Consistency, amplitudes, and probabilities in quantum theory

Ariel Caticha
Department of Physics, State University of New York at Albany, Albany, New York 12222
(Received 6 October 1997)

Quantum theory is formulated as the only consistent way to manipulate probability amplitudes. The crucial ingredient is a very natural consistency constraint: if there are two different ways to compute the amplitude for a given process, the two answers must agree. This constraint is expressed in the form of functional equations the solution of which leads to the usual sum and product rules for amplitudes. An immediate consequence is that the Schrödinger equation must be linear; i.e., nonlinear variants of quantum mechanics are inconsistent. The physical interpretation of the theory is given in terms of a single natural rule. This rule, which does not itself involve any probabilities, is used to obtain a proof of Born's statistical postulate: we show that the probability of a certain outcome in an experiment is given by the square of the modulus of the corresponding amplitude. Thus, consistency leads to indeterminism. [S1050-2947(98)04103-1]

PACS number(s): 03.65.Bz, 03.65.Ca

I. INTRODUCTION

In 1946 Cox gave an argument showing that once degrees of probability are represented by real members there is a unique set of rules for inductive reasoning that is for reasoning under conditions of insufficient information [1]. The crux of the argument is a consistency requirement: if a probability can be computed in two different ways, the two answers must agree. Cox expressed this consistency requirement in the form of functional equations, the solution of which showed that the rules for inductive reasoning coincide with the well-known rules of probability theory. The importance of this achievement is twofold: First, it legitimized viewing probability theory as an extended form of logic, a point of view that goes back to Bernoulli and Laplace, was arguably held by Gibbs, and that, more recently, has been forcefully advocated by Jaynes [2]. Second, Cox's argument provides an explanation for the uniqueness of probability theory, for its inevitability; any modifications of the rules of probability theory will necessarily lead to inconsistencies, and therefore will be unsatisfactory.

This latter feature, the robustness of probability theory, is also shared by quantum theory. The quest to explain the strange behavior of quantum systems has, since the beginning, led to all sorts of attempts to modify the theory. Two of the most central attributes of quantum theory, indeterminism and linearity, have been the target of many such unsuccessful attempts. There has been considerable progress on issues related to the possibility of hidden variables and to the nature of statistical correlations [3]. Linearity violations have been considered as a means to resolve the difficulties associated with the quantum mechanics of macroscopic objects [4]. Some authors have been motivated by the fact that many linear physical theories are mere approximations to more fundamental nonlinear theories [5], while others were led by the desire to either test quantum mechanics ever more stringently [6], or just to explore the curious implications of nonlinearity [7]. Such extensive theoretical investigations have prompted several increasingly precise experimental tests [8,9] that have confirmed, at least for the time being, the robustness of quantum mechanics.

In this work we propose an approach to quantum theory using ideas inspired by Cox's, although in a very different context. The result is the standard quantum theory [10,11]. The crux of our argument is also a consistency requirement: if a probability amplitude can be computed in two different ways, the two answers must agree. This requirement is expressed in the form of functional equations, the solution of which leads to the usual sum and product rules for quantum probability amplitudes. In other words, *quantum theory emerges as the unique way to manipulate probability amplitudes consistently*.

Next we obtain two important consequences. The first is that the equation for time evolution, the Schrödinger equation, is necessarily linear. The implication is that the question of whether nonlinear versions of quantum mechanics are at all possible should not be posed at the dynamical level of the Schrödinger equation but rather at a much deeper kinematical level requiring a reexamination of the use and utility of the concept of amplitude.

The second result addresses the issue of how the knowledge of the numerical value of an amplitude assists us in predicting the outcomes of experiments. This question of the physical interpretation of an otherwise abstract formalism is handled by proposing a very natural general rule, which applies to situations in which the result of an experiment is predicted with certainty. Using this rule, which involves no probabilities, we obtain a proof of Born's statistical postulate. The implication here is that a quantum theory formulated in terms of consistently assigned amplitudes must be indeterministic.

These two results, the proof of linearity and of Born's probability interpretation, are not new. They have been anticipated within various axiomatic approaches to quantum mechanics [12–15]. For example, the fact that Born's postulate is actually a theorem was independently discovered long ago by Gleason [14], Finkelstein [15], Hartle [16], and Graham [17]. What is different here is the manner in which the results are obtained; the emphasis is on consistency in a formulation where amplitudes, rather than states or observables, are the central concept. From the pedagogical point of view there is an advantage in sharing the intuitive appeal of Feynman's path integrals [11]. Axiomatic methods, on the other hand, tend to be considerably more abstract and mathematically sophisticated; this is not in itself a defect but it does hinder their accessibility.

To limit the risk that too general a treatment might obscure the simplicity of the main ideas we will focus our attention on a simple example: a particle with no spin or other internal structure; its only attribute is its position. Furthermore, to avoid distractions with mathematical technicalities (which might, in other contexts, be very relevant) we will restrict the positions of the particle to sites on a discrete lattice. We emphasize that these simplifications are not necessary. The generalization to other systems involving more complicated configuration spaces is straightforward. If one wants to describe a particle moving in a continuum the modifications are rather trivial, a mere replacement of sums by integrals; the case of a quantum field theory might not be as easy, but in principle it should be doable as well.

In Sec. II we consider various idealized experimental setups that will test whether a particle moves from an initial starting point to a final destination point. The use of these setups defines what statements or propositions about the particle we are allowed to make. No mention of observables beyond position is ever made, but there is a possibility of combining simple setups into more complex ones. This is described by introducing two operations, which we call AND and OR, that allow us to construct complex setups (or propositions) from simpler ones.

At first sight this approach to quantum theory might resemble other axiomatic approaches. For example, in the quantum logic approach [12,13,15] propositions are also defined operationally in terms of the setups that will test them. But there are major differences, for example, an operation of central importance in quantum logic is that of negating a proposition. In our approach negation is never introduced. A comparison with the complex probability approach [18] also shows a similarity that on further analysis has, again, proved superficial. Unlike the latter theory, our AND and OR operations are not the usual Boolean ones, although they do enjoy a sufficient measure of associativity and distributivity to justify their names. In fact, the set of statements allowed here is much more restricted than in either of the two approaches mentioned.

In Sec. III we seek a quantitative representation of the AND and OR operations, i.e., a representation of the possible relations among various experimental setups. This is done by assigning a complex number to each setup in such a way that relations among setups translate into relations among the corresponding complex numbers. The assignment is highly constrained by a consistency requirement, expressed in terms of functional equations, that if the complex number associated to a given setup can be computed in more than one way, the various answers must agree. Solving the consistency constraints shows that all representations of AND and OR are actually equivalent to each other and that there is a particular choice that is singled out by its convenience. With this choice the AND and OR operations are represented by the product and sum of complex numbers, i.e., the product and sum of probability amplitudes [19].

After introducing, in Sec. IV, the concept of a state described by a wave function, we show how the product and sum rules imply the linearity of the Schrödinger equation. Then we address the issue of the physical interpretation of the formalism, that is, of how probability amplitudes are to be used in the prediction of outcomes of experiments and, in Secs. V and VI, we give a proof of Born's statistical postulate. Final comments appear in Sec. VII.

II. WHAT CAN WE SAY ABOUT A SIMPLE PARTICLE?

Suppose the only experiments we can perform are those that can detect the presence or absence of the particle in a sufficiently small region of space-time around an event $x=(\vec{x},t)$. Later, in Sec. VII, we will argue that this is not as restrictive as one might at first think. The propositions in which we are interested will typically describe motion. The simplest statement of this sort, "the particle moves from x_i to x_f," which we will denote by $[x_f,x_i]$, can be tested by preparing a particle at x_i and placing a detector at x_f. Our eventual goal is that of predicting the likelihood of a positive outcome of a test of $[x_f,x_i]$.

To analyze this problem we consider placing various obstacles in the path of the particle. This leads to propositions involving various constraints intermediate between x_i and x_f. Consider, for example, "the particle goes from x_i to x_f via the intermediate point x_1" (we assume that $t_i<t_i<t_f$) or, equivalently, "the particle goes from x_i to x_1 and from there to x_f". This we will denote by $[x_f,x_1,x_i]$. How could we test this proposition? We will certainly have to prepare the particle at the starting point x_i and place a detector at the final destination point x_f, but we cannot place a second detector at x_1; our particle is a delicate microscopic object, and our detectors are clumsy macroscopic devices that will totally alter the nature of the motion. All detections should be kept to the bare minimum: just one detection at the final destination point.

To carry out a test of $[x_f,x_1,x_i]$ we will imagine an experimental setup with a source at x_i, a detector at the final destination x_f and some sort of device that implements the constraint at x_1. Needless to say, we deal here with a highly idealized conceptual device used as an aid for reasoning rather than for actual experimentation; such devices are not unusual in theoretical physics. We imagine first an extended obstacle that blocks all paths that the particle could have taken through some arbitrary spacetime region [see Fig. 1(a)]. This already represents a considerable idealization; a more realistic obstacle would have fuzzy edges, regions of partial transparency rather than total opacity, and so on, but let us nevertheless proceed. The complications due to the arbitrary shape can be alleviated by imagining our obstacle as a succession of simpler obstacles each operating at a single time [see Fig. 1(b)]. The next step in idealization, shown in Fig. 1(c), is one of these single-time obstacles of infinitesimal spatial extent: It blocks all the paths passing through the spacetime point x_1.

In Figs. 2(a) and 2(b) we show the setups needed to test $[x_f,x_i]$ and $[x_f,x_1,x_i]$. The obstacle that implements the constraint at x_1 is the complement of the infinitesimal obstacle of Fig. 1(c). This idealized device we will call a "filter." It suddenly appears at time t_1, blocking the particle everywhere in space except for a small "hole" around the point $\vec{x}_1$, through which the particle may pass undisturbed.

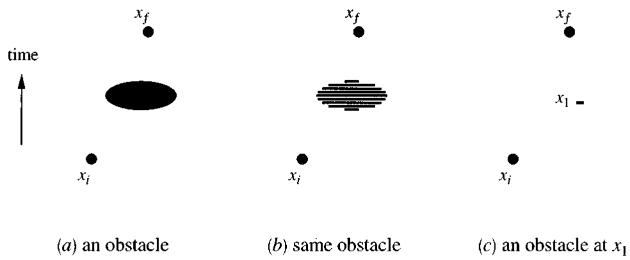

FIG. 1. (a) A generic obstacle of irregular shape and duration is placed in the path of the particle from x_i to x_f. (b) The obstacle can be envisioned as a succession of idealized obstacles of irregular shapes but essentially no duration. (c) An obstacle of infinitesimal extent; it blocks all the paths passing through a given spacetime point.

The filter lasts an infinitesimally short interval and then, just as suddenly, it disappears. The net result is that the filter prevents any motion from x_i to x_f except via the intermediate point x_1. If we want to impose more constraints, as in $[x_f,x_2,x_1,x_i]=$ ''the particle goes from x_i to x_f via the intermediate events x_1 and x_2,'' we introduce two filters, one at time t_1 and another at t_2 with holes at $\vec{x}_1$ and $\vec{x}_2$, respectively. The use of an infinite number of filters would allow us to specify completely the path followed by the particle.

The conceptual device of using these idealized filters allows us to introduce yet another kind of setup or proposition. Suppose that instead of having one hole in the filter at t_1 we open two holes, one at $\vec{x}_1$ and another at $\vec{x}_1'$ [see Fig. 2(c)]. This physical situation is one that one might *classically* describe as ''the particle goes from x_i to x_f via point x_1 or x_1'.'' Such a proposition we will denote by $[x_f,x_1,x_1',x_i]$. (We will generally use subscripts to label the times at which events or filters occur and superscripts or primes to distinguish events or holes that happen at the same time but at different locations.) Although it is not quite necessary, for the sake of clarity, we may wish to write $[x_f,(x_1,x_1'),x_i]$ where we have grouped together events that, being simultaneous, represent holes in the same filter.

The propositions we will consider will all be of the general form

$$a=[x_f,s_N,s_{N-1},\ldots,s_2,s_1,x_i], \tag{1}$$

where $s_n=(x_n,x_n',x_n'',\ldots)$ denotes a filter at time t_n, intermediate between t_i and t_f, with holes at $\vec{x}_n,\vec{x}_n',\vec{x}_n'',\ldots$. Statements such as $[s_N,s_{N-1},\ldots,s_2,s_1,x_i]$ or $[x_f,s_N,s_{N-1},\ldots,s_2,s_1]$ are not allowed. Two propositions will be considered equal when they represent the same experimental setup, i.e., the same distribution of filters and holes.

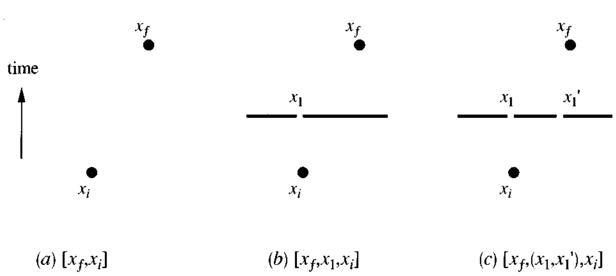

FIG. 2. Examples of simple propositions: (a) ''the particle moves from x_i to x_f,'' (b) ''the particle goes from x_i to x_1 and from there to x_f,'' and (c) ''the particle goes from x_i to x_f via x_1 or x_1'.''

Notice that Eq. (1) incorporates two crucial features of quantum theory: First, the allowed setups involve a single initial and a single final event where we can place a source and a detector [20]. This is a recognition that measurements and other interactions with macroscopic devices that induce uncontrollable disturbances must be avoided. Second, there is a one-to-one correspondence between the allowed statements and the idealized experimental setups with which we could test those propositions: All propositions are testable. In fact, we are identifying propositions with setups; this is a recognition that no statements can be made about the particle by itself independently of the experimental context.

From now on the words ''proposition'' and ''setup'' will be used interchangeably. In fact, we prefer the latter and will use it more often because, first, its use helps emphasize that the goal is to find out whether the detector at x_f will fire or not. Second, by avoiding statements about the particle itself we hope to eliminate misconceptions about what the particle is and what it is actually doing between source and detector. We are not saying that the particle is either a point particle or a wave, or both, or neither. We are not saying that it went through either one hole or through another, or even that it went through both holes at the same time. In fact, beyond the fact that the particle is capable of being emitted and detected we are not assuming much at all.

In attempting to predict the result of tests it seems reasonable to assume that if two propositions are related in some way (one proposition might, for example, be testable using a part of the setup used for the other), then information about one should be relevant to predictions about the other. Our next step will be to exhibit relations of this sort. This will allow us to use simple setups to build more complex ones, and conversely, also to analyze complex setups into simpler ones.

The basic relations that we wish establish are of two kinds. The first kind arises when two setups a and b can be placed in immediate succession. This results in a third setup, obviously ''related'' to the first two, which we will denote by ab. Notice that this operation, which we will call AND, cannot be used to combine any two arbitrarily chosen setups a and b. It is only when the destination point of the earlier setup coincides with the source point of the later setup that the combined ab is an allowed setup; a and b must be consecutive. The simplest instance of this is

$$[x_f,x_1][x_1,x_i]=[x_f,x_1,x_i], \tag{2}$$

and another example is shown in Fig. 3(a). In general,

$$[x_f,s_N,\ldots,s_{n+1},x_n][x_n,s_{n-1},\ldots,s_1,x_i]$$
$$=[x_f,s_N,\ldots,x_n,\ldots,s_1,x_i]. \tag{3}$$

Conversely, any proposition with a filter containing a single hole can be decomposed into two consecutive propositions. In the left member of Eq. (3) it is important that all

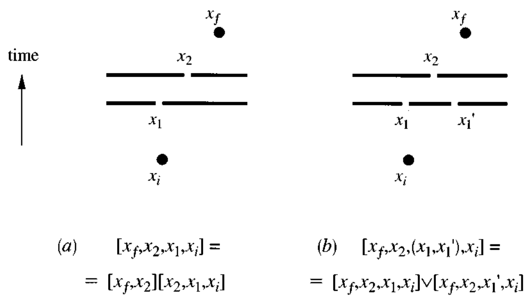

FIG. 3. Two examples of using AND and OR to construct complex propositions out of simpler ones.

$t_1,\ldots,t_{n-1}$ happen before t_n and that $t_{n+1},\ldots,t_N$ happen after t_n, otherwise the two setups overlap and are not consecutive.

The second useful kind of relation we consider arises when two setups a' and a'' are identical except on one single filter where none of the holes of a' overlap any of the holes of a''. We may then form a third setup a, denoted by $a'\vee a''$, which includes the holes of both a' and a''. A simple instance of this operation, which we will call OR, occurs in a "two-slit" experiment [see Fig. 2(c)].

$$[x_f,x_1',x_i]\vee[x_f,x_1'',x_i]=[x_f,(x_1',x_1''),x_i], \tag{4}$$

and another example is shown in Fig. 3(b). The general case is

$$[x_f,\ldots,s_n',\ldots,x_i]\vee[x_f,\ldots,s_n'',\ldots,x_i]=[x_f,\ldots,s_n,\ldots,x_i], \tag{5}$$

where

$$s_n'=(x_n'^1,x_n'^2,\ldots,x_n'^j),\quad s_n''=(x_n''^1,x_n''^2,\ldots,x_n''^k),$$

and

$$s_n=(x_n'^1,\ldots,x_n'^j,x_n''^1,\ldots,x_n''^k).$$

Again, notice that it is only for special choices of setups a and b that this OR operation will result in an allowed setup $a\vee b$.

The two symbols we have introduced, AND and OR, represent our presumed ability to construct more complex setups out of simpler ones. If one considers them as operations it is natural to ask if there are any rules that should be followed to manipulate them consistently. To obtain these rules we follow the principle, mentioned earlier, that two propositions are equal when they represent the same setup of filters and holes. The first rule is that the OR operation is commutative

$$a\vee b=b\vee a. \tag{6}$$

For the AND operation, however, there is an asymmetry implicit in the idea of placing setups in succession, one setup is the earlier one. It is convenient to incorporate this feature into the notation: if ab is an allowed setup,

$$ab\neq ba, \tag{7}$$

because ba is not allowed.

Next, we can see that both AND and OR enjoy a certain amount of associativity. For example, given three consecutive setups a, b, and c we can write

$$(ab)c=a(bc)\equiv abc, \tag{8}$$

provided ab and bc are allowed. In this case $(ab)c$ and $a(bc)$ are automatically allowed but ac is not because a and c are not consecutive. Similarly, for the OR operation we have

$$(a\vee b)\vee c=a\vee(b\vee c)\equiv a\vee b\vee c, \tag{9}$$

provided all four setups $(a\vee b)$, $(b\vee c)$, $(a\vee b)\vee c$, and $a\vee(b\vee c)$ are allowed. In this case Eq. (9) is also equal to $(a\vee c)\vee b$. Notice that any differences between the three setups a, b, and c must be found in one and the same filter. Otherwise $(a\vee b)\vee c\neq a\vee(b\vee c)$ because if the member on the left is allowed the one on the right is not.

The last important rule is that of distributivity. This may take the form

$$a(b\vee c)=(ab)\vee(ac)\quad\text{or}\quad(b\vee c)a=(ba)\vee(ca). \tag{10}$$

Which equality holds, if any, will depend on whether the relevant propositions are allowed. Both equalities cannot hold simultaneously.

An important illustration of the use and utility of the AND and the OR operations arises from the observation that a single filter that is totally covered with holes is equivalent to having no filter at all. In other words, the absence of a filter at time t_1 is a special kind of filter σ_1, which may be freely introduced into any proposition (provided $t_i<t_1<t_f$). For example, using Eqs. (3) and (5), we have

$$\begin{aligned}[x_f,x_i]&=[x_f,\sigma_1,x_i]=\bigvee_{\text{all }\vec{x}_1}[x_f,x_1,x_i]\\&=\bigvee_{\text{all }\vec{x}_1}([x_f,x_1][x_1,x_i]),\end{aligned} \tag{11}$$

and, introducing additional σ filters at times $t_2,\ldots,t_N$ we get

$$[x_f,x_i]=\bigvee_{\text{all }\vec{x}_N}\cdots\bigvee_{\text{all }\vec{x}_2}\bigvee_{\text{all }\vec{x}_1}([x_f,x_N]\cdots[x_2,x_1][x_1,x_i]). \tag{11$'$}$$

This shows how motion over a long distance can be analyzed in terms of motion over shorter steps.

One cannot fail to see some similarity between our quantum AND and OR operations with the Boolean operations AND and OR, which also happen to be commutative, associative, and distributive and are also used to construct more complex propositions out of simpler ones. But the similarity ends there: The quantum AND and OR introduced here are not logical but rather physical connectives, they are used to describe the relative dispositions of various pieces of equipment. Also, and perhaps more important, is the fact that the Boolean operations (and this applies as well to the AND and OR introduced in most quantum logics) can connect any two arbitrary propositions, while the conditions on allowed setups impose severe restrictions on the action of the quantum AND and OR.

III. AMPLITUDES: THE SUM AND PRODUCT RULES

Our goal is to predict the outcomes of experiments and the strategy is to establish a network of relations among setups in the hope that information about some setups might be helpful in making predictions about others. Our next step will be to obtain a quantitative representation of these relations.

Suppose each setup a is assigned a complex number $\phi(a)$. By a ''representation'' we mean that the assignment of ϕ is such that relations among physical setups should translate into relations among the complex numbers associated to them. Why should such a representation exist? It need not, but all of physics consists of representing elements of reality, or relations among these elements, or our information about them, by mathematical objects. The existence of such representations may be mysterious, but it is not surprising; there are too many examples. A second, simpler question is why do we seek a representation in terms of complex numbers? Again, there is no answer here; this is an unexplained feature of quantum theory. It seems that a single complex number is sufficient to convey the physically relevant information about a setup.

To be specific consider a ''double-slit'' experiment. The relation between $[x_f,(x_1,x_1'),x_i]$ and its components $[x_f,x_1,x_i]$ and $[x_f,x_1',x_i]$ will be represented as a relation between the corresponding complex numbers $\phi(x_f,(x_1,x_1'),x_i)$, $\phi(x_f,x_1,x_i)$, and $\phi(x_f,x_1',x_i)$. What we require is that there exist a function S such that

$$\phi(x_f,(x_1,x_1'),x_i)=S(\phi(x_f,x_1,x_i),\phi(x_f,x_1',x_i)), \quad (12)$$

and that this same function applies to any other setups that are similarly related. More generally, if the setups associated to a and to a' are such that $a\vee a'$ is an allowed setup then

$$\phi(a\vee a')=S(\phi(a),\phi(a')). \quad (13)$$

Thus, the function S is a representation of the relation OR.

The requirement that S should exist is a strong constraint on the allowed assignment of ϕ. Consider, for example, the number ϕ assigned to $a\vee a'\vee a''$. Using associativity this can be calculated in two different ways, either as $\phi((a\vee a')\vee a'')$ or as $\phi(a\vee(a'\vee a''))$. Consistency requires that the two ways agree,

$$S(\phi(a\vee a'),\phi(a''))=S(\phi(a),\phi(a'\vee a'')). \quad (14)$$

Using S once again one obtains the following consistency constraint:

$$S(S(u,v),w)=S(u,S(v,w)), \quad (15)$$

where we have let $\phi(a)=u$, $\phi(a')=v$, and $\phi(a'')=w$.

One can check, by substitution, that the associativity constraint, Eq. (15), is satisfied if

$$S(u,v)=\xi^{-1}(\xi(u)+\xi(v)) \quad \text{or} \quad \xi(S(u,v))=\xi(u)+\xi(v), \quad (16)$$

where ξ is an arbitrary function. In the Appendix we give a proof (similar to Cox's [1]) that this is also the general solution. In other words, Eq. (16) tells us what forms the function S may take, and conversely, that if the function S exists then there must also exist another function ξ, calculable from S, such that

$$\xi(\phi(a\vee a'))=\xi(\phi(a))+\xi(\phi(a')). \quad (17)$$

This is remarkable. It immediately suggests that instead of the original representation in terms of the complex numbers $\phi(a)$, we should opt for an equivalent, simpler, and more convenient representation in terms of the numbers $\xi(\phi(a))$. In other words, the consistent assignment of complex numbers $\xi(a)$ to propositions a can always be done so that the OR operation is represented by a simple sum rule,

$$\xi(a\vee a')=\xi(a)+\xi(a'). \quad (18)$$

In this representation S is addition.

Next we turn our attention to the representation of the AND operation. From this point onward Cox's treatment and ours differ. Cox focused on the operation of negating a proposition, and was thus led to consider the consistency requirement ensuing from the possibility of double negation. Negation is not an operation available to us, we rather choose to concentrate on the associative and distributive properties of AND.

Consider for example, a particle that goes from an initial x_i to a final x_f via an intermediate point x. We want to represent the relation between $[x_f,x,x_i]$ and its components $[x_f,x]$ and $[x,x_i]$ as a relation between the complex numbers $\xi(x_f,x,x_i)$, $\xi(x_f,x)$, and $\xi(x,x_i)$. We then require that there exists a function P such that

$$\xi(x_f,x,x_i)=P(\xi(x_f,x),\xi(x,x_i)), \quad (19)$$

and that the same function P applies to any other propositions that are similarly related. Specifically, if ab is any allowed proposition we require that

$$\xi(ab)=P(\xi(a),\xi(b)), \quad (20)$$

so that the function P is a representation of the AND operation.

The functional form of P is highly constrained by the requirement that P should exist for any a and b such that ab is allowed. We can repeat the argument we used earlier for the OR operation: the number ξ associated to the (allowed) proposition abc can be computed in two ways, either as $\xi((ab)c)$ or as $\xi(a(bc))$, and these should agree. Therefore, P must satisfy the associativity constraint

$$P(P(u,v),w)=P(u,P(v,w)). \quad (21)$$

Furthermore, AND and OR are not unrelated: using distributivity, the number ξ associated to the (allowed) proposition $a(b\vee c)$ can also be computed in two ways, either as $\xi(a(b\vee c))$ or as $\xi((ab)\vee(ac))$. Therefore, using Eq. (18),

$$P(\xi(a),\xi(b\vee c))=\xi(ab)+\xi(ac), \quad (22)$$

and, using P and S once again, we conclude that left distributivity leads to the following constraint:

$$P(u,v+w)=P(u,v)+P(u,w), \quad (23)$$

where $\xi(a)=u$, $\xi(b)=v$, and $\xi(c)=w$. Similarly, from propositions of the form $(a\vee b)c$ for which right distributivity holds we obtain

$$P(u+v,w)=P(u,w)+P(v,w). \tag{23$'$}$$

The solution of these distributivity constraints is trivial. Differentiating Eq. (23) with respect to v and w and letting $v+w=z$ gives

$$\frac{\partial^2}{\partial z^2}P(u,z)=0, \tag{24}$$

so that P is linear in its second argument, $P(u,v)=A(u)v+B(u)$. Substituting back into Eq. (23) gives $B(u)=0$. Similarly, Eq. (23′) implies that P is linear in its first argument, therefore

$$P(u,v)=Cuv \quad \text{or} \quad \xi(ab)=C\xi(a)\xi(b). \tag{25}$$

The associativity constraint, Eq. (21), is automatically satisfied. The constant C can be absorbed into yet a new number $\psi(a)=C\xi(a)$, so that the AND operation can be conveniently represented by a simple product rule,

$$\psi(ab)=\psi(a)\psi(b), \tag{26}$$

while the sum rule remains unaffected,

$$\psi(a\vee b)=\psi(a)+\psi(b). \tag{27}$$

Complex numbers assigned in this particularly convenient way will be called ‘‘amplitudes.’’

Let us summarize the results of this section: A quantitative representation of the relations between setups can be obtained by assigning a complex number $\psi(a)$ to each proposition a. Because of the crucial requirement of consistency the considerable arbitrariness in the actual choice of $\psi(a)$ is largely illusory; it turns out that all representations are equivalent to each other; i.e., they are obtained from each other by mere ‘‘changes of variables.’’ Although all consistent assignments are equally correct in the sense that they serve our purpose of providing the desired representation, some are singled out by their sheer convenience. They lead to representations of the quantum AND and the OR operations that take very simple forms: products and sums. This is the central result of this paper.

IV. WAVE FUNCTIONS AND THE LINEARITY OF TIME EVOLUTION

Amplitudes have been introduced in the last section as the natural concept to describe experiments quantitatively but we have not yet indicated how the knowledge of an amplitude is to be used in predicting the outcomes of experiments. In order to suggest, in the following section, how amplitudes are to be interpreted, we will first explore, along conventional lines [11], the properties of the amplitude $\psi(x_f,x_i)$ associated to the basic proposition $[x_f,x_i]$.

We had seen earlier, in Eq. (11), how to analyze a motion from x_i to x_f in terms of motion over shorter steps from x_i to x and from there to x_f. Now we can express this in terms of probability amplitudes; using the sum and product rules, we get

$$\psi(x_f,x_i)=\sum_{\text{all }\vec{x}\text{ at }t}\psi(x_f,x,x_i)=\sum_{\text{all }\vec{x}\text{ at }t}\psi(x_f,x)\psi(x,x_i). \tag{28}$$

The sums are a consequence of restricting the positions $\vec{x}$ to sites on a discrete lattice; the generalization to a more realistic continuum where the sums are replaced by integrals is straightforward. Equation (28) or, more explicitly,

$$\psi(\vec{x}_f,t_f;\vec{x}_i,t_i)=\sum_{\text{all }\vec{x}\text{ at }t}\psi(\vec{x}_f,t_f;\vec{x},t)\psi(\vec{x},t;\vec{x}_i,t_i), \tag{28$'$}$$

describes time evolution and *therefore* holds the key to the question of the physical interpretation. To see this it is convenient to introduce the notion of a state described by a wave function.

Suppose that a particle starts at $(\vec{x}_i,t_i)$ and prior to time t it undergoes various interactions the net result of which is that the amplitude to reach the point $\vec{x}$ at time t is given by $\Psi(\vec{x},t)$. Of course, $\Psi(\vec{x},t)$ is numerically equal to $\psi(\vec{x},t;\vec{x}_i,t_i)$, but situations are common where we know $\Psi(\vec{x},t)$ and either we have no interest or have lost track of what happened before t. In these cases we may streamline the notation and replace $\psi(\vec{x},t;\vec{x}_i,t_i)$ by $\Psi(\vec{x},t)$. Using this knowledge of $\Psi(\vec{x},t)$ we can calculate $\psi(x_f,x_i)$ directly. From Eq. (28),

$$\psi(x_f,x_i)\equiv\Psi(\vec{x}_f,t_f)=\sum_{\text{all }\vec{x}\text{ at }t}\psi(\vec{x}_f,t_f;\vec{x},t)\Psi(\vec{x},t). \tag{29}$$

The function $\Psi(\vec{x},t)$, called the wave function, represents all those features of interactions previous to time t that are relevant to the prediction of evolution after t. We might, by abuse of (classical) language, say that Ψ describes the state of the particle at time t, and that the effect of those interactions prior to t has been to *prepare* the particle in state Ψ.

Let us return to the description of time evolution implicit in Eq. (29). Differentiating with respect to t_f and evaluating at $t_f=t$ we get

$$\frac{\partial\Psi(\vec{x}_f,t)}{\partial t}=\sum_{\text{all }\vec{x}\text{ at }t}\left.\frac{\partial\psi(\vec{x}_f,t';\vec{x},t)}{\partial t'}\right|_{t'=t}\Psi(\vec{x},t).$$

The derivative on the right is a function of $\vec{x}_f$, $\vec{x}$, and of t. If we define

$$\left.\frac{\partial\psi(\vec{x}_f,t';\vec{x},t)}{\partial t'}\right|_{t'=t}\equiv-\frac{i}{\hbar}H(\vec{x}_f,\vec{x},t),$$

then

$$i\hbar\frac{\partial\Psi(\vec{x}_f,t)}{\partial t}=\sum_{\text{all }\vec{x}\text{ at }t}H(\vec{x}_f,\vec{x},t)\Psi(\vec{x},t), \tag{30}$$

which is recognized as the Schrödinger equation. We might not yet know what Ψ means, nor what the Hamiltonian H should be, but we have obtained an important result: Once certain natural consistency requirements are accepted, the time evolution of quantum states is given by a Schrödinger equation, which is necessarily *a linear* equation.

The conclusion is clear: Nonlinear variants of quantum mechanics that preserve the notion of amplitudes violate natural requirements of consistency. The question of whether it is possible to formulate nonlinear versions of quantum mechanics should not be formulated as a dynamical question about which nonlinear terms one is allowed to add to the Schrödinger equation, but rather it should be phrased as a deeper kinematical question about whether quantum mechanics should be formulated in terms of mathematical objects other than amplitudes and wave functions. However, whatever the nature of those mathematical objects the requirement that they be manipulated in a consistent manner should be maintained.

V. PHYSICAL INTERPRETATION: BORN'S STATISTICAL POSTULATE

Having established rules for the consistent manipulation of amplitudes we can, finally, address the question of how to use them to predict the outcomes of experiments. The key to finding a physical interpretation for wave functions or, equivalently, for amplitudes, is the time evolution equation [Eq. (28) or (29)]. We reason as follows:

Consider first a special case. Suppose that as a result of a very special preparation procedure (between times t_i and t) the wave function $\Psi(\vec{x},t)$ at time t vanishes everywhere except at a single point $\vec{x}_0$,

$$\Psi(\vec{x},t)=A\,\delta_{\vec{x},\vec{x}_0}. \tag{31}$$

Next, place a filter at time t with a single hole at $\vec{x}_0$. It is easy to see [from Eq. (28) or (29)] that the presence or absence of this filter has absolutely no influence on the subsequent evolution of the wave function or on the amplitude to arrive at any final destination point x_f. Since relations among amplitudes are meant to reflect corresponding relations among setups, it seems natural to assume that the presence or absence of the filter will have no effect on whether detection at x_f occurs or not. This, in turn, suggests that even in the absence of the filter, at time t the particle must have been at $\vec{x}_0$ and nowhere else. Therefore the special case where $\Psi(\vec{x},t)\propto\delta_{\vec{x},\vec{x}_0}$ is interpreted as ''at time t the particle is located at $\vec{x}_0$.''

Essentially the same argument can be applied in a variety of other cases. For example, if as a result of the preparation procedure $\Psi(\vec{x},t)$ vanishes at a certain point $\vec{x}\,'$ then placing a filter with holes everywhere except at $\vec{x}\,'$ should have no effect on the subsequent evolution of Ψ. Therefore $\Psi(\vec{x}\,',t)=0$ is interpreted as ''at time t the particle is not at $\vec{x}\,'$.''

This leads naturally to the following general interpretative postulate: Consider a filter the action of which is to block out those components of the wave function characterized by a certain feature (what that feature is should be obvious to whoever built the filter). Suppose that at time t the system is in a state of wave function $\Psi(t)$. If the introduction or removal of the filter at t has no effect on the future evolution of the wave function then the rule of interpretation is that the system in state $\Psi(t)$ does not exhibit the feature in question. The rule applies to amplitudes in general.

The power of this rule becomes apparent when applied to a particle with a generic wave function

$$\Psi(\vec{x},t)=\sum_i A_i\,\delta_{\vec{x},\vec{x}_i}, \tag{32}$$

where the number and location of the $\vec{x}_i$'s is arbitrary. We want to predict the outcome of an experiment in which a detector is placed at a certain $\vec{x}_k$. If $\vec{x}_k$ differs from all of the $\vec{x}_i$'s in the sum in Eq. (32) the interpretative rule directly implies that the particle will definitely not be detected.

The interesting problem arises when $\vec{x}_k$ coincides with one of the $\vec{x}_i$'s. In this case, as expected, one cannot predict the actual outcome of the experiment. What is predictable, and quite precisely, in fact, is the probability of various outcomes. At this point we will make an important assumption about the wave function: We will assume that it can be normalized. For convenience we will from now on assume that Ψ in Eq. (32) has been appropriately normalized,

$$\|\Psi\|^2=(\Psi,\Psi)\equiv\sum_{\vec{x}}|\Psi(\vec{x},t)|^2=\sum_i|A_i|^2=1. \tag{33}$$

Next we show that the probability of detection at $\vec{x}_k$ is $|A_k|^2$. Thus Born's statistical interpretation is actually a theorem: It follows from a simpler interpretative rule that only refers to situations of absolute certainty. This remarkable fact was discovered long ago by Finkelstein [15], Hartle [16], and Graham [17]. The proof below is particularly suited to the approach to quantum theory being developed in this paper.

Consider an ensemble of N identically prepared, independent replicas of our particle; later we will take $N\rightarrow\infty$. In the next section we will show that the wave function for this N-particle system is the product

$$\begin{aligned}\Psi_N(\vec{x}_1,\ldots,\vec{x}_N,t)&=\prod_{\alpha=1}^{N}\Psi(\vec{x}_\alpha,t)\\&=\sum_{i_1\cdots i_N}A_{i_1}\cdots A_{i_N}\delta_{\vec{x}_1,\vec{x}_{i_1}}\cdots\delta_{\vec{x}_N,\vec{x}_{i_N}}.\end{aligned} \tag{34}$$

Suppose that in the N-particle configuration space we place a special filter, denoted by P_n^k, the action of which is to block all components of Ψ_N except those where exactly n of the N replicas are at $\vec{x}_k$. The wave function right after this filter is

$$P_n^k\Psi_N=\sum_{i_1\cdots i_N}\delta_{n,n_k}A_{i_1}\cdots A_{i_N}\delta_{\vec{x}_1,\vec{x}_{i_1}}\cdots\delta_{\vec{x}_N,\vec{x}_{i_N}}, \tag{35}$$

where

$$n_k=\sum_{\alpha=1}^{N}\delta_{k,i_\alpha}. \tag{36}$$

Actually this filter is too strict, it selects a single sharply defined fraction $f=n/N$. What we need is a more lenient filter (presumably built by opening additional ''holes'' in

P_n^k) that allows passage of all fractions in a range from $f-\epsilon$ to $f+\epsilon$. The action of this filter is described by

$$P_{f,\epsilon}^k \Psi_N = \sum_{i_1 \cdots i_N} \left(\sum_{n=(f-\epsilon)N}^{(f+\epsilon)N} \delta_{n,n_k} \right) A_{i_1} \cdots A_{i_N} \delta_{\vec{x}_1,\vec{x}_{i_1}} \cdots \delta_{\vec{x}_N,\vec{x}_{i_N}}. \tag{37}$$

We are now ready to apply our interpretative rule: If, as $N \rightarrow \infty$, the presence of the filter $P_{f,\epsilon}^k$ is found to have no influence whatsoever on the future evolution of the wave function Ψ_N we will interpret Ψ_N as representing a state for which the fractions of replicas at $\vec{x}_k$ must lie in the range from $f-\epsilon$ to $f+\epsilon$. To show that this is actually the case we must show that as $N \rightarrow \infty$ the wave function right after the filter, $P_{f,\epsilon}^k \Psi_N$, becomes more and more similar to the wave function right before the filter, Ψ_N. To this end we calculate the norm

$$\| P_{f,\epsilon}^k \Psi_N - \Psi_N \|^2 \equiv \sum_{\vec{x}_1 \cdots \vec{x}_N} | P_{f,\epsilon}^k \Psi_N - \Psi_N |^2. \tag{38}$$

Since filters act as projectors, $PP=P$, we get

$$\| P_{f,\epsilon}^k \Psi_N - \Psi_N \|^2 \equiv 1 - (\Psi_N, P_{f,\epsilon}^k \Psi_N). \tag{39}$$

The calculation of the scalar product is straightforward,

$$\begin{aligned}(\Psi_N, P_{f,\epsilon}^k \Psi_N) &= \sum_{\vec{x}_1 \cdots \vec{x}_N} \Psi_N^* P_{f,\epsilon}^k \Psi_N \\ &= \sum_{n=(f-\epsilon)N}^{(f+\epsilon)N} \left(\sum_{i_1 \cdots i_N} \delta_{n,n_k} |A_{i_1}|^2 \cdots |A_{i_N}|^2 \right).\end{aligned} \tag{40}$$

The sum over $i_1,\ldots,i_N$ is done as follows: Suppose we satisfy the Kronecker δ_{n,n_k} constraint by choosing n of the N indices $i_1,\ldots,i_N$ and setting them to the value k. Since the individual Ψ's are normalized each sum over the remaining $N-n$ indices gives

$$\sum_{i \neq k} |A_i|^2 = 1 - |A_k|^2. \tag{41}$$

But there are $\binom{N}{n}$ ways to choose which n indices are set equal to k, therefore

$$(\Psi_N, P_{f,\epsilon}^k \Psi_N) = \sum_{n=(f-\epsilon)N}^{(f+\epsilon)N} \binom{N}{n} (|A_k|^2)^n (1-|A_k|^2)^{N-n}. \tag{42}$$

For large N this binomial sum tends to the integral of a Gaussian,

$$(\Psi_N, P_{f,\epsilon}^k \Psi_N) = \int_{f-\epsilon}^{f+\epsilon} \frac{1}{\sqrt{2\pi\sigma_N^2}} \exp\left(\frac{(f' - \bar{f}\,)^2}{2\sigma_N^2} \right) df', \tag{43}$$

with mean $\bar{f} = |A_k|^2$ and variance $\sigma_N^2 = \bar{f}(1-\bar{f}\,)/N$. In the limit $N \rightarrow \infty$ this is more concisely written as a δ function, therefore,

$$\lim_{N\rightarrow\infty} \| P_{f,\epsilon}^k \Psi_N - \Psi_N \|^2 \equiv 1 - \int_{f-\epsilon}^{f+\epsilon} \delta(f' - |A_k|^2) df'. \tag{44}$$

The interpretation is clear: As $N \rightarrow \infty$ the filter $P_{f,\epsilon}^k$ will have no effect on the wave function Ψ_N provided f lies in a range 2ϵ about $|A_k|^2$, and according to our interpretative rule Ψ_N cannot contain any fractions outside this range. Choosing stricter filters with $\epsilon \rightarrow 0$ we conclude that as $N \rightarrow \infty$, Ψ_N is a state for which the fraction of replicas at $\vec{x}_k$ is exactly $|A_k|^2$.

Returning to the original single-particle system, we see that we cannot predict whether a detection at $\vec{x}_k$ will occur or not. In fact, we have a very definite prediction of indeterminism: for large N detection will certainly occur for a fraction $|A_k|^2$ and, equally significant, detection will certainly not occur for a fraction $1-|A_k|^2$. Once the assumption is made that the relations among possible experimental setups are quantitatively represented by consistently assigned amplitudes, the general interpretative rule implies indeterminism. The best one can do is assign a probability p to this detection. Given the information that for a large number of identically prepared systems the fraction of successful detections is $|A_k|^2$ the only assignment consistent with the law of large numbers is the value

$$p = |A_k|^2. \tag{45}$$

To complete our proof of Born's postulate we must prove that the wave function for N independent particles is the product of the wave functions for each one of the particles. This is the topic of the next section.

VI. SEVERAL INDEPENDENT PARTICLES

We want to show that the wave function of a system $\alpha\beta$ composed of two independent particles α and β is the product of the wave functions for each particle, $\Psi_{\alpha\beta} = \Psi_\alpha \Psi_\beta$. In the spirit of the previous sections the first step must be that of defining the statements about composite systems in terms of the experimental setups designed to test them and of providing a representation in terms of amplitudes of the relations among those setups.

Our system is composed of two independent particles. The notion of independence imposes highly nontrivial constraints. Suppose that the allowed propositions about particle α by itself and the corresponding setups designed to test them are denoted by a, as in Eq. (1), and similarly, that the allowed propositions and setups about particle β are denoted by b. Then the first condition implied by independence is that the statements c about the composite $\alpha\beta$ are restricted to those that can be tested by composite setups that separately test a about α and b about β. Thus the setups allowed for $\alpha\beta$ are of the form $c = \{a;b\}$.

The various ways in which the composite setups c can be combined can be derived from the various ways in which the a's and the b's can be combined among themselves. Thus, if $c_1 = \{a_1;b_1\}$ and $c_2 = \{a_2;b_2\}$ and if $a_1 \vee a_2$ is allowed and $b_1 = b_2$ then we define OR by $c_1 \vee c_2 \equiv \{a_1 \vee a_2, b_1\}$. On the other hand if it is $b_1 \vee b_2$ that is allowed and $a_1 = a_2$ then $c_1 \vee c_2 \equiv \{a_1, b_1 \vee b_2\}$. In the general case $c_1 \vee c_2$ will be an allowed setup only if c_1 and c_2 differ in one and only one filter; if the b setups differ then the a's must be identical and vice versa. Similarly, if both $a_1 a_2$ and $b_1 b_2$ are allowed then we define $c_1 c_2 \equiv \{a_1 a_2, b_1 b_2\}$. Commutativity, associativity, and distributivity for the AND and OR relations among the c setups follow from the corresponding properties for the a and b setups. The argument of Sec. III can now be repeated: the relations among different composite setups can be conveniently represented quantitatively by assigning a complex amplitude $\psi(c)$ to each setup c in such a way that the sum and product rules hold.

The second crucial condition implicit in the notion of independence, one that goes beyond the mere capability of independently placing filters in the paths of α and β, is the requirement that changing the filters acting on α, i.e., changing a to a', shall have no influence whatsoever on the outcome of b, and vice versa. Since relations among amplitudes are meant to reflect corresponding relations among setups, and the physically relevant information about setups a and b is contained in $\psi(a)$ and $\psi(b)$ this second condition can be quantitatively expressed by the requirement that the physically relevant information about setup c, expressed by $\psi(c)$, be some function of $\psi(a)$ and $\psi(b)$ and nothing else. Thus,

$$\psi(c) = F(\psi(a), \psi(b)). \tag{46}$$

This is what we mean by independence [21].

The function F is determined from the fact that not only $\psi(a)$ and $\psi(b)$ but also $\psi(c)$ must satisfy the sum and product rules. For example, if $c_1 = \{a_1 ; b_1\}$ and $c_2 = \{a_2 ; b_2\}$ and if $a_1 \vee a_2$ is allowed and $b_1 = b_2$ then $\psi(c_1 \vee c_2) = \psi(c_1) + \psi(c_2)$ implies

$$F(\psi(a_1) + \psi(a_2), \psi(b_1)) = F(\psi(a_1), \psi(b_1)) + F(\psi(a_2), \psi(b_1)) \tag{47}$$

so that

$$F(u+v, w) = F(u,w) + F(v,w). \tag{48}$$

Similarly, if $b_1 \vee b_2$ is allowed and $a_1 = a_2$ then

$$F(u, v+w) = F(u,v) + F(u,w). \tag{48$'$}$$

Finally, if $c_1 c_2 \equiv \{a_1 a_2, b_1 b_2\}$ is allowed, from the product rule $\psi(c_1 c_2) = \psi(c_1)\psi(c_2)$ we get

$$F(uv, wz) = F(u,w) F(v,z). \tag{49}$$

Equations (48) and (48′) are formally identical with equations (23) and (23′). Therefore $F(u,v) = Cuv$. Substituting into Eq. (49) we get $C = 1$, $F(u,v) = uv$.

Therefore, if particles α and β are independent, the amplitude associated to $c = \{a;b\}$ is the product of the amplitudes associated to a and to b,

$$\psi(c) = \psi(a)\psi(b). \tag{50}$$

The generalization to N independent particles is straightforward.

VII. CONCLUSIONS AND SOME COMMENTS

Let us summarize our main conclusions: After having identified a restricted set of allowed propositions in terms of the experimental setups designed to test them we introduce amplitudes as the essentially unique tool to carry out consistent calculations. The rules of manipulation are necessarily such that time evolution is described by a linear Schrödinger equation and that the Born probability interpretation holds.

It might seem surprising that a substantial amount of quantum mechanics has been reproduced without any mention of commutation relations among incompatible observables and of the corresponding uncertainty relations; in fact, we have only discussed the measurement of position. However, observables other than position are useful concepts in that they facilitate the description of experiments and the manipulation of information. Although they have not been central in this formulation of quantum theory it may be worthwhile to remark on how they arise. In general, they originate from the idea that by placing various filters, diffraction gratings, magnetic fields, etc., prior to the final position detection at x_f one can effectively build a more complex detector. This raises two questions; the first is what does such a complex device actually measure. The answer [11] is that what is measured is the extent to which the actual wave function Ψ resembles another wave function Φ, which is a characteristic of the detector. Should the resemblance be complete the particle would be detected at x_f with absolute certainty. The second question is what properties are we actually interested in measuring. Typically, interesting measurements will result in information useful for prediction in other experiments, and foremost among these is the measurement of properties that have some lasting value, i.e., conserved quantities. A related question is that of deciding how the Hamiltonian should be chosen; this choice, like that of most other observables, is dictated by symmetries [12,22] and will not be further pursued here.

For clarity we have focused our attention on the special case of a single particle moving in a discrete lattice. But it should be clear that the argument can be generalized to considerably more complicated configuration spaces. The crucial feature is to identify the relevant propositions or, equivalently, the idealized experimental setups designed to test them, and verify that the appropriate rules of associativity and distributivity hold. It is of some interest that in order to implement the associativity constraint one requires a configuration space that consists at least of three values. Remarkably, a similar restriction to spaces of three dimensions or more appears also in the work of Gleason [14]. This does not, of course, represent a problem: two-valued configuration spaces are unphysical. For example, realistic spin-$\frac{1}{2}$ systems also have translational and a variety of other degrees of freedom.

It is possible that the use of more complex detectors (i.e., other observables) will permit extending the set of allowed propositions. Perhaps this would bridge the gap between the present formulation and the quantum logic approach [12,13,15]. However, whether such an extension is advantageous is not at all clear. It would surely spoil the distributivity property (of AND and OR) that has played such a crucial role here.

Many are the questions left open. Most, like the application of quantum mechanics to the detectors themselves, as well as to other macroscopic objects, the nature of the classical limit, and other issues associated with decoherence through interaction with the environment, are common to all approaches to quantum theory. But some questions seem more urgent in this formulation. A particularly glaring one is the following: Why complex numbers? Perhaps other mathematical objects with the appropriate associative and distributive algebras (e.g., quaternions, multivectors, etc.) [23] could be used to obtain representations of the AND and OR operations.

We conclude with a couple of brief comments. The first concerns the similarity between quantum motion and Brownian motion, or between the Schrödinger equation and the diffusion equation. There is a natural reticence to dismiss it as a mere coincidence, and it has been suggested that perhaps there is some underlying stochastic physical agent responsible for the peculiar features of quantum motion. Our results suggest that such a physical agent need not exist, that the similarities between quantum and Brownian theories arise from formalisms that are strongly constrained by similar logical requirements of consistency, which force one to manipulate amplitudes in one case, and probabilities in the other, using similar sum and product rules.

The second comment addresses another aspect of the robustness of quantum theory, its universality. Quantum mechanics applies to a wide variety of systems over a broad range of energy and distance scales. Were new exotic objects (say, new excitations in condensed matter, or new particles, or strings) to be discovered, could we expect them to be described by quantum mechanics? Classical mechanics fails at atomic scales; how short a distance can we go and still expect quantum mechanics to hold? Or, in other words: What are the accepted features of today's physics that could reasonably be expected to hold in the future physics of objects that are yet to be discovered, or of energy and distance scales that are yet to be explored? It seems natural to assume that any list of such features should prominently include those derivable from mere consistency requirements; linearity and indeterminism are likely to be among them.

ACKNOWLEDGMENTS

I am indebted to C. Rodriguez, P. Zambianchi, A. Inomata, and J. Kimball for valuable discussions and many insightful remarks.

APPENDIX: SOLUTION OF THE CONSISTENCY EQUATIONS

Our approach to solving the associativity equation (15),

$$S(S(u,v),w)=S(u,S(v,w)), \tag{A1}$$

is essentially that due to Cox [1]. A minor difference is that we deal with complex rather than real variables. Let $r=S(u,v)$, $s=S(v,w)$, $S_1(u,v)=\partial S(u,v)/\partial u$, and $S_2(u,v)=\partial S(u,v)/\partial v$. Then Eq. (A1) and its derivatives with respect to u and v are

$$S(r,w)=S(u,s), \tag{A1$'$}$$

$$S_1(r,w)S_1(u,v)=S_1(u,s), \tag{A2}$$

and

$$S_1(r,w)S_2(u,v)=S_2(u,s)S_1(v,w). \tag{A3}$$

Eliminating $S_1(r,w)$ from these last two equations we get

$$G(u,v)=G(u,s)S_1(v,w), \tag{A4}$$

where

$$G(u,v)=\frac{S_2(u,v)}{S_1(u,v)}. \tag{A5}$$

Multiplying Eq. (A4) by $G(v,w)$ we get

$$G(u,s)G(v,w)=G(u,s)S_2(v,w). \tag{A6}$$

Differentiating the right hand side of Eq. (A6) with respect to v and comparing with the derivative of Eq. (A4) with respect to w, we have

$$\begin{aligned}\frac{\partial}{\partial v}\,(G(u,s)S_2(v,w))&=\frac{\partial}{\partial w}\,(G(u,s)S_1(v,w))\\&=\frac{\partial}{\partial w}\,(G(u,v))=0.\end{aligned} \tag{A7}$$

Therefore

$$\frac{\partial}{\partial v}\,(G(u,v)G(v,w))=0, \tag{A8}$$

or

$$\frac{1}{G(u,v)}\frac{\partial G(u,v)}{\partial v}=-\frac{1}{G(v,w)}\frac{\partial G(v,w)}{\partial v}\equiv h(v). \tag{A9}$$

Integrating, we get

$$G(u,v)=G(u,0)\exp\left(\int_0^v h(v')dv'\right), \tag{A10}$$

and also

$$G(v,w)=G(0,w)\exp\left(-\int_0^v h(v')dv'\right), \tag{A11}$$

so that

$$G(u,v)=c\,\frac{H(u)}{H(v)}, \tag{A12}$$

where $c=G(0,0)$ is a constant and

$$H(u)=\exp\left(-\int_0^u h(u')du'\right). \tag{A13}$$

On substituting back into Eqs. (A4) and (A6) we get

$$S_1(v,w)=\frac{H(s)}{H(v)}\quad\text{and}\quad S_2(v,w)=c\,\frac{H(s)}{H(w)}. \tag{A14}$$

But $s=S(v,w)$, so substituting Eq. (A14) into $ds=S_1(v,w)dv+S_2(v,w)dw$ we get

$$\frac{ds}{H(s)}=\frac{dv}{H(v)}+c\,\frac{dw}{H(w)}. \tag{A15}$$

This is easily integrated. Let

$$\xi(u)=\xi(0)+\int_0^u \frac{du'}{H(u')}, \tag{A16}$$

so that $du/H(u)=d\xi(u)$. Then

$$\xi(S(v,w))=\xi(v)+c\xi(w), \tag{A17}$$

where a constant of integration has been absorbed into $\xi(0)$. Substituting this function ξ back into Eq. (A1) we obtain $c=1$. This completes our derivation of Eq. (16).

[1] R. T. Cox, Am. J. Phys. **14**, 1 (1946).

[2] E. T. Jayne, *Papers on Probability, Statistics and Statistical Physics*, edited by R. D. Rosenkrantz (Reidel, Dordrecht, 1983).

[3] See, for example, J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987); D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys. **58**, 1131 (1990), and references therein.

[4] L. de Broglie, *Non-Linear Wave Mechanics—A Causal Interpretation* (Elsevier, Amsterdam, 1950); P. Pearle, Phys. Rev. D **13**, 857 (1976).

[5] I. Bialynicki-Birula and J. Mycielski, Ann. Phys. (N.Y.) **100**, 62 (1976); A. Shimony, Phys. Rev. A **20**, 394 (1979).

[6] S. Weinberg, Phys. Rev. Lett. **62**, 485 (1989); Ann. Phys. (N.Y.) **194**, 336 (1989).

[7] N. Herbert, Found. Phys. **12**, 1171 (1982); W. K. Wootters and W. Zurek, Nature (London) **299**, 802 (1982); N. Gisin, Helv. Phys. Acta **62**, 363 (1989); Phys. Lett. A **143**, 1 (1990); J. Polchinski, Phys. Rev. Lett. **66**, 397 (1991); H. Scherer and P. Busch, Phys. Rev. A **47**, 1647 (1993).

[8] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **49**, 91 (1982).

[9] C. G. Shull, D. K. Atwood, J. Arthur, and M. A. Horne, Phys. Rev. Lett. **44**, 765 (1980); R. Gahler, A. G. Klein, and A. Zeilinger, Phys. Rev. A **23**, 1611 (1981); J. J. Bollinger *et al.*, Phys. Rev. Lett. **63**, 1031 (1989).

[10] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, 1958).

[11] R. P. Feynman, Rev. Mod. Phys. **20**, 267 (1948); R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[12] J. M. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, New York, 1968).

[13] *The Logico-Algebraic Approach to Quantum Mechanics*, edited by C. A. Hooker (Reidel, Dordrecht, 1979), Vols. I and II.

[14] A. M. Gleason, J. Rat. Mech. Anal. **6**, 885 (1957), reprinted in Ref. [13], Vol. I.

[15] D. Finkelstein, Trans. NY Acad. Sci. **25**, 621 (1963).

[16] J. B. Hartle, Am. J. Phys. **36**, 704 (1968).

[17] N. Graham, in *The Many-Worlds Interpretation of Quantum Mechanics*, edited by B. S. DeWitt and N. Graham (Princeton University Press, Princeton, 1973).

[18] S. Youssef, Mod. Phys. Lett. A **6**, 225 (1991); **9**, 2571 (1994); Phys. Lett. A **204**, 181 (1995).

[19] This result is briefly described without proof in A. Caticha (unpublished).

[20] These propositions are special instances of what are usually called histories. Since the word ‘‘consistency’’ also appears in this paper the reader might be led to suspect a connection with Griffith’s ‘‘consistent histories.’’ See, e.g., R. B. Griffiths, Phys. Rev. A **54**, 2759 (1996); R. Omnes, Rev. Mod. Phys. **64**, 339 (1992) and references therein. We hasten to point out that there is no connection. The ‘‘consistent histories’’ approach seeks to clarify issues of interpretation by introducing rules (the consistency conditions) which restrict discourse to a subset of histories within which *probabilities* can be assigned. Our goal is different, we seek to justify the formalism itself; the propositions we consider are those to which *amplitudes* can be assigned. Our use of special histories does not reflect a lack of generality.

[21] The two conditions implicit in the notion of independence address the two ways in which the wave function of two systems can become entangled: either during the preparation or through interaction. The challenge here is to implement these conditions without using tools that are available in the usual formulations (ideas of entanglement, interaction potentials, etc.) but are not yet available to us.

[22] See, e.g., S. Weinberg, *The Quantum Theory of Fields* (Cambridge University Press, Cambridge, 1995).

[23] See, e.g., D. Finkelstein, J. M. Jauch, S. Schiminovich, and D. Speiser, J. Math. Phys. **3**, 207 (1962); **4**, 788 (1963); S. L. Adler, Phys. Rev. Lett. **55**, 783 (1985); Commun. Math. Phys. **104**, 611 (1986); D. Hestenes, *Spacetime Algebra* (Gordon and Breach, New York, 1966); *Clifford (Geometric) Algebras*, edited by W. E. Baylis (Birkhäuser, Boston, 1996).