NON-KOLMOGOROVIAN PROBABILISTIC MODELS AND QUANTUM THEORY

LUIGI ACCARDI
Princeton University
Department of Statistics*

February 1985

SUMMARY
(after the bibliography)

Invited talk at the 45-th ISI session.
Amsterdam, Netherlands August 12-15, 1985

* Permanent address: Universita di Roma II
Dipartimento di Matematica
Via Orazio Raimondo
00173 Rome, Italy

1.) Statement of the problems.

The sixth, among the twenty-three problems stated by Hilbert at the Paris international congress of Mathematicians in 1900 was [25] "...to treat axiomatically those physical disciplines in which already todays mathematics plays a predominant role...these ones are in the first place the calculus of probability and mechanics..."

Twentyfive years after Hilbert's Paris address, the new quantum mechanics was discovered and it was soon clear that this theory was as much a new mechanics as a new probability calculus. Hilbert himself was immediately aware of these features of the new theory as well as of the changes of perspective that this discovery caused on his sixth Paris problem. In fact he devoted a seminar (in 1926) to the description of the mathematical structure of quantum theory, in which the non-classical features of the new probabilistic calculus were clearly described. The notes of this seminar, collected by von Neumann and Nordheim, were later published in a joint paper [24]. Thus the Kolmogorov axiomatization of the classical probabilistic model came to light in a time (1933) when already the most advanced physical theories were making extensive use of a completely different mathematical formalism. The challenge posed to all probabilists by this new probability calculus was pointed out in Feynman's communication in the 2nd Berkely Symposium on Probability and Statistics (1954) [19].

In fact, from a theoretical point of view, it would be very unsatisfactory for contemporary probability theory to be unable to answer questions as the following: imagine you want to handle a set of statistical data (say- transition probabilities, or correlation functions,...). Which type of probabilistic models will you try to fit your data? The one based on the usual Kolmogorovian calculus or the one based on the new calculus used by the quantum physicists?

Of course from an empirical point of view there is no problem: if your data come from some experiments on - say - populations biology, then you use the classical model; if they come from CERN, then you use the quantum one. (Some people are perfectly happy with this level of understanding).

Another natural question is: is the new probabilistic model really necessary? or maybe is it the result of an historical misunderstanding, and the new physics can be completely described within the context of the classical probabilistic model? There is a large and subtle literature on the problem. The theories which try to fit the statistical data of quantum mechanics within a classical probabilistic model, are called "hidden variables" theories. We will not discuss them here.

But assume that the new probabilistic model is not superfluous and has an intrinsic necessity. Then, since the qualitative mathematical features of the classical probabilistic calculus are entirely determined by its axioms, it follows that the new calculus must be the expression of different axioms, and in particular that some axiom of the classical probability calculus must be false in the new one. In other words the Kolmogorovian probabilistic model should have its "parallel axiom", like the euclidean geometric model. Once the analogy between euclidean and non euclidean geometries and Kolmogorovian and non-Kolmogorovian models is accepted one can go further, and look for "statistical invariants". By analogy with "geometrical invariants" these would provide a rigorous mathematical distinction between the different models, rather than an empirical one, based on the fact that "it works". Moreover, one has to single out the axioms which lead to the new probabilistic model or, to say it again with Hilbert's words: "...to formulate the physical requirements so completely that the mathematical formalism becomes uniquely determined by them..."[24].

In my talk I will report on some results obtained in the last years which have lead to a complete clarification of the fundamental relations between the classical and quantum probabilistic models and which have opened the way to the construction of an entirely new class of probabilistic models whose mathematical structure has revealed some unsuspected connection between probability and geometry (c.f.the author's paper in [13]). Finally, the inner development of the quantum probabilistic models has given rise to a whole crop of new results concerning quantum Markov chains [1],[2],[3],[12], quantum central limit theorems [15], [26],[36], the subtle technical and conceptual problems connected with the notion of quantum conditional expectation [8], the quantum Feynman-Kac formula [7], quantum infinitely divisible processes [35], and the newly developed quantum stochastic calculus [32],[11],[27],[28],[9],[10]. The whole body of these results constitutes a new and rapidly developing branch of probability theory: QUANTUM PROBABILITY. Unfortunately I will not be able here even to mention some of the results listed above. The interested reader is referred to the two volumes [13],[14] which present a fairly complete description of the state of the art of quantum probability. In the present talk I will limit myself to discussing the necessity of enlarging the horizons of our discipline keeping into account the stimuli which arose from quantum physics, but whose perspectives go far beyond it. In this respect it is a pleasure for me to thank the organizers, and in particular M. Keane, for giving me the opportunity to expose these ideas at the 45-th session of the ISI. I want also to thank G. Watson for several discussions which lead to formal and substantial improvements of the present paper.

2. The axioms of probability theory.

For the discussion which follows it will be of the utmost importance to distinguish the probabilistic (or physical) notions, such as "events", "probabilities", "conditional probabilities", "observables", "states", ... from their representative in the mathematical model. In fact, following the example of geometry, our strategy to distinguish mathematically between different probabilistic models will be first to show that the possibility of describing a certain family of empirically given and model indendent data (e.g. angles in geometry; transition probabilities or correlations in probability) within a given probabilistic model imposes some constraints on these data. And then to produce examples of such experimentally given statistical data which do not fulfill the constraints of the classical probabilistic model. We assume as primitive the notion of "event" and denote them by capital letters A,B,C,.... The probability of an event A given that an event C is known to happen is denoted P(A|C). Even if not necessary for the analysis which follows, it will be convenient to keep in mind the physical interpretation and to think of the probability P(A|C) as the approximate relative frequency of the event A in an ensemble of systems prepared so that the event C is certainly verified for each of them. The physical counterpart of the probabilistic notion of "conditioning" is that of "preparation "

The axioms underlying any known probabilistic model can be subdivided into five groups:

I). Normalization axioms.
II). Structure axioms.
III). The finite additivity axiom.
IV). Continuity axioms.
V). Conditioning axioms.

In the debate on the foundations of classical probability, the continuity axioms have been questioned, while in the debate on the foundations of quantum theory each of the first three groups of axioms have been questioned by different authors. The fifth group of axioms doesn't seem to have been explicitly mentioned in the literature. It will constitute the main topic of our analysis.

The normalization axioms are common to all known probabilistic models:

I.1) For any pair of events A, C, one has:

$$0 \leq P(A|C) \leq 1$$

I.2) $P(C|C) = 1$

The structure axioms describe the structure of the family of events in whose probabilities we are interested. Here the difference between the classical and the non-classical (quantum) context begins to appear. In fact structure axiom common to all classical probabilistic models is:

II.C) The family of events, whose probabilities are considered, is closed under the action of the logical connectives (conjunction, disjunction, negation).

Here, when referring the logical connectives to events, we are implicitly identifying an event A with the proposition "the event A happens". By Stone's representation theorem, Postulate (II.C) uniquely determines the mathematical model of the family of events as a boolean algebra of sub-sets of a given set. Kolmogorov used a stronger version of (II.C), namely:

II.C1.) (II.C) holds and the family of events is closed under countable disjunctions.

However the Heisenberg principle forced us to admit that in many physically meaningful situations, the family of events satisfies the following "negative postulate":

II.Q) There exist pairs of events whose conjunction (or disjunction) does not represent events whose physical realizability can be checked experimentally.

In other terms: there are in nature pairs of events A,B such that the probabilities $P(A \cap B)$, $P(A \cup B)$ cannot be given an experimental meaning. Such pairs of events will be called <u>incompatible</u>, otherwise we say they are compatible. The existence of incompatible events does not imply, per se, the impossibility of using the classical probabilistic model. In fact, following a commonly used

procedure, one might introduce, in the mathematical formalism, some expressions (such as $P(A \cap B)$ for incompatible A,B) which have no experimental meaning but which are used only in intermediate steps - the final result being expressed only in terms of experimentally measurable quantities. There have been a large number of attempts to turn the "negative postulate" (II.Q) into a positive statement which either fixes uniquely the model of the family of events (as axiom (II.C) does for the classical model), or at least isolates a class of reasonable models for this family. From this line of research the interesting mathematical theory of Quantum Logics arose, which is a structure theory for a wide class of non-distributive lattices (e.g.[17]), and which, through Gleason's theorem [31], is connected to quantum probability, roughly as classical measure theory is connected to integration theory.

The postulate of finite additivity is also common to all known probabilistic models. We formulate it as follows:

III). If two events A,B are mutually exclusive then

$$P(A \cup B) = P(A) + P(B)$$

Two events A,B are called mutually exclusive if they are compatible and if the fact that A is verified implies that B is not and conversely. The prototype of continuity axiom is the axiom of countable additivity. We will not insist on this group of axioms here, since our goal is to show that the occurrence of non-Komogorovian models can be proved even in the most elementary probabilistic setting: when only a finite number of events is involved.

By a "conditioning axiom" we mean a rule which provides an answer to questions of the following type: assume taht, knowing that a certain event C is realized, we evaluate the probability of the event A by the number P(A|C). How shall the probability of A be evaluated if our information is changed by the knowledge of the fact that also the event B has happened? The answer, given by classical probability to this question, is provided by the "Bayes formula" which we formulate as an axiom:

V.1) If A,B,C are three events such that A is compatible with B and B with C, and if $P(B|C) \neq 0$, then

$$P(A|B \cap C) = \frac{P(A \cap B|C)}{P(B|C)}$$

In the probabilistic model this formula is usually taken as the definition of conditional probability, meaning by this that the right hand side defines the left hand side. There are however cases in which the right and the left hand side of Bayes formula can be measured independently, i.e. by different physical experiments. For example we can imagine an apparatus which produces a statistical ensemble of particles for which the event C is surely verified and which then measures the relative frequencies of the events B and $A \cap B$ in this ensemble. But we can also imagine another apparatus which prepares a statistical ensemble of particles for which both the events C and B are surely verified, and which then measures the relative frequencies of the event A. The former apparatus corresponds to the right hand side of Bayes formula, the latter to the left hand side. Since the two experimental situations are different, the equality of the two sides in Bayes formula in this case is certainly not a question of definitions, but of experiments.

3.) Experiments to test the validity of Bayes' formula.

In order to relate the two probabilities $P(A|C)$ and $P(A|B\cap C)$, it is important to take into account how the information on B has been acquired; in particular to distinguish the case when the acquisition of new information does not alter the previous information, from the case when it does. When speaking of the probability $P(A|B\cap C)$ in a classical context, it is always implicitly assumed that the new information is acquired without destroying the old. But quantum physicists have taught us that there are cases in which this cannot in principle be achieved. An idealized example of this situation might be the following. A biologist wants to study the correlation between the field of view of a given population of animals with the chemical and the electrical activity of the cells of their eyes. Assume he has succeeded in preparing a sample population in such a way that the chemical activity of the eyes is known (event C). It might well be that, in order to acquire some information on the electrical activities of the cells (event B), he will alter in an apriori uncontrollable way their chemical activity. Then the information obtained by performing in series the two experiments, is not $B\cap C$ but only B. We express this statement with the equality:

$$P(A|B\cap C) = P(A|B)$$

The fact that in some cases the process of acquisition of information is not cumulative has an interesting consequence. Consider two mutually exclusive events B and B' which "exhaust all possibilities", i.e. such that in any measurement done to ascertain which of the two events B and B' happens, one finds that at least one (and only one) of them is realized. In such a situation, according to classical probability, the event $B\cup B'$ is a trivial one, i.e. it happens with probability 1 and therefore the knowledge that it happened does not alter our information.

However, keeping in mind that the acquisition of information on B and B' might have altered some previous information we had on the system, we are induced to distinguish the case (denoted $B\cup B'$) in which no experiment has been done to discriminate between the two alternatives, from the case (denoted $B\vee B'$) in which an experiment to discriminate between the two alternatives has been done, but we do not know the result of it. Clearly, in the case when no experiment has been done one has:

$$(P(A|(B\cup B')\cap C) = P(A|C) \qquad (3.1)$$

because in this case we are merely stating a tautological fact, and neither the physical situation nor our information has changed.

In the latter case however, even if we do not know the result of the experiment, at least we know that it was done, namely that at a certain moment the particles of our experimentally prepared ensemble interacted with an apparatus (possibly being disturbed by this interaction). We can thus say that the act of measuring B and B' has split the original ensemble into two: one for which $B\cap C$ is verified, the other one for which $B'\cap C$ is verified. Since there are no other possibilities, it follows that, in any experiment done to verify the occurencies of A, after the experiment on the alternatives B,B', one must have:

$$N(A|(B\vee B')\cap C) = N(A|B\cap C) + N(A|B'\cap C) = \qquad (3.2)$$
$$\frac{N(A|B\cap C)}{N(B|C)} \cdot N(B|C) + \frac{N(A|B'\cap C)}{N(B'|C)} \cdot N(B|C)$$

where $N(A|(B\vee B')\cap C)$ is the number of occurencies of A in the original ensemble, $N(A|B\cap C)$ is the number of occurencies of A among the particles for which $B\cap C$ was known to happen, and $N(B|C)$, $N(B'|C)$, $N(A|B'\cap C)$ are defined in a similar way. Thus, identifying probabilities and relative frequencies:

$$P(A|(B\vee B')\cap C) = P(A|B\cap C)\cdot P(B|C) + P(A|B'\cap C)\cdot P(B'|C) \quad (3.3)$$

Equation (3.3) is not a postulate. It follows only from the identification of probabilities with relative frequencies and from the assumption that all the particles of the initial ensemble can be split into two mutually exclusive classes: one for which the event B is verified, another in which the event B' is verified.

The crucial fact about the identity (3.3) is that it involves only experimentally measurable quatities. This gives us the possibility to device a simple experiment to test the range of applicability of the classical probabilistic model.

In fact, assume that the conditional probabilities $P(A|B\cap C)$, $P(A|B'\cap C)$ can be described according to the prescription of Bayes formula i.e.:

$$P(A|B\cap C) = \frac{P(A\cap B|C)}{P(B|C)}; \quad P(A|B'\cap C) = \frac{P(A\cap B'|C)}{P(B'|C)} \qquad (3.4)$$

In this case, inserting (3.4) into (3.3), we should find:

$$P(A|(B\vee B')\cap C) = \frac{P(A\cap B|C)}{P(B|C)} \cdot P(B|C) + \frac{P(A\cap B'|C)}{P(B'|C)} \cdot P(B'|C) =$$
$$= P(A|C) \qquad (3.5)$$

(Note that, for the validity of (3.5), it is not required that the probabilities $P(A\cap B|C)$ and $P(A\cap B'|C)$ be experimentally measurable quantities: it is sufficient that two such numbers exist, that they belong to the interval [0,1], and that the axiom of finite additivity is satisfied).

It follows that, if we can produce experimentally realizable examples of events A,B,B',C for which:

$$P(A|(B\vee B')\cap C) \neq P(A|C) \qquad (3.6)$$

then the conditional probabilities $P(A|B\cap C)$ and $P(B|C)$ cannot be related by the Bayes formula (3.4) for whatever choice of the probabilities $P(A\cap B|C)$, $P(A\cap B'|C)$. This would therefore imply that the set of statistical data $P(A|C)$, $P(A|B\cap C)$, $P(A|B'\cap C)$, $P(B|C)$, $P(B'|C)$ cannot be described by any classical probabilistic (or Kolmogorovian) model.

In the next section we describe an experiment which has been effectively done and which has confirmed the inequality (3.6).

4.) The two-slit experiment.

This is the experiment discussed by Feynman in his talk to the II-d Berkeley Symposium on probability and statistics [19] and which, according to Feynman [21] "...is formulated so to include all the mysteries of quantum mechanics...". A source S emits identically prepared particles (nowadays this experiment is done using neutrons) towards a screen Σ_1 with two slits on it, denoted 1 and 2. The particles which pass the screen Σ_1 are collected on another screen Σ_2, parallel to Σ_1 and one measures the relative frequency of the particles which hit a small region X of the screen Σ_2. This measurement is done in three different physical situations: (i) with both slits 1 and 2 open; (ii) only slit 1 open; (iii) only slit 2 open. Denoting C - the event corresponding to the common preparation of the electrons emitted by the source S (say - a fixed value of the energy); A - the event that an electron falls in the region X of the screen Σ_2; B - the event that the electron passes through slit 1; and B' - the event that the electron passes through slit 2, the following probabilities can be experimentally measured (by approximating them with relative frequencies): $P(A|C)$, $P(A|B\cap C)$, $P(A|B'\cap C)$, $P(A|(B\vee B')\cap C)$, $P(B|C)$, $P(B'|C)$ (where $B\vee B'$ is defined in section(3.)). Applying implicitly Bayes' formula, Feynman concludes that one should have:

$$P(A|C) = P(B|C)\cdot P(A|B\cap C)+P(B'|C)\cdot P(A|B'\cap C) \qquad (4.1)$$

But the experiments give that, for any choice of $P(B|C)$ and $P(B'|C)$, one has:

$$P(A|C) \neq P(B|C)\cdot P(A|B\cap C)+P(B'|C)\cdot P(A|B'\cap C) \qquad (4.2)$$

Moreover, in agreement with the theoretical analysis of section (3.), the experiments give:

$$P(A|(B\vee B')\cap C) = P(B|C)\cdot P(A|B\cap C)+P(B'|C)\cdot P(A|B'\cap C) \qquad (4.3)$$

That is: if an experiment is done to discriminate through which of the two slits the particles pass, then the theorem of composite probability holds. According to the analysis of section (3.), the inequality (4.2) only means that the statistical data $P(A|C)$, $P(A|B\cap C)$, $P(A|B'\cap C)$ cannot be described by a Kolmogorovian model. This however is not Feynman's conclusion. Feynman claims that the identity (4.1) must be true as soon as we accept that, even when nobody is looking at them, the particles which reach the screen Σ_2 either have passed through slit 1 or through slit 2, but not both. Explicitly he says ([19], pg. 538) "...we concluded on logical bases that, since (4.2) is not true [Feynman's numbering of the formulas, as well as his notations, are different from ours], it is not true that the electron passes either through hole 1 or through hole 2...". Our analysis in section (3.) shows that this conclusion is not justified on probabilistic grounds. Let us now show that this conclusion is not justified on mathematical grounds. Feynman's implicit assumption in claiming that if B and B' are disjoint events then (4.1) must hold, is the validity of Bayes' formula; i.e. the existence of four positive numbers (not necessarily experimentally measurable): $P(B|C)$, $P(B'|C)$, $P(A\cap B|C)$, $P(A\cap B'|C)$ such that:

$$P(B|C) + P(B'|C) = 1 \qquad (4.4)$$
$$P(A\cap B|C) + P(A\cap B'|C) = P(A|C)$$
$$\frac{P(A\cap B|C)}{P(B|C)} = P(A|B\cap C)$$
$$\frac{P(A\cap B'|C)}{P(B'|C)} = P(A|B'\cap C)$$

Thus Feynman's implicit assumption is equivalent to postulating that the above system of four equations in four unknowns always admits a solution or equivalently, with a simple computation, that:

$$0 < \frac{P(A|C) - P(A|B'\cap C)}{P(A|B\cap C) - P(A|B'\cap C)} < 1 \qquad (4.5)$$

But why should it be evident "...on a logical grounds..." that if the events B and B' are mutually exclusive then the probabilities $P(A|C)$, $P(A|B\cap C)$, $P(A|B'\cap C)$, which have been measured in three completely different physical situations, should satisfy the inequalities (4.5)?

Therefore the first answer provided by quantum probability to Feynman's analysis is: the inequality (4.2) means that the three probabilities $P(A|C)$, $P(A|B\cap C)$, $P(A|B'\cap C)$ do not admit a Kolmogorovian model and no mysterious properties of the particles involved are needed to justify it (cf. the discussion in section (10) below).

The next "mystery" of quantum mechanics is the correct substitute for the therem of composite probabilities (4.1). The rules, found by the physicists, say that you have to deal with complex numbers $\psi(A|C)$, $\psi(B|C)$,... (called conditional probability amplitudes) and related to the probabilities by:

$$P(X|Y) = |\psi(X|Y)|^2; \; X,Y \text{ - events} \qquad (4.6)$$

and that these amplitudes satisfy the analogue of the relation (4.1), i.e.

$$\psi(A|C) = \psi(B|C)\cdot\psi(A|B\cap C) + \psi(B'|C)\cdot\psi(A|B'\cap C) \qquad (4.7)$$

This relation is called the "theorem of composite amplitudes" and it is an empirical fact that, in the two slits experiment, it leads to results in agreement with the experiment.

Can the strange probabilistic calculus based on the above prescriptions be in some sense explained starting from some more fundamental prescription with an immediate probabilistic or physical meaning? Or shall we accept it as a "truth of nature", something not reducible to more fundamental requirements, and which can only be used, but not explained? Sections (8.) and (9.) below are devoted to answering this question.

5.) A more recent example: Bell's inequality.

The two slits experiment, discussed in the previous section shows how, since the early days of quantum theory, one came across some simple sets of statistical data which could not be described by any Kolmogorovian model. More recently J.S.J. Bell has discussed another example of this situation different from those considered in the early days of quantum theory, in which the statistical data were some transition probabilities. In Bell's example the statistical data which do not admit a description within the Kolmogorovian model are a set of correlation functions. The following theorem gives a proof of the celebrated

"Bell's inequality" and of some equivalent formulations of it, which have been discussed in the literature (§. [22]).

Theorem (5.1) Let A,B,C,D denote four random variables on a probability space $(\Omega, ,P)$ with values in $\{+1,-1\}$. Then the following, equivalent, inequalities are satisfied:

$$|E(A \cdot B) - E(B \cdot C)| \leq 1 - E(A \cdot C) \tag{5.1}$$

$$|E(A \cdot B) - E(B \cdot C)| \leq 1 + E(A \cdot C) \tag{5.2}$$

$$|E(A \cdot B) - E(B \cdot C)| + |E(A \cdot D) - E(D \cdot C)| \leq 2 \tag{5.3}$$

Notation For any function $X: \Omega \longrightarrow \mathbb{R}$, we use notation

$$\int_\Omega X \, dP = E(X) \tag{5.4}$$

Proof. Clearly, since $B^2 = |A \cdot B| = 1$, one has:

$$|E(A \cdot B) - E(B \cdot C)| \leq E(|AB - BC|) = E(|1 - AC|) = 1 - E(A \cdot C)$$

and this proves (5.1). Substituting $-B$ for B and $-C$ for C, the inequalities (5.1) and (5.2) are interchanged, hence they are equivalent. Substituting A for D in (5.3) gives (5.1). Finally, if (5.1) (hence also (5.2)) hold, then substituting in (5.2) D for B and $-C$ for C one obtains:

$$|E(A \cdot D) + E(D \cdot C)| \leq 1 + E(A \cdot C)$$

and adding this inequality to (5.1) gives back (5.3).

Corollary (5.2) There exist triples a,b,c of unit vectors in $\mathbb{R}^3$ for which it is not possible to find six random variables $S_x^j (x = a,b,c \; ; \; j = 1,2)$ on some probability space $(\Omega,\mathfrak{F},P)$ with values in $\{-1,+1\}$ such that their correlations are given by:

$$E(S_x^1 \cdot S_y^2) = -x \cdot y \; ; \; x,y = a,b,c \tag{5.5}$$

(where, for x,y $\mathbb{R}^3$, $x \cdot y$ denotes the euclidean scalar product.)

Proof. A consequence of (5.5) is that:

$$E(S_x^1 \cdot S_x^2) = -x \cdot x = -\|x\|^2 = -1 \; ; \; x = a,b,c$$

and, since $|S_x^1 S_x^2| = 1$ this is possible if and only if $S_x^1 = -S_x^2$ $(x = a,b,c)$. Thus, using the inequality (5.1), we obtain:

$$|E(S_a^1 S_b^2) - E(S_b^2 S_c^1)| \leq 1 - E(S_a^1 S_c^1) = 1 + E(S_a^1 S_c^2) \, .$$

Or equivalently, using (5.5):

$$|a \cdot b - b \cdot c| \leq 1 - a \cdot c \tag{5.6}$$

Thus if we choose the three vectors a,b,c to be coplanar and such that a is perpendicular to b and c lies between a and b, forming an angle θ with a, then the inequality (5.6) becomes:

$$\sin\theta + \cos\theta \leq 1 \; ; \; 0 \leq \theta \leq \pi/2 \tag{5.7}$$

But the maximum of the function $\theta \mapsto \sin\theta + \cos\theta$ in $[0,\pi/2]$ is $\sqrt{2}$ (obtained for $\theta = \pi/4$). So, for θ near to $\pi/4$, the left hand side of (5.7) will be near to $\sqrt{2} > 1$. Therefore, for such a choice of θ, the triple of unit vectors a,b,c will not satisfy the inequality (5.6) hence, by Theorem (5.1) it cannot admit any Kolmogorovian model.

Corollary (5.3) There exist triples a,b,c of unit vectors in $\mathbb{R}^3$ for which it is not possible to find three random variables S_a,S_b,S_c on a probability space $(\Omega,\mathfrak{F},P)$ with values in $\{+1,-1\}$ such that their correlations are given by:

$$E(S_x \cdot S_y) = x \cdot y \; ; \; x,y = a,b,c \tag{5.8}$$

Proof. As in Corollary (5.2)

The conclusion is exactly the same as in the two-slits experiment: you start with a set of statistical data and you show that they do not admit a Kolmogorovian model. The interest of the corollaries above lies in the fact that statistical correlations of this kind can be obtained from a small set of related experiments.

Let us now describe two simple non-Kolmogorovian models which can account for the correlations described in Corollaries (5.2) and (5.3).

A Non-Kolmogorovian model for the correlations of Corollary (5.3). Let M be the algebra of all 2×2 matrices; denote $\tau: (a_{ij})_{i,j} \longrightarrow 1/2(a_{11} + a_{22})$ -the normalized trace; consider the vector space basis of M given by the identity matrix and:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \; ; \; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \; ; \; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.9}$$

For each unit vector $a = (a_1,a_2,a_3)$ $\mathbb{R}^3$, define

$$S_a = a \cdot \sigma = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 \tag{5.10}$$

then S_a is hermitian with eigenvalues ± 1 and if a,b are two unit vectors then

$$\tau(S_a S_b) = a \cdot b \tag{5.11}$$

Thus, if we represent the random variables as the 2×2 matrices (5.10), their values by the eigenvalues of the corresponding matrices, and the correlation function of two observables by the trace of the product of the corresponding matrices, we obtain a (non Kolmogorovian) model by which we can describe the correlations (5.8) not only for three given unit vectors, but for any triple of unit vectors.

A Non-Kolmogorovian model for the correlations of Corollary (5.3)

Denote H -the Hilbert space $\mathbb{C}^2$ with the usual scalar product defined with respect to the basis:

$$\psi_+ = (0,1) \; ; \; \psi_- = (1,0) \tag{5.12}$$

For a unit vector a in $\mathbb{R}^3$, let S_a be defined by (5.9), (5.10) and define

$$S_a^1 = S_a \otimes 1 \; ; \; S_a^2 = 1 \otimes S_a \tag{5.13}$$

i.e. the (4×4) matrices S_a^1, S_a^2 act on $H \otimes H$ (1 -is the identity 2×2 matrix, in (5.13)). Let now σ denote the vector in $H \otimes H$ defined by:

$$\sigma = 1/\sqrt{2}(\psi_+ \otimes \psi_- - \psi_- \otimes \psi_+) \tag{5.14}$$

One then easily verifies that

$$\langle \sigma, (S_a^1 \otimes S_b^2)\sigma \rangle = - a \cdot b \tag{5.15}$$

Therefore, if we represent the observables S_a^1, S_b^2 (a,b -unit vectors) as in (5.13), and their correlations by the left hand side of (5.15), again we obtain a (non-Kolmogorovian) model of the correlations (5.5) for any triple of unit vectors a,b,c $\mathbb{R}^3$.

Remark 1.) The term "correlations" referred to expressions such as (5.11) or (5.15), is justified by the fact that it can be shown ([3]) that the values assumed by each observable S_x must be equiprobable (if they admit at all a Kolmogorovian model with the quantum mechanical transition probabilities). So they have mean zero, and therefore by expressing the left hand sides of (5.11), (5.15) in terms of transition probabilties (using the formalism described in section (10) below), one finds an expression which is formally identical to the expression for the correlations in the classical Kolmogorovian models. This is a general fact for pair correlations. Already for triple correlations (e.g. $E(S_a S_b S_c)$) this will be true only under special statistical assumptions on the observables.

Remark 2.) One could hardly believe that such trivial remarks as Theorem (5.1) and its Corollary (5.2) gave rise to a really huge literature. In section (11) I have tried to explain in a clear and condensed way why so many discussions about "non-locality" and other mysterious and never clearly explained notions (such as "collapse of the wave packet", "physical superpositions of states", "non-separability",---) could attract the attention of distinguished scientists.

6.) The statistical invariants.

Let T be a set and, for each $x \in T$, let $A(x)$ be an observable quantity whose values will be denoted $a_1(x), ---, a_n(x)$ ($n < +\infty$, for all x). We assume that the transition probabilities

$$P(A(y) = a_j(y) | A(x) = a_i(x)) = P_{ij}(x,y) \qquad (6.1)$$

are given for each $x,y \in T$; $i,j = 1, ---, n$ (they are considered to be the experimentally given statistical data). The matrix $(P_{ij}(x,y))$ will be denoted $P(x,y)$.

Definition (6.1) The family of transition probability matrices $\{P(x,y) : x,y \in T\}$ is said to admit a Kolmogorovian model if there exist:

- A probability space (Ω, F, P)
- For each $x \in T$ a measurable partition $A_1(x), \dots, A_n(x)$ of Ω such that for each $x,y \in T$ and each $i,j = 1, \dots, n$:

$$P_{ij}(x,y) = \frac{P((A_i) \cap A_j(y))}{P(A_i(x))} \qquad (6.2)$$

Definition (6.2) The family of transition probability matrices $\{P(x,y) : x,y \in T\}$ is said to admit a complex n-dimensional Hilbert space model if there exist:

- a complex n-dimensional Hilbert space H.
- for each $x \in T$, an orthonormal basis $\phi_1(x), \dots, \phi_n(x)$ of H such that, for each $x,y \in T$ and each $i,j = 1, \dots, n$:

$$P_{ij}(x,y) = |\langle \phi_i(x), \phi_j(y) \rangle|^2 \qquad (6.3)$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in H. In a similar way one defines the notion of real or quaternion Hilbert space model.

The notion of "statistical invariant" is illustrated by the following theorem (concerning the particular case in which $n = 2$ and T is a set containing three elements).

Theorem (6.3) Let p,q,r $(0,1)$ be three given numbers, and let

$$P = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix} ; \quad Q = \begin{pmatrix} q & 1-q \\ 1-q & q \end{pmatrix} ; \quad R = \begin{pmatrix} r & 1-r \\ 1-r & r \end{pmatrix} \qquad (6.4)$$

be three bi-stochastic matrices. Then:

i) A Kolmogorovian model for P,Q,R exists if and only if
$$|p + q - 1| \leq r \leq 1 - |p - q| \qquad (6.5)$$

ii) A complex Hilbert space model for P,Q,R exists if and only if:
$$-1 \leq \frac{p + q + r - 1}{2\sqrt{pqr}} \leq 1 \qquad (6.6)$$

iii) A real Hilbert space for P,Q,R, exists if and only if:
$$\frac{p + q + r - 1}{2\sqrt{pqr}} = \pm 1 \qquad (6.7)$$

iv) A quaternion Hilbert space model for P,Q,R exists if and only if a complex Hilbert space model exists.

Theorem (6.3) above (§. [4] for a proof) suggests the following general definition: a set of statistical invariants for the family of transition probabilities $\{P(x,y) : x,y \in T\}$ with respect to a given probabilistic model is defined by:

- a family $(F_i)_{i \in I}$ of real valued functions depending on the elements $P_{ij}(x,y)$ of the matrices $P(x,y)$.
- a family $(B_i)_{i \in I}$ of subsets of $\mathbb{R}$ such that the conditions

$$F_k(\{P_{ij}(x,y)\}_{i,j,x,y}) \in B_k \; ; \forall k \in I \qquad (6.8)$$

are necessary and sufficient for the existence of the given model.

Similarly to what happens for the geometrical, or topological invariants, there is no general algorithm to calculate the statistical invariants of a set of transition probability matrices with respect to a given probabilistic model. Several particular cases have been worked out [13].

Even with these limitations, Theorem (6.3) provides a general answer to another of the "mysteries" of quantum theory, to which a large leterature has been devoted since the early sixities (cf. [4] for references) i.e. : why just the complex numbers? It is clear from (6.5), (6.6), (6.7) that not only the non-Kolmogorovianity of the model, but also the discrimination between real and complex numbers is built into the statistical data.

Let us conclude with an historical remark: an old theorem, due to von Neumann, states that within the Hilbert space model n self-adjoint operators admit a joint distribution in any quantum state if and only if they commute. There have been many beautiful generalizations of this theorem to structures more general than operators in Hilbert space - The theory of statistical invariants provides for the first time a model independent Solution of the problem: the existence or not of a Kolmogorovian model can be now decided (at least in principle) uniquely in terms of the statistical data.

7.) The danger of mixing the two models: the so-called quantum Zeno paradox.

Let A be an observable and let A_t denote the observable A at time t. Under certain assumptions which we will assume to be fulfilled without making them explicit, quantum theory predicts that for any $n \in \mathbb{N}$, for any set $a_0, a_1, ---, a_{n+2}$ of values of A and for any $t_0 < t_1 < t_2 < --- < t_n < s < t$, one has:

$$\begin{aligned} P(A_t = a_{n+2} | (A_s = a_{n+1}) \cap (A_{t_n} = a_n) \cap --- \cap (A_{t_0} = a_0)) \qquad (7.1) \\ = P(A_t = a_{n+2} | A_s = a_{n+1}) \end{aligned}$$

i.e., the family $(A_t)_{t \geq 0}$ behaves like a Markov process. We say "like", because from the analysis in section (3.) it is clear that we should distinguish the case when acquisition of information on A_t alters previous information acquired on $A_s (s < t)$, from the case in which this does not happen. Now, suppose that we uncritically apply the rules of the Kolmogorovian model to the identity (7.1). Then we necessarily find the familiar identity for Markov chains, i.e. letting $a_1 = a_2 = --- = a$, and assuming that the transition probabilities are stationary:

$$\begin{aligned} & P([A_t = a] \cap [A_{t_n} = a] \cap --- \cap [A_{t_1} = a] | [A_{t_0} = a]) \qquad (7.2) \\ & = P(A_t = a | A_{t_n} = a) \cdot P(A_{t_n} = a | A_{t_{n-1}} = a) \cdot --- \cdot P(A_{t_1} = a | A_{t_0} = a) \\ & = P(A_{t-t_n} = a | A_0 = a) \cdot P(A_{t_n - t_{n-1}} = a | A_0 = a) \cdot --- \cdot P(A_{t_1 - t_0} = a | A_0 = a) \end{aligned}$$

Thus, if $t_{j+1} - t_j = t/n$ ($j = 0, ---, n$; $t_{n+1} = t$), then:

$$\begin{aligned} & P([A_t = a] \cap [A_{t_n} = a] \cap --- \cap [A_{t_1} = a] | [A_{t_0} = a]) \qquad (7.3) \\ & = [P(A_{t/n} = a | A_0 = a)]^n \end{aligned}$$

Now, it happens that the von Neumann theory of the quantum measurement process, as generalized first by Zumino and Lüders and then by Wigner, leads exactly to the identity (7.3), where

$$P(A_s = a | A_0 = a) = |\langle \psi_a, e^{iHs} \psi_a \rangle|^2 \qquad (7.4)$$

(ψ_a -a unit vector in a Hilbert space; H -a self-adjoint operator on that space which can be assumed to contain ψ_a in its domain). Using (7.4) to evaluate the right hand side of (7.3), one obtains:

$$P([A_t = a] \cap [A_{t_n} = a] \cap \cdots \cap [A_{t_1} = n] \,|\, A_{t_0} = a) \tag{7.5}$$

$$= |\langle \psi_a , e^{\frac{itH}{n}} \psi_a \rangle|^{2n} = \left| 1 + \frac{it}{n} \langle \psi_a , H\psi_a \rangle - \frac{t^2}{2n^2}\langle \psi_a , H^2\psi_a \rangle + \ldots \right|^{2n}$$

$$= \sim\left(1 - \frac{t^2}{n^2}\langle \psi_a , H^2\psi_a \rangle + \frac{t^2}{n^2} \langle \psi_a , H\psi_a \rangle^2 \right)^n$$

$$\sim \exp(- t^2\Delta_a(H)/n) \longrightarrow 1 \qquad ; \Delta_a(H) = \langle \psi_a , H^2\psi_a \rangle - \langle \psi_a , H\psi_a \rangle^2$$

as $n \longrightarrow \infty$.

Thus: the more we measure the observable A, the more the probability of finding the value a as a result of each measurement, approaches 1: i.e. observation prevents change states.

Hence the name quantum Zeno paradox". Rather than "paradox", we should say "contradiction with the experiments", in fact the relation (7.5) can be shown to be experimentally false.

The error here consists in the application of Bayes' formula (7.4)(in the deduction of (7.2) from (7.1)) to a set of statistical data (i.e. the transition probabilities (7.1)) which do not admit a Kolmogorovian model (§.[6] for a more detailed discussion). Let us mention, incidentally, that the theory of quantum Markov chains provides a formula for joint probabilities on the left hand side of (7.2), which is free from these incoherences [2].

8.) How to set up the new structure axioms.

In this section we look for a constructive formulation of the structure axioms which keeps into account the "negative postulate" (II.2) of section (2.). As we have shown, the notion of conditioning is crucial in order to understand the rise of non-Kolmogorovian models. In particular, the distinction between conditioning with or without alteration of previous information. Since possible alterations can occur because we act upon a system by a measurement, it is natural to set up our mathematical model as an idealization of the various operations which are present in the measurement procedures. The usual boolean-Kolmogorovian structure will be recovered as the limiting case in which the information acquired in a measurement process does not affect the information acquired by previous measurements. The language we adopt is an extension of one introduced by J. Schwinger [34] and called "the algebra of measurements". Let A be an observable with values $a_j (j=1,\ldots,n)$ and let t be an instant of time. To the triple (A,a_j,t) we associate an idealized measurement apparatus, denoted $A_j(t)$, which from an ensemble of independent similar systems selects those for which the value of the observable A at time t is a_j. Such an apparatus will be called an <u>elementary filter</u>.

The symbol $A_j(t) \cdot B_k(t)$ $(s < t)$ will be associated with the apparatus corresponding to the consecutive application, to the same ensemble, first of the filter $A_j(t)$ and then of $B_k(t)$. This operation is associative, with an identity consisting of the apparatus which does not filter away anything. Elementary filters are idealizations of the so-called first-kind measurements, meaning by this that, if ε is "very small" then each particle which passed the filter $A_j(t)$ will also pass the filter $A_j(t + \varepsilon)$ (i.e. -if the observable A has the value a_j at time $t - 0$, it will also have this value at $t + 0$). Thus elementary filters act both as measurement apparatus and as preparing apparatus. Not all measurements in nature are of first kind, but our mathematical analysis will be limited to this class. Since the measurements which do not disturb the system at all are of first kind, it follows that all the events considered in the classical theory are included in the present discussion.

Elementary filters are mutually exclusive, in the sense that if an observable A has the value a_j at time t, it cannot have any other value in a short time after t. In symbols:

$$\text{"}\lim_{\varepsilon \to 0}\text{"} A_j(t) \cdot A_k(t+\varepsilon) = \delta_{jk} A_j(t) \; ; \tag{8.1}$$

meaning by this that, if ε is small enough and $j \neq k$, then practically no partcles will emerge from the apparatus $A_j(t) \cdot A_k(t+\varepsilon)$. In (8.1), $\delta_{jk} = 0$ if $j \neq k$ and $=1$ if $j=k$, and 0 means the apparatus which does not allow any particle to pass. Two observables A,B are called <u>compatible</u> if

$$\text{"}\lim_{\varepsilon \to 0}\text{"} A_j(t) \cdot B_k(t+\varepsilon) = \text{"}\lim_{\varepsilon \to 0}\text{"} B_k(t) \cdot A_j(t+\varepsilon) \tag{8.2}$$

for any pair of values a_j of A and b_k of B. This will be written

$$A_j(t) \cdot B_k(t) = B_k(t) \cdot A_j(t) \tag{8.3}$$

The commutativity of the observables A,B means that acquisition of information on the observable A does not alter previous information acquired on B and conversely.

Another natural operation on filters is the time reversal, which will be denoted $*$ and which corresponds to the applications of the filters $A_j(s)$ and $B_k(t)$ $(s < t)$ in reversed time order:

$$[A_j(s) \cdot B_k(t)]^* = B_k(s) \cdot A_j(t) \tag{8.4}$$

$$A_j(s)^* = A_j(s) \tag{8.5}$$

If p is a number in $[0,1]$ a and $A_j(t)$, an elementary filter, the symbol $p \cdot A_j(t)$ will be associated to the apparatus with the following properties:

i) $(p \cdot A_j(t)) \cdot A_j(t) = A_j(t) \cdot (p \cdot A_j(t)) = p \cdot A_j(t)$ (8.6)

ii) from an ensemble of particles, identically prepared, at time $t-\varepsilon$, so that $A=a_j$, the apparatus $p \cdot A_j(t)$ chooses at random a fraction p of particles.

(formula (9.5) shows how to realize experimentally such apparatus). Compatible filters $A_j(t)$, $B_k(t)$ can be applied in parallel, and the resulting apparatus, denoted $A_j(t)+B_k(t)$, is characterized by the fact that it will allow those particles for which either $A=a_j$ or $B=b_k$. Finally, elementary filters corresponding to different values of the same observable A are always compatible and, moreover:

$$A_1(t) + \cdots + A_n(t) = 1 \tag{8.7}$$

By standard mathematical procedures (quotienting a free algebra by certain relations) this structure of "partial algebra" -i.e. in which addition, multiplication, and scalar multiplication are not everywhere defined - can be embedded in a real associative $*$-algebra with identity.

These algebras generalize the usual boolean structure of the Kolmogorovian model. In the following section we show that this generalization is not too wide, in the sense that in some significant cases it allows a complete classification of the new models which can arise.

9.) Schrödinger's equation as a compatibility condition for a non-Kolmogorovian model.

In this section we prove that from the physically meaningful assumptions described in section (8.) all the main mathematical features of the quantum formalism (the Schrödinger equation included) can be rigorously deduced. The problem (originally posed by D. Hilbert - cf. section (1) of the present paper) of deducing the quantum formalism from physically meaningful requirements, has been studied in the last sixty years by many authors, including von Neumann, Mackey, Segal, ... The approach considered here makes use of some standard terminology developed in the studies on the foundations of quantum theory, however, it is based on a strategy of proof which is new with respect to the existing literature.

We start from a set T (index set) and we assume that, for each $x \in T$ it is given an observable quantity $A(x)$ with values $a_1(x), \dots, a_n(x)$ ($n < +\infty$, – the same for all x). For the moment we limit ourselves to the structure axioms, so we do not introduce probabilities. The family of events in whose probabilities we will be eventually interested is

$$\{[A(x) = a_j(x)] : x \in T, \quad j = 1, \dots, n\}$$

where $[A(x) = a_j(x)]$ denotes the event that $A(x)$ takes the value $a_j(x)$. The mathematical model for this family is defined by:

- an associative real *-algebra with identity $\mathcal{A}$.
- the assignment of a correspondence which, to the event $[A(x) = a_j(x)]$ associates an element $A_j(x)$ of $\mathcal{A}$ satisfying

$$A_j(x)^* = A_j(x) \ ; \ A_j(x) \cdot A_k(x) = \delta_{jk} A_j(x) \qquad 9.1$$

$$\sum_{j=1}^{n} A_j(x) = 1 \qquad 9.2$$

- it is also required that the algebra $\mathcal{A}$ is algebraically spanned by the set of projections $\{A_j(x) : x \in T, \ j = 1, \dots, n\}$.

An algebra $\mathcal{A}$ as above will be called a <u>Schwinger algebra of measurements</u> associated to the set of observables $\{A(x) : x \in T\}$. Two elements of $\mathcal{A}$ are called compatible if they commute. Two typical examples are:

<u>Example 1.)</u> The requirement that all the observables $A(x)$ $(x \in T)$ are mutually compatible (i.e. any pair of projectors $A_j(x)$ $(j=1,\dots,n \ ; \ x \in T)$ commute) allows to associate in a canonical way a boolean algebra to the Schwinger algebra $\mathcal{A}$. Conversely every boolean algebra defines a commutative Schwinger algebra.

<u>Example 2.)</u> Let $H = \mathbb{C}^n$ with the usual hermitean product; let, for each $x \in T$, $\phi_1(x), \dots, \phi_n(x)$ be an orthonormal basis of H. To the event $[A_j(x) = a_j]$ we associate the rank one projector

$$A_j(x) : \psi \in H \to \langle \phi_j(x), \psi \rangle \cdot \phi_j(x) = A_j(x) \cdot \psi$$

($\langle \cdot , \cdot \rangle$ – the scalar product).

Clearly (9.1) and (9.2) are verified. Here the Schwinger algebra is the algebra of all $n \times n$ matrices.

The two examples considered above are extreme, in the sense that while in the former for each $x \in T$, the observable $A(x)$ is compatible with any other observable $A(y)$, in the latter we have that if an element of $\mathcal{A}$ $(= M(n, \mathbb{C}))$ commutes with each one of the $A_1(x), \dots, A_n(x)$, then it must be a linear combination of them (with complex coefficients), so that the observable $A(x)$ presents the maximal degree of compatibility. Motivated by this remark we introduce the following:

<u>Definition (9.1).</u> In the above notations the observable $A(x)$ $(x \in T)$ is called <u>maximal</u> if an element of $\mathcal{A}$ commutes with each of the $A_1(x), \dots, A_n(x)$ when and only when it is a linear combination of them with coefficients in the center of $\mathcal{A}$ (i.e. the family of elements of $\mathcal{A}$ commuting with each element of $\mathcal{A}$).

In the following the center of the Schwinger algebra $\mathcal{A}$ will be denoted κ and we will assume the validity of the following two technical conditions (they are much stronger than we need, but assuming their validity greatly simplifies the exposition):

$$X \in \mathcal{A} \ ; \quad \gamma \in \kappa \ ; \ \gamma \cdot X = 0 \ \text{ if and only if } \ \gamma = 0 \ \text{ or } \ X = 0 \qquad 9.3$$

$$\gamma \cdot A_j(x) \geq 0 \ ; \ \text{ with } \ \gamma \in \kappa; \ \text{ if and only if } \ \gamma \geq 0 \qquad 9.4$$

(in a *-algebra an element is called positive if it can be written as a sum of elements of the form b^*b).

In the classical case the same boolean structure is a model for the family of events both for a deterministic and a statistical theory. Therefore the boolean algebra model cannot be intrinsically related to a single set of probabilities. Converseley, it is intuitively clear that if we are dealing with two different maximal observables, then the acquirement of an exact information on the values of one of them implies that the information on the values of the other one cannot be but statistical. Therefore, in this case we expect a set of "privileged" probabilities intrinsically associated to the observables. It is a remarkable feature of the Schwinger algebra of measurements that they provide a precise mathematical support for this intuition:

<u>Theorem (9.2)</u> In the above notations, let $A(x)$ and $A(y)$ $(x,y \in T)$ be two maximal observables. Then there exists a bi-stochastic matrix $P(x,y) = (p_{ij}(x,y))$; $(p_{ij}(x,y) \geq 0 \ ; \ \Sigma_{i=1}^n p_{ij}(x,y) = \Sigma_{j=1}^n p_{ij}(x,y) = 1)$ such that:

$$A_i(x) \cdot A_j(y) \cdot A_i(x) = p_{ij}(x,y) \cdot A_i(x) \qquad 9.5$$

$$p_{\alpha\beta}(x,y) = p_{\beta\alpha}(y,x) \qquad 9.6$$

Thus: not only do two maximal observables in a Schwinger algebra canonically define a transition probability matrix, but this matrix has necessarily the symmetry property (9.6) which is found in the usual $\mathbb{C}$-Hilbert space model (cf. the identity (6.3)).

Now, just as Stone's theorem provides a standard mathmatical model for the boolean algebras of classical probability, we would like to have a standard representation theorem for Schwinger algebras in order to obtain a classification of all the possible models which can be obtained in that way. In the classical case (all observables are compatible) we recover all the classical models. So let us consider the opposite extreme, i.e. the case in which each observable $A(x)$ $(x \in T)$ is maximal. An important corollary of Theorem (9.2) is that the Schwinger algebra generated by a finite number of maximal observables is necessarily finite-dimensional over its centre κ. Moreover, if we add to our assumptions (9.1), (9.2) the assumption that:

$$p_{ij}(x,y) > 0 \ ; \quad x,y \in T \ ; \ i,j = 1,\dots,n \qquad 9.7$$

((9.7) will be implicitly assumed, in the following), then for any $x,y \in T$ the n^2 elements of $\mathcal{A}$, $\{A_i(x) \cdot A_j(y) : i,j = 1,\dots,n\}$ are linearly independent over the centre of $\mathcal{A}$. Thus a Schwinger algebra with each $A(x)$ maximal has at least dimension n^2 over its center if card(T) 2.

<u>Definition (9.3)</u> An <u>Heisenberg algebra</u> associated with the family of observables $\{A(x) : x \in T\}$ is a Schwinger algebra $\mathcal{A}$ associated with $\{A(x) : x \in T\}$ such that:

- each observable $A(x)$ is maximal.
- $\mathcal{A}$ has minimum dimensions over its centre.

The following problem arises now quite naturally: given a set $\{P(x,y) : x,y \in T\}$ of $n \times n$ bi-stochastic matrices, when does there exist a Schwinger algebra $\mathcal{A}$ associated with a family $\{A(x) : x \in T\}$ of n-valued observables, such that (i) each $A(x)$ is maximal in $\mathcal{A}$ (ii) for each $x,y \in T$ the bi-stochastic matrix canonically associated to the pair $A(x)$, $A(y)$ according to Theorem (9.2) is $P(x,y)$?

This theorem is the quantum probabilistic analogue of the well known classical problem: given a family of transition probability matrices $\{P(s,t) : s < t, \ s, t \in \mathbb{R}_+\}$, when does there exist a markovian process

$(A(t))$ such that for each $s < t$, the transition matrix canonically associated with the pair of random variables $A(s)$, $A(t)$ is $P(s,t)$? It is well known that the classical probabilisitic problem has a positive solution if and only if the family of transition probability matrices $(P(s,t))$ satisfies the Chapman-Kolmogorov equation:

$$P(r,s) \cdot P(s,t) = P(r,t) \; ; \; r < s < t \qquad 9.8$$

A rather surprising fact is that the quantum-probabilistic problem has a positive solution if and only if to each transition matrix $P(x,y)$ one can associate an "amplitude matrix" $U(x,y)$ so that the family $\{U(x,y):x,y \in T\}$ satisfies a generalization of the Schrödinger equation (in integral form). More precisely:

Theorem (9.4) (cf. [13]) The following assertions are equivalent:

i) There exists an Heisenberg algebra $\mathcal{A}$ with center κ satisfying the conditions of the problem stated above.

ii) For each $x,y \in T$ there exists a κ-valued matrix $U(x,y)$ such that for each $x,y,z \in T$, $i,j,k = 1,\dots,n$.

$$\Sigma_{i=1}^{n}\left(\frac{P_{ij}(x,y)}{U_{ij}(x,y)}\right) \cdot U_{ik}(x,y) = \delta_{jk} \qquad 9.9$$

$$\Sigma_{j=1}^{n}\left(\frac{P_{ij}(x,y)}{U_{ij}(x,y)}\right) \cdot U_{kj}(x,y) = \delta_{ik} \qquad 9.10$$

$$U(x,x) = 1 \qquad 9.11$$

$$U(x,y) \cdot U(y,z) = U(x,z) \qquad 9.12$$

Moreover, in this case, the Heisenberg algebra $\mathcal{A}$ can be identified with the algebra of all $n \times n$ matrices with coefficients in κ.

Remark 1.) The Hilbert space model of quantum theory is recovered when $\kappa = \mathbb{C}$ and

$$p_{ij}(x,y) = |U_{ij}(x,y)|^2 \qquad 9.13$$

one immediately recognizes in (9.9), (9.10) the conditions of unitarity of the matrix $U(x,y)$. Moreover, considering the case $T = \mathbb{R}$ and assuming a translation invariant situation, (9.11) and (9.12) become equivalent to:

$$U(s) \cdot U(t) = U(s+t) \; ; \; U(0) = 1 \qquad 9.14$$

where $(U(s))$ is a 1-parameter unitary group of $n \times n$ matrices. Denoting by H its generator, (9.14) becomes equivalent to:

$$\frac{d}{dt}\psi(t) = iH\psi(t) \; ; \; \psi(0) = \psi_0 \in \mathbb{C}^n \qquad 9.15$$

which is the usual finite-dimensional form of the Schrödinger equation.

Remark 2.) It can be shown that equation (9.12) includes also the notion of group representation and that a natural modification of it (making it "path-dependent") leads to the notion of connection in a Hilbert bundle and of the associated gauge theories.

Many problems remain open in the clarification of the structure of Schrödinger algebras of measurements and their connections to probabilistic models. Theorem (3.4) above shows however that essentially all the features of the quantum mechanical model follow from a natural and physically

meaningful idealization of the measurement procedures. It remains to determine the nature of the centre of $\mathcal{A}$. But this is done with the technique of statistical invariants described in section (6.).

The transition to observables with infinitely many (possibly continuous) values can be achieved starting from Theorem (9.4), through the application of known mathematical techniques. The idea of the procedure is the following: The problem of classifying (in the $\mathcal{N}$-dimensional $(N < +\infty)$ Hilbert space model) the pairs of observables which realize the maximal mutual indeterminacy (i.e. the transition probabilities between their values are always $1/\mathcal{N}$), leads naturally to the introduction of the finite form of the Heisenberg commutation relations (cf. the author's paper in [13]', section (2.)). It is well known that, taking the limit of these for $n \to \infty$, one obtains the Weyl form of the Heisenberg commutation relations hence, via the Stone-von Neumann uniqueness theorem, the usual $L^2(\mathbb{R},dx)$-model of quantum theory. A final remark: the definition (6.2) of "Hilbert space model" is limited to maximal observables (i.e. each of the observables considered represents a complete set of compatible observables in the physicist's terminology). In the mathematical model, this assumption is reflected through the fact that to each value of the given observable corresponds a vector in the Hilbert space or, more precisely, a rank one projection. If the observable is not maximal, then to each of its values will correspond a projection which is not of rank 1. However the probabilistic analysis of non-maximal observables is reduced to that of maximal ones through the theorem of composite probabilities, whose application is now justified by the considerations made in section (3.). The following section provides a further clarification of this point.

10.) The theorem of composite probabilities and the measurement process.

Let A, B, C be three different n-valued observables. We denote by $X(t)$ the observable X at time $t (X = A,B,C)$ and $A_a(t)$ - the event corresponding to $A(t) = a$ (a - a value of A). Similar notations will be used for B and C. The set of values of the observables will be assumed time independent. Following (6.1) we denote:

$$P_{ab}(t,s) = P(A_a(t)|B_b(s)) \qquad 10.1$$

(and similarly for BC, CA). We assume A, B, C to be maximal, in the sense that for $r < s < t$ and values a, b, c:

$$P(A_a(t)|B_b(s) \cap C_c(r)) = P(A_a(t)|B_b(s)) \qquad 10.2$$

The identity (10.2) corresponds to the identity (3.0) in section (3.) and the probability on the left hand side of (10.2) is measured as approximate relative frequency of the event $A_a(t)$ in an ensemble of particles obtained by selection of the particles for which $B(s) = b$ from an ensemble prepared so that $C(r) = c$. While the probability on the right hand side of (10.2) is evaluated as the approximate relative frequency of the event $A_a(t)$ in an ensemble of particles prepared so that $B(s) = b$. Thus the identity (10.2) should be considered as an experimental fact, and there are many physical systems in which it is satisfied. Formally (10.2) is a weak form of Markov property, but it must be kept in mind that in a classical context (10.2) means that the information on $C_c(r)$ is not necssary while in the present context - that this information has lost any value. The mathematical difference is big, in fact in the classical markovian case equality.

$$P_{ca}(r,t) = \Sigma_b P_{cb}(r,s) \cdot P_{ba}(s,t)$$

holds, while in the present case it doesn't. A selective measurement such as the one described to define the left hand side of (10.2) is often called, in the quantum mechanical literature, a complete measurement. A non selective measurement is called incomplete . According to the notation introduced in section (3.), we denote

$$V_b B_b(s) \qquad \text{(b runs over all values of B)} \qquad 10.3$$

the event corresponding to having done an incomplete measurement of $B(s)$. According to the general formula (3.3), deduced in section (3.), (extended in an obvious way to n disjoint events) and keeping into account (10.2), the probability of the event $A_a(t)$ given the preparation $C_c(r)$ and the incomplete measurement $V_b B_b(s)$, is given by:

$$P(A_a(t)\,|\,[V_b B_b(s)] \cap C_c(r)) = \qquad 10.4$$

$$= \Sigma_b P(A_a(t)\,|\,B_b(s)) \cdot P(B_b(s)\,|\,C_c(r)) =$$

$$= \Sigma_b P_{cb}(r,s) \cdot P_{ba}(s,t)$$

(which is usually different from $P_{cb}(r,t)$).
Assume now that the transition probabilities (10.1) are described by the usual quantum mechanical model. This means (according to Definition (6.2), that there exist: an Hilbert space $\mathcal{H} \equiv \mathbb{C}^n$, and for each observable $A(t)$, $B(t)$, $C(t)$, an orthonormal basis $(\phi_a^A(t))$, $(\phi_b^B(t))$, $(\phi_c^C(t))$ of $\mathcal{H}$, such that:

$$P_{ba}(s,t) = |\langle \phi_a^A(t), \phi_b^B(s)\rangle|^2 = \mathrm{Tr}(P_a^A(t) \cdot P_b^B(s)) \qquad 10.5$$

(similarly for B, C, r, s) where $\mathrm{Tr}(\,\cdot\,)$ denotes the (non normalized) trace on the space of all the bounded operators on $\mathcal{H}$ and $P_a^A(t)$ is the rank one projector:

$$P_a^A(t)\psi = \langle \phi_a^A(t), \psi\rangle\, \phi_a^A(t) \qquad 10.6$$

Inserting (10.6) into (10.2) and (10.4) respectively, one finds:

$$P(A_a(t)\,|\,B_b(s) \cap C_c(r)) = P(A_a(t)\,|\,B_b(s)) = \mathrm{Tr}(P_a^A(t) \cdot P_b^B(s)) \qquad 10.7$$

$$P(A_a(t)\,|\,[V_b B_b(s)] \cap C_c(r)) = \qquad 10.8$$

$$= \Sigma_b \mathrm{Tr}(P_a^A(t) \cdot P_b^B(s)) \cdot P_{cb}(r,s) =$$

$$= \mathrm{Tr}(P_a^A(t) \cdot W(r,s))$$

where the operator $W = W(r,s) = W_{CB}(r,s)$ is defined by:

$$W(r,s) = \Sigma_b P_{cb}(r,s) \cdot P_b^B(s) \qquad 10.9$$

and has the following properties:

W is a positive hermitean operator 10.10

$\mathrm{Tr}(W) = 1$ 10.11

In the quantum mechanical literature an operator W with these properties is called a density matrix (or a mixture, or a state). The rank one projections are particular density matrices called pure states.

Summing up: when no measurement is done at time s (or, more generally, when the particles of the ensemble can be considered as isolated in the time interval $(0,t)$) the probability of the event $A_a(t)$ shall be evaluated by:

$$P(A_a(t)\,|\,C_c(r)) = \mathrm{Tr}(P_a^A(t) \cdot P_c^C(r)) \qquad 10.12$$

When a selective measurement of $B(s)$ is performed, corresponding to the value b, we shall use formula (10.7). (It is not necessary that the particles for which $B(s) \neq b$ are materially discarded. The only relevant thing is that the relative frequency of the event $A_a(t)$ is evaluated only over the ensemble of those particles for which $B(s) = b$).
Finally, in a non selective measurement of $B(s)$, one shall use formula (10.8). The fact that in general the right hand sides of (10.8) and of (10.12) are different corresponds mathematically to the non validity of Bayes' axiom in the quantum probabilistic model, and physically to the fact that the physical situations in which the two probabilities of $A_a(t)$ are evaluated are different.
So, for example, if a friend of mine makes a selective measurement of $B(s)$ but doesn't tell me which value of $B(s)$ he selected, then I will use formula (10.8) to evaluate the probability of $A_a(t)$, while he will use (10.7). His results will be experimentally more precise than mine, which is not surprising since he has got more information.
In general I must use formula (10.8) when there is some information on the system which in principle I could obtain but which I have not. Formula (10.7) instead, is used when I dispose of the maximal amount of exact information obtainable (in agreement with the indeterminacy principle) at a given moment. There is nothing surprising in the fact that the mere possibility of obtaining in principle some information on the system alters my way of evaluating probabilities of events concerning it. This mere possibility is in fact not of a psychological nature, but corresponds to different physical and experimental conditions to which the system has been subjected. These differences in the experimental conditions of the system can arise as "distnrbances" occurring between the original preparation and the measurement of the event in question - as it is the case when we use the formula (10.4) (incomplete measurements) - as when we constrain the system to be related to another system by a conservation law (in the mathematical model of quantum theory this situation is described by a wave vector of the form (5.14)). The "physical state" of the system changes in time, in the sense that the values of all the observables of the system are functions of t. However, by the Heisenberg principle, we cannot obtain exact information on all the observables of a system, but only on a maximal observable (a complete set of compatible observables - in the physicists' terminology). Thus the knowledge of the "physical state" of the system, in the sense of the mechanicistic theories, is ruled out a priori by Heisenberg principle. In the mathematical model the maximal amount of information, attainable on a given system at a given moment s, is representated by a vector $\phi_b^B(s)$ in a Hilbert space (more precisely - by the corresponding rank one projection $P_b^B(s)$). This gives two types of information on the system:

i) that the exact value of $B(s)$ is b .
ii) that the probabilities of the values of any observable $A(t)$ $(t \geq s)$ should be evaluated according to the formula (10.7).

For reasons made explicit in section (11.) some physicists have claimed that, if the mathematical model of the state of a quantum system at time s is $\phi_b^B(s)$, then all the observables incompatible with B do not assume any value but are in a physically undefined (and in principle unobservable) "state of physical superposition". These "states" have the peculiar property of existing only if nobody tries to ascertain whether they exist or not: as soon as one tries they disappear (this is the "collapse of the wave packet" - cf.

section (11)) and one faces the familiar situation in which an observable takes one and only one of its values. But why, one might ask, physicists have introduced such a peculiar notion? It is not easy to find an answer to this question in the physical literature, because few contemporary physicists had the scientific courage of writing down explicitly the precise definition of a "state of physical superposition" and, most of all, the rational reasons which determined the introduction of such a bizarre notion. The papers of R.P. Feynman constitute a notable exception, and that is the reason why the answer to the questions arisen in those papers have been the starting point of many investigations on the relationship between probability and quantum theory. In section (11) we will give a brief outline of the reasons which lead to the notion of "physical superposition of states" (which is to be distinguished by the corresponding mathematical notion which is perfectly well defined and corresponds to the familiar operation of taking linear combinations of vectors in a complex Hilbert space).

11) <u>An historical digression on the so-called paradoxes of quantum mechanics</u>

As stated section (1), since the late twenties the physicists were aware that quantum theory was a new probabilistic calculus. The point of view advocated in the present paper is that this new caluculus reflects the non universal applicability of some of the elementary rules of classical probability (in particular, Bayes' definition of conditional probability).

This possibility has not been taken into account in the physical literature where, on the contrary, starting from the (implicit) assumption of the universal applicability of the rules of the classical probabilistic model, one was lead to the necessity of introducing some strange physical properties to justify the discrepancies between the experimental date and the theoretical predictions (deduced using these rules).

In the following we briefly outline the considerations which lead to the introduction of such strange physical notions as: non-locality, non-separability, physical superpositions, collapse of the wave packet, ... It should be underlined that the correct application of the rules of this new probability calculus do not lead to contradictions (at least not more than any other mathematical theory). It is only when the rules of the new probabilisitic calculus are uncritically mixed up with hose of the old one that contradictions arise. In section (4) we have illustrated this situation with a particular example. Now we will illustrate, in its generality, the mechanism which lead to the so called "paradoxes of quantum theory". The identity (3.5), i.e.:

$$P(A|C) = P(A|B\cap C) \cdot P(B|C) + P(A|B'\cap C) \cdot P(B'|C) \qquad 11.1$$

was considered by the physicists a tautology on relative frequencies, as long as the events B,B' are disjoint and exhaust all the possibilities. Therefore, in order to explain the fact that some experiments (as the one described in section (4.)) contradict (11.1) they were led to the conclusion that in such cases they had to do with events B, B' which were not mutually exclusive. However this conclusion also contradicted the experiments, which unequivocally showed that the events B and B' were mutually exclusive. However strange may be the way out chosen by the physicist from this empasse was the following: "since the identity (11.1) is falsified by the experiments in a situation in which no attempt is made to verify which of the alternatives B, B' was fulfilled; and since, whenever we do an experiment to verify such a thing, we find that one and only one of the events B, B' can happen at each time, then we conclude that <u>the events B and B' are mutually exclusive when we look at them but they are not mutually exclusive when we do not look at them</u>".

So, for example, if we consider an observable which, when measured, gives as result either the value +1 or the value -1, in appropriate units, but not others, then when we are not looking at this observable, we cannot say that it either assumes the value +1 or the value -1 but we do not know which one (otherwise we will find a contradiction with the experiments).

This statement constitutes the essence of what, after N. Bohr, W. Heisenberg and many others has called "the Copenhagen interpretation of quantum mechanics". Let us quote, for example, a statement taken from a famous paper of R.P. Feynman ([20], pg. 369): "... Looking at probability from a frequency point of view [the identity] (11.1) simply results from the statement that in each experiment giving A and C, $\bar{B}$ had some value. The only way (11.1) could be wrong is the statement "$\bar{B}$ had some value", must sometimes be meaningless. Noting that (4.7) replaces (11.1) only under the circumstance that we make no attempt to measure $\bar{B}$, we are led to say that the statement "$\bar{B}$ had some value", may be meaningless whenever we make no attempt to measure $\bar{B}$ ....." . (here we have slightly changed Feynman's notations: $\bar{B}$ denotes a 2-valued observable and B, B', in (11.1), are the events corresponding to the fact that $\bar{B}$ assumes one of its 2 values. Also the numeration of the formulae is different from Feynman's original one).

Feynman's discussion of this statement [19], [21], makes clear that the term "meaningless" means that, if one assumes that the observable B had some value when one does not look at it, then necessarily one arrives to a contradiction with the experiments. One might quote many statements equivalent to Feynman's, taken from articles or books by Heisenberg, Wigner, Ragge, Dyson, ... (to quote only some very well known physicists, not suspectable of having been "deviated" by an excessive interest for the foundational problems of quantum mechanics). It should also be said that not all physicists accepted the Copenhagen interpretation of quantum theory. On the contrary, the controversy about this interpretation gave rise to one of the most fascinating and subtle conceptual debates of our century which, from the stature and the number of persons involved as well as for the depth and the wealth of the scientific production it stimulated in different fields such as physics, mathematics, philosophy, logic, epistemology, ... can only be compared to the great conceptual controversies which have always accompanied major development of scientific thought (e.g. the debates concerning the notions of action at distance or at contact, the wave or particle nature of light, determinism or indeterminism, evolutionism-creationism, ...) . A well informed exposition of this debate can be found in the book by M. Jammer [29]. Those who opposed the view that observables cannot assume any of their values when nobody looks at them included among others Einstein, Schrödinger, de Broglie, ... while the "orthodox front", aligned with the Copenhagen interpretation, included the overwhelming majority of the leading physicists. The opponents of the orthodox interpretation did not succeed in finding a convincing alternative, but at least they succeeded in giving some consequences of it which contradict either some basic principles of physics (such as the principle of special relativity) or some well established experiences about the behaviour of macroscopic objects.

In fact, according to the Copenhagen interpretation, an observable "not looked at" is in a strange new type of physical state - called a superposition state - in which no one of its values is assumed, but each one is virtually present (Heisenberg, to describe such a state, used the word "potentia" as opposed to the "actum" of assuming a simple value). This "superposition state" is in principle not observable, because as soon as one does an experiment on the system to check which value the observable has, the "potentia" becomes instantaneously "actum" and the result is always a well defined value. This abrupt change from "potentia" to "actum" is called, in the physical literature, "the collapse of the wave packet". For example an electron in a definite energy state, when nobody looks at it, is virtually everywhere in the space (i.e. its position observable is in a "superposition state"). However as soon as somebody looks at it suddenly materializes in a region of a few microns. To describe this point of view Schrödinger, who opposed it, used the term "witchery". The most famous objection to this statement is known under the name of "Einstein-Podolsky-Rosen paradox" [18] and (one of the many equivalent versions of it, which is due to Heisenberg [23]) can be described as follows: an electron is in a region of space delimited by two communicating boxes. The communication between the two boxes is closed and they are separated by an arbitrarily large distance. According to the orthodox interpretation, it is wrong to say that the electron is in one and only one of the two boxes with a certain probability. The correct statement is that the electron is "virtually" spread in both boxes, but "actually" in neither of them. Now, if an experimenter opens one of the two boxes, the electron "collapses" - according to the orthodox view - into one and only one of them (not necessarily the open one), with a certain probability which is correctly predicted by quantum theory. But, Einstein, Podolsky and Rosen asked, how does the "part of the "electron virtually contained in box 2" know that box 1 has been opened? It must "know" it in a time shorter than that needed to the experimenter of box 1 to take conscience of the presence or absence of the electron in box 1; however, the two boxes can be so far apart that the time needed to any physical signal (according to the theory of special relativity) to propagate itself from box 1 to box 2 would be much greater than the "relaxation time" needed by the experimenter to ascertain whether the electron is in box 1 or not.

Thus, if the orthodox interpretation of quantum theory is correct, the principle of special relativity, according to which no physical signal can propagate itself faster than the light in the vacuum, must be wrong. One might take comfort from the fact that this "superluminal propagation" has no observable consequences. But the theoretical contradiction remains. This is the root of the notorious debate on "non-locality" in quantum theory. "Non-separability" is roughly the same thing: if instead of our electron I consider two electrons (or balls) then until I do not open the box I can't say that they have a separate physical identity: they are in a state of physical superposition" and I create their individuality by the act of opening the box: in this sense some people have spoken of "non-separability".

Many other "paradoxes" have been deviced by exploiting the notion of "collapse of the wave packet". Some of them have been taken so seriously by some physicists to lead them to abandon the dream of Boltzmann, Einstein, ... of the unification of physics under a few general principles and to try to construct two separate physical theories: one for the description of the "microscopic world" and one for the description of the "macroscopic world" (cf. [33]). A description of these "paradoxes" can be found in [29], the proof that they are all based on the inappropriate use of a single formula of classical probability theory (Bayes' formula) to deal with sets of statistical data not admitting a Kolmogorovian model is in [6]. The conclusion is that, given for granted that quantum theory represents the deepest level of our present understanding of nature, nevertheless some parts of its currently accepted interpretation should be changed. In particular the statement that the introduction of notions such as "collapse of the wave packet", "physical superposition of states", ... is necessary to avoid contradictions between theory and experiments, is mathematically wrong: the only contradiction is between experimental (statistical) data and the pretense to fit them within a single Kolmogorovian model.

One might still ask "which are the physical reasons that prevent the use of the Kolmogorovian model?". I hope that the analysis of section (3.) has

convinced the reader that the symmetric question "which are the physical reasons that suggest the use of the Kolmogorovian model?" is at least as legitimate -

BIBLIOGRAPHY

1.) Accardi, L. Non commutative Markov chains. in: Proceedings School of Math. Phys., Università di Camerino 1974.

2.) " " Non commutative Markov chains associated to a pre-assigned evolution: an application to the quantum theory of measurement. Adv. in Math. 29 (1978), 226-243.

3.) " " Topics in Quantum Probability. Physics Reports. 77 (1981) 169-192.

4.) " ", Fedullo, A. On the statistical meaning of complex numbers in quantum theory. Lett. Nuovo Cimento 34(1982)161

5.) " " Foundations of quantum probability Rend. Sem. Mat. Univ. e Poli. Di Torino (1982) 249-270.

6.) " " The probabilistic roots of the quantum mechanical paradoxes. in: The wave particle dualism, eds. S. Diner, G. Lochak, F. Selleri: Reidel (1983)

7.) " " On the quantum Feynman-Kac formula. Rendiconti del Seminario Mat. e Fis. di Milano 48 (1978) 135-179.

8.) " ", Cecchini, C. Conditional expectations in von Neumann algebras and a theorem of Takesaki. Journ. of Funct. Anal. 45 (1982) 245-273.

9.) " ", Parthasarathy, K.R. Quantum stochastic calculus. in ref. (14).

10.) " :, Quaegebeur, J. Some representation theorems in quantum stochastic calculus. To appear.

11.) " ", Frigerio, A., Lewis, J.T. Quantum Stochastic processes. Publications of the R.I.M.S. Kyoto 18 (1982) 37-133.

12.) " ", Frigerio, A. Markovian cocycles. Proc. Roy. Irish Acad. 1982.

13.) " ", Frigerio, A., Gorini, V. (eds) Quantum probability and application to the quantum theory of irreversible processes. Springer LNM N° 1055.

14.) " ", Von Waldenfels, W. (eds.) Quantum probability and Applications II. Springer LNM (to appear).

15.) " ", Bach, A. Quantum central limit theorems for strongly mixing random variables. Zeitsch. Wahrscheinlichkeintstheorie Verw. Geb. 68 (1985) 393-402.

16.) Bell, J.S. On the Einstein-Podolsky-Rosen Paradox. Physics 1 (1965) 105

17.) Beltrametti, E., Cassinelli G. The logic of Quantum Mechanics Encl. of Math., Addison-Wesley (1982).

18.) Einstein, A., Podolsky, B., Rosen, N. Can quantum mechanical description of physical reality be considered complete? Phys. Rev. 47 (1935) 777-780.

19.) Feynman, R. The concept of probability in Quantum Mechanics, Proc. II-d Berkeley Symp. on Math. Stat. and Prob., Univ. of California Press, Berkeley (1951) 533-541.

20.) " " Space-time approach to non-relativistic quantum mechanics, Rev. of Mod. Phys. 20 (1948) 367-385.

21.) " " The character of physical law (1965) (it. ed. Boringhierei 1971).

22.) Garuccio, A., Selleri, F., Systematic derivation of all the inequalities of Einstein locality. Found. of Phys. 10 (1980), 209-216.

23.) Heisenberg, W. Physics and Philosophy: the revolution in modern science. New York, Harper 1958).

24.) Hilbert, D., Nordheim, L., von Neumann, J. Über die Grundlagen der Quantenmechanik. Math. Ann. 98 (1927) 1-30.

25.) " " Mathematical developments arising from Hilbert problems. AMS, Proceedings Symp. in Pure Math. 28 (1976).

26.) Hudson, R.L., Cushen, C.D. A quantum mechanical central limit theorem. J. Appl. Prob. 8 (1971) 454-469.

27.) " ", Parthasarathy, K.R. Quantum Ito formula and stochastic evolutions. Commun. Math. Phys. 93 (1984) 301-323.

28.) " ", Lindsay, M. A noncommutative martingale representation theorem for non-Fock quantum Brownian motion. J. Funct. Anal. (to appear).

29.) Jammer, M. The philosophy of Quantum Mechanics. Wiley (1974).

30.) Jauch, J.M. Foundations of quantum mechanics. Addison-Wesley (1968).

31.) Kruszynski, P. Extensions of Gleason Theorem. in Ref. (13.).

32.) Lewis, J.T., Thomas, L.C. On the existence of a class of stationary quantum stochastic processes. Ann. Inst. H. Poincaré Sect. A22 (1975) 241-248.

33.) Ludwig, G. Foundations of Quantum Mechanics. Springer 1983.

34.) Schwinger, J. Quantum kinematics and dynamics. Academic Press 1970.

35.) Quaegebeur, J., Fannes, M. Infinitely divisible completely positive maps in Ref. (14.).

36.) von Waldenfels, W., Giri, N. An algebraic version of the central limit theorem. Z. Wahrscheinlichkeilstheorie Verw. Geb. 42 (1978) 123-134.

SUMMARY

Table of Contents:

In section (1.) one analyses the problems posed to classical probability by the existence of a probability calculus used by the physicists and quite different from the usual one: when to use on model instead of another one? Is the new model really necessary? And, if so, in what sense? Can we explain the mathematical features of the new model - which are far from having an intuitive meaning - starting from simple probabilistic assumptions? In section (3.) the notion of "conditioning" is analyzed in situations in which acquisition of new information on a system might destroy some previously acquired one. The result of the analysis is the description of possible experiments to check the universal applicability of the Kolmogorovian model.

In section (4.) one describes an effectively performed experiment leading to the conclusion that some very simple sets of statistical data of physical significance cannot be described within the classical Kolmogorovian model. Another class of such examples is described in section (5.).

In section (6.) the notion of "statistical invariant" is introduced, which allows a mathematical distinction between the classical and the quantum probabilistic model. The conclusion is that this distinction can be read into the experimentally measurable probability.

In section (7.) the conclusions of the previous analysis are used to show that a "paradox" well known in the physical literature has his roots in the inappropriate application of some rules of the Kolmogorovian model (in the case in question - Bayes' formula for the conditional probability) to a set of statistical data not admitting a Kolmogorovian model.

In section (8.) one discusses how to formulate the axioms concerning the structure of the family of events in a context where incompatible events are present (for which the usual boolean opertions are meaningless).

In section (9.) it is shown that from the new structure axioms it is possible to deduce all the essential "kinematical" features of the quantum formalism, just as from the boolean axioms of classical probability the usual probabilistic model is recovered via Stone's theorem.

In section (9.) the problems of the quantum measurement process are explained making use of the probabilistic analysis developed in section (3.).

Finally, in section (11) a (very) short description is given of the line of thought which, starting from the (implicit) assumption that the rules of classical probability had a universal application, has lead the physicists to a number of paradoxes.