A possible explanation for the probabilities of quantum mechanics

Dirk Aerts
Theoretische Natuurkunde, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
(Received 4 October 1984; accepted for publication 10 July 1985)

It is shown that a lack of knowledge about the measurements of a physical system gives rise to a nonclassical probability calculus for this physical system. It is also shown that the nonclassical probability calculus of quantum mechanics can be interpreted as being the result of a lack of knowledge about the measurements. Examples are given of macroscopic real systems that have a nonclassical probability calculus. A macroscopic real system that has a quantum probability calculus is also given, and more specifically a model for the spin of a spin-\(1/2\) particle is constructed. These results are analyzed in light of the old hidden variable problem.

I. INTRODUCTION

It is an opinion that the probabilities appearing in quantum mechanics do not arise as a consequence of our lack of knowledge, but are inherent in nature and hence ontological. This opinion was present from the advent of quantum mechanics as a consequence of one of the possible interpretations of the Heisenberg uncertainty relations. Some physicists, however, did not agree with this interpretation and wondered whether it would not be possible to imbue quantum mechanics into a classical theory. They had in mind what happened with thermodynamics. Indeed, the theory of thermodynamics is independent of classical mechanics, and has its own set of observables such as pressure, volume, temperature, energy, and entropy and its own set of states. It was, however, possible to introduce an underlying theory of classical mechanics. To do this one assumes that every thermodynamical system consists of a large number of molecules and the real pure state of the system is determined by the positions and momenta of all these molecules. A thermodynamic state of the system is then a mixed state of the underlying theory. It was possible and it was a great success to derive the laws of thermodynamics in this way from Newtonian mechanics. The theory that resulted was called statistical mechanics.

Is it possible to do something similar for quantum mechanics? Is it possible to introduce extra variables into quantum mechanics such that these extra variables would define new states, and the description of the system based on these new states and the same observables would be classical? Moreover, quantum mechanics would be the statistical theory that results by averaging over these variables. A theory that has such extra variables and a procedure for averaging over these extra variables is usually called a "hidden variable theory." Because we want to have the same observables and only new, better defined states, the algebraic structure of the observables has to be conserved in this hidden variable theory.

Von Neumann gave the first proof of the impossibility of hidden variables for quantum mechanics. One of the assumptions made by Von Neumann is that the expectation value of a linear combination of two observables is the linear combination of the expectation values of the observables. As remarked by Bell in Ref. 2, this assumption is not justified for noncompatible observables, such that, indeed, Von Neumann's proof cannot be considered to be conclusive. Bell constructs in the same reference a hidden variable model for the spin-\(1/2\), and shows that indeed Von Neumann's assumption is not satisfied in this model. Bell also criticizes in the same paper two other proofs of the nonexistence of hidden variables, the proof by Jauch and Piron and the proof by Gleason. Bell correctly points out the danger of demanding extra assumptions to be satisfied without knowing exactly what these assumptions mean physically. The extra mathematical assumptions criticized by Bell were introduced in all these approaches to express the physical idea that it must be possible to find, in the hidden variable description, the original observables and their basic algebra. We think that this physical idea was expressed correctly, without extra mathematical assumptions, and used in the impossibility proof of Kochen-Specker. Gudder gave an impossibility proof along the same lines as the one by Jauch and Piron, but now carefully avoiding the assumption criticized by Bell.

One could conclude by stating that every one of these impossibility proofs consists of showing that a hidden variable theory gives rise to a certain mathematical structure for the observables (in Refs. 1, 4, and 5) or for the properties (in Refs. 3 and 6) of the physical system under consideration. The observables or the properties of a quantum system do not have this mathematical structure. Therefore it is impossible to replace quantum mechanics by a hidden variable theory. To be more specific, if one works in the category of observables, then a hidden variable theory always has a commutative algebraic structure for the set of observables, while the algebra of observables of a quantum system is never commutative. If one works in the category of properties (yes-no observables) then a hidden variable theory always has a Boolean lattice structure for the set of properties, while the lattice of properties of a quantum system is never Boolean.

Recently this structural difference between classical systems and quantum systems has been studied by Accardi in another category, namely the category of probability models. Accardi gives a definition of a Kolmogorovian probability model, which is the probability model of a classical system, and a quantum probability model, which is the probability model of a quantum system. Again these two probability models have completely different mathematical structures. What is more powerful in this probability ap-
approach, compared to the algebraic or lattice theoretic approach, is that probability models can be compared with experimental results. To make this possible, Accardi derives a set of inequalities that characterizes the Kolmogorovian model and shows that these inequalities can always be violated by experiments with a quantum system. Accardi even manages to derive inequalities that can discriminate between a complex and a real Hilbert-space model, which shows the power of this approach. Again this probability approach shows the fundamental difference between a classical theory and a quantum theory.

A lot of physicists, once aware of this fundamental structural difference between a classical theory and a quantum theory, gave up hope that it would ever be possible to replace quantum mechanics by a hidden variable theory; and I admit that I was one of them. I should like to show in this paper that the state of affairs is, however, more complicated.

Some time ago I managed to build a macroscopic classical system that violates Bell inequalities. On the other hand, Accardi had shown that Bell inequalities are equivalent to his inequalities characterizing a Kolmogorovian probability model. Since my example did violate Bell inequalities, it should also violate Accardi’s inequalities characterizing a Kolmogorovian probability model. This is indeed the case. But then I had given an example of a macroscopic “classical” system having a non-Kolmogorovian probability model. This was very amazing, and the classification made by a lot of physicists of a micro world described by quantum mechanics and a macro world described by classical physics was challenged completely. This system violating Bell inequalities and having a non-Kolmogorovian probability model is presented in Ref. 10. The reason macroscopic systems can have non-Kolmogorovian probability models is what I want to analyze in this paper. I shall show that the state of affairs is the following: If we have a physical system $S$ and we have a lack of knowledge about the state of $S$, then a theory describing this situation is necessarily a classical statistical theory having a classical Kolmogorovian probability model. If we have a physical system $S$ and a measurement $e$ on this physical system $S$, and the situation is such that we do not have a lack of knowledge about the state of $S$, but we do have a lack of knowledge about the measurement $e$, then we cannot describe this situation by a classical statistical theory, because the probability model which arises is non-Kolmogorovian. Hence, lack of knowledge about the measurements leads to a non-Kolmogorovian probability model. What do we mean by “lack of knowledge about the measurement $e$”? Well, we mean that the measurement $e$ is in fact not a “pure” measurement, in the sense that there are hidden measurements $e_4$ such that the measurement $e$ consists of choosing in one way or another between the measurements $e_4$ and then performing the chosen measurement.

We can ask now whether it is possible to get a quantum probability model in this way. That this can indeed be done is shown in Sec. III, based on a macroscopic example. The example in Sec. III has a non-Kolmogorovian probability model of the spin of a spin-$\frac{1}{2}$ particle, and is constructed by introducing a lack of knowledge about the measurements.

It is shown in Sec. IV, based on another macroscopic example, that this physical situation of lack of knowledge about the measurements does not only give rise to a quantum probability model, but can also deliver a probability model that is neither Kolmogorovian nor quantum.

We can wonder now whether every quantum system can be described by a model with a lack of knowledge about the measurements. That this is indeed the case is shown in Sec. V.

In Ref. 5 Kochen and Specker give a classical statistical model for the spin of a spin-$\frac{1}{2}$ particle. Because we show in Sec. III that a spin model is a non-Kolmogorovian model, this seems to contradict the result of Kochen and Specker. We analyze this situation in Sec. VI.

Now, what is the relation of our model with lack of knowledge about the measurement to the attempts to build a hidden variable model? Well, our model with lack of knowledge about the measurements can also be regarded as a hidden variable model. The hidden variables are not then hidden variables of the system, but hidden variables of the measuring apparatus. But, for each measuring apparatus we then have a different set of hidden variables. As I came to know recently, it was shown by Gudder that a hidden variable model, where a different set of variables is allowed for each measurement, can reconstruct the probabilities of quantum mechanics. It seems, however, that nobody wanted to take into consideration these kinds of hidden variables. This is certainly because no interpretation was given to these hidden variables, and the theorem of Gudder was only considered to be an interesting mathematical result.

If we accept our explanation for the probabilities of quantum mechanics, namely that they are due to a lack of knowledge about the measurements, then these probabilities are not more ontological than ordinary probabilities. They form a nonclassical probability model because they correspond to a different physical situation, namely the physical situation where we lack knowledge about the measurements and not about the state of the system. It is clear that such a physical situation can be found, as well, in the micro world. We have shown this based on our examples. We can now ask why nonclassical probabilities only appeared in the micro world. In light of our hypothesis, the answer would be that the type of measurement, introducing nonclassical probabilities, is never used to describe a macroscopic system, because we have enough other measurements to replace them. This is no longer the case in the micro world.

II. CLASSICAL AND QUANTUM PROBABILITY MODELS

We will, in this paper, always consider the following situation: We have a physical system $S$. This physical system can be in different states $p, q, r, \ldots$. We denote the set of states by $\Sigma$. We can perform measurements $e, f, g, \ldots$ on this physical system. We denote the set of measurements by $\mathcal{M}$. We suppose for sake of simplicity that all measurements have an outcome set which is a discrete set of real numbers. Hence a measurement $e$ has possible outcomes $\{e_1, e_2, \ldots\}$.

If the system $S$ is a classical system, then for a state $p$, a measurement $e$ has a determined outcome $e(p)$, and measurements can, in this case, be represented by real valued functions (random variables) on $\Sigma$. If we have a lack of knowledge...
about the state of this classical system, then this lack of knowledge is described by a probability measure \( \mu \) on \( \Sigma \) such that if \( D \) is a measurable subset of \( \Sigma \), then \( \mu(D) \) is the probability that the state of \( S \) is in \( D \). This is the description of a classical system by means of a statistical theory, and often \( \mu \) is called a mixed state of the system \( S \). We then have

\[
P(e = e_i) = \mu(E_{e_i}),
\]

where \( P(e = e_i) \) means the probability of finding the outcome \( e_i \) if the measurement \( e \) is performed, and \( E_{e_i} \) is the set of states for which the measurement \( e \) gives outcome \( e_i \).

If the system \( S \) is a quantum system, then measurements are represented by self-adjoint operators on a Hilbert space \( H \). Since we considered only measurements with a discrete set of outcomes, we will have corresponding self-adjoint operators with a discrete spectrum. If \( e \) is such a measurement, with possible outcomes \( \{e_1, e_2, \ldots \} \), then the corresponding self-adjoint operator always has an orthonormal basis \( \{\phi_1, \phi_2, \ldots \} \) of eigenvectors. The state \( p \) of the system is represented by a normalized vector \( \psi \). We then have

\[
P(e = e_i | \psi) = (\langle \phi_i, \psi \rangle)^2,
\]

where \( P(e = e_i | \psi) \) means the probability of finding the outcome \( e_i \) when the measurement \( e \) is performed and when the system is in a state \( \psi \).

These two descriptions, classical and quantum, both give rise to a probability model. We would like to compare these two models. To do this, we have to find a concept that exists in both descriptions. We would also like to compare both models with physical experimental examples. Therefore, the common concept must have a physical meaning independent of the theories. For this purpose, we will use the concept of conditional probability \( P(e = e_i | f = f_j) \).

We will attribute the following physical meaning to \( P(e = e_i | f = f_j) \): The probability of finding the outcome \( e_i \) when we perform the measurement \( e \) when the state of the system is such that if we performed the measurement \( f \) we would find the outcome \( f_j \). Hence we do not have to perform the measurement \( f \). There is only a condition on the state of the system such that an eventual measurement of \( f \) would give us the outcome \( f_j \). We insist on making this remark, because often the conditional probability \( P(e = e_i | f = f_j) \) is given the following meaning: The probability of finding the outcome \( e_i \) for a measurement \( e \) if a measurement \( f \) has been performed and has given outcome \( f_j \). If the performance of the measurement \( f \) creates the good conditioning on the state of the system, namely the condition that if we performed \( f \) (again) we would find the outcome \( f_j \), then the two meanings are equivalent. This is the case for measurements of the first kind in quantum mechanics. But it is certainly not true for general measurements. Indeed, often a measurement destroys the system if one of the outcomes is obtained (e.g., measurement of the polarization of a photon). In such cases it usually makes no sense to perform the measurement \( e \) after we have performed \( f \).

Let us see how we can find this conditional probability in both theories.

If \( S \) is the classical system with lack of knowledge about the states described by the probability measure \( \mu \), then

\[
P(e = e_i | f = f_j) = \mu(E_{e_i} \cap F_{f_j}) / \mu(F_{f_j}),
\]

where \( E_{e_i} \) is the set of states for which the measurement \( e \) gives the outcome \( e_i \), and \( F_{f_j} \) is the set of states for which the measurement \( f \) gives the outcome \( f_j \). This equality is often introduced as a definition of the conditional probability in Kolmogorovian probability theory. We have, however, given a physical meaning to the conditional probability such that this conditional probability can be derived from the experimental results. Therefore we shall consider this equality as one of the characterizations of the Kolmogorovian probability model. If considered in this way, the equality is often called Bayes' axiom.

If \( S \) is a quantum system, then the condition that the state of the system \( S \) must be such that a measurement of \( f \) would give the outcome \( f_j \) can be expressed by asking that the state of \( S \) must be an eigenstate \( \psi_j \) of the self-adjoint operator corresponding to \( f \), an eigenstate corresponding to the eigenvalue \( f_j \). Then

\[
P(e = e_i | f = f_j) = (\langle \phi_i, \psi_j \rangle)^2.
\]

Let us now give the mathematical definitions for a classical and a quantum probability model as proposed by Accardi in Refs. 7 and 8.

**Definition 1:** A Kolmogorovian [classical] model for the family of conditional probabilities \( \{P(e = e_i | f = f_j) ; e, f, e \in \mathcal{A}\} \) is defined by a set \( \Sigma \), a set \( \mathcal{B} \) of measurable subsets of \( \Sigma \) which has the structure of a \( \sigma \)-algebra, and the probability measure \( \mu: \mathcal{B} \to [0,1] \). For every \( e \in \mathcal{A} \) and for every outcome \( e_i \) of \( e \), there exists a set \( E_{e_i} \in \mathcal{B} \) such that for every \( f \)

\[
P(e = e_i | f = f_j) = \mu(E_{e_i} \cap F_{f_j}) / \mu(F_{f_j}).
\]

**Definition 2:** A complex (resp. real or quaternionian) Hilbert-space model for the family of conditional probabilities \( \{P(e = e_i | f = f_j) ; e, f, e \in \mathcal{A}\} \) is defined by a complex (resp. real or quaternionian) Hilbert space \( H \). For every \( e \in \mathcal{A} \) there exists an orthonormal basis \( \{\phi_i\} \) such that

\[
P(e = e_i | f = f_j) = (\langle \phi_i, \psi_j \rangle)^2,
\]

if \( \phi_i \) and \( \psi_j \) are the bases corresponding to \( e \) and \( f \).

In Sec. III we will give an example of a macroscopic system with lack of knowledge about the measurements that has a quantum probability model.

### III. EXAMPLE OF A MACROSCOPIC PHYSICAL SYSTEM WITH A QUANTUM DESCRIPTION; CONSTRUCTION OF A SPIN-\(1\) MODEL

We will give, in this section, an example of a macroscopic physical system that gives a model for the spin-\(1\).

We consider a particle with positive charge \( q \) that is located on a sphere with radius \( r \) at a point \( (r, \theta, \phi) \). The measurement \( e_{\alpha\beta} \) consists of the following operation: We choose two particles with negative charges \( q_1 \) and \( q_2 \) such that \( q_1 + q_2 = Q \). The charge \( q_1 \) is chosen at random between 0 and \( Q \). This represents the lack of knowledge about the measurement. Once the charges \( q_1 \) and \( q_2 \) are chosen we put the two particles diametrically on the sphere, such that \( q_1 \) is in the point \( (r, \alpha, \beta) \) and \( q_2 \) is in the point \( (r, -\alpha, -\beta) \). Let us call \( F_1 \) and \( F_2 \) the Coulomb forces of \( q_1 \) on \( q_2 \) and of \( q_2 \) on \( q_1 \). If the magnitude of \( F_1 \) is bigger than the magnitude of \( F_2 \), we give the outcome \( e_1 \) to the measurement \( e_{\alpha\beta} \). If the
magnitude of $F_2$ is bigger than the magnitude of $F_1$, we give the outcome $e_2$ to the measurement $e_{\alpha\beta}$. (See Fig. 1.)

The forces $F_1$ and $F_2$ are in the plane through $|r, \theta, \phi\rangle$, $|r, \alpha, \beta\rangle$, and $(r, \pi - \alpha, \pi + \beta)$. (See Fig. 2.)

Let us call $\gamma$ the angle between $(r, \theta, \phi)$ and $(r, \alpha, \beta)$. Then

$$\|F_1\| = \frac{q_1 \cdot q}{4 \pi \epsilon_0 r^2 \sin^2 (\gamma/2)},$$

and

$$\|F_2\| = \frac{q_2 \cdot q}{4 \pi \epsilon_0 r^2 \cos^2 (\gamma/2)}.$$

Let us now calculate the probability that we get the outcome $e_1$ for $e_{\alpha\beta}$ if the particle $q$ is in state $(\theta, \phi)$.

$$P(\|F_1\| > \|F_2\|) = P \left( \frac{q_1 \cdot q}{4 \pi \epsilon_0 r^2 \sin^2 (\gamma/2)} > \frac{q_2 \cdot q}{4 \pi \epsilon_0 r^2 \cos^2 (\gamma/2)} \right)$$

$$= P(q_1 \cos^2 (\gamma/2) > q_2 \sin^2 (\gamma/2))$$

$$= P(q_1 \cos^2 (\gamma/2) > (Q - q_1) \sin^2 (\gamma/2))$$

$$= P(q_1 > Q \sin^2 (\gamma/2))$$

$$= (Q - Q \sin^2 (\gamma/2))/Q = \cos^2 (\gamma/2).$$

This is exactly the probability that we would find if $e_{\alpha\beta}$ represented the measurement of the spin of a spin-1 particle in the $(\alpha, \beta)$ direction while the particle had spin in the $(\theta, \phi)$ direction.

We can describe this system by means of a two-dimensional complex Hilbert space. We then represent the state of the particle $q$ in the $(\theta, \phi)$ direction by means of the vector

$$x_{\alpha, \beta} = (e^{-i \theta/2} \cos (\theta/2), e^{i \theta/2} \sin (\theta/2)).$$

and the measurement $e_{\alpha\beta}$ by means of the self-adjoint operator

$$S_{\alpha, \beta} = \frac{1}{2} \left( \cos \alpha \quad e^{i \alpha} \sin \alpha \right).$$

We can then apply the calculus of quantum mechanics to the description of our system. Let us remark again that the state of the particle $q$ is a pure state and the probability only comes from a lack of knowledge about the measurement $e_{\alpha\beta}$.

Let us now consider the physical situation where we also have a lack of knowledge about the state of the system. More specifically, we suppose that the charge $q$ (or the spin of the spin-1 particle) is in every direction $(\theta, \phi)$ with equal probability. We consider three measurements $e$, $f$, and $g$ such that $e = e_{\theta, 0, 0}$, $f = e_{\pi/3, 0}$, and $g = e_{2\pi/3, 0}$ (see Fig. 3). Then clearly

$$P(e = e_1) = P(f = f_1) = P(g = g_1) = P(e = e_2) = P(f = f_2) = P(g = g_2) = \frac{1}{3}.$$ 

Let us now show that there does not exist a classical Kolmogorovian probability model for this system.

If there does exist a classical description, then we must have a probability measure $\mu$ and

$$\mu(E_i) = \mu(E_2) = \mu(F_1) = \mu(F_2) = \mu(G_1) = \mu(G_2) = \frac{1}{3}.$$ 

Using Bayes axiom and the properties of the probability measure we have

$$1P(f = f_1 | g = g_1) = \mu(F_1 \cap G_1) = \mu(E_2 \cap F_1 \cap G_1) + \mu(E_2 \cap F_1 \cap G_1),$$

$$1P(e = e_1 | g = g_1) = \mu(E_1 \cap G_1) = \mu(E_1 \cap F_1 \cap G_1) + \mu(E_1 \cap F_1 \cap G_1).$$

Now $P(f = f_1 | g = g_1)$ is the probability that a measurement of $g$ gives the outcome $f$, if the state of the system is such that an eventual measurement of $g$ would give the outcome $g_1$. The only state of the system with the property that a measurement of $g$ would always give $g_1$ is the state where the charge $q$ is at $(r, 2\pi/3, 0)$ (or the spin of the spin-1 particle is in direction $(2\pi/3, 0)$).

Hence

FIG. 1. A positive charge $q$ is located on the sphere at $(r, \theta, \phi)$ and two negative charges $g_1$ and $g_2$ are chosen as explained in the text and located on the sphere at points $(r, \alpha, \beta)$ and $(r, r - \alpha, r + \beta)$.

FIG. 2. We consider the three charges of Fig. 1 as they are located in one plane.

FIG. 3. The three measurements $e$, $f$, and $g$ that are considered to show that the system of Fig. 1 does not allow a Kolmogorovian probability model.
we will consider the macroscopic system consisting of the same charge \( q \) on the sphere with the same measurements \( \epsilon_{ab} \) and the same lack of knowledge about these measurements, but now we define the outcomes as follows: \( \epsilon_{ab} \) has outcome \( e_1 \) if \( q \) moves towards \( q_1 \) on the sphere (not on the line) and \( \epsilon_{ab} \) has outcome \( e_2 \) if \( q \) moves towards \( q_2 \) on the sphere. We will then show that this macroscopic system admits neither a classical Kolmogorovian nor a quantum Hilbert-space probability model.

**IV. EXAMPLE OF A MACROSCOPIC PHYSICAL SYSTEM THAT ADMITS NEITHER A CLASSICAL NOR A QUANTUM PROBABILITY MODEL**

We consider again a particle with positive charge \( q \) that is located on a sphere with radius \( r \) at a point \((r, \theta, \phi)\). The measurement consists of the following operation: We choose two particles with negative charges \( q_1 \) and \( q_2 \) such that \( q_1 + q_2 = Q \). The charge \( q_1 \) is chosen at random between \( O \) and \( Q \). This represents the lack of knowledge about the measurement. Once the charges \( q_1 \) and \( q_2 \) are chosen we put them diametrically on the sphere at points \((r, \alpha, \beta)\) and \((r, \pi - \alpha, \pi + \beta)\). We call \( F_1 \) and \( F_2 \) the Coulomb forces of \( q \) on \( q_1 \) and \( q_2 \) on \( q \). The charge \( q \) can move on the sphere. (See Fig. 5.) We call \( F_1 \) and \( F_2 \) the projections of \( F_1 \) and \( F_2 \) on the tangent plane at \((r, \theta, \phi)\). If \( ||F_1|| > ||F_2|| \), \( q \) will move towards \( q_1 \), and then we will give outcome \( e_1 \) to \( \epsilon_{ab} \). If \( ||F_1|| < ||F_2|| \), \( q \) will move towards \( q_2 \), and then we will give outcome \( e_2 \) to \( \epsilon_{ab} \).

One could ask why we had to make such a rather complicated mechanical picture for the motion of the charge \( q \). Why not just let it move on the sphere? Well, the forces that control the motion on the sphere are not \( F_1 \) and \( F_2 \), but the projections of \( F_1 \) and \( F_2 \) on the tangent plane at \( q \). In Sec. IV...
Let us now calculate the probability of getting $e$, for $e_{up}$ if the particle $g$ is in state $(\theta, \phi)$. Again we call $\gamma$ the angle between $(r, \theta, \phi)$ and $(r, \alpha, \beta)$. Then

$$\|F'_1\| = \frac{g \cdot q}{4 \pi \varepsilon_0 \varepsilon_r^2 \sin^2(\gamma/2)} \cdot \cos \frac{\gamma}{2},$$

$$\|F'_2\| = \frac{g \cdot q}{4 \pi \varepsilon_0 \varepsilon_r^2 \cos^2(\gamma/2)} \cdot \sin \frac{\gamma}{2},$$

$$P(\|F'_1\| > \|F'_2\|) =\frac{P\left(\frac{g \cdot q}{4 \pi \varepsilon_0 \varepsilon_r^2 \sin^2(\gamma/2)} \cdot \cos \frac{\gamma}{2} > \frac{g \cdot q}{4 \pi \varepsilon_0 \varepsilon_r^2 \cos^2(\gamma/2)} \cdot \sin \frac{\gamma}{2}\right)}{P\left(\frac{g \cdot q}{4 \pi \varepsilon_0 \varepsilon_r^2 \sin^3(\gamma/2)} > \frac{q}{4 \pi \varepsilon_0 \varepsilon_r^2 \cos^3(\gamma/2)} \cdot \sin \frac{\gamma}{2}\right)}$$

$$= \frac{P\left(\frac{g \cdot q}{4 \pi \varepsilon_0 \varepsilon_r^2 \sin^3(\gamma/2)} > q \cdot \cos^3(\gamma/2) > q \cdot \sin^3(\gamma/2)\right)}{P\left(\frac{g \cdot q}{4 \pi \varepsilon_0 \varepsilon_r^2 \cos^3(\gamma/2)} \cdot \sin \frac{\gamma}{2}\right)}$$

$$= \frac{P\left(g \cdot \cos^3(\gamma/2) > q \cdot \sin^3(\gamma/2)\right)}{P\left(g \cdot \cos^3(\gamma/2) > q \cdot \sin^3(\gamma/2)\right)}$$

$$= \frac{P\left(g \cdot \cos^3(\gamma/2) > q \cdot \sin^3(\gamma/2)\right)}{P\left(g \cdot \cos^3(\gamma/2) > q \cdot \sin^3(\gamma/2)\right)}$$

$$= \frac{\cos^3(\gamma/2) + \sin^3(\gamma/2)}{\cos^3(\gamma/2) + \sin^3(\gamma/2)} \cdot \cos \frac{\gamma}{2}.$$

We can, in the same way as for the macroscopic system of Sec. III, prove that this macroscopic system cannot be described by means of a Kolmogorovian probability model. We will not repeat this proof because it is completely analogous to the proof in Sec. III.

The macroscopic system in Sec. III could be given a quantum description. Let us now show that for the physical system in this section this cannot be done anymore. Hence this physical system has a probability model which is neither classical nor quantum.

We consider the measurements $e$, $f$, and $g$ such that $e = e_{0,0}$, $f = e_{1/2}$, and $g = e_{-1/2}$ (see Fig. 6). Clearly with our choice of measurements

$$P(g = g_1 | e = e_1) = P(g = g_2 | f = f_1) = P(f = f_1 | e = e_1)$$

$$= \frac{1}{\cos^3(\gamma/3) + \sin^3(\gamma/3)} \cdot \frac{1}{3\sqrt{3} + 1},$$

$$P(g = g_1 | e = e_2) = P(g = g_2 | f = f_2) = P(f = f_2 | e = e_2)$$

$$= \frac{\sin^3(\gamma/3)}{\cos^3(\gamma/3) + \sin^3(\gamma/3)} \cdot \frac{3\sqrt{3}}{3\sqrt{3} + 1}.$$
The example of Sec. III shows that we can find the quantum probability calculus of the spin-\(1\) by supposing a lack of knowledge about the measurements. The example of Sec. IV shows that a lack of knowledge about the measurements can also, however, give rise to a nonclassical probability calculus which is nonquantal. In both examples there is no lack of knowledge about the state of the system. So we could say that the states are pure states; and it is clear that it will not be possible to build a hidden variable model if one wants the hidden variables to be hidden variables of the systems.

V. A MODEL WITH LACK OF KNOWLEDGE ABOUT THE MEASUREMENTS FOR A GENERAL QUANTUM SYSTEM

The example of Sec. III shows that lack of knowledge about the measurements can lead to a quantum mechanical probability model. We wonder whether we can make a model for a general quantum mechanical system. We will give in this section a construction that shows that this can indeed be done. In this construction we will only consider quantum mechanical systems described in an \(n\)-dimensional Hilbert space. An analogous reasoning can be made, however, for the case of an infinite-dimensional Hilbert space.

So we have a physical system \(S\) described in an \(n\)-dimensional Hilbert space \(H\). A measurement \(e\) on this physical system is represented by a self-adjoint operator \(A\). With this self-adjoint operator correspond \(n\) eigenvectors \(v_1,...,v_n\) and eigenvalues \(a_1,...,a_n\). If \(w\) is an arbitrary state of the system, we can write

\[
w = \sum_{j=1}^{n} \langle w, v_j \rangle v_j,
\]

because \([v_1,...,v_n]\) is chosen to be orthonormal. Now \(x_i = |\langle w, v_i \rangle|^2\) is the probability that by measurement of \(e\) we find the value \(a_i\), if the system is in state \(w\). So we can represent this state \(w\) by means of these \(n\) probabilities \([x_1,...,x_n]\).

Hence all the states can be represented in this way by \(n\)-tuples

\[x = (x_1,...,x_n),\]

such that \(\sum_{i=1}^{n} x_i = 1\) and \(0 < x_i < 1\).

These are the points of the simplex \(S_n\) in \(\mathbb{R}^n\) spanned by the canonical base vectors \(h_1 = (1,0,...,0),\ h_2 = (0,1,0,...,0),...,\ h_n = (0,...,1)\). Hence

\[x = \sum_{i=1}^{n} x_i h_i.
\]

We now have to show that it is possible to construct a set of measurements \(e_1,...,e_n\) such that the measurement \(e\) consists of (i) choosing at random one of the measurements \(e_i\), and (ii) performing this chosen measurement. Moreover the measurements \(e_i\) have to be classical measurements, which means that they give a determined outcome for a given state of the system.

We will construct these hidden measurements as follows: We will label the measurements by \(\lambda\) where \(\lambda\) is an \(n\)-tuple \(\lambda = (\lambda_1,...,\lambda_n)\) such that \(\sum_{i=1}^{n} \lambda_i = 1\,\text{and}\,0 < \lambda_i < 1\).

Suppose that we have a given state represented by the \(n\)-tuple \(x = (x_1,...,x_n)\). Let us call \(A_i\) the convex closure of \([x, h_1,...,h_{i-1}, h_{i+1},...,h_n]\). Then clearly \(S_i = \bigcup_{A_i} A_i\) (see Fig. 7). We define the measurements \(e_\lambda\) as follows: If \(\lambda \in A_i\), then the measurement \(e_\lambda\) gives the outcome \(a_i\) if the system is in the state \(x\). If \(\lambda\) is a point of the boundary of \(A_i\) and, hence, also a point of \(A_{i-1}\) or of \(A_{i+1}\), then the outcome of \(e_\lambda\) is indeterminate, but indeterminate in the classical sense (as, for example, in the case of a classical unstable equilibrium). The probability of choosing \(\lambda\) on such a boundary is, however, zero. So these boundary situations do not contribute to the final probabilities.

Let us now calculate the probability of choosing \(\lambda\) in the simplex \(A_i\). This probability is given by

\[m^* (A_i) / m^* (S_i),\]

where \(m^*\) is the trace on \(S_n\) of the Lebesgue measure on \(\mathbb{R}^n\). For example, if \(n = 3\) (see Fig. 7), then we have to calculate the surface of \(A_i\) and divide by the surface of \(S_i\). Clearly

\[m^* (A_i) = [1/(n-1)] m^* (S_{i-1}),\]

where \(S_{i-1}\) is the convex closure of \(h_1,...,h_{i-1}, h_{i+1},...,h_n\), and \(I_i\) is the distance from \(x\) to \(S_{i-1}^\perp\). A point of \(S_{i-1}^\perp\) can be written as follows: \((y_1,...,y_{i-1}, 0, y_{i+1},...,y_n)\), with \(\sum_{i=1}^{n} y_i = 1\) and \(0 < y_i < 1\). The line through this point and the point \((x_1,...,x_n)\) is carried by the vector \((x_1-y_1,...,x_{i-1}-y_{i-1}, x_{i+1}-y_{i+1},...,x_n-y_n)\). To find \(I_i\), on this line, this vector has to be orthogonal on the vectors \(h_k-h_m\) for all \(k\) and \(m\) different from \(i\). This means that \(x_k - y_k = x_m - y_m\) for all \(k\) and \(m\) different from \(i\). We also have

\[y_1 + ... + y_{i-1} + y_{i+1} + ... + y_n = 1,
\]

\[x_1 + ... + x_{i-1} + x_{i+1} + ... + x_n = 1 - x_i.
\]

Hence

\[(y_1-x_1) + ... + (y_{i-1}-x_{i-1})
\]

\[+ (y_{i+1}-x_{i+1}) + ... + (y_n-x_n) = x_i.
\]

From this, it follows that \(y_k - x_k = x_n / (n-1)\) for all \(k \neq i\).

Hence \(y_k = x_n + x_n / (n-1)\) for all \(k \neq i\). So

\[I_i = \left| \left| (x_1,...,x_n) - \left( x_1 + \frac{x_1}{n-1},...x_{i-1} + \frac{x_{i-1}}{n-1}, 0,\right) \right| \right|
\]

\[= \left| \left| \left( -\frac{x_{i+1}}{n-1}, ..., -\frac{x_n}{n-1} \right) \right| \right|
\]

\[= \sqrt{\frac{n}{n-1}} |x_i|.
\]

FIG. 7.
Hence \( m(A_n) = \{1/(n-1)\} m^{n-1}(S_{n-1}) \cdot \sqrt{n/(n-1)} \cdot x_i \),
and as a consequence,
\[ m(A_n)/m(S_n) = x_i. \]

This shows that if the system is in the state \( x_i \), and we perform the measurement \( e_i \), then the probability of obtaining the result \( a_i \) is exactly given by \( x_i \). For each measurement we can construct such a collection of hidden measurements \( f_j \). Let us call these hidden measurements the “pure measurements.” Is it possible to characterize the collection of all pure measurements? We can do this in the following way.

We have a physical system \( S \). The pure states of \( S \) are represented by normalized vectors of an \( n \)-dimensional complex Hilbert space \( H \). The pure measurements can also be represented by the normalized vectors of the same \( n \)-dimensional complex Hilbert space \( H \). Let us denote by \( \rho_{w} \) a pure state represented by the vector \( w \), and by \( e_w \), a pure measurement represented by the vector \( w \). To derive quantum mechanics we adopt the following rules.

If we want to perform a measurement of a system, then first we choose a system and a measurement. This choice corresponds to choosing an orthonormal basis \( \{w_1, \ldots, w_n\} \) of the Hilbert space \( H \), and a set of possible outcomes \( \{a_1, \ldots, a_n\} \) (or a self-adjoint operator). This choice is of course not governed by a probability rule, it is just the choice of the context of the measurement. The system will, however, be in a certain pure state \( \rho_w \) represented by the vector \( w \), and the measurement will be a certain pure measurement \( e_w \) represented by the vector \( w \). The outcome of this pure measurement \( e_w \) is determined for the state \( \rho_w \) in the following way.

Let us put
\[ b_i = |\langle w, w_i \rangle|^2, \]
and let us consider the set of real numbers \( B = \{ b_1, \ldots, b_n \} \). If \( \langle w, w_i \rangle = 0 \) and \( \langle w, w_i \rangle \neq 0 \), then \( b_i = +\infty \). If \( \langle w, w_i \rangle = 0 \) and \( \langle w, w_i \rangle = 0 \), then \( b_i \) is not taken into consideration. If \( B \) has a maximum, for example the number \( b_{i'} \), then the outcome of the measurement \( e_a \) when the system is in the state \( \rho_w \) is always \( a_{i'} \). If \( B \) has no maximum, then the outcome of the measurement \( e_a \) when the system is in the state \( \rho_w \) is indeterminate in the classical sense (as, for example, in the case of a classical unstable equilibrium).

If we now want to know the probability that in the measurement context, represented by the orthonormal basis \( \{w_1, \ldots, w_n\} \) and the system being in the state \( \rho_w \) we find the outcome \( a_{i'} \), then this probability will be given by \( |\langle w, w_{i'} \rangle|^2 \).

Indeed, if we put \( \Lambda_i = |\langle w, w_i \rangle|^2 \), then \( \Lambda = (\Lambda_1, \ldots, \Lambda_n) \) will be a point of the simplex \( \Lambda \) if and only if \( b_i \) is a maximum of \( B \).

VI. COMPARING OUR SPIN MODEL WITH THE MODEL OF KOCHEN AND SPECKER

In Ref. 5 Kochen and Specker construct a spin model which has a classical statistical mathematical structure. This seems to contradict our result of Sec. III where we explicitly show that every spin model will have a non-Kolmogorovian probability model. Let us analyze this situation. The model of Kochen and Specker is mathematically more complicated than ours, but we can construct a model equivalent to the one of Kochen and Specker which is very similar to our model and can easily be compared. Let us do this first. In the model that we proposed in Sec. III, the lack of knowledge that gives rise to the probabilities is about the measurements. We can, however, in the same spirit build a model where the lack of knowledge is about the state of the system. To obtain this we consider the measuring apparatus of the example in Sec. III as the physical system and we consider the system as measuring apparatus. The model that we get in this way is completely equivalent to the model proposed by Kochen and Specker. But let us see now that it is in fact not really a spin model; so the physical system consists of two negative charges \( q_1 \) and \( q_2 \) such that \( q_1 + q_2 = Q \), that are located diametrically on a sphere with radius \( r \) at points \( (r, \alpha, \beta) \) and \( (r, \pi - \alpha, \pi + \beta) \). The pure state of the system can be described by the direction \( (\alpha, \beta) \) and the charge \( q_1 \). Hence the pure states can be represented by a point of the sphere and a point in \([0, Q]\). We now suppose that we have a lack of knowledge about these states, in the sense that we do not know the charges \( q_1 \) and \( q_2 \). The mixed states are described by the directions \( (\alpha, \beta) \), and \( q_1 \) is a hidden variable. The measurement \( e_{\alpha,\beta} \) consists of putting a positive charge \( q \) in a point \( (r, \theta, \phi) \). Let us call \( F_1 \) and \( F_2 \) the two Coulomb forces between \( q_1 \) and \( q_2 \) and between \( q_1 \) and \( q_2 \) (see Fig. 8). If the magnitude of \( F_1 \) is bigger than the magnitude of \( F_2 \) we give the outcome \( e_1 \) to the measurement \( e_{\alpha,\beta} \) and if the magnitude of \( F_1 \) is smaller than the magnitude of \( F_2 \) we give the outcome \( e_2 \) to the measurement \( e_{\alpha,\beta} \). If we call \( \gamma \) the angle between \( (r, \alpha, \beta) \) and \( (r, \theta, \phi) \), then we find with the same calculation as in Sec. III that

\[ P(\|F_1\| > \|F_2\|) = \cos^2 \left( \frac{\gamma}{2} \right). \]

So for one measurement \( e_{\alpha,\beta} \) we find the good probability of the spin-I. But let us see what happens if we make a second measurement \( e_{\alpha',\beta'} \) after \( e_{\alpha,\beta} \) in the direction \( (\theta', \phi') \) which makes an angle \( \delta \) with the direction \( (\theta, \phi) \). We also suppose that \( (r, \alpha, \beta) \), \( (r, \theta, \phi) \), and \( (r', \theta', \phi') \) lie in the same plane (see Fig. 9). So, we now suppose that the state (mixture) of the system was \( (\alpha, \beta) \) and we have performed \( e_{\alpha,\beta} \) and gotten the outcome \( e_1 \). After this we perform the measurement \( e_{\alpha',\beta'} \) and we want to calculate the probability that we get the outcome \( e_1 \) for \( e_{\alpha',\beta'} \). Let us denote this probability as follows:

\[ P(e_{\alpha',\beta'} = e_1 | e_{\alpha,\beta} = e_1). \]

Then, because we have a classical statistical situation,
does not agree with the spin-$\frac{1}{2}$ model. This explains the apparent contradiction between our results and those of Kochen and Specker.

The reason the model proposed in this section is not a spin model is because the hidden variable $q_i$ does not get distributed at random again after one measurement. Kochen and Specker were clearly aware of this difficulty in their model, and, therefore, they add the extra condition that after a measurement the hidden variable must in one way or another get distributed at random again. If this extra condition is added, then the model is a spin model, but of course it is easy to see that if one adds the extra condition that the hidden variable gets distributed at random again after the measurement, then the hidden variable is in fact a hidden variable of the measurement, and hence with this extra condition we really are in a physical situation as the one described in this paper, namely the physical situation that we lack knowledge about the measurement, and this is no classical statistical situation.