the identity operator acting on the space of all degrees of freedom of the experiment, except the one referring to the center of mass of the pointer. The above expression is a consequence of the well-known formula typical of DRMs [2]:

$$\mathbb{E}_{\mathbb{P}}[\langle \psi_t | O | \psi_t \rangle] = \text{Tr}[O \rho_t] \quad \text{with} \quad \rho_t \equiv \mathbb{E}_{\mathbb{P}}[|\psi_t\rangle\langle\psi_t|], \tag{32}$$

where O is any suitable operator. We remind the reader that the state vector $|\psi_t\rangle$ (thus also ρ_t) refers both to the (whole) state of the measuring device as well as to the state of the microscopic system. Regarding the initial state $|\psi_{t_0}\rangle$ (or, equivalently, ρ_{t_0}), right before the experiment begins, we make the following assumption:

Assumption 1. At the beginning of the experiment ($t = t_0$), the state of the micro-system and that of the apparatus are factorized:

$$\rho_{t_0} = \rho_{t_0}^S \otimes \rho_{t_0}^A, \tag{33}$$

where $\rho_{t_0}^S = |\psi_{t_0}^S\rangle\langle\psi_{t_0}^S|$ represents the initial state of the microscopic system, which we assume to be pure, while $\rho_{t_0}^A$ represents the initial state of the apparatus.

This initial factorization of the two states is a very natural assumption to make, since otherwise the microscopic system would not be in any defined state, whose property the experiment should detect. Moreover, we recall the following important property characterizing dynamical reduction models:

Property 3. The dynamical evolution $\Sigma_{(t_0, t_F)}$ mapping the density matrix ρ_{t_0} describing the global state prior to the experiment, into the final state ρ_{t_F} after the experiment, is of the quantum dynamical semigroup type, thus, in particular, linear and trace-preserving.

According to assumption 1 and property 3, we can write \mathbb{Q}_t , computed at time $t = t_F$, as follows:

$$\mathbb{Q}_{t_F}[V] \equiv \text{Tr}[\Sigma_{(t_0, t_F)}(\rho_{t_0}^S \otimes \rho_{t_0}^A) \mathbb{I}_{\{q\}} \otimes P_{\Delta_V}] \quad V \in \mathcal{G}, \tag{34}$$

for any fixed $V \in \mathcal{G}$. Assuming now that $|\psi_{t_0}^S\rangle$ can run over the entire Hilbert space \mathcal{H}^S , the above expression defines the diagonal part of the following bounded sesquilinear form:

$$\begin{aligned} \mathcal{H}^S \otimes \mathcal{H}^S &\longrightarrow \mathbb{C} \\ (|\psi_{t_0}^S\rangle, |\phi_{t_0}^S\rangle) &\longrightarrow \text{Tr}[\Sigma_{(t_0, t_F)}(|\psi_{t_0}^S\rangle\langle\phi_{t_0}^S| \otimes \rho_{t_0}^A) \mathbb{I}_S \otimes P_{\Delta_V}]; \end{aligned} \tag{35}$$

which, according to the Riesz representation theorem, can be written as follows:

$$\mathbb{Q}_{t_F}[V] = \langle \psi_{t_0}^S | O_V | \psi_{t_0}^S \rangle = \text{Tr}[\rho_{t_0} O_V], \tag{36}$$

where O_V is a bounded linear operator in \mathcal{H}^S . In our case, O_V turns out to be also self-adjoint and defines a POVM from the measurable space of the possible outcomes $(\mathcal{O}, \mathcal{G})$ to the Hilbert space \mathcal{H}^S of the micro-system. This is the desired result, which we formalize in the following theorem:

Theorem 2. *According to the properties of DRMs stated before (properties 1–3), according to assumption 1 and, within the limits set by theorem 1, one can write*

$$\mathbb{P}_{f(q_F)}[V] \simeq \langle \psi_{t_0}^S | O_V | \psi_{t_0}^S \rangle \quad \forall V \in \mathcal{G}. \tag{37}$$

In other words, the probability that the outcome of a given experiment belongs to the measurable subset V of the set \mathcal{O} of possible outcomes can be (with very high accuracy)

expressed as the average value of the effect O_V , $V \in \mathcal{G}$, of a POVM, over the initial state $|\psi_{t_0}^S\rangle$ of the microscopic system.

We have thus recovered the operator formalism of standard quantum mechanics. It is interesting to compare equation (36) with equation (34):

$$\mathbb{Q}_{t_F}[V] = \text{Tr}[\Sigma_{(t_0,t_F)}(\rho_{t_0}^S \otimes \rho_{t_0}^A)\mathbb{I}_S \otimes P_{\Delta_V}] = \text{Tr}[\rho_{t_0} O_V], \tag{38}$$

the middle term of the above equation provides the true physical description of the experiment: it gives the probability for the pointer to lie within a well-defined region along the graduate scale at time t_F ; on the other hand, the last term provides the compact (and very handy) quantum way of expressing such probabilities in terms of the initial state of the micro-system. A couple of further comments are at order.

- (1) Clearly, since a wavefunction has always a spatial extension, the width of the intervals Δ_n cannot be set equal to zero. This means that, according to DRMs, only experiments having at most a countable number of outcomes can be performed. This of course includes all physically realizable experiments, while those having a continuous number of outcomes represent a mathematical idealization.
- (2) Nowhere in our analysis we have explicitly used the fact that the probabilities of the possible outcomes of the experiment, as predicted by DRMs, practically coincides with quantum probabilities. However, such a feature of DRMs is implicitly contained in property 3, i.e. in the fact that the evolution law for the statistical operator is linear; the reason is the following. When a jump process of the form (1) occurs on the n th particle of a many-particle system, a density matrix $\rho \equiv \sum_i c_i |\psi_i\rangle\langle\psi_i|$ changes, in accordance to axiom B, as follows:

$$\rho \equiv \sum_i c_i |\psi_i\rangle\langle\psi_i| \longrightarrow \sum_i c_i \int d^3\mathbf{x} p(\mathbf{x}) \frac{L_n(\mathbf{x})|\psi_i\rangle\langle\psi_i|L_n(\mathbf{x})}{\|L_n(\mathbf{x})\psi_i\|^2}, \tag{39}$$

where $p(\mathbf{x})$ is the probability density for a jump to occur in \mathbf{x} . As we see, the above evolution is *not* linear in ρ , unless we require $p(\mathbf{x})$ to be of the form (2), which agrees with the Born probability rule.

4. Reproducibility and PVM

The example we have discussed in section 2 belongs to a particular subclass of experiments, because it can be associated with a PVM, while the general theorem of the previous section shows that experiments are in general associated with POVMs, which are more general than PVM. In this section, we show that the possibility of associating a PVM with an experiment is strictly connected to the *reproducibility* of the experiment itself: since on standard books on quantum mechanics it is (implicitly) assumed that experiments are reproducible, the analysis shows why usually experiments are associated with projection operators.

We say that an experiment is reproducible if it can be performed many times on the same physical system and, each time we perform two runs in a row, one right after the other, the second one always yields the same outcome as the first one. In order to give a more rigorous definition, we first have to recall another important feature of DRMs:

Property 4. At the end of a measurement process, the final state $|\psi_{t_F}\rangle$ of the whole system is, to a great accuracy, a factorized state of the system and the apparatus¹⁶: $|\psi_{t_F}\rangle \simeq |\psi_{t_F}^S\rangle \otimes |\psi_{t_F}^A\rangle$. Accordingly, the vector $|\psi_{t_0}^S\rangle$ representing the state of the microscopic system right before

¹⁶ See [10] for a quantitative analysis of this feature of DRMs.

the experiment began, changes to a new (normalized) state $|\psi_{t_F}^S\rangle$, which depends in general both on $|\psi_{t_0}^S\rangle$ and on the stochastic dynamics of the interaction between the system S and the apparatus A .

In the following, when convenient we will write $|\psi_{t_F}^S/\psi_{t_0}^S\rangle$ in place of $|\psi_{t_F}^S\rangle$, to stress the dependence of the final state on the initial one; moreover, we will often write $|\psi_{t_F}^{S,v_n}\rangle$ in place of $|\psi_{t_F}^S\rangle$ to signify that $|\psi_{t_F}^S\rangle$ is the state to which the systems S is reduced at the end of the first experiment in which we suppose that the outcome v_n has been obtained. Let $|\psi_t^{2:v_n}\rangle$ describe the time evolution during the second experiment, when the initial state at time t_0' at which the second experiment takes place is assumed to be $|\psi_{t_0'}^{2:v_n}\rangle = |\psi_{t_F}^{S,v_n}\rangle \otimes |\psi_{t_0'}^A\rangle$. A reproducible experiment is defined as follows.

Definition 2 (Reproducible experiment). *An experiment on a physical system S is said to be reproducible if and only if:*

- (i) *The experiment can be performed on S at least twice; moreover, the Hilbert space \mathcal{H} of vectors describing the possible states of S before the measurement is the same as the Hilbert space of vectors describing the possible states of the system after the measurement; in other words, the totality of the possible final states $|\psi_{t_F}^S\rangle$ must span \mathcal{H} :*

$$\text{span}\{|\psi_{t_F}^S/\psi_{t_0}^S\rangle, |\psi_{t_0}^S\rangle \in \mathcal{H}\} = \mathcal{H}. \quad (40)$$

- (ii) *Let us suppose that two experiments are done, one immediately after the other; let $\mathbb{P}_{f(q_{t_F}')}^{v_n}$ be the probability distribution of the outcomes of the second experiment, assuming that $|\psi_{t_F}^{S,v_n}\rangle$ has been taken as the initial state of the microscopic system for the second experiment, i.e. assuming that the outcome of the first experiment is v_n . We then require that*

$$\mathbb{P}_{f(q_{t_F}')}^{v_n}[V] = 1, \quad \forall V : v_n \in V, \quad (41)$$

i.e. that the outcome of the second run of the experiment belongs to V with certainty.

Let us briefly comment on the above definition. The first request in the definition above excludes experiments which alter the nature of the physical system, e.g. because they destroy it, or one of its parts, or because they transform it in a new physical system with a space of states \mathcal{H}' different from the original space \mathcal{H} . The second request just embodies the idea of reproducibility, by assuring that the same outcome is obtained with certainty when the experiment is performed twice. Now we can state the following theorem:

Theorem 3 (Reproducible experiment). *Let us consider an experiment which, according to theorem 2, is associated with a POVM $\{O_V\}_V$. If the experiment is reproducible, and within the limits of replacing the probability $\mathbb{P}_{f(q_{t_F}')}^{v_n}$ with the probability $\mathbb{Q}_{t_F'}^{v_n}[V] = \mathbb{E}_{\mathbb{P}}[\langle \psi_{t_F'}^{2:v_n} | P_{\Delta_V} | \psi_{t_F'}^{2:v_n} \rangle]$, then the POVM is a PVM.*

The proof of the theorem is given in [9]; here we propose a simplified version of it. We first of all note that, given a self-adjoint positive semidefinite operator O with bounds $0 \leq O \leq I$, then

$$\langle \psi | O | \psi \rangle = 1 \Rightarrow O | \psi \rangle = | \psi \rangle \quad (42)$$

$$\langle \psi | O | \psi \rangle = 0 \Rightarrow O | \psi \rangle = | \omega \rangle, \quad (43)$$

where $|\omega\rangle$ is the null vector. To show this, let us write $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_1^\perp$, where \mathcal{H}_1 is the subspace of all eigenstates of O corresponding to the eigenvalue 1 and \mathcal{H}_1^\perp its orthogonal

complement. Let $|\psi\rangle$ be a normalized vector, which we decompose in: $|\psi\rangle = |\psi_{\parallel}\rangle + |\psi_{\perp}\rangle$, where $|\psi_{\parallel}\rangle \in \mathcal{H}_1$ and $|\psi_{\perp}\rangle \in \mathcal{H}_1^{\perp}$. It follows that

$$\langle \psi | O | \psi \rangle = 1 \Rightarrow \langle \psi_{\perp} | O | \psi_{\perp} \rangle = \langle \psi_{\perp} | \psi_{\perp} \rangle. \tag{44}$$

On the other hand, the property $0 \leq O \leq I$ implies that O acts on the subspace \mathcal{H}_1^{\perp} like a contraction, i.e. that $\|O|\psi_{\perp}\rangle\| < \| |\psi_{\perp}\rangle \|$, unless $|\psi_{\perp}\rangle = |\omega\rangle$. This contraction property, together with the Cauchy–Schwarz inequality, gives

$$\langle \psi_{\perp} | O | \psi_{\perp} \rangle < \langle \psi_{\perp} | \psi_{\perp} \rangle \quad \text{unless} \quad |\psi_{\perp}\rangle = |\omega\rangle. \tag{45}$$

Now, equations (44) and (45) are incompatible unless $|\psi_{\perp}\rangle = |\omega\rangle$, which proves equation (42). In a similar way one can prove also equation (43).

Coming back to the probability measure $\mathbb{Q}_{t_F}^{v_n}$, according to the analysis of the previous section and to the hypotheses of our theorem, we can write, for any $V \subseteq \mathcal{G}$:

$$\mathbb{Q}_{t_F}^{v_n}[V] = \langle \psi_{t_F}^{S:v_n} | O_V | \psi_{t_F}^{S:v_n} \rangle = \begin{cases} 1 & \text{if } v_n \in V, \\ 0 & \text{else.} \end{cases} \tag{46}$$

Let \mathcal{H}_1^V be the subspace of all eigenstates of O_V corresponding to the eigenvalue 1, and \mathcal{H}_0^V be the subspace of all eigenstates corresponding to the eigenvalue 0; according to (46), and to the fact that $\text{span}\{|\psi_{t_F}^{S:v_n}\rangle\} \equiv \mathcal{H}$, we can write $\mathcal{H} = \mathcal{H}_1^V \oplus \mathcal{H}_0^V$. This result proves that O_V is an orthogonal projector on \mathcal{H} , with support in \mathcal{H}_1^V . Hence the POVM $\{O_V\}_V$ is a PVM.

The *ideal measurements* usually considered in many QM textbook, i.e., those obeying the rule of the wave-packet reduction (WPR) postulate, are reproducible, and therefore the POVM associated with them reduces to a PVM. When this happens, the experiment can be entirely characterized, with respect to its statistical properties, by a single self-adjoint operator, as is usually done in quantum mechanics.

5. The ‘tail problem’: the Stern–Gerlach experiment revisited

One of the reason why, with reference to theorem 1, the two probabilities \mathbb{P}_{q_t} and \mathbb{Q}_{t_F} are not strictly equal is that, even when $q_t \in \delta_n$, the wavefunction for the center of mass of the pointer has not a compact support contained in the interval Δ_n , but it has tails spreading out to infinity; such tails, being extremely small (in the sense that the integral of the square modulus of the wavefunction over the region laying outside Δ_n , assuming that its center is contained within δ_n , is very small), give not rise to any problem in connection with the interpretation of the theory and its physical predictions.

In this section, we want to point out that precisely the same problem with tails occur also in standard quantum mechanics. Let us take as an example the Stern–Gerlach experiment: this is often used in textbooks as *the* paradigmatic experiment which illustrates the correspondence between observables and self-adjoint operators. What textbooks usually provide is just a simplified description of the truly observed experimental results; a more realistic analysis which takes into account also the spatial (beside spin) degrees of freedom of the atoms sent through the apparatus would show that tails emerge also here, which (just in principle) give rise to some potential problems in interpreting the outcome of the experiment, and in associating a self-adjoint operator with it. In this section, we perform such a kind of analysis, we discuss the role of the tails of the wavefunction of the atoms and we make precise the sense in which it is legitimate to associate the usual spin operator S_z with the experiment.

As it is well known, Stern and Gerlach used an oven to produce and send a beam of silver atoms through an inhomogeneous magnetic field, letting it eventually impinge on a glass plate. In order to analyze the effect of the magnetic field on the beam, two separate experiments were

originally [11] performed: one with the magnet generating the field turned on, with a run time of 8 h, another with the magnet turned off, with a run time of 4.5 h. In the magnet-off case, a single bar of silver approximately 1.1 mm long and 0.06–0.1 mm wide was deposited on the glass plate. In the magnet-on case, a pair-of-lips shape appeared on the glass: the shape was 1.1 mm long, one lip was 0.11 mm wide, the other was 0.20 mm wide; both lips appeared deflected with respect to the position of the magnet-off bar, and the maximum gap between the upper and lower lips was approximately of the order of magnitude of the width of the lips. Stern and Gerlach made only visual observations through a microscope, with no statistics on the distributions: they did not obtain ‘two spots’ as it is usually stated on textbooks, and though the beam was clearly split in two distinguishable parts, these were not disjoint. They accounted for the experiment as exhibiting the property of ‘space quantization in a magnetic field’.

Let us give a simplified mathematical description of the experiment, taking however into account not only the spin, but also the spatial degrees of freedom of (the center of mass of) the silver atoms [12]. The silver atoms of a Stern–Gerlach experiment can be treated as spin one-half elementary particles to a very high degree of accuracy; let us assume that they are initially prepared in an (arbitrary) spin state

$$|c_+, c_-\rangle = c_+|+\rangle + c_-|-\rangle, \quad (|c_+|^2 + |c_-|^2 = 1), \quad (47)$$

where $|+\rangle$ and $|-\rangle$ denote the usual normalized eigenstates of S_z . The wavefunction describing the center of mass of the atom at the initial time $t = 0$ can be taken to be a normalized Gaussian wave packet

$$G(\mathbf{x}; \mathbf{x}_0, \mathbf{p}_0, \Delta_{\mathbf{x}}) = \langle \mathbf{x} | \mathbf{x}_0, \mathbf{p}_0, \Delta_{\mathbf{x}} \rangle \quad (48)$$

centered around the position \mathbf{x}_0 , traveling along the x -axis towards the region where the magnet is located, with mean value of momentum \mathbf{p}_0 and spatial spread $\Delta_{\mathbf{x}}$, extremely well localized with respect to the dimensions of the region where the magnetic field is different from 0.

Let us model the interaction between the atom and the magnetic field, which we treat as a static external field, by the usual Hamiltonian operator $H_{\text{SG}} = \boldsymbol{\mu} \cdot \mathbf{B}$ with $\boldsymbol{\mu} = k\boldsymbol{\sigma}$, \mathbf{B} the magnetic field, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ the three Pauli operators and k a known constant whose value is irrelevant for the following discussion. Denoting by $\hat{\mathbf{z}}$ the unit vector along the z -axis, let us assume that the inhomogeneous magnetic field is¹⁷ $\mathbf{B} = (B_0 - bz)\hat{\mathbf{z}}$ inside the spatial region where the magnet is located and with a negligible gradient outside it. Finally, to simplify the matter as far as possible, let us adopt the ‘impulsive measurement’ assumption, which amounts to ignoring the free evolution of the silver atom while it is interacting with the magnetic field, and let us suppose that the time interval τ between the emission of the silver atom from the oven and the moment when it impinges on the glass plate is so small that the spread $\Delta_{\mathbf{x}}$ of its wave packet does not vary appreciably.

With these premises, a simple calculation shows that the state of the silver atom at the time $t = \tau$, after it went through the inhomogeneous magnetic field region and just before the detection by the glass plate, is

$$|\psi_\tau\rangle = c_+|\mathbf{x}_+, \mathbf{p}_+, \Delta_{\mathbf{x}}\rangle \otimes |+\rangle + c_-|\mathbf{x}_-, \mathbf{p}_-, \Delta_{\mathbf{x}}\rangle \otimes |-\rangle, \quad (49)$$

¹⁷ This assumption is inconsistent with the Maxwell equations, in particular, with $\vec{\nabla} \times \mathbf{B} = 0$. A more realistic assumption would be $\mathbf{B} = bx\hat{\mathbf{x}} + (B_0 - bz)\hat{\mathbf{z}}$. A simple analysis shows, however, that this second assumption gives rise to correcting terms which are inessential for our present discussion.

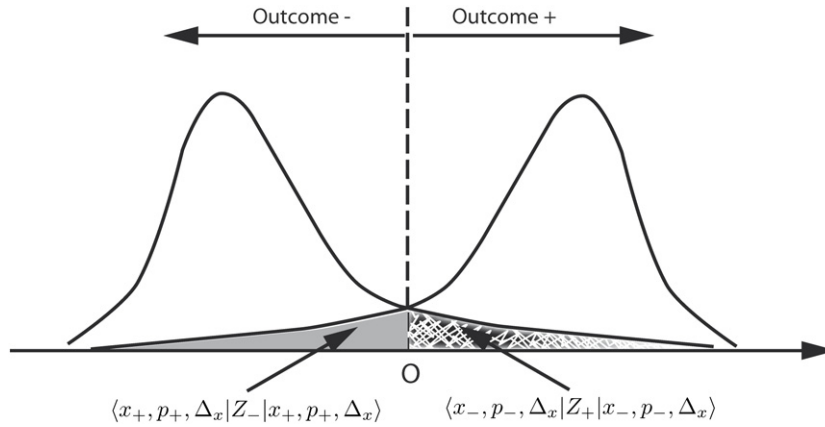


Figure 2. Typical (schematic) spatial distribution of the superposition of two Gaussian wavefunctions in a Stern–Gerlach experiment. The wavefunction corresponding to the outcome ‘+’ has a tail spreading in the region corresponding to the outcome ‘–’ and vice-versa.

where the wavefunctions $G(\mathbf{x}; \mathbf{x}_\pm, \mathbf{p}_\pm, \Delta_x) = \langle \mathbf{x} | \mathbf{x}_\pm, \mathbf{p}_\pm, \Delta_x \rangle$ are normalized Gaussian packets of mean momentum \mathbf{p}_\pm and mean position \mathbf{x}_\pm given by¹⁸

$$\mathbf{p}_\pm = \mathbf{p}_0 \pm kb\tau\hat{z}, \quad \mathbf{x}_\pm \simeq \mathbf{x}_0 + \frac{\mathbf{p}_0}{m}\tau \pm \frac{kb}{m}\tau^2\hat{z}. \quad (50)$$

The detection process by means of the glass plate can be modeled by associating with the plate the operator $Z \equiv Z_+ - Z_-$, acting on the spatial degrees of freedom, where Z_+ and Z_- are the two projection operators corresponding to the localization in the upper ($z \geq 0$) or lower ($z \leq 0$) half parts of the plate. Recalling that we choose a normalized vector to represent the initial state, the corresponding probabilities according to standard quantum mechanical rules are then easily seen to be

$$\begin{aligned} \mathbb{P}_{(c_+, c_-)}(\pm 1) &= \langle \psi_\tau | Z_\pm \otimes I_{\mathbb{C}^2} | \psi_\tau \rangle \\ &= |c_+|^2 \langle \mathbf{x}_+, \mathbf{p}_+, \Delta_x | Z_\pm | \mathbf{x}_+, \mathbf{p}_+, \Delta_x \rangle + |c_-|^2 \langle \mathbf{x}_-, \mathbf{p}_-, \Delta_x | Z_\pm | \mathbf{x}_-, \mathbf{p}_-, \Delta_x \rangle \end{aligned} \quad (51)$$

By looking at equation (51), we immediately see that the probability to observe an outcome in, say, the upper part of the screen, is not $|c_+|^2$, as one can find in the textbook descriptions of the experiment, because the two wavefunctions $G(\mathbf{x}; \mathbf{x}_\pm, \mathbf{p}_\pm, \Delta_x)$ are not localized sharply in (respectively) the upper and lower parts of the glass plate. Indeed, the tail of the ‘+’ wavefunction $G(\mathbf{x}; \mathbf{x}_+, \mathbf{p}_+, \Delta_x)$ which lies in the lower part does not contribute at all to $\mathbb{P}_{(c_+, c_-)}(+1)$, while the tail of the ‘–’ wavefunction $G(\mathbf{x}; \mathbf{x}_-, \mathbf{p}_-, \Delta_x)$ laying in the upper part does contribute (see figure 2). Needless to say, this is not a disfeature due to the use of a Gaussian wave packet: the whole analysis could be repeated practically unaltered for an arbitrary initial wave packet, because even if one starts with a wave packet with compact support, the free evolution would immediately and unavoidably spread it all over space.

However, provided the initial wave packet is sufficiently well localized both in position as well as in momentum (within the limits allowed by the Heisenberg principle, of course), and the spatial separation of the two final wave packets is much greater than their spatial spreads, i.e. $2kb\tau^2/m \gg \Delta_x$ the ‘crossed’ contributions of the tails, namely $\langle \mathbf{x}_+, \mathbf{p}_+, \Delta_x | Z_- | \mathbf{x}_+, \mathbf{p}_+, \Delta_x \rangle$ and $\langle \mathbf{x}_-, \mathbf{p}_-, \Delta_x | Z_+ | \mathbf{x}_-, \mathbf{p}_-, \Delta_x \rangle$ are small compared with the uncrossed contributions, and the textbook approximation $\mathbb{P}_{(c_+, c_-)}(\pm 1) = |c_\pm|^2$ can be meaningfully recovered.

¹⁸ In deriving equation (50) the realistic assumption $\tau kb \ll p_0 \Rightarrow |\mathbf{p}_\pm| \simeq |\mathbf{p}_0|$ (i.e., that the transverse variation of the momentum is negligible with respect to its modulus) has also been taken into account.

It is interesting to restate the previous argument in the following way, by observing that equation (51) can be rewritten in terms of the initial spin state $|c_+, c_-\rangle$ as

$$\mathbb{P}_{(c_+, c_-)}(\pm 1) = \langle c_+, c_- | O_{\pm} | c_+, c_- \rangle, \quad (52)$$

where the two effects O_{\pm} are

$$O_{\pm} \equiv \langle \mathbf{x}_+, \mathbf{p}_+, \Delta_{\mathbf{x}} | Z_{\pm} | \mathbf{x}_+, \mathbf{p}_+, \Delta_{\mathbf{x}} \rangle |+\rangle\langle +| + \langle \mathbf{x}_-, \mathbf{p}_-, \Delta_{\mathbf{x}} | Z_{\pm} | \mathbf{x}_-, \mathbf{p}_-, \Delta_{\mathbf{x}} \rangle |-\rangle\langle -| \quad (53)$$

acting on the spin Hilbert space \mathbb{C}^2 . We thus obtain a result we are now used to: the probability distribution of the outcomes has been expressed as the average value of the effects of a POVM over the initial state of the microscopic quantum system. However, since the wave packets are spread out in space, we see that the effects (53) are *not* projection operators, i.e., $O_{\pm}^2 \neq O_{\pm}$. Concluding, the observable associated with the Stern–Gerlach experiment is not exactly represented by the operator S_z (or, equivalently, by the PVM given by $|+\rangle\langle +|$ and $|-\rangle\langle -|$): it is rather a generalized observable described by the POVM formed by the two effects O_{\pm} . Thus, to be rigorous, also within standard quantum mechanics the association of the operator S_z with a Stern–Gerlach experiment is just an approximation, whose range of validity has been made clear by the previous discussion: indeed, when the crossed contributions of the tails of the two Gaussian wavefunctions are negligible with respect to the uncrossed ones, we have

$$O_+ \simeq |+\rangle\langle +| \quad \text{and} \quad O_- \simeq |-\rangle\langle -|, \quad (54)$$

so that the POVM $\{O_+, O_-\}$ associated with the experiment approximately becomes the spectral PVM $\{|+\rangle\langle +|, |-\rangle\langle -|\}$ of the spin operator S_z .

6. Observables as self-adjoint operators in classical mechanics: an example from classical electromagnetic theory

From the previous analysis one should have grasped that the operator formalism for describing the outcome of experiments is not a peculiar feature of quantum mechanics but, according to the Riesz representation theorem, it can be applied whenever the states of physical systems are represented by vectors of a linear vector space and the outcome of an experiment depends ‘quadratically’ on the initial state of the system being measured. As such, there is no reason why it should not be possible to apply such a formalism also to classical systems: to illustrate this fact, we now give an example taken from classical electromagnetic theory.

Let us confine our attention to a monochromatic plane wave of frequency ω traveling in the vacuum in some given direction¹⁹ $\hat{\mathbf{n}}$ (i.e., with wave vector $\mathbf{k} = \omega\hat{\mathbf{n}}$). Such a wave is completely characterized by the electric field

$$\mathbf{E}(t, \mathbf{x}) = \text{Re}(\epsilon \exp [i(\omega t - \mathbf{k} \cdot \mathbf{x})]), \quad (55)$$

the complex wave amplitude ϵ expresses both the intensity as well as the polarization state of the wave. Accordingly, taking into account the linearity of Maxwell’s equations and the fact that the wave intensity²⁰ I can be written in terms of ϵ by means of the canonical complex scalar product,

$$I = \epsilon^* \cdot \epsilon, \quad (56)$$

we see that, if we limit our attention to the degrees of freedom contained in the wave amplitude alone, ignoring the spatial ones, a monochromatic plane wave qualifies as a physical system

¹⁹ According to the customary practice in classical theories, in this subsection we will denote vectors with the traditional bold notation \mathbf{v} rather than the Dirac notation $|v\rangle$.

²⁰ The wave intensity we are referring here to is of course the mean value of the Poynting vector of the wave over a time interval long enough with respect to the wave period.

whose states belong to the Hilbert space \mathbb{C}^2 and on which experiments (typically detection of the intensity) represented by quadratic forms can be performed.

One of the simplest examples of such a kind of experiment on a monochromatic plane wave can be constructed by letting the wave impinge on a linear Polaroid and by measuring the transmitted intensity. To further simplify the matter, let us restrict to the case of a linearly polarized monochromatic plane wave, i.e., let us assume $\epsilon \in \mathbb{R}^2$. Experience shows that in such a situation the transmitted wave intensity I_f depends on the angle between the directions of polarization of the wave and the filter according to the Malus law:

$$I_f = I_i \cos^2 \theta, \quad (57)$$

where $I_i = \epsilon \cdot \epsilon$ is of course the incident wave intensity and, if $\hat{\eta}$ is the polarization direction of the filter, $\cos \theta = \epsilon \cdot \hat{\eta} / \|\epsilon\|$. Accordingly, Malus law (57) can be re-written as

$$I_f = (\hat{\eta} \cdot \epsilon)^2, \quad (58)$$

and we easily recognize that the outcome of the experiment (the output intensity I_f) is a quadratic form of the initial state of the wave (its polarization ϵ). Then, according to the Riesz representation theorem, we can express I_f in terms of a bounded linear operator O as follows (we now pass to the Dirac notation to highlight the conclusion):

$$I_f(\hat{\eta}) = \langle \epsilon | O_{\hat{\eta}} | \epsilon \rangle, \quad O_{\hat{\eta}} = |\hat{\eta}\rangle\langle\hat{\eta}|. \quad (59)$$

Of course, in the present example, the average value of the operator $O_{\hat{\eta}}$ over the initial state of the micro-system does not represent the average of the possible outcomes weighted with their probabilities, but it is simply the (deterministic) value of the output intensity.

Let us note that this example also shows that *non-commutativity*, which is usually considered as a characteristic trait of quantum mechanics, can and does arise in classical and deterministic contexts as well. Indeed, by taking into account the fact that the transmitted wave is polarized along the direction of the filter:

$$|\epsilon_{\text{out}}\rangle = |\hat{\eta}\rangle\langle\hat{\eta}|\epsilon_{\text{in}}\rangle, \quad (60)$$

and by considering two successive arbitrarily oriented filters, we see that the transmitted intensity

$$I_f = I_i \cos^2 \theta_1 \cos^2(\theta_2 - \theta_1) \quad (61)$$

of the wave after the two filters *depends*, in general, on their respective order. The mathematical counterpart of this property is that the two operators $O_{\hat{\eta}_1}$ and $O_{\hat{\eta}_2}$ associated with the two experiments do *not* commute, a part from the very special orientations which correspond to $|\hat{\eta}_1\rangle = \pm|\hat{\eta}_2\rangle$ or $\langle\hat{\eta}_1|\hat{\eta}_2\rangle = 0$.

7. Conclusions

In this paper, we have shown how the Hilbert-space operator formalism—i.e., the use of self-adjoint operators and, more generally, of POVM to describe experiments on quantum systems—can be recovered within the context of DRMs: it can be derived from the dynamical laws of DRMs satisfying the very general assumptions that we have analyzed in section 3.2 as a compact way to express the statistical properties of the outcomes of *measurement-like* experiments. It is worthwhile stressing once again that within the context of DRMs the operator formalism has no special ontological meaning: as this paper shows in detail, it is merely a convenient tool that can be used to describe certain experiments that we are accustomed to think of as ‘measurements’.

This and the previous analysis prove that DRMs represent a well-grounded theory, whose ontology is clearly specified and consistent with our macroscopic perceptions, deprived from the typical paradoxes of quantum mechanics which are connected to the special roles that the theory attributes to ‘measurement processes’ and ‘observables’.

We consider the results of this paper as an important step which completes, at the non-relativistic level, the general line of thought and the world view which has inspired the dynamical reduction program, both at the formal and interpretational level.

Acknowledgments

We are indebted with D Dürr and R Tumulka for many stimulating comments. DGMS gratefully acknowledges support by the Department of Theoretical Physics of the University of Trieste. The work of AB has been supported partly by the EU grants MEIF CT 2003–500543 and ERG 044941-STOCH-EQ, and partly by DFG (Germany).

References

- [1] Ghirardi G C, Rimini A and Weber T 1986 *Phys. Rev. D* **34** 470
- [2] Bassi A and Ghirardi G C 2003 *Phys. Rep.* **379** 257
- [3] Pearle P 1999 *Open Systems and Measurement in Relativistic Quantum Theory (Lecture Notes in Physics)* vol 526 ed H-P Breuer and F Petruccione (Berlin: Springer-Verlag)
- [4] Pearle P 1989 *Phys. Rev. A* **39** 2277
- [5] Ghirardi G C, Rimini A and Weber T 1990 *Phys. Rev. A* **42** 78
- [6] Adler S L 2004 *Quantum Theory as an Emergent Phenomenon* (Cambridge: Cambridge University Press) chapter 6
- [7] Ghirardi G C, Grassi R and Benatti F 1995 *Found. Phys.* **25** 5
- [8] Bell J S 2004 *Speakable and Unsayable in Quantum Mechanics* 2nd edn (Cambridge: Cambridge University Press) p 166
- [9] Dürr D, Goldstein S and Zanghí N 2004 *J. Stat. Phys.* **116** 959
- [10] Bassi A and Salvetti D G M 2007 *J. Phys. A: Math. Theor.* **40** 9859
- [11] Gerlach W and Stern O 1922 *Z. Phys.* **9** 349
- [12] Busch P, Grabowski M and Lahti P J 1995 *Operational Quantum Physics* (Berlin: Springer)