Physics Letters A 160 (1991) 107–115
North-Holland

PHYSICS LETTERS A

Conditioning in quantum mechanics

Nicola Cufaro Petroni
Dipartimento di Fisica dell'Università and Istituto Nazionale di Fisica Nucleare, via Amendola 173, 70126 Bari, Italy

Received 9 July 1991; accepted for publication 13 September 1991
Communicated by J.P. Vigier

It is shown that Kolmogorovian probability models, like stochastic mechanics, are compatible with the nature of quantum conditioning if we admit that the classical processes of the model cannot be all observable at once.

Conditioning is the acquisition of new information. However, while in classical physics this acquisition process is always cumulative, in quantum physics it is not. In fact in the quantum world there are incompatible informations, so that when you try to put two or more of them together (namely by conditioning) you cannot always cumulate them. Often you can retain only the last (in a time sequence) information and, by doing so, you destroy the memory of the former informations. This feature, based on the uncertainty principle, is the physical reason for the different way of calculating probabilities and expectation values in quantum mechanics.

However, there is a classical approach to quantum physics where probabilities can be calculated in the usual way and that, notwithstanding, gives rise to the correct quantum predictions: stochastic mechanics. In its framework, of course, all the informations are compatible since it is a classical theory, but they are not all available at the same time for observation. Formally that means that in stochastic mechanics you have the right to do the usual classical conditioning, but not all the conditional probabilities that you can calculate are observable in the same experimental set-up (namely under the same experimental conditions). Generally speaking, only those conditional probabilities turn out to be observable which can also be calculated from the usual quantum formalism by means of the square modulus of the probability amplitudes obtained from state vectors.

Since the possibility of calculating transition probabilities is connected to the possibility of speaking of space–time trajectories for quantum particles as recently pointed out [1], the discussion on these topics has been revived [2] by the recent advances in the experiments directed to find empirical evidence for the existence of real paths even in critical situations (like two-slit experiments or other interfering devices) [3]. The aim of this Letter is to review some particular topics of this discussion in order to stress both the possibility of speaking of trajectories given by the actual researches and the danger of making too naive statements on this particularly delicate point.

Let us start by showing, by means of a very simple example, that in fact the rules for computing and combining probabilities and conditional probabilities [#1] are not unique, or *natural*, as we in general tend to believe, and that we must distinguish between a set of statistical (empirical) data and the mathematical model that we use in order to systematize them and to build a theory of the natural phenomena. It must also be stressed here that this formulation of the new probabilistic features of quantum theory does not deal with the usual terms of the debate about quantum paradoxes (non-locality, action at a distance, completeness ...) even if it can be said that the roots of all these paradoxes are here [4]. The important point is that, if there is more than one

[#1] In fact all probabilities are conditional probabilities in the sense that probabilities for every event in an experimental situation are given only after a *preparation*, i.e. a preliminary measurement followed by a selection of the outcomes, namely a conditioning.

0375-9601/91/$ 03.50 © 1991 Elsevier Science Publishers B.V. All rights reserved.

mathematical model available to organize our experimental results, we should be very careful in using the correct one in order to avoid paradoxes and errors.

Let us suppose to consider three dichotomic physical quantities X, Y, Z, taking values ± 1, and suppose that, by measuring relative frequencies, we have

$$P(X=+1)=P(X=-1)=\tfrac{1}{2}\,,$$
$$P(Y=+1)=P(Y=-1)=\tfrac{1}{2}\,,$$
$$P(Z=+1)=P(Z=-1)=\tfrac{1}{2}\,. \tag{1}$$

Of course to complete the statistical data we must measure conditional relative frequencies by selecting the subset of systems with a given value for the conditioning quantity and then by calculating the frequencies of the second quantity with respect to this preselected subset. Let us suppose now that the transition matrices are the following bistochastic matrices (the symmetry properties are supposed here only for the sake of simplicity),

$$P(X,Y)=P(Y,X)=\|P(X=i\,|\,Y=j)\|=\begin{pmatrix}\frac{2}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{2}{3}\end{pmatrix},$$
$$P(Y,Z)=P(Z,Y)=\begin{pmatrix}\frac{2}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{2}{3}\end{pmatrix},$$
$$P(Z,X)=P(X,Z)=\begin{pmatrix}\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{1}{3}\end{pmatrix}. \tag{2}$$

Eqs. (1) and (2) form a set of statistical data which shows the usual coherence properties, as for example

$$P(X=+1\,|\,Y=+1)+P(X=-1\,|\,Y=+1)=1\,,$$

and so on. It is easy to see that these data can be systematized in a very simple model: let us consider a *fair* die with the elementary outcomes all equiprobable, $\Omega=\{\omega_k\}$, $P(\omega_k)=\frac{1}{6}$, $k=1, ..., 6$, and consider the events $A=\{\omega_1, \omega_2, \omega_3\}$, $B=\{\omega_2, \omega_3, \omega_4]$, $C=\{\omega_3, \omega_4, \omega_5\}$, and the random variables $\xi=2I_A-1$, $\eta=2I_B-1$, $\zeta=2I_C-1$, where I_A is the indicator of the set A, and so on. It is immediately seen that these random variables have exactly the same statistical properties as the quantities X, Y, Z measured before. Of course, even if we accept this simple model to systmatize our measurements, we should not immediately deduce from this that *the real world behind our empirical data is a die*: at present it is only one, maybe the simplest, out of a number of possible models.

Let us suppose now that our situation be a bit more complicated since our quantities X, Y, Z satisfy (1) but have the following transition matrices,

$$P(X,Y)=P(Y,X)=\begin{pmatrix}\frac{10}{18} & \frac{8}{18}\\ \frac{8}{18} & \frac{10}{18}\end{pmatrix},$$
$$P(Y,Z)=P(Z,Y)=\begin{pmatrix}\frac{13}{18} & \frac{5}{18}\\ \frac{5}{18} & \frac{13}{18}\end{pmatrix},$$
$$P(Z,X)=P(X,Z)=\begin{pmatrix}\frac{5}{18} & \frac{13}{18}\\ \frac{13}{18} & \frac{5}{18}\end{pmatrix}. \tag{3}$$

It can be seen that now our model should be that of an *biased* die with outcomes not equiprobable: $P(\omega_1)=P(\omega_4)=\frac{8}{36}$, $P(\omega_2)=P(\omega_3)=P(\omega_5)=P(\omega_6)=\frac{5}{36}$, and with random variables ξ, η, ζ defined exactly as before.

Of course our model can be complicated in such a way that it can systematize less simple data such as that obtained from dichotomic quantities with probabilities (1) and bistochastic transition matrices

$$P(X,Y)=P(Y,X)=\begin{pmatrix}p & 1-p\\ 1-p & p\end{pmatrix},$$
$$P(Y,Z)=P(Z,Y)=\begin{pmatrix}q & 1-q\\ 1-q & q\end{pmatrix},$$
$$P(Z,X)=P(X,Z)=\begin{pmatrix}r & 1-r\\ 1-r & r\end{pmatrix}. \tag{4}$$

This is not the more general situation but it is suitable for our purposes. Indeed from the preceding discussion we could be led to think that for all values of p, q, $r\in[0, 1]$ we can find a die model, or in any case a complicated enough Kolmogorovian model [5], namely a classical probabilistic one, able to systematize our empirical data. The fact is that this is simply not true and that this situation is far from being a mathematical curiosity.

Let us suppose indeed that the measurements of our dichotomic quantities X, Y, Z give results satisfying (1) and (4) with

$$p=q=\frac{\sqrt{2}+1}{2\sqrt{2}}\,,\quad r=\frac{1}{2}\,. \tag{5}$$

It can be seen that there is no die and no other more

complicated classical Kolmogorovian model that can account for these data [5]. In fact, if such a model would exist, we should be able to build a probability space $(\Omega, \mathscr{F}, \mu)$ containing three events $A=\{X=+1\}$, $B=\{Y=+1\}$, $C=\{Z=+1\}$, such that relations like

$$P(X=+1\,|\,Y=+1)=\frac{\mu(A\cap B)}{\mu(B)}\,,$$

and similar, with all other possible combinations of our events and their complements $\bar{A}$, $\bar{B}$, $\bar{C}$ must be satisfied. This implies that the eight joint probabilities $\mu(A\cap B\cap C)$, $\mu(A\cap B\cap \bar{C})$, ..., $\mu(\bar{A}\cap\bar{B}\cap\bar{C})$ must satisfy a system of twelve equations of the following type (from the symmetry conditions in (4) we can see that $\mu(A)=\mu(B)=\mu(C)=\frac{1}{2}$),

$$P(X=+1\,|\,Y=+1)=\frac{\mu(A\cap B)}{\mu(B)}$$
$$=2[\mu(A\cap B\cap C)+\mu(A\cap B\cap\bar{C})]\,,$$

and similar. But this system not always admits a solution, and in particular for the statistical data (5) this solution does not exist. Hence no coherent Kolmogorovian model can be built to accommodate our measurements.

However, we could wonder, why should we worry about a situation that seems to be only a mathematical game showing that in some unfortunate cases we cannot build a Kolmogorovian model. The fact is that these unfortunate cases are neither uncommon nor physically irrelevant. For instance our example has been elaborated according to the statistical data that we would have obtained from the measurements of the following quantum spin observables, $\xi=\hat{\alpha}\cdot\boldsymbol{S}$, $\eta=\hat{\beta}\cdot\boldsymbol{S}$, $\zeta=\hat{\gamma}\cdot\boldsymbol{S}$, where $\boldsymbol{S}=(\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices, the unit vectors $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$ are defined respectively by the polar angles $(\frac{1}{2}\pi, 0)$, $(\frac{1}{4}\pi, 0)$, $(0, 0)$, and the initial state is an eigenstate of σ_y. Of course a non-Kolmogorovian probability model in Hilbert space for this set of experimental data is very well known and can be easily built. By paraphrasing a very famous statement, we could say that "God does not play dice" (or any other Kolmogorovian game) in cases like this, but in a sense very different from that supposed by the author of the original sentence.

More general necessary and sufficient conditions on the so-called *statistical invariants* that must be satisfied by p, q, r of (4) in order to assure the existence of the different probability models have been elaborated long time ago [5] and we refer to the literature on the foundation of quantum probability for a discussion of that problem. However, we want to remark that up to now a theory of these statistical invariants exists only for a few simple cases, and that in particular the Bell inequalities and their generalizations [6] can be seen as necessary conditions on a set of statistical data for the existence of Kolmogorovian models.

This preliminary discussion puts in evidence the fact that it is always very dangerous to approach these problems from an aprioristically Kolmogorovian standpoint [7], since that could lead directly to paradoxical situations. For instance we could look at the Feynman position on the two-slit experiment [8]: if we denote with C the event corresponding to the preparation of the particles; with A the event corresponding to the arrival of a particle on a given region of the screen; B_1 and B_2 respectively the events corresponding to the passage of the particle through the slits 1 and 2; and if we uncritically apply the ideas of the Kolmogorovian probability model, we conclude that one should have

$$P(A\,|\,C)=P(B_1\,|\,C)P(A\,|\,B_1\cap C)$$
$$+P(B_2\,|\,C)P(A\,|\,B_2\cap C)\,. \tag{6}$$

But the experiments say that there are interference terms so that

$$P(A\,|\,C)\neq P(B_1\,|\,C)P(A\,|\,B_1\cap C)$$
$$+P(B_2\,|\,C)P(A\,|\,B_2\cap C)\,. \tag{7}$$

Since Feynman claims that (6) *must* be true if we suppose that (even when nobody looks at them) the particle passes through 1 or 2, his conclusion is that, since (6) is not true "... it is not true that the electron passes either through hole 1 or through hole 2 ..." [8]. The fact not considered by Feynman in drawing this conclusion is that it is possible to show immediately that the empirical data $P(A\,|\,C)$, $P(A\,|\,B_1\cap C)$, $P(A\,|\,B_2\cap C)$ simply do not admit a Kolmogorovian model [7], so that there is no need whatsoever to declare *meaningless* the statement "... B had some value ... whenever we make no attempt to measure B ..." [8]. And, as a further consequence, there is no

need to say that when a particle travels between two positions we cannot think of it as being somewhere with a given probability, at intermediate times. In other words there is no need to rule out the possibility of speaking of particle trajectories in space and time, a problem that will be discussed later.

We want to emphasize here that a crucial point to understand the difference between classical and quantum probabilistic models lies in clarifying the differences between the two ways of adding new information by conditioning. It is already clear from the preceding examples that we should be very careful in discussing particular empirical sets of conditional probabilities. Here we will remember some other particular features of quantum conditioning.

(I) Let us consider in a classical probabilistic model on a probability space $(\Omega, \mathscr{F}, P)$ an event A such that $P(\mathrm{A})=0$ and let us ask if conditioning can modify the likelihood of the realization of A. If B is a conditioning event with $P(\mathrm{B})\neq 0$, we can immediately see that even $P(\mathrm{A}|\mathrm{B})=0$. Indeed in a Kolmogorovian model $P(\mathrm{A}|\mathrm{B})=P(\mathrm{A}\cap\mathrm{B})/P(\mathrm{B})$ and, since $\mathrm{A}\cap\mathrm{B}\subseteq\mathrm{A}$, we have $0\leqslant P(\mathrm{A}\cap\mathrm{B})\leqslant P(\mathrm{A})=0$, namely $P(\mathrm{A}\cap\mathrm{B})=0$. Hence no conditioning can change the fact that the event A happens with probability zero. A similar conclusion can be drawn for events which happen with probability one. The situation in quantum mechanics is completely different: let us show it with a simple example. Let $\mathscr{H}$ be a Hilbert space, X, Y two non-compatible observables with non-commuting operators $\hat{X}$, $\hat{Y}$ and x_1, y_1 two non-degenerate normalized eigenvectors respectively for the eigenvalues ξ_1, η_1 of $\hat{X}$, $\hat{Y}$. Let us suppose also that

$$|(x_1, y_1)|^2\neq 0 \text{ or } 1 . \tag{8}$$

We initially prepare our state (preparation $\mathscr{U}$) by measuring many times X and by selecting the systems which give ξ_1 as a result. We know that after this preparation

$$P(X=\xi_1 \mid \mathscr{U})=|(x_1, x_1)|^2=1 ,$$

$$P(X\neq\xi_1 \mid \mathscr{U})=1-P(X=\xi_1 \mid \mathscr{U})=0 .$$

We now want to add the information that Y has the value η_1 by carrying on a subsequent preparation $\mathscr{V}$ consisting of a measurement of Y on our preselected set $\mathscr{U}$ of systems and in a second selection of the systems giving η_1 as a result; because of (8) we are sure that the ensemble of the final systems is not empty. However, again because of (8), we have now

$$P(X=\xi_1 \mid \mathscr{V}, \mathscr{U})=P(X=\xi_1 \mid \mathscr{V})=|(x_1, y_1)|^2\neq 1 ,$$

$$P(X\neq\xi_1 \mid \mathscr{V}, \mathscr{U})=1-P(X=\xi_1 \mid \mathscr{V}, \mathscr{U})\neq 0 ,$$

namely the former probability-zero event becomes possible and the former probability-one event becomes less than certain because of our second conditioning $\mathscr{V}$.

(II) This particular form of interaction among quantum informations is also put in evidence by an analysis of the entropy behaviour. When in a Kolmogorovian case the measure of a quantity X has only several possible outcomes ξ_k with probabilities p_k, the average entropy of this scheme in bits of information is defined as

$$H=\sum_k p_k \log_2 p_k ,$$

and it can be shown that every subsequent conditioning can only lessen this quantity [9]. This is not always the case for a quantum system. The effect of a preparation $\mathscr{U}$ is represented here by a statistical operator $\hat{U}$ so that the entropy of our ensemble is proportional to $H(\mathscr{U})=\mathrm{Tr}(\hat{U}\ln\hat{U})$ [10]. When we condition our measurements by performing the two preparations $\mathscr{U}$, $\mathscr{V}$ in sequence we have $H_{\mathscr{V}}(\mathscr{U})=H(\mathscr{V})$ so that [#2]:

(i) if the two preparations lead to pure states $\hat{U}=\hat{P}_x$, $\hat{V}=\hat{P}_y$, where $\hat{P}_x$, $\hat{P}_y$ are the projections of two vectors x, y, the entropy remains constant since for pure states we have $H(\mathscr{U})=H(\mathscr{V})=0$;

(ii) if on the contrary $\hat{U}=\hat{P}_x$ is a pure state, but $\mathscr{V}$ is a mixture ($\mathscr{V}$ could be for example a measurement of an observable $\hat{Y}$, for which x is not an eigenstate, plus a selection giving rise to a mixture and not to a pure state) the entropy increases since $H(\mathscr{V})\geqslant 0=H(\mathscr{U})$;

(iii) if finally $\mathscr{U}$ is a mixture and $\mathscr{V}$ is a pure state the final entropy is lessened by the second preparation since $H(\mathscr{V})=0\leqslant H(\mathscr{U})$.

[#2] A simple quantum mechanical measurement always increases the entropy [10]; however, it must be recalled that a preparation is not only a measurement, but a measurement followed by a selection of the outcomes so that we could keep the entropy constant or even let it decrease as it is shown in the following.

The crucial difference with the classical case is in the fact that in general the second preparation cancels every memory of the previous informations. While in the classical case it is possible to say that the information contained in a probability measure $P(\ |\mathscr{U})$ is conserved when we add more information in $P(\ |\mathscr{V}, \mathscr{U})$ by conditioning, in the quantum case we always have $P(\ |\mathscr{V}, \mathscr{U}) = P(\ |\mathscr{V})$. This looks like a Markov property, but we must always remember that this analogy can in fact be misleading since we must distinguish the case when the acquisition of new information alters previous information (quantum) from the case in which this does not happen, even if the initial information becomes irrelevant (classical) [7].

(III) Let us thirdly remark that in the quantum probabilistic case, while a conditioning is always possible, it can be meaningless to make a joint statement. In the Kolmogorovian models we are used to think that joint probabilities and conditional probabilities are strictly related concepts; we even know that an elementary definition of conditional probability is inherently based on the concept of joint probability. This is not the case in the non-Kolmogorovian models, as we already know since the previous proof of the non-existence of Kolmogorovian models for particular sets of statistical data was based on remarks on the non-existence of a coherent set of joint probabilities. We will devote a subsequent paper to a deeper analysis of these questions, but here we want to remind of some simple facts. In a Hilbert space model the events are closed subspaces M of the Hilbert space $\mathscr{H}$ and the relative projectors $\hat{P}_{\mathrm{M}}$ play the role of the indicators of subsets in a classical probability space. However, there are formal and substantial differences between these two descriptions. For instance it is always possible to calculate both joint and conditional probabilities for classical events; but, given two projectors $\hat{P}_{\mathrm{M}}$, $\hat{P}_{\mathrm{N}}$ on $\mathscr{H}$, even if we can always calculate the conditional probability of the event M given the event N as $P(\mathrm{M}|\mathrm{N}) = \mathrm{Tr}(\hat{P}_{\mathrm{M}}\hat{P}_{\mathrm{N}})$, it has no meaning to speak of the joint probability of these events unless

$$[\hat{P}_{\mathrm{M}}, \hat{P}_{\mathrm{N}}] = 0\,, \tag{9}$$

which is verified only in particular cases. In fact if we try to calculate the joint probability as $P(\mathrm{M}, \mathrm{N}) = \mathrm{Tr}(\hat{U}\hat{P}_{\mathrm{M}}\hat{P}_{\mathrm{N}})$ for a given initial preparation $\mathscr{U}$ we must immediately say that the definition is not coherent since $\hat{P}_{\mathrm{M}}\hat{P}_{\mathrm{N}}$ is neither a self-adjoint operator nor a projection unless (9) is verified and hence our joint probability will take in general complex values. If we substitute $\hat{P}_{\mathrm{M}}\hat{P}_{\mathrm{N}}$ with $\frac{1}{2}\{\hat{P}_{\mathrm{M}}, \hat{P}_{\mathrm{N}}\}$ we get a self-adjoint operator which is not a projector so that our would-be joint probability will be real but its values will not always belong to [0, 1]. Then we could think to use the definition $P(\mathrm{M}, \mathrm{N}) = P(\mathrm{MN}) = \mathrm{Tr}(\hat{U}\hat{P}_{\mathrm{MN}})$, where now $\hat{P}_{\mathrm{MN}}$ is the projector on the intersection subspace MN. But even in this case we would have disturbing properties; for example, since the intersection is not distributive with respect to the linear sum of two closed subspaces, namely since in general $\mathrm{MN} + \mathrm{M}\bar{\mathrm{N}} \neq \mathrm{M}(\mathrm{N} + \bar{\mathrm{N}}) = \mathrm{M}$, we will have $P(\mathrm{M}, \mathrm{N}) + P(\mathrm{M}, \bar{\mathrm{N}}) \neq P(\mathrm{M})$, where $\bar{\mathrm{N}}$ is the orthogonal complement of N. Finally we could hope to use the conditional probability to define the joint probability as $P(\mathrm{M}, \mathrm{N}) = P(\mathrm{M}|\mathrm{N})\,P(\mathrm{N})$. However, even this last definition is not coherent since the theorem of total probability does not hold in general in quantum mechanics, as (7) and the previous examples have shown, and hence we will have again

$$P(\mathrm{M}, \mathrm{N}) + P(\mathrm{M}, \bar{\mathrm{N}}) = P(\mathrm{M}|\mathrm{N})P(\mathrm{N}) + P(\mathrm{M}|\bar{\mathrm{N}})P(\bar{\mathrm{N}}) \neq P(\mathrm{M})\,.$$

One could wonder at this point, why are we ready to accept the fact that a set of probabilities does not verify the theorem of total probability, namely

$$P(\mathrm{M}|\mathrm{N})P(\mathrm{N}) + P(\mathrm{M}|\bar{\mathrm{N}})P(\bar{\mathrm{N}}) \neq P(\mathrm{M})\,, \tag{10}$$

as we did in (7), but cannot accept that

$$P(\mathrm{M}, \mathrm{N}) + P(\mathrm{M}, \bar{\mathrm{N}}) \neq P(\mathrm{M})\,. \tag{11}$$

The difference lies in the fact that in (11) the probabilities are all calculated with respect to the same probability measure (we are always in the same preparation $\mathscr{U}$) so that this inequality appears as a logical incoherence [#3]; on the contrary in (10) the probabilities are calculated with respect to two different preparations, $\mathscr{U}$, $\mathscr{V}$ so that this inequality only indicates the impossibility of collecting all these data in a unique classical probability space, namely it in-

[#3] N and $\bar{\mathrm{N}}$ are a partition of our event space and there is no third possibility beyond them, so that in no coherent frequentistic interpretation our probability can verify (11).

dicates the incompatibility of the two preparations.

Let us finally come back to a discussion of the two-slit experiment in the light of all these observations on the conditioning in quantum mechanics. It has been remarked in recent papers [1,2] that an interpretation of the experimental results in terms of space–time trajectories is in fact not forbidden (as even here recalled in the discussion about Feynman's argument). However, it must be emphasized that a correct appreciation of the subtleties of this problem can be achieved only if the discussion is done by taking into account all the constraints that the quantum empirical results impose on the mathematical probability models. In this sense we think that in the light of previous remarks it is rather misleading to say simply that the interference term of $P(\mathrm{A}|\mathrm{C})$ in (7) "... has no direct physical significance in terms of the particle properties and it also lacks meaning in the framework of probability theory" [2], or to state without further specifications that the standard Bohrian statistical interpretation "... deviates from the general spirit used in standard mathematical probability theory, where the main task is to construct the space of elementary events which happen under the same conditions under which the event considered happens" [2]. Even to show that relation (6) can always be *forced* on the experimental data [2] could lead to the only half true idea that a Kolmogorovian model of the quantum mechanical empirical results is at hand and that it should only be picked up by everyone who wants to see it. In fact that situation looks not so linear: we agree with Bozić and Marić on the fact that the quantum probabilistic model deviates sharply from the usual classical one (and we have devoted some discussion to elucidate the need to do that); we agree also on the fact that a Kolmogorovian model for the quantum results is always possible; however, we think that something more has to be said about this model and that this can be said only if we build it not only by hand in some particular cases, as in the two-slit experiment, but in the general framework of a systematic theory.

This theory in fact exists: stochastic mechanics [11] simulates all the quantum mechanical results in the framework of a completely classical probabilistic formalism and we will devote some final remarks to discuss how this is possible without contradictions with what has been argued until now on the need of a non-Kolmogorovian model of quantum probability. We simply refer to the existing literature [12] for the details of the formulation, here adopted, of this model on the basis of a stochastic variational principle, and we will limit ourselves only to say that the quantum behaviour of a scalar non-relativistic particle can be simulated by means of a stochastic process $\boldsymbol{\xi}(t)$, describing the evolution of the particle position, which is a solution of a stochastic differential equation of the form

$$\mathrm{d}\boldsymbol{\xi}(t)=\boldsymbol{v}_{(+)}(\boldsymbol{\xi}(t),t)\,\mathrm{d}t+\mathrm{d}\boldsymbol{\beta}(t)\,, \tag{12}$$

where $\boldsymbol{\beta}(t)$ is a Wiener process with diffusion constant $\nu=\hbar/2m$. Here $\boldsymbol{\xi}(t)$ does not contain all the information about the state of our system, but is only a sort of configurational variable, the rest of the information being stored elsewhere. Indeed stochastic mechanics introduces the so-called *forward and backward derivatives* of $F(\boldsymbol{\xi}(t),t)$, where $F(\boldsymbol{r}(t),t)$ is an arbitrary regular function, as

$$(\mathrm{D}_{(\pm)}F)(\boldsymbol{r},t)=\pm\lim_{\Delta t\downarrow 0}\frac{1}{\Delta t}E(\Delta^{(\pm)}F|\boldsymbol{\xi}(t)=\boldsymbol{r})\,,$$

where $\Delta^{(\pm)}F=F(\boldsymbol{\xi}(t\pm\Delta t),t\pm\Delta t)-F(\boldsymbol{\xi}(t),t)$, and hence the *forward and backward velocities* $\boldsymbol{v}_{(\pm)}(\boldsymbol{r},t)=(\mathrm{D}_{(\pm)}\boldsymbol{\xi})(\boldsymbol{r},t)$, and takes both $\boldsymbol{\xi}(t)$ and $\boldsymbol{v}_{(+)}(\boldsymbol{r},t)$ as the dynamical variables of our problem in the sense that $\boldsymbol{v}_{(+)}$ will not be given a priori but will be determined, as a part of the problem, by means of a *stochastic variational principle* which will select the physically realized (*measurable*) processes among all the possible (*virtual*) processes described by (12). By doing so we will have a model which is a stochastic control theory [13] and that, by means of a suitable choice of the Lagrangian [12] that we cannot discuss here, leads directly to a perfect reproduction of the quantum results.

We can also look at this procedure from another standpoint: starting from the Schrödinger equation

$$\mathrm{i}\hbar\partial_t\psi(\boldsymbol{r},t)=\left(-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r},t)\right)\psi(\boldsymbol{r},t)\,, \tag{13}$$

stochastic mechanics associates a different process to every $\psi(\boldsymbol{r},t)$ in the following way: from the decomposition $\psi(\boldsymbol{r},t)=R(\boldsymbol{r},t)\exp[\mathrm{i}S(\boldsymbol{r},t)/\hbar]$ we can calculate the forward velocity as

$$\boldsymbol{v}_{(+)}(\boldsymbol{r},t)=2\nu\nabla W_{(+)}(\boldsymbol{r},t)\,, \tag{14}$$

where $W_{(+)}(\boldsymbol{r}, t) = \ln\{R(\boldsymbol{r}, t)\exp[\mathrm{i}S(\boldsymbol{r}, t)/\hbar]\}$, we can separate (13) into its real and imaginary parts

$$\partial_t R^2 + \nabla\cdot\left(R^2\frac{\nabla S}{m}\right) = 0\,,$$

$$\partial_t S + \frac{m}{2}\left(\frac{\nabla S}{m}\right)^2 - \frac{\hbar^2}{2m}\frac{\nabla^2 R}{R} + V = 0\,,$$

and finally we can cast the continuity equation in the form of a forward Fokker–Planck equation for the density of our process,

$$\partial_t \rho = -\nabla\cdot(\rho \boldsymbol{v}_{(+)}) + \nu\nabla^2\rho\,. \tag{15}$$

However, we must remember that now (15) is not a Fokker–Planck equation in the usual sense since $\boldsymbol{v}_{(+)}$, as remarked before, is not an a priori given function but depends on the solution ρ of (15) through (14). In fact, if we fix a solution ψ of (13), the form of (15) (namely $\boldsymbol{v}_{(+)}$) will also be fixed. However, (15) has an infinity of solutions ρ, and among them only one satisfies the stochastic variational principle. This solution verifies $\rho = R^2 = |\psi|^2$, so that it corresponds to the probability density of the extremal process and can be considered as selected through the initial condition $\rho_i(\boldsymbol{r}) = R(\boldsymbol{r}, t_i)$. Of course other solutions corresponding to different initial conditions are formally available: for example there are transition probabilities $p(\boldsymbol{r}, t; \boldsymbol{r}_i, t_i)$ solutions of (15) for an initial condition like

$$\lim_{t\downarrow t_i} p(\boldsymbol{r}, t; \boldsymbol{r}_i, t_i) = \delta^3(\boldsymbol{r} - \boldsymbol{r}_i)\,,$$

but, if the initial density does not coincide with the $R(\boldsymbol{r}, t_i)$ of our given wave function ψ, these solutions of (15) are in some sense virtual since they do not satisfy the stochastic variational principle, so that we cannot associate to them a direct meaning of physical observability. Of course these solutions do not correspond to the square modulus of a wave function solution of (13).

Let us suppose now that we have a screen, with two holes located in $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, between a source of particles and a detector located in $\boldsymbol{r}$, and let us ask for the probability densities of the detected particles in the three situations given in table 1.

Table 1

Situation	Hole 1	Hole 2
"1"	open	closed
"2"	closed	open
"1, 2"	open	open

From the initial conditions

$$\psi^{(1)}(\boldsymbol{r}, t_i) = \delta^3(\boldsymbol{r} - \boldsymbol{r}_1)\,,$$

$$\psi^{(2)}(\boldsymbol{r}, t_i) = \delta^3(\boldsymbol{r} - \boldsymbol{r}_2)\,,$$

$$\psi^{(1,2)}(\boldsymbol{r}, t_i) = c_1\delta^3(\boldsymbol{r} - \boldsymbol{r}_1) + c_2\delta^3(\boldsymbol{r} - \boldsymbol{r}_2)\,,$$

(where $|c_1|^2 + |c_2|^2 = 1$) we can calculate the wave functions $\psi^{(1)}(\boldsymbol{r}, t)$, $\psi^{(2)}(\boldsymbol{r}, t)$, $\psi^{(1,2)}(\boldsymbol{r}, t)$, respectively for the three proposed situations, and we can verify directly that the probability densities do not add (a relation perfectly analogous to (7)),

$$|\psi^{(1,2)}(\boldsymbol{r}, t)|^2 \neq |c_1|^2|\psi^{(1)}(\boldsymbol{r}, t)|^2 + |c_2|^2|\psi^{(2)}(\boldsymbol{r}, t)|^2\,, \tag{16}$$

the difference between the two sides being, of course, in the interference terms. However, from the wave equation (13) we can determine the Fokker–Planck equations corresponding to every situation. In the cases "1" and "2" the transition probabilities are

$$p^{(1)}(\boldsymbol{r}, t; \boldsymbol{r}_1, t_i) = |\psi^{(1)}(\boldsymbol{r}, t)|^2\,,$$

$$p^{(2)}(\boldsymbol{r}, t; \boldsymbol{r}_2, t_i) = |\psi^{(2)}(\boldsymbol{r}, t)|^2\,. \tag{17}$$

As for the situation "1, 2" the solution selected by quantum mechanics (or equivalently by the stochastic variational principle) for the corresponding Fokker–Planck equation is $|\psi^{(1,2)}|^2$ giving the probability density on the screen (interference) when initially the state is $\psi_i^{(1,2)}(\boldsymbol{r})$. Of course in our scheme $|\psi^{(1,2)}|^2$ is the probability distribution function of a stochastic process and hence it follows trajectories in the space–time. But if we try to consider this probability distribution function as the statistical superposition of observable transition probabilities, we get into trouble. In fact we can calculate the solutions $p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{t}_1, t_i)$ and $p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{r}_2, t_i)$ of (15) in the case "1, 2", namely the transition probability densities respectively either from "1" or from "2" to the screen *when two holes are open*, but we should bear in mind that now

$$p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{r}_1, t_i) \neq p^{(1)}(\boldsymbol{r}, t; \boldsymbol{r}_1, t_i)\,,$$

$$p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{r}_2, t_i) \neq p^{(2)}(\boldsymbol{r}, t; \boldsymbol{r}_2, t_i)\,, \tag{18}$$

and that, differently from $p^{(1)}(\boldsymbol{r}, t; \boldsymbol{r}_1, t_i)$ and $p^{(2)}(\boldsymbol{r}, t; \boldsymbol{r}_2, t_i)$ (which correspond to processes that are extremal separately in the situations "1" and "2"), the transition probabilities $p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{r}_1, t_i)$ and $p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{r}_2, t_i)$ cannot be calculated as the square modulus of a wave function (since they do not correspond to processes that are extremal in the situation "1, 2"); namely that they are the probability densities of virtual processes. Of course, since $|\psi^{(1,2)}(\boldsymbol{r}, t)|^2$, $p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{r}_1, t_i)$ and $p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{r}_2, t_i)$ are solutions of the same classical Fokker–Planck equation, and classical probabilities superpose as usual, we will have

$$|\psi^{(1,2)}(\boldsymbol{r}, t)|^2 = |c_1|^2 p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{r}_1, t_i) + |c_2|^2 p^{(1,2)}(\boldsymbol{r}, t; \boldsymbol{r}_2, t_i) , \tag{19}$$

a relation analogous to (6), and this result is not in contradiction with (16) because we have to take into account (17) and (18) [14]. It has also been shown that all these transition probability densities can be calculated by means of path integrals [1,15] if all the processes (measurable and virtual) are taken into account. In this way every transition from an initial to a final position can always be built in a probabilistic way by taking into account all the possible trajectories with real, positive statistical weights, but our trajectories can also be virtual. In other words: in the process corresponding to the interfering wave function $\psi^{(1,2)}(\boldsymbol{r}, t)$ the trajectories going to $\boldsymbol{r}$ either from $\boldsymbol{r}_1$ or $\boldsymbol{r}_2$ are endowed with a well-defined probability and we can *calculate* the mathematical transition probability densities satisfying (19); but if we want to observe these paths and *measure* the corresponding transition probability densities we must perform a *physical conditioning* (namely we must measure and select) in a way that changes the wave functions and the corresponding extremal process. In doing so we obtain the measurable transition probability densities (17), but, since the second preparation is not compatible (in the sense discussed in the first part of this Letter) with the initial one, we have (16) and (18) and the interference disappears.

As a consequence if we want to describe the quantum interference effect by means of space–time trajectories we get the following ambiguous situation: the interference pattern can always be considered as the sum of suitable transition probability densities for particles coming either from hole "1" or from hole "2", but only in the sense of the virtual processes and of the mathematical conditioning. On the one hand we can always calculate suitable classical conditional probabilities and use them in a classical way, even in interference experiments; on the other hand we should remember that these processes are virtual; namely not directly physically measurable: if we try to observe them they get modified by the new preparation in such a way that the interference disappears. Indeed it is a characteristic of stochastic mechanics [16] to allow one to calculate not only all the results of quantum mechanics (as in the stochastic interpretation of quantum mechanics [17]), but also some extra quantities (as the transition probabilities for virtual processes) which, however, seem to be not directly observable.

The author wants to thank Professors L. Accardi, F. Guerra and J.P. Vigier for invaluable discussions and suggestions.

References

[1] N. Cufaro Petroni, Phys. Lett. A 141 (1989) 370.
[2] M. Bozić and Z. Marić, Phys. Lett. A 158 (1991) 33.
[3] G. Badurek, H. Rauch and D. Tuppinger, Phys. Rev. A 34 (1986) 2600;
H. Rauch and J.P. Vigier, Phys. Lett. A 151 (1990) 269;
H. Rauch and J.P. Vigier, Possible utilization of a Mach–Zehnder interferometer to test Einstein's Welcherweg assumption with single photon sources (Paris, May 1991), preprint.
[4] L. Accardi, The probabilistic roots of the quantum mechanical paradoxes, in: The wave–particle dualism, eds. S. Diner et al. (Reidel, Dordrecht, 1984).
[5] L. Accardi and A. Fedullo, Lett. Nuovo Cimento 34 (1982) 161.
[6] J.S. Bell, Physics 1 (1965) 195;
D. Gutkowski and G. Masotto, Nuovo Cimento 22B (1974) 121;
C. Corleo, D. Gutkowski, G. Masotto and M.V. Valdes, Nuovo Cimento 25B (1975) 413;
A. Garuccio and F. Selleri, Found. Phys. 10 (1980) 209.
[7] L. Accardi, Non-Kolmogorovian probabilistic models and quantum theory (Rome, 1985), preprint.
[8] R.P. Feynman, Rev. Mod. Phys. 20 (1948) 367.
[9] A.I. Khinchin, Mathematical foundations of information theory (Dover, New York, 1957).
[10] J. von Neumann, Mathematical foundations of quantum mechanics (Princeton Univ. Press, Princeton, 1955).

[11] E. Nelson, Dynamical theories of Brownian motion (Princeton Univ. Press, Princeton, 1967); Quantum fluctuations (Princeton Univ. Press, Princeton, 1985); F. Guerra, Phys. Rep. 77 (1981) 263.
[12] F. Guerra and L.M. Morato, Phys. Rev. D 27 (1983) 1774; F. Guerra and R. Marra, Phys. Rev. D 28 (1983) 1916; 29 (1984) 1647.
[13] W.H. Fleming and R. Rishel, Deterministic and stochastic optimal control (Springer, Berlin, 1975).
[14] F. Guerra, Probability and quantum mechanics. The conceptual foundations of stochastic mechanics, in: Quantum probability and applications to the quantum theory of the irreversible processes, eds. L. Accardi et al. (Springer, Berlin, 1984).
[15] F.W. Wiegel, Physica 33 (1967) 734; 37 (1967) 105; M.S. Wang, Phys. Rev. A 38 (1988) 5401.
[16] F. Guerra, Meccanica stocastica (Rome, 1986), preprint.
[17] D. Bohm and J.P. Vigier, Phys. Rev. 96 (1954) 208; J.P. Vigier, Lett. Nuovo Cimento 24 (1979) 258, 265.