The transactional interpretation of quantum mechanics

John G. Cramer

*Department of Physics FM-15, University of Washington, Seattle, Washington 98195**
and Bereich Kern- und Strahlenphysik, Hahn-Meitner-Institut für Kernforschung, D-1000 Berlin 39

The interpretational problems of quantum mechanics are considered. The way in which the standard Copenhagen interpretation of quantum mechanics deals with these problems is reviewed. A new interpretation of the formalism of quantum mechanics, the *transactional interpretation*, is presented. The basic element of this interpretation is the *transaction* describing a quantum event as an exchange of advanced and retarded waves, as implied by the work of Wheeler and Feynman, Dirac, and others. The transactional interpretation is explicitly nonlocal and thereby consistent with recent tests of the Bell inequality, yet is relativistically invariant and fully causal. A detailed comparison of the transactional and Copenhagen interpretations is made in the context of well-known quantum-mechanical *Gedankenexperimente* and "paradoxes." The transactional interpretation permits quantum-mechanical wave functions to be interpreted as real waves physically present in space rather than as "mathematical representations of knowledge" as in the Copenhagen interpretation. The transactional interpretation is shown to provide insight into the complex character of the quantum-mechanical state vector and the mechanism associated with its "collapse." It also leads in a natural way to justification of the Heisenberg uncertainty principle and the Born probability law ($P=\psi\psi^*$), basic elements of the Copenhagen interpretation.

CONTENTS

*Present address.

I. INTRODUCTION

"There was a time when the newspapers said that only twelve men understood the theory of relativity. I do not believe that there ever was such a time. ... On the other hand, I think it is safe to say that no one understands quantum mechanics. ... Do not keep saying to yourself, if you can possibly avoid it, 'But how can it be like that?' because you will get 'down the drain' into a blind alley from which nobody has yet escaped. Nobody knows how it can be like that."

R. P. Feynman
The Character of Physical Law
(1967a, p. 129)

It has been over half a century since that remarkable period of 1925—1927 when modern quantum mechanics suddenly emerged from the work of Heisenberg (1925,1927), de Broglie (1926,1927a, 1927b), Schrödinger (1926a,1926b,1926c,1926d,1927a,1927b), and Born (1926a, 1926b,1927) and quickly replaced Newtonian mechanics and the "old quantum theory" of Planck, Einstein, and Bohr as the standard theory for dealing with all microscopic phenomena. The mathematical *formalism* of quantum mechanics, although refined and generalized in the intervening decades, has never been seriously challenged either theoretically or experimentally and remains as firmly established today as it was in the 1930s.

Yet over the entire period since the original development of quantum mechanics there has been controversy surrounding its *interpretation.* The questions of the meaning of the mathematics and of the underlying reality behind the laws and procedures of quantum mechanics have been a battlefield for 5 decades, and no truce is yet in sight. The controversy has, in fact, recently been intensified. The "spooky actions at a distance" that Einstein (1947) perceived in quantum mechanics seem to have been demonstrated by the theoretical work of Bell (1964,1966) and the experimental work that has followed from it (Freedman and Clauser, 1972; Clauser and Shimony, 1978; Aspect *et al.,* 1982a,1982b). This body of work (Stapp, 1975,1982; Clauser and Shimony, 1978) makes a compelling case that quantum mechanics (and nature) cannot simultaneously have the properties of "locality" and "contrafactual definiteness," but rather must lack one or the other (or both).

The term *contrafactual definiteness*[1] (CFD) used here was introduced by Stapp (1971; see also Herbert and Karush, 1978) as a minimal assumption. It means that for the various alternative possible measurements (perhaps of noncommuting variables) which might have been performed on a quantum system, each would have produced a definite (but unknown and possibly random) observational result and further that this set of results is an appropriate matter for discussion. CFD is actually a rather weak assumption and is often employed by practicing physicists in investigating and discussing quantum systems. It is completely compatible with the mathematics of quantum mechanics but is in some conflict with the positivistic element of the Copenhagen interpretation (Sec. II) and with certain other interpretations (see the Appendix).

The term *locality*[2] means that the separated parts of the system described are assumed to remain correlated only so long as they retain the possibility of speed-of-light contact and that when isolated from such contact the separated parts can retain correlations only through "memory" of previous contact. The term *nonlocality* implies the converse of this, e.g., correlations established faster-than-light across spacelike or negative timelike intervals. One should make the distinction between nonlocal *enforcement of correlations,* which is at issue here, and nonlocal *communication,* which (although sometimes confused with the former) is a far stronger condition. This distinction will be clarified later.

The mathematics of quantum mechanics does not deal explicitly with such nonlocal correlations. It does, however, require that any separated measurements of the properties of an extended system be treated as parts of the same quantum-mechanical "state," regardless of the degree of separation of the measurements in time and/or space. This common-state requirement could be interpreted as a kind of *de facto* nonlocality, but that association is not conventionally made in applying the Copenhagen interpretation to the mathematics.

Mermin (1985) has suggested that on the question of whether there is some fundamental problem with quantum mechanics signaled by tests of Bell's inequality, physicists can be divided into a majority who are "indifferent" and a minority who are "bothered." If there were a prevailing view among this concerned minority as to the resolution of the above dichotomy, CFD versus locality, it would probably be that CFD, although pragmatically useful in practical applications and discussions of quantum mechanics, must be philosophically abandoned to positivism because the alternative of nonlocality is unacceptable. It is perceived by some that nonlocality must be in direct conflict with special relativity because it could be used, at least at the level of *Gedankenexperimente,* for "true" determinations of relativistic simultaneity and must be in conflict with causality[3] because it offers the possibility of backward-in-time signaling. But this view is at best questionable. While it is clear that nonlocal communication between *observers* could lead to such conflicts, the minimum nonlocal correlations required to invalidate the Bell locality postulate are compatible with both relativity and causality.

The alternative approach to the dichotomy, advocated in this work, is to retain CFD while abandoning locality. Contrary to what might be expected, this does not require any revision of the mathematical formalism of quantum mechanics, but only a revision of the interpretation of the formalism. The *transactional interpretation* of quantum mechanics, presented below, is explicitly nonlocal but is

---

[1]There have been several attempts in the literature to answer the question "what is the minimum assumption about the physical world that one must relinquish in order to retain the locality assumption in the face of the Bell inequality experimental results?" d'Espagnat (1976,1979) has suggested that the minimum assumption is the existence of an objective external reality that is independent of the knowledge of observers. Clauser and Shimony (1978) have suggested the slightly weaker assumption of "realism," i.e., that external reality exists and has definite objective properties whether we measure them or not. Stapp's (1975) CFD assumption described in the text is, in the opinion of the author, a considerably weaker assumption than others proposed in the literature and is therefore the minimum assumption of choice.

[2]Some authors use the terms "separable" and "separability," which are synonymous with "local" and "locality" in the present context.

[3]Here we must distinguish between *causality* in the sense that an effect must follow its cause in time sequence, and *causality* as used, for example, by Heisenberg (1927) to mean that the effect is completely and uniquely determined by its cause(s). The former concept as applied to macroscopic systems is assumed here to be a valid and fundamental law of nature: the latter concept is presumably invalid because it is in conflict with the statistical interpretation of quantum mechanics (see Sec. II). In this paper we shall take *causality* to have the former meaning and reserve the term *determinism* to refer to the latter concept.

also relativistically invariant and fully causal. It is consistent with all of the familiar theoretical predictions and experimental demonstrations of conventional quantum mechanics,[4] and indeed provides new insights into some of the more counterintuitive aspects of the quantum-mechanical formalism, as will be discussed in Sec. IV.

In the body of this paper we review the Copenhagen interpretation and the interpretational problems of the quantum-mechanical formalism which it is designed to resolve. We then present the transactional interpretation and examine the way in which it deals with the same problems. Finally, we consider a number of new and traditional *Gedankenexperimente* and interpretational paradoxes as illustrative examples of the applications and power of the transactional interpretation. We find that the transactional interpretation deals with these problems in a deeper and more intuitive way.

In the main body of this paper we shall not consider other alternatives to the Copenhagen interpretation that have been proposed, but rather reserve discussion of some of these for an appendix. Furthermore, we shall not consider the rather orthogonal approach of quantum logic, which would bypass considerations of interpretation and supply instead a revision of conventional logic better suited to the quantum-mechanical formalism. We find here that standard logic can meet the needs of the quantum-mechanical formalism when a proper interpretation of that formalism is provided.

In this paper, when we explicitly examine a formalism of quantum mechanics in the context of interpretation, we shall restrict our consideration to the Schrödinger-Dirac formalism (Dirac, 1930) of wave mechanics. Although that formalism is perhaps less elegant than some of its alternatives, we find it to be the most transparent to interpretation. Because of the complete equivalence (Schrödinger, 1926c) between the wave mechanics formalism and its principal alternatives, no loss of generality is incurred through this restriction. For reasons that will be discussed later, we shall assume that the wave equations describing the system under consideration are relativistically invariant.

The task that we have undertaken here is a critical comparison of the Copenhagen interpretation with the new transactional interpretation presented below. Interpretations of a physical theory cannot normally be subjected to experimental verification. For this reason, it will be necessary to use criteria other than appeal to experiment to make any sort of critical comparison. We should like to list those criteria explicitly here.

(1) *Economy* (Occam's razor). It is preferable in constructing the interpretation to use a minimum number of independent postulates.

(2) *Compatibility.* It is preferable that the nonobservable construction of the interpretation be compatible with physical laws, even when such laws are not directly related to the theory being interpreted, i.e., quantum mechanics. In the present case we shall employ the laws of relativistic invariance, macroscopic causality, and time-reversal invariance in this context. (The violation of this criterion, i.e., the violation of a physical law by an interpretational construction is what is sometimes called an "interpretational paradox." These are to be avoided.)

(3) *Plausibility.* It is preferable that the mechanisms, if any, employed by the interpretation should be physically plausible. Common sense is not always a reliable guide in physics, but it can often help make a relative choice between otherwise equal alternatives.

(4) *Insightfulness.* It is preferable that an interpretation provide insight into the underlying mechanism of nature behind the mathematical formalism. Providing insight into the fundamental processes of nature is an important function of an interpretation. For example, the interpretational concept of field lines introduced by Faraday, while unnecessary to the formalism of electrodynamics, provides a rich and powerful medium for gaining insights into the operation of electromagnetic phenomena.

II. WHAT IS THE COPENHAGEN INTERPRETATION?

As was implied in the Introduction, we consider the theory of quantum mechanics to be divisible into a *formalism* and an *interpretation.* We shall assume for the purposes of this work that the formalism of quantum mechanics is correct and is well supported by experimental evidence. We shall therefore focus on the interpretational part of the theory and, in particular, on the Copenhagen interpretation.

Despite an extensive literature that refers to, discusses, and criticizes the Copenhagen interpretation of quantum mechanics, nowhere does there seem to be any concise statement that *defines* the full Copenhagen interpretation. We require a definition of this interpretation for the discussion that follows, and so we have attempted to provide a definitive statement by summarizing the extensive discussions by Jammer (1966) and Audi (1973) in a few sentences, identifying what we consider to be the key concepts. We have been able to identify five principal elements.

(C-1) The uncertainty principle of Heisenberg (1927): this includes wave-particle duality, the role of canonically conjugate variables, and the impossibility of simultaneously measuring pairs of such variables to arbitrary accuracy.

(C-2) The statistical interpretation of Born (1926b): this includes the meaning of the state vector (see Sec.

---

[4]There are experimental predictions associated with the Wheeler-Feynman approach which differ from those of orthodox quantum mechanics, but only for special situations involving very weak and anisotropic absorption (Partridge, 1973; Pegg, 1975; Cramer, 1980,1983). However, even in the unlikely event that such effects were observed, no revision of the quantum-mechanical formalism would be required. Rather, the same formalism would be used in a slightly different way involving altered boundary conditions for such cases.

II.A) given by the probability law ($P=\Psi\Psi^*$) and the predictivity of the formalism only for the average behavior of a group of similar events.[5]

(C-3) The complementarity concept of Bohr (1928): this includes the "wholeness" of the microscopic system and macroscopic measurement apparatus, the complementary nature of wave-particle duality, and the character of the uncertainty principle as an intrinsic property of nature rather than a peculiarity of the measurement process.

(C-4) Identification of the state vector with "knowledge of the system" by Heisenberg:[6] this includes the identification itself and the use of this concept to explain the collapse of the state vector[7] (see Sec. II.C) and to eliminate simple nonlocality problems (see Sec. II.D).

(C-5) The positivism of Heisenberg:[8] this includes declining to discuss "meaning" or "reality" and focusing interpretive discussions exclusively on observables.[9]

These five elements comprise, for the purposes of the present discussion,[10] the Copenhagen interpretation.

Two distinct functions are performed by the Copenhagen interpretation (or, for that matter, by any physical interpretation of a mathematical formalism). First, as many authors have emphasized, the interpretation must provide a connection between the mathematics of the formalism and the physical world. This connection makes it possible to test the formalism by confronting its predictions with experimental results. Without some interpretation of the symbols of the formalism in terms that can be related to experimental observables, the formalism remains as abstract mathematics without a physical context. It is perhaps in this sense that Bohr maintained (Popper, 1967) that the Copenhagen interpretation had been "proven by experiment."

However, there is another function of the interpretation that is sometimes overlooked. This function relates to the question of how the theory deals with unobserved objects (Reichenbach, 1944). While participating in a colloquium at Cambridge, von Weizsaecker (1971) denied that the Copenhagen interpretation asserted "What cannot be observed does not exist." He suggested instead that it follows the principle "What is observed certainly exists; about what is not observed we are still free to make suitable assumptions. We use that freedom to avoid paradoxes." This principle does not, of course, uniquely define the Copenhagen Interpretation, but it does give an important criterion for developing a consistent interpretation of a formalism. The interpretation must not only relate the formalism to physical observables. It must also define the domain of applicability of the formalism and must interpret the nonobservables in such a way as to avoid paradoxes and contradictions.

It may seem surprising that the *interpretation* of a physical theory can perform the function of avoiding "paradoxes," i.e., internal contradictions and conflicts with other established theories. It is therefore useful to

[5]Some authors (Ballentine, 1970) extend the statistical interpretation further by asserting that the formalism of quantum mechanics is *applicable* only to groups of similar events and should not be applied to isolated events. In our opinion this extreme view is unwarranted as long as it is appreciated that the predictivity of the quantum-mechanical formalism is severely limited in its application to isolated events. We note that the discovery of an important particle in the development of particle physics, the Ω^- baryon, was accomplished with the observation of a single isolated quantum event.

[6]The "our knowledge" element of the Copenhagen interpretation was not explicitly stated in the early interpretational papers of Bohr and Heisenberg. It apparently played an important role in the general discussion that followed Einstein's criticism of quantum mechanics at the 5th Solvay Conference in 1927 (Jammer, 1966). It was also clearly articulated in the later writings of Heisenberg (1958): "The laws of nature which we formulate mathematically in quantum theory deal no longer with the particles themselves but with *our knowledge* of the elementary particles. ... The conception of objective reality ... evaporated into the ... mathematics that represents no longer the behavior of elementary particles but rather *our knowledge* of this behavior" (italics by author). See also Hartle (1969) and Peierls (1979) for discussions of the "our knowledge" interpretation of the state vector.

[7]Heisenberg states "the act of recording, on the other hand, which leads to the reduction of the state, is not a physical, but rather, so to say, a mathematical process. With the sudden change of our knowledge also the mathematical presentation of our knowledge undergoes of course a sudden change." [As translated by Jammer (1974) from a letter to Renninger dated February 2, 1960.]

[8]Heisenberg (1927) states "... it is possible to ask whether there is still concealed behind the statistical universe of perception a 'true' universe in which the law of causality [=determinism] would be valid. But such speculations seem to us to be without value and meaningless, for physics must confine itself to the description of the relationships between perceptions."

[9]The strict positivism of (C-5) is the most "detachable" element of the Copenhagen interpretation. It was later softened by many of the proponents of the Copenhagen interpretation, including Heisenberg (1960). For example, consider the statement by Feynman *et al.* (1965): "Just because we cannot *measure* position and momentum precisely does not *a priori* mean that we *cannot* talk about them. It only means that we *need not* talk about them. The situation in the sciences is this. A concept or an idea which cannot be measured or cannot be referred directly to experiment may or may not be useful. It need not exist in a theory."

[10]To our list of Copenhagen interpretation elements some might add Bohr's correspondence principle, the reduction of quantum-mechanical predictions to those of classical mechanics in the limit of large principle quantum numbers. In our opinion this is a useful property of the quantum-mechanical formalism rather than an aspect of its interpretation.

consider some examples. Newton's second law, $\mathbf{F}=m\mathbf{a}$, is of no physical significance until the symbol $\mathbf{F}$ is identified as a vector representing force, $\mathbf{a}$ as a vector representing acceleration, and m as a scalar representing mass. Furthermore, while $\mathbf{F}$ and $\mathbf{a}$ can have any (real) magnitude and direction, the formalism is interpreted as meaningful only when $m>0$. This is because zero and negative masses lead to unphysical (or paradoxical) results, e.g., infinite acceleration or acceleration in a direction opposite that of the force vector.

As another example, consider the Lorentz transformations of special relativity for the case $v>c$. Until fairly recently, physicists had always applied to this case interpretation (A): "The transformations with $v>c$ produce unphysical imaginary values for the transformed variables and are therefore meaningless." But recently an alternative has been suggested by Feinberg (1967,1978) as interpretation (B): "The transformations in the $v>c$ domain describe a new kind of particle called the *tachyon,* which has the characteristic of imaginary mass, which always travels at velocity $v>c$, and which approaches the $v=c$ limit asymptotically from above when it is given additional kinetic energy."

While the tachyons of interpretation (B) are by no means an established physical phenomenon, this example illustrates how a change in interpretation can alter the meaning of a formalism, can extend the range of its application, and can deal with "paradoxical" or unphysical results," e.g., $v>c$ and imaginary mass. A study of the debate over interpretation in the early history of quantum mechanics (Jammer, 1966) will show a similar process at work in early attempts to interpret the quantum-mechanical formalism. It is this process that produced the Copenhagen interpretation.

In the present context it should be clear that elements (C-1) and (C-2) fulfill the function of relating the formalism to experiment, while elements (C-3) through (C-5) perform the function of avoiding paradoxes, and particularly those associated with the collapse of the state vector and with nonlocality (see Secs. II.C and II.D). Moreover, it is only elements (C-1) and (C-2) that are employed by working physicists in using quantum mechanics. Indeed (C-1) and (C-2) are represented in many quantum-mechanics textbooks as "the Copenhagen interpretation." Elements (C-3) through (C-5) are held in reserve and usually employed only in pedagogical and philosophical discussions. Thus Bohr's contention that the Copenhagen interpretation has been "proven by experiment" is perhaps correct as it applies to elements (C-1) and (C-2), but not as it applies to (C-3) through (C-5). Moreover, (C-4) has, in effect, been tested by experiment (see Sec. II.D) and found wanting, in that it has failed to neutralize the manifest nonlocality exhibited by carefully designed Bell inequality experiments.

In the remainder of this section we shall list the interpretational problems presented by the quantum-mechanical formalism and examine these problems from the point of view of the Copenhagen interpretation as defined above.

A. Identity: What is the state vector?

In the formalism of quantum mechanics the possible states of a system are described by a *state vector* (SV), a function (usually complex) that depends on position, momentum, time, energy, spin, and isospin variables, etc. The SV (which will be represented as $|S\rangle$ in the notation of Dirac) is the most general form of the quantum-mechanical wave function Ψ. The central problem of the interpretation of the quantum-mechanical formalism is to explain the physical significance of the SV. This we shall call the problem of identity.

The early semiclassical interpretations of de Broglie (1926,1927a,1927b) and of Schrödinger (1927c) attempted to make the obvious and straightforward analogy between the matter waves of quantum mechanics and the classical waves of Maxwellian electrodynamics. This approach asserts that the state vector of an electron, for example, is the quantum-mechanical equivalent of the electric field of an electromagnetic wave. Thus the SV of an electron would be considered to start at the point of emission and to travel physically through space as a wave. It would exhibit the properties of a particle only when (and if) it interacted with a scatterer or an absorber.

This apparently simple interpretation was found to lead to many conceptual problems. In particular, severe problems were found with the intrinsic nonlocality of such an interpretation (see Sec. II.D). Heisenberg recognized these problems and argued strongly and successfully against the semiclassical interpretation.[11] He devised (C-4) and (C-5) specifically to avoid any association of nonlocal implications with the formalism.

The Copenhagen interpretation approaches the problem of identity through elements (C-2) and (C-4). The statistical interpretation and the probability law of (C-2) give limited meaning to the SV by representing it as the vehicle for describing the probabilities of various possible outcomes in a quantum event. This provides the needed connection between quantum-mechanical calculations and experimental observations. Element (C-2) is, however, vague on the question of whether there is some *unique* SV that describes the present and evolving state of the system and on the question of whether the SV has a physical location in space, as the semiclassical interpretation would imply.

Element (C-4) is a more radical departure from the semiclassical interpretation in its description of the SV. According to (C-4) the SV is not analogous to the electric field of a classical light wave or indeed to any other directly observable entity. Rather it is a mathematical representation of "our knowledge of the system" (see foot-

[11]Schrödinger's visit to Copenhagen in the summer of 1926 and his discussions with Heisenberg and Bohr over his semiclassical interpretation represented a crucial test in the development of the Copenhagen interpretation. See Jammer (1974), pp. 56 and 57 and Wheeler and Zurek (1983), pp. 50 and 51.

note 6) or more properly, that knowledge which is obtained by an ideal observer in an optimum experiment, the latter qualification covering the possibility that the actual experiment performed may be less than optimum due to noise, insensitivity, or other instrumental problems. The SV is approachable only through the results of a physical measurement. The observations from measurements, in an average and statistical way, determine the values of the absolute square of components of the SV. When a measurement is performed, our knowledge of the system changes, and therefore the SV also changes. It instantaneously changes all of its components, even those which describe the quantum state in regions of space quite distant from the site of the measurement.

The instantaneous "propagation" of this change gives the appearance of action at a distance, but it is accommodated by (C-4) by associating it with a change in knowledge. According to (C-4), when the SV describing the state of a particle (perhaps an electron) has a nonzero value at some position in space at some particular time, this does *not* mean that the SV is physically present at that point, but only that our knowledge (or lack of knowledge) of the system allows the particle the *possibility* of being present at that point at that instant. Therefore, in (C-4) the wave function that the Schrödinger equation or its relativistic equivalent provides as a solution is not a physical entity, but rather an encoded mathematical message describing *our knowledge* of a physical entity.

The identification of the SV in this way raises a number of questions about the phrases "our knowledge" and "the system." This type of language begs the questions "Whose knowledge?" and "What is meant by 'the system'?" The notion that the solution of a simple second-order differential equation (particularly, an equation that is only an operator relationship between mass, momentum, and energy) is somehow a mathematical representation of "knowledge" is a very curious and provocative one. The concept of knowledge implies an observer who is the recipient of that knowledge, and because the results of a given experiment often contain information about the state of the system only in a very indirect and highly encoded way, that knowledge may be accessible only to a *conscious* and *intelligent* observer. Therefore the observer implicit in the Copenhagen interpretation has degrees of freedom that are not any explicit part of the quantum-mechanical formalism and that are not characteristics required of the observers used, for example, in the interpretation of special relativity.

Furthermore, the concept of knowledge implies stored information, i.e., a memory to store the knowledge, a time sequence before and after the creation of the memory in the mind of the observer, and a flow of information representing a time-dependent change in knowledge. Thus the Copenhagen interpretation implicitly associates with quantum events a time directionality that, while appropriate to macroscopic observers, is quite alien to and inconsistent with the even-handedness with which microphysics deals with the flow of time. Somehow the thermodynamic irreversibility of the macroscopic observer is intruding into the description of a fully reversible microscopic process.

Moreover, the assertion that knowledge is changed by measurement is not free of ambiguity. Measurements performed on any real physical system invariably contain an element of noise, which partially obscures the knowledge obtained from the measurement. (C-4) makes no provision for such noise, but treats all measurements in the same way, even when the actual signal-to-noise ratio would be such as to preclude any real gain in knowledge from the measurement. The "measurement" that changes the "knowledge" is not the real measurement actually performed, but an ideal measurement from which optimum information is assumed to have been extracted. Furthermore, the measurement event is implicitly given a special status, which distinguishes it from otherwise identical interaction events, presumably because the measurement interaction effects the knowledge of the observer, while otherwise similar interactions do not.

The question of the uniqueness of the SV is not directly addressed by (C-4). This leads to two possible ways of applying the (C-4) "knowledge" interpretation when more than one observer makes observations (perhaps simultaneously) on the same quantum-mechanical system. These are the following: (C-4a) There is one unique SV that describes the overall state of knowledge of the system, and this SV is changed when any observer makes a measurement of the state of the system; or (C-4b) there are several nonunique SV's for a given system, each describing the knowledge of some particular observer of the system, and the SV for one such observer is different and distinguishable from the SV for any other observer of the system. In Sec. II.D we shall see that each of these alternatives has its own problems.

The seemingly innocuous phrase "the system" has also been found to provide semantic difficulties. Attempts to formulate a quantum-mechanical version of general relativity and to employ the Copenhagen interpretation for its interpretation have foundered in attempting to treat the universe as a whole as a quantum-mechanical "system" in the sense of (C-4). In such a system there are (presumably) no external observers and no "knowledge of the system" that can be changed by experiments external to the system. Therefore (C-4) cannot be used for a SV describing the universe as a whole. This calls into question the whole concept.

Moreover, Wigner (1962) has demonstrated (see Sec. IV.C and the Appendix) that severe conceptual problems arise when (C-4) is applied to the SV of any system that includes a conscious observer within it, particularly when measurements on this system are performed by a second conscious observer external to the system. This has led him and others to conclude that the Copenhagen interpretation implicitly must give a special role to consciousness in the application of (C-4).

It is our conclusion from the above considerations that the approach of (C-4) to the problem of identity is a relatively superficial one. It has raised as many problems as it has solved and has led its practitioners into very deep philosophical waters. We suspect that the broad acceptance of the Copenhagen interpretation's identification of the state vector with knowledge is attributable more to the lack of a satisfactory alternative than to its compelling logic.

B. Complexity: Why is the state vector a complex quantity?

One of the serious objections to Schrödinger's (1927c) early semiclassical interpretation of the SV, as recounted by Jammer (1974), is that the SV is a complex quantity. Complex functions are also found in classical physics, but are invariably interpreted either (1) as an indication that the solution is unphysical, as in the case of the Lorentz transformations with $v > c$, or (2) as a shorthand way of dealing with two independent and equally valid solutions of the equations, one real and one imaginary, as in the case of complex electrical impedance. In the latter case the complex algebra is essentially a mathematical device for avoiding trigonometry, and the physical variables of interest are ultimately extracted as the real (or imaginary) part of the complex variables. Never in classical physics is the full complex function "swallowed whole" as it is in quantum mechanics. This is the problem of complexity.

Born's (1926b) probability law ($P = \Psi\Psi^*$) is the basis of the statistical interpretation that is embodied in (C-2). Together with (C-4) it provides a way of dealing with the problem of complexity. The SV is not directly observable and is not a real physical entity, and therefore its complex character is irrelevant. All physical observables depend on the absolute squares of the components of the SV, which are always real. (C-4) interprets the SV as an encoded mathematical representation of "knowledge" removed from the domain of physical reality and thus makes its complex character more acceptable.

However, this solution of the problem raises some questions of its own. Why is the probability equal to the absolute square of SV elements, rather than to the absolute value, or to the real part [as Born (1926a) first suggested], or to the square of the real part, or some other similar quantity? Why, moreover, is this mathematical representation of "our knowledge of the system" characterized by complex quantities that are very remote from our knowledge? In particular, who does the SV involve an overall complex phase that can never, by any conceivable experiment, become a part of "our knowledge"?

Some insight into these questions can be gained from the observation that the time-reversal operator of Wigner (1950) is the operation of complex conjugation, i.e., reversing the sign of the imaginary part or the complex phase of the SV elements. Thus the complex character of the SV is a manifestation of its time structure. The real part of the SV is time-reversal even, and the imaginary part is time-reversal odd. Moreover, a reversal of the complex phase of the SV reverses its time sense and the signs of its energy and frequency observables. Thus (C-2), Born's probability law, implicitly tells us that the probability of a particular observation is obtained by taking the product of a component of the SV with its time reverse. However, the Copenhagen interpretation provides us with no insight into why this should be the case. Why should probability be compounded of "knowledge" and the time reverse of knowledge ("information loss")?

C. Collapse: How and why does the state vector abruptly change?

The SV of a system before a measurement is performed is very different from the SV immediately after the measurement, even when the measurement is not the final state of the system but rather one of a series of sequential measurements or operations, e.g., transmission through a polarizing filter or Stern-Gerlach apparatus. Wigner (1962), following von Neumann (1932), has pointed out that there are two distinctly different types of changes that the SV undergoes: Type (1) changes the SV smoothly and continuously with time as the system evolves; Type (2) changes the SV abruptly and discontinuously with time in accordance with the laws of probability when (and only when) a measurement is made on the system. He further observed that from the point of view of classical physics these changes seem to be inverted: one would expect classically that the laws of probability and uncertainty would assert themselves in the time evolution of a wave but not in the act of measurement.

A change in the SV of type (2) described above is conventionally referred to as the "collapse of the state vector," and we shall use this terminology.[12] It is an aspect of the formalism of quantum mechanics (von Neumann, 1932) rather than its interpretation, and it is the source of many of the most severe interpretational problems. As will be discussed in Sec. IV, *Gedankenexperimente* have been devised to demonstrate that, for example, the collapse can be precipitated by the *absence* of an interaction with experimental apparatus (Sec. IV.A), but on the other hand, that the SV must remain uncollapsed after a photon has interacted with a pair of slits on the way to an experiment that may determine through which slit the photon has passed (Sec. IV.B).

Element (C-4) deals with the problem of collapse by identifying the SV with "our knowledge of the system," so that measurements which alter such knowledge will produce an abrupt change of type (2), described above, in the SV as a direct consequence of this change in knowledge. Since the SV is not physically present at the locations in space where it has a nonzero value, an abrupt change in these values does not lead to any problems with

[12]In this paper we shall employ the term "collapse of the state vector." Other authors sometimes use the terms "reduction" in place of "collapse" and/or "wave packet" or "wave function" in place of "state vector" with the same meaning. Bohr (1935a,1935b) employs the term "rupture of description." von Neumann (1932) describes the phenomenon mathematically with his projection postulate.

propagation times or speed-of-light delays in information transfer. On the other hand, Schrödinger's (1927c) interpretation of the SV as a real semiclassical wave physically present in space has severe intrinsic problems with SV collapse.

However, the (C-4) account of collapse is not without its own problems. Wigner (1962) has pointed out (see Sec. IV.C) the conceptual difficulties implicit in the Copenhagen description of collapse when the SV describes a system containing an intelligent observer. Wigner and others have suggested that the process of collapse should involve a special role for consciousness (Wigner, 1962), for permanent recording of experimental results (Schrödinger, 1935) or for entry of the system into the domain of thermodynamic irreversibility (Heisenberg, 1960). In fact, most of the efforts to revise or replace the Copenhagen interpretation have focused on the problem of collapse, which remains the most puzzling and counterintuitive aspect of the interpretation of quantum mechanics.

D. Nonlocality: How are correlations of separated parts of the state vector arranged?

The problem of nonlocality is closely related to that of the collapse of the SV. The problem in a simple form was first raised by Einstein (1928) at the 5th Solvay Conference. Later it was presented in a more subtle form as one of the criticisms of quantum mechanics by Einstein, Podolsky, and Rosen (1935). Einstein (1949) stated the problem as follows: "But on one point we should, in my opinion, absolutely hold fast: the real factual situation of system S_1 is independent of what is done with system S_2, which is spatially separated from the former." Since these words were written the thrust of the nonlocality problem has been sharpened considerably through theoretical and experimental investigations, but the issue remains essentially the same.

For the purposes of the present discussion we shall distinguish between two kinds of nonlocality. *Nonlocality of the first kind* arises from the interpretation of the SV as a physical wave. When the SV collapses the change implicit in the collapse occurs at all positions in space described by the SV at the same time. A physical wave undergoing such a change would seem to require faster-than-light propagation of information. Indeed, even the phrase "at the same time" is only meaningful relativistically in a particular inertial reference frame. It was this kind of problem that was the basis of Einstein's (1928) original objections to quantum mechanics. Similar nonlocality problems brought about the rejection of Schrödinger's semiclassical interpretation (see footnote 11).

(C-4) was constructed to avoid difficulties with nonlocalities of the first kind by denying the physical reality of the SV and identifying it instead with "our knowledge of the system." Therefore, when a measurement is made showing that a photon is located at point A (and not at B or C), our knowledge of the photon's location abruptly changes and the magnitude of the SV's value must suddenly drop to zero at B and C, although no spatial propagation, according to (C-4), is associated with that abrupt change.

(C-4) works well in this context. Its effectiveness may, however, reflect the naive statement of the nonlocality problem, which seems to require attribution of physical reality to the SV. The intrinsic nonlocality of the quantum-mechanical formalism runs deeper than this. This becomes clear when more complicated situations are considered which involve separated measurements of parts of a correlated system. In that situation, definitions of the SV become irrelevant because real measurements are involved. This leads to a *nonlocality of the second kind,* which is associated with the enforcement of correlations in spatially separated measurements.

This kind of nonlocality is demonstrated by the Freedman-Clauser experiment (1972) illustrated by Fig. 1(a). In this figure, calcium atoms undergo a $0^+ \rightarrow 1^- \rightarrow 0^+$ atomic cascade and provide a pair of photons, assumed to be emitted back-to-back, which are in a relative $L=0$ angular-momentum state. Because of angular-momentum conservation these photons are required to have identical helicities or linear combination of helicities, i.e., they must be in identical states of circular or linear polarization. For this reason the SV of the two-photon system permits the photons to be in *any* polarization state, provided only that both are in the *same* state. Experimentally this means that if the photons are transmitted through perfect polarizing filters before detection, they must be transmitted with 100% probability if the polarizations of the filters select matching states and 0% if the filters select orthogonal states, *no matter what orientation or polarization selectivity the filters have.*

The Freedman-Clauser (FC) experiment employs linear polarizing filters and measures the coincident transmission yield of the two photon detectors when the principal axes of the two filters are set at angles θ_A and θ_B, which are varied independently. Quantum mechanics predicts that the experimentally observed yield will depend only on the relative angle $\theta_{\text{rel}} \equiv \theta_A - \theta_B$ between the two principal axes, and further that for ideal filters the yield will have the normalized angular dependence

$$R(\theta_{\text{rel}}) = \cos^2(\theta_{\text{rel}}) \ . \qquad (1)$$

Note that this is just the expression for the Malus law, which gives the transmission probability of a single photon (or a beam of unpolarized light) through two crossed linear polarizers with an angle θ_{rel} between their principle axes. Thus the coincidence rate predicted when there is one polarizing filter in arm A and one in arm B of the experiment [Fig. 1(a)] is the same as if the photon in arm A went direct and unfiltered to its detector while the photon in arm B went through *both* polarizing filters in succession before reaching its detector [Fig. 1(b)]. The coincidence rate is also the same as if one of the photons, on encountering its polarizer, reached across with "spooky action at a distance" and placed the other photon in the same state.

The FC experiment (1972) was the first definitive experimental test of the Bell inequality (1964,1966), which for local theories with CFD places limits on the strength of changes in the polarization correlation function when the polarimeters differ in alignment by an increasing amount. A detailed discussion of the Bell inequality is beyond the scope of the present review, and we refer the reader to the original papers of Bell (1964,1966), the review by Clauser and Shimony (1978), Herbert (1975), d'Espagnat (1979), and Mermin (1981,1985). The actual FC experiment used nonideal filters and consequently had a rather more complicated expression for the correlation function than that given in Eq. (1). The measured function was found to be in excellent agreement with the quantum-mechanical prediction and to show a 6σ violation of the limit imposed by the Bell inequality. A more recent series of similar experiments by Aspect *et al.* (1982a,1982b) has demonstrated consistency with quantum mechanics and a 46σ violation of the Bell inequality. These results indicate, assuming CFD, that the predictions of *all* local theories (see Sec. I) are inconsistent with experimental observation.

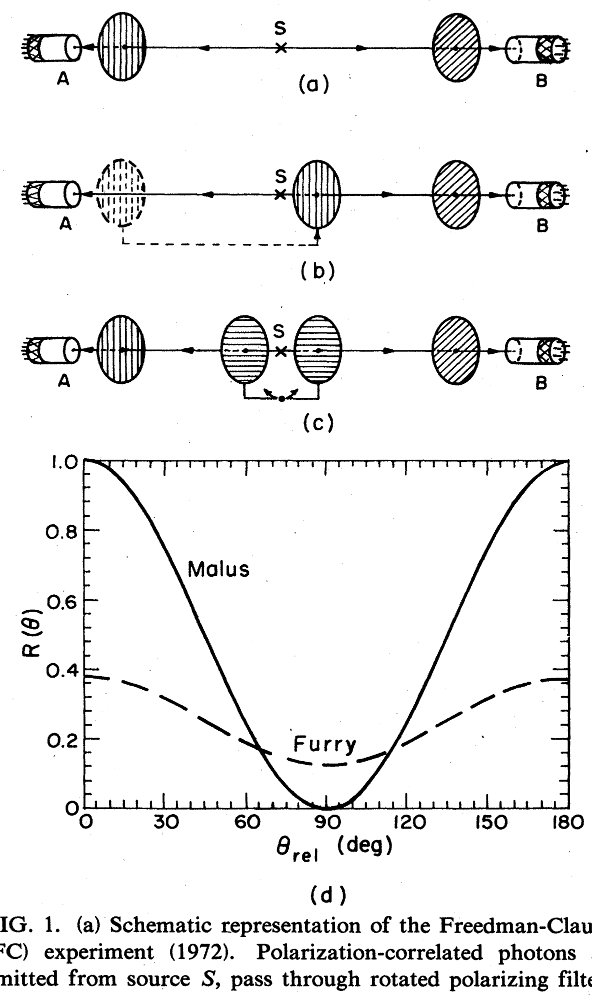


FIG. 1. (a) Schematic representation of the Freedman-Clauser (FC) experiment (1972). Polarization-correlated photons are emitted from source S, pass through rotated polarizing filters, and are detected by quantum-sensitive photomultiplier tubes A and B. (b) The "Malus" modification of the FC experiment, in which the photon in arm A passes through no filters while the photon in arm B passes through two filters. (c) The "Furry" modification of the FC experiment, in which both photons are placed in a definite but random polarization state on leaving the source S through the use of auxiliary filters rotated through random angle φ. (d) Coincidence-rate angular dependence $R(\theta_{\rm rel})$ in the above three experiments, with "Malus" corresponding to (a) and (b) and "Furry" corresponding to (c).

To illustrate that nonlocality of the second kind is exhibited by the FC result, let us consider the local modification of quantum mechanics. Furry (1936a,1936b) suggested this modification as a way of clarifying the content of the Einstein-Podolsky-Rosen (1935) criticism of quantum mechanics. Furry suggested that quantum mechanics would become a local theory if, when two parts of the system (like the two photons in the FC experiment) separate and become isolated from the possibility of speed-of-light contact, the SV describing them immediately collapses into a *definite but random* state. In the FC case the SV would collapse into a definite but random state of linear polarization shared by the two oppositely directed photons. This modified version of quantum mechanics would be a local theory because the Furry condition would satisfy the definition of locality given in Sec. I. The correlated state of the two photons would only be the result of "memory" of the correlation that had existed before they became separated. The Furry modification has no effect on many of the predictions of conventional quantum mechanics. However, does it significantly modify the predictions for the FC experiment that are predicted by quantum mechanics and observed in the experiment? Also, do the Furry predictions obey Bell's inequality? The answers to both questions are yes.

We cannot readily modify quantum mechanics so that it becomes local in this way. We can, however, simulate the Furry modification within the FC experiment by placing near the source an additional pair of aligned linear polarizing filters, which are rapidly and randomly changed. By this mechanism each pair of photons emerging from the source will be placed in definite and identical but sequentially random states of linear polarization as the photons are transmitted through these filters near the source. This arrangement is illustrated in Fig. 1(c).

The quantum-mechanical prediction for this case can easily be obtained by calculating the predicted rate of two-photon detection for a particular orientation angle φ of the randomizing filters and then averaging over all possible values of φ. The result of this calculation is

$$R_f(\theta_{\rm rel})=\tfrac{1}{8}[1+2\cos^2(\theta_{\rm rel})]\ . \qquad (2)$$

Figure 1(d) compares the functions $R(\theta_{\rm rel})$ (labeled "Malus") and $R_f(\theta_{\rm rel})$ (labeled "Furry"). The angular dependence of $R_f(\theta_{\rm rel})$ is weaker than that of $R(\theta_{\rm rel})$. In particular, $R_f(\theta_{\rm rel})$ has a maximum value of $\frac{3}{8}$, a minimum value of $\frac{1}{8}$, and cannot go to zero for any value of $\theta_{\rm rel}$. The $R_f(\theta_{\rm rel})$ correlation satisfies Bell's inequality but is inconsistent with the FC results and with quantum mechanics.

Thus the SV of the photons cannot be described as in a definite but random state. Rather, the SV must contain components that describe the photons as being in *all possible* states of polarization. Only when at least one of the two photons is detected is the SV allowed to collapse into a definite state of polarization, which must be the same for both photons. Until the detection(s) takes place, the polarizations of the photons must remain in states that are *connected* but not *specified,* in a way that is inconsistent with locality. It is this connectedness that is addressed by the Bell inequality and that cannot be explained away by the "our knowledge" definition of the SV. It is this which we have called nonlocality of the second kind.

The Furry modification of quantum mechanics is only one example of a local theory. Other local theories can give a variety of predictions, including transmissions of 100% when $\theta=0°$ and 0% when $\theta=90°$, but none can reproduce fully the FC result or the quantum-mechanical prediction. Bell's theorem demonstrates that no realistic local theory, no matter how cleverly contrived, can reproduce the experimental result of the FC experiment.

One might be tempted to think that the connectedness between the two measurements of the FC experiment, i.e., their nonlocal *correlation,* might be exploited for nonlocal *communication* to transmit messages instantaneously from one arm of the experiment to the other. Perhaps, for example, one observer could telegraph a message in Morse code by rotating his polarimeter. It has been demonstrated (Eberhard, 1977,1978; Ghirardi and Weber, 1979; Ghirardi *et al.,* 1980; Mittelstaedt, 1983) that no such observer-to-observer communication is possible, essentially because the quantum-mechanical operator corresponding to any measurement done on the right photon commutes with the operator for any measurement on the left photon. The nonlocal character of the connectedness is a subtle one, which permits the instantaneous enforcement of correlations across spacelike separations but does not permit signaling. See Sec. IV.E for further discussion of this point.

As previously mentioned in Sec. II.A, the (C-4) "knowledge" interpretation of the SV may be applied in two different ways.

(C-4a) There is a unique SV that describes the overall state of knowledge of the quantum-mechanical system and that changes when any observer changes that state of knowledge by making a measurement of the state of the system.

(C-4b) For each possible observer there is a nonunique SV that describes his knowledge of the system and that changes only when his knowledge of the system changes.

Application of these variants of the "knowledge" interpretation to the FC experiment allows us to demonstrate that alternative (C-4a) above, which may appear to be the more reasonable of the two, leads to a relativity and/or causality paradox.

As a *Gedankenexperiment* consider a "stretched" version of the Freedman-Clauser experiment in which the two arms of the apparatus are lengthened to very large distances. Let us also assume the use of 100% efficient linear polarimeters of the type used by the Aspect group, that split the incident beam into two orthogonal polarization states, so that a given photon is always detected by one or the other of a pair of photomultiplier detectors sensing the two orthogonal states, i.e., linear polarization states parallel and perpendicular to the principal axis of the polarimeter.

We assume that a previous arrangement has been made with an assistant at the light source to direct a pair of correlated photons to the two measurement sites, with the photons leaving the source apparatus at a well-defined time T. The left photon travels to the location of the left observer, who sets his polarimeter angle θ_1 and makes a measurement that we shall call M_1. Similarly, the right photon travels to the location of the right observer, who sets his polarimeter to another angle θ_2 and makes measurement M_2. Each observer always knows in advance when a photon will arrive and always obtains a definite result from each polarization measurement. Let us consider the SV collapse event that occurs as a result of one or the other of these measurements, under assumption (C-4a) that there is some universal SV which is a representation of the overall knowledge of the system.

We choose to describe the experiment as it occurs in some inertial reference frame F_1 in which the measurement event M_1 occurs earlier in the time sequence than does event M_2. Accordingly (C-4a) tells us that event M_1 alters the SV describing the entire system because that measurement alters "our knowledge of the system." The formalism of quantum mechanics requires that the SV collapse to a state that is consistent with the result measurement M_1, and from (C-4a) this collapse is triggered by a local event occurring at the location and time of M_1. Later, when the other photon reaches the right polarimeter and measurement M_2 is made, the system is already in a definite quantum-mechanical state, determined by the result of M_1. Therefore, measurement M_2 produces no further SV collapse because the knowledge gained is redundant with that already obtained by M_1.

On the other hand, relativity tells us that since the two detection events are separated by a spacelike interval, either detection event can be made to precede the other in time sequence by an appropriate choice of reference frames. Therefore, suppose that we describe the *same* experiment from some second reference frame F_2 in which measurement event M_1 occurs *after* M_2 in the time sequence. Now (C-4a) tells us that event M_2 alters the SV describing the entire system. The SV collapses to a state that is consistent with the results of measurement M_2, and this collapse is a local event occurring at the location and time of M_2. When the other photon reaches the left polarimeter and measurement M_1 is made, the system is already in a definite quantum-mechanical state, determined by the result of M_2. Therefore, measurement M_1 produces no further SV collapse because the knowledge gained is redundant with that already obtained by M_2.

Clearly, these two histories of the collapse of the overall SV are mutually exclusive and contradictory. Moreover, they conflict with the principle of relativistic invariance, because the collapse event is not a phenomenon that is independent of the reference frame in which it is viewed. Thus relativity is inconsistent with (C-4a).

One might try to avoid this conflict with relativity by making the *ad hoc* assumption that one of these descriptions (say that describing M_1 as producing the SV collapse) is the correct one independent of the reference frame. This assumption becomes troublesome because it favors one measurement and one observer over the other for no apparent reason. But it does reduce the level of conflict with relativity. However, it has another problem. In the reference frame F_2 it permits a *cause*, the collapse event at M_1, to occur after its *effect*, the arrival of the other photon at M_2 in a definite quantum-mechanical state.

Thus (C4-a) leads to conflicts at the interpretational level with either special relativity or the principle of causality. There has been some recognition of this dilemma among the founders of quantum mechanics. For example, Dirac (Hiley, 1981) said the following with reference to this problem: "It is against the spirit of relativity, but it is the best we can do . . . . We cannot be content with such a theory."

We should emphasize that the contradictions discussed above do not apply to (C-4b), which uses a different SV for each observer. In any case, these contradictions do *not* have consequences at the observational level because state vector collapse is not an observable event. Collapse is a construct perceived in the formalism (von Neumann, 1932), a pseudoevent that is asserted by the Copenhagen interpretation to occur when the state of knowledge changes. It is only when we require that the Copenhagen interpretation give an account of the collapse of some unique overall state vector and require that this account be interpretationally consistent with other established laws of physics that we reveal an interpretational paradox. The paradox is not a new one. It is the Einstein-Podolsky-Rosen paradox, but it is restated here in the language of the Copenhagen interpretation itself.

If (C-4a) is to be rejected because it leads to interpretational paradoxes, is (C-4b) an acceptable alternative? In our opinion (C-4b) is acceptable in the sense that it successfully dodges nonlocality problems of the second kind. But it does this ostrich fashion, retreating behind a solipsistic blindfold of local knowledge and positivism. The most serious criticism of (C4-b), in the view of the author, is that the account of the SV given by (C-4b) bears little resemblance to the SV most physicists *think* they are calculating (Weisskopf, 1980) when they perform quantum-mechanical calculations implicitly involving SV collapse. Nonlocality is dealt with by (C-4b) in an airtight but counterintuitive way.

E. Completeness: Do canonically conjugate variables have simultaneous reality?

Another problem raised in the Einstein-Podolsky-Rosen (EPR) paper (1935) is that of the correspondence between the quantum-mechanical formalism and reality for the case of pairs of canonically conjugate variables, i.e., pairs of variables like position and momentum having quantum-mechanical operators that do not commute. The EPR paper argues that "every element of the physical reality must have a counterpart in the physical theory" and points out that, in terms of the quantum-mechanical formalism, "when the operators corresponding to physical quantities do not commute, they cannot have simultaneous reality." Thus (goes the argument) there is a lack of correspondence between quantum mechanics and reality, and the former must be "incomplete." It is this part of the EPR criticism of quantum mechanics which received the most subsequent discussion in the literature. It became the central focus of the debate over quantum mechanics and its interpretation for a long period thereafter.

Yet, from one point of view, the quantum-mechanical formalism contains the solution to the completeness problem. The uncollapsed SV of the formalism that describes a particle (say an electron) is clearly complete in the sense that it contains components or projections which can localize either of a pair of conjugate variables. For the case of position and momentum, the SV contains projections which localize the position of an electron to arbitrary precision and other components which will similarly localize its momentum. When a measurement is made that collapses the SV, only one of these two kinds of components can be projected out by the collapse, so the simultaneous measurement of both variables can only be made to the precision specified by the uncertainty principle. Thus the variables do have "simultaneous reality" in the uncollapsed SV but can never have simultaneous reality in a single component of the SV which results from the collapse. This should satisfy the EPR criterion of completeness.

The above resolution of the EPR completeness criticism is, however, demolished by the Copenhagen interpretation itself, since (C-4b) denies the objective reality of the SV and associates it instead with the "knowledge" of an observer. If the SV is not a physical entity, but rather an ephemeral construction existing only as "knowledge" in the mind of one observer (as beauty in the eye of the beholder), then the "reality" of the conjugate variables becomes only a subjective one arising from the observer's lack of information, in support of the EPR criticism.

This leads us to the conclusion that there is indeed a completeness problem associated with quantum mechanics as the EPR paper (1935) asserted. It is not, as was supposed, a problem with the quantum-mechanical formalism, however, but with the interpretation of the formalism. An interpretation that gives physical reality to the SV of the formalism provides a *de facto* solution to the problem of completeness.

F. Predictivity: Why can we not predict the outcome of an individual quantum event?

The third criticism of quantum mechanics by the EPR paper (1935) was that a proper theory should enable the user to, "without in any way disturbing the system, ... predict with certainty ... the value of a physical quantity." Quantum mechanics, on the other hand, provides the user with a way of predicting only *average* behavior of an ensemble of quantum events but not the behavior of a particular particle in a particular event.[13] This is the problem of predictivity.

Born's statistical interpretation as embodied in (C-2) meets the problem of predictivity head on. It asserts that there is an intrinsic randomness in the microcosm which precludes the kind of predictivity we have come to expect in classical physics, and that the quantum-mechanical formalism provides the only predictivity possible, the prediction of average behavior and of probabilities as obtained from Born's probability law ($P=\Psi\Psi^*$).

While the element of the Copenhagen interpretation may not satisfy the desires of some physicists for a completely predictive and deterministic theory, it must be considered as at least an adequate solution to the problem unless a better alternative can be found. Perhaps the greatest weakness of (C-2) in this context is not that it asserts an intrinsic randomness but that it supplies no insight into the nature or origin of this randomness. If "God plays dice," as Einstein (1932) has declined to believe, one would at least like a glimpse of the gaming apparatus that is in use.

G. The Copenhagen interpretation and the uncertainty principle

Element (C-1), the uncertainty principle of Heisenberg (1927), is one of the most important aspects of the Copenhagen interpretation. It is also an interpretational aspect of quantum mechanics that has received a large amount of attention in the literature. It has been the subject of books and symposia, and it was the focus of the famous Bohr-Einstein debate.

Yet it is the aspect of quantum-mechanical interpretation that has perhaps the best grounding in analogous classical phenomena and that is easiest to understand from the viewpoint of classical physics. Heisenberg's uncertainty relations are a direct consequence of the character of the solutions of the Schrödinger equation and its relativistic equivalents, solutions that are functions of products of conjugate variables such as $\mathbf{k}\cdot\mathbf{r}$ and Et. In fact, Heisenberg's original derivation of the uncertainty principle dealt directly with this property of the wave equation solutions by showing that the Fourier transform of a localized Gaussian position wave function is a localized Gaussian momentum-space wave function, with the momentum width of the latter Gaussian proportional to the reciprocal of the position width of the former Gaussian. This property of Gaussian distributions under Fourier transformations is well known. Perhaps more important, it has many analogs in classical physics.

As an example, consider the representations of fast electrical pulses in the time and frequency domains. Such a pulse can be represented either in the time domain as a set of voltages varying continuously as a function of time, or in the frequency domain as a continuous set of Fourier components, i.e., a set of voltages varying continuously as a function of frequency. These representations of fast electrical pulses have exactly the Bohr-Heisenberg complementary relationship and exhibit their own "uncertainty principle." The localization of a fast pulse in the time domain (by making it extremely short in duration) requires a corresponding delocalization in the frequency domain, since the Fourier frequency spectrum of such a pulse must include a broader range of frequencies including very high ones. Conversely, one can increase the localization of the pulse in the frequency domain by passing the pulse through an electrical "bandpass filter," which eliminates the Fourier components that do not fall within the frequency "window" of the filter. The observable result of this frequency localization is a corresponding broadening of the pulse in the time domain. Here then is a purely classical phenomenon which exhibits an "uncertainty principle." This fast pulse uncertainty principle can be observed directly on an oscilloscope screen in any well-equipped electronics laboratory.

However (C-4b) asserts that the SV that is the carrier of these canonically conjugate quantities is not a real wave. Rather, according to (C-4b), the SV is a mathematical representation of the knowledge of some observer. This renders more questionable any association of the uncertainty principle of quantum mechanics with similar phenomena of classical physics. If it is not a physical wave but the observer's knowledge that is being localized or delocalized, one is less secure in associating that behavior with classical analogs that show an "uncertainty principle" directly in the object itself.

III. THE TRANSACTIONAL INTERPRETATION OF QUANTUM MECHANICS

In the preceding section we applied the criteria of Sec. I to the Copenhagen interpretation as it deals with the interpretational problems of the quantum-mechanical formalism. This exercise has shown that several interpretational problems are handled only superficially by the Copenhagen interpretation. The problem area is centered on (C-4), the association of the state vector with subjective knowledge of the system by an observer. In this section we shall present the transactional interpretation of quantum mechanics, an alternative to the Copenhagen interpretation, which retains the interpretational links between formalism and experiment but which replaces the subjec-

[13]The formalism of quantum mechanics *can* predict the outcome of a single isolated quantum event in the unusual circumstance when one particular outcome of the event has a predicted probability of 1.0 and all other outcomes have predicted probabilities of zero. An example is the measurement of a particular variable after the system has been prepared in a unique state of that variable, e.g., transmission of a photon through successive aligned linear polarizers.

tivity and the nonlocality evasions of (C-4) with an objective and explicitly nonlocal description of quantum processes.

Our survey has illuminated several interpretational problems intrinsic in the Copenhagen interpretation. Based on this discussion, we can now define as goals a set of characteristics desirable in a more "ideal" interpretation: (1) it should permit the operation of the microcosm to be isolated from the macrocosm and particularly from intrinsically complicated macroscopic concepts, e.g., knowledge, intelligent observers, consciousness, irreversibility, and measurement; (2) it should account for the nonlocal correlations of the Bell inequality tests in a way consistent with relativity and causality; (3) it should account for the collapse of the state vector without subjective "collapse triggers" (e.g., consciousness); and (4) it should give added meaning to the state vector and provide insights into the problems of complexity, completeness, and predictivity.

With these goals in mind, we now present the transactional interpretation of quantum mechanics. We shall find that this interpretation, which is objective and explicitly nonlocal, satisfies each of these goals. The interpretation provides a description of the state vector as an actual wave physically present in real space and provides a mechanism for the occurrence of nonlocal correlation effects through the use of advanced waves. The collapse of the state vector of the transactional interpretation is the formation of a *transaction,* which occurs by an exchange of retarded and advanced waves. The transaction model provides a way of clearly visualizing and developing intuition about the quantum phenomena that have remained mysterious and counterintuitive for half a century.

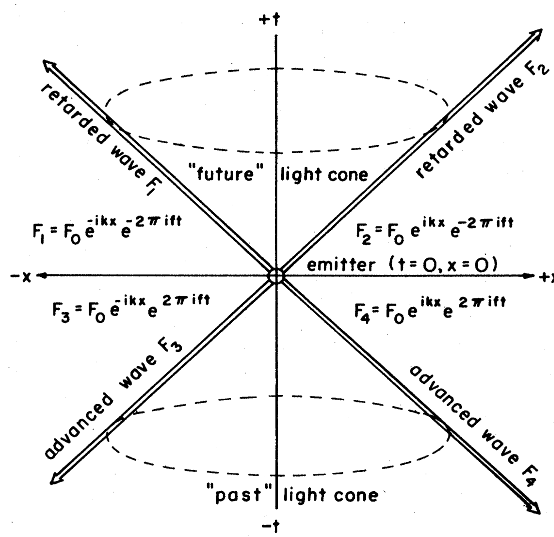


FIG. 2. Minkowski diagram showing the propagation of advanced and retarded waves from an emission locus at $(x,t)=(0,0)$.

A. Advanced waves and Wheeler-Feynman absorber theory

The basic element of the transactional interpretation is an emitter-absorber transaction through the exchange of advanced and retarded waves, as first described by Wheeler and Feynman (1945,1949; see also Feynman, 1967b). Advanced waves are solutions of the electromagnetic wave equation and other similar wave equations that contain only the second time derivative. Advanced waves have characteristic eigenvalues of negative energy and frequency, and they propagate in the negative time direction. Figure 2 illustrates the propagation of advanced and retarded waves. The advanced-wave solutions of the electromagnetic wave equation are usually ignored as unphysical because they seem to have no counterpart in nature.

The classical electrodynamics described by Wheeler and Feynman (WF) was intended to deal with the problem of the self-energy of the electron in an innovative way. Assuming the time-symmetric formalism of Dirac (1938) combined with the *ad hoc* assumption that an electron does not interact with its own field, Wheeler and Feynman were able to formally eliminate the self-energy term from their electrodynamics. But along with self-energy these assumptions also removed the well-observed energy loss and recoil processes (i.e., radiative damping) arising from the interaction of the radiating electron with its own radiation field.

Wheeler and Feynman accounted for these well-known damping effects by allowing the emitting electron to interact with the advanced waves sent by other electrons, which would ultimately, at some future time, absorb the retarded radiation. Thus the energy loss and recoil of the emitter were accounted for without having it interact with its own field. Moreover, the calculation succeeded in describing electrodynamic interactions in a completely time-symmetric way. To account for the observed asymmetric dominance of retarded radiation, Wheeler and Feynman invoked the action of external boundary conditions arising from thermodynamics. Thus, they avoided resorting to the *ad hoc* "causality" condition usually needed to eliminate the advanced-radiation solutions.

Regrettably, the WF paper (1945), while mathematically correct, proved to be an invalid way of dealing with self-energy. As Feynman (1949) later pointed out, the self-interaction is a necessary part of electrodynamics, needed, for example, to account for the Lamb shift. It is relevant that the WF *ad hoc* assumption of noninteraction is not needed in their recoil calculations because, as later authors have pointed out (Pegg, 1975; Cramer, 1980), the electron cannot undergo energy loss or recoil, which are intrinsically time-unsymmetric processes, as a result of interacting with its own (or any other) time-symmetric field.

When the offending assumption of noninteraction is removed from the WF formalism, what remains is a classical self-consistent and time-symmetric electrodynamics that cannot be used to deal with the problem of self-energy. Furthermore, this WF formalism is not particularly useful as an alternative method of calculating the electrodynamics of radiative processes because the mathematical description of radiation explicitly involves the interaction of the emitter with the entire future universe. Thus a simple integration over local coordinates in the conventional formalism is replaced by an integral over all future space-time in the light cone of the emitter in the WF formalism.

However, this "difficulty" can be viewed as asset. The WF mathematics can be used to investigate the properties of cosmological models describing the future state of the universe by relating such models to radiative processes. In essence this approach provides a way of linking the cosmological arrow of time (the time direction in which the universe expands) to the electromagnetic arrow of time (the complete dominance of retarded over advanced radiation in all radiative processes). There is a considerable literature in this field, which the author has reviewed in a previous publication (Cramer, 1983).

Although the original WF work dealt exclusively with classical electrodynamics, later authors (Hoyle and Narlikar, 1969,1971; Davies, 1970,1971,1972) have developed equivalent time-symmetric quantum-electrodynamic (QED) versions of the same approach. The predictions of these QED theories have been shown to be completely consistent with those predictions of conventional QED which can be compared with experimental observation (see footnote 4). It has also been shown (Davies, 1972) that despite this similarity of prediction, the time-symmetric QED provides a qualitatively different description of electromagnetic processes. It is essentially an action-at-a-distance theory with no extra degrees of freedom for the radiation fields and no second quantization. The field in effect becomes a mathematical convenience for describing action-at-a-distance processes.

There may also be another advantage to the WF approach to electrodynamics. Dirac's (1938) work on time-symmetry electrodynamics, on which the WF theory is based, was introduced as a way of dealing with singularities in the radiation field in the conventional theory near a radiating electron. Konopinski (1980), in his Lorentz covariant treatment of the radiating electron, has pointed out that this time-symmetric "Lorentz-Dirac" approach eliminates such singularities and therefore amounts to a self-renormalizing theory. This formulation may have applications in eliminating related singularities in QCD and in quantum field theory in curved space-time.

B. The emitter-absorber transaction model

There is a second application of the Wheeler-Feynman approach, which was introduced by the author in a previous publication (Cramer, 1980). The WF description of radiative processes can be applied to the microscopic exchange of a single quantum of energy, momentum, etc., between a present emitter and a future absorber through the medium of a *transaction*, a Wheeler-Feynman exchange of advanced and retarded waves. Figure 3 illustrates a simplified form (one space dimension and one time dimension) of the transaction process.

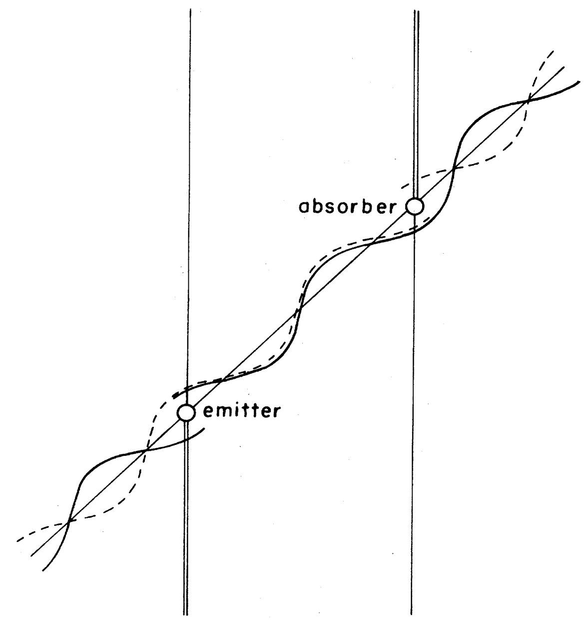


FIG. 3. A plane-wave transaction between emitter and absorber through the exchange of advanced and retarded waves (see the text). Waves from emitter are indicated by solid lines and waves from absorber by dashed lines. Relative phase of waves is indicated schematically by sinusoids inscribed on lightlike world lines. Double timelike world lines for emitter and absorber indicate higher energy state. Wave amplitudes have the value $\Psi+\Psi^*=2\,\mathrm{Re}(\Psi)$ at emitter and absorber loci and are therefore real.

The emitter, e.g., a vibrating electron or atom in an excited state, attempts to radiate by producing a field. This field, according to the Wheeler-Feynman description, is a time-symmetric combination of a retarded field which propagates into the future and an advanced field which propagates into the past. For simplicity let us first consider the net field to consist of a retarded plane wave of the form $F_1 \simeq \exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)]$ for $t \geq T_1$ (T_1 is the instant of emission) and an advanced plane wave of the form $G_1 \simeq \exp[-i(\mathbf{k}\cdot\mathbf{r}-\omega t)]$ for $t \leq T_1$. Since the retarded wave F_1 has eigenvalues characteristic of positive energy $\hbar\omega$ and momentum $\hbar\mathbf{k}$, while the advanced wave G_1 has eigenvalues of negative energy $-\hbar\omega$ and momentum $-\hbar\mathbf{k}$, the net loss of energy and momentum by the emitter in producing the pair of waves (F_1+G_1) is zero, as might be expected from the time symmetry of the composite wave.

Let us for the moment set aside consideration of the advanced wave G_1 and follow the retarded wave F_1. This wave will propagate in the positive time direction $(t \geq T_1)$ until it encounters an absorber. The process of absorption, as is well known, can be described as a movement of the absorbing electron (or atom) in response to the incident retarded field F_1 in such a way as to gain energy, recoil, and to produce a new retarded field $F_2=-F_1$, which exactly cancels the incident field F_1. Thus the retarded wave from the absorber exactly cancels the retarded wave from the emitter, and there is no net field present after the instant of absorption T, i.e.,

$$F_{\rm net} \equiv (F_1 + F_2) = 0 \quad \text{for } t > T_1 \ . \tag{3}$$

The Dirac-Wheeler-Feynman assumption of time-symmetric radiative processes requires that the absorber only produce the canceling retarded field F_1 for $t \geq T_2$ if it also produces an advanced field G_2 for $t \leq T_2$. This field G_2 will propagate in the negative time direction (i.e., into the past) from the instant of absorption T_2, traveling back down the track of the incident wave F_1 to the instant of emission T_1. There it interacts with the radiating electron (or atom) at the instant of emission, causing it to recoil and to lose energy. Furthermore, the advanced wave G_2 continues to times such that $t > T_1$, where it is superimposed on the advanced wave from the emitter G_1 to produce a net advanced field:

$$G_{\rm net} \equiv G_1 + G_2 \ . \tag{4}$$

The condition that $F_2 = -F_1$ at the absorber for $t < T_2$ brings with it a similar condition for the advanced fields, so that $G_2 = -G_1$ at the emitter and for $t < T_1$, and $G_{\rm net} = 0$ for $t < T_1$. The result of the cancellation of the preemission and postabsorbtion waves is that only in the interval $T_1 \leq t \leq T_2$ is there a nonzero field:

$$F_{\rm net} = F_1 + G_2 \ . \tag{5}$$

From this we see that even under the Dirac assumption of time-symmetric radiation of retarded and advanced waves the advanced field G_1 cannot produce "advanced effects" such as backward-in-time signaling and the emission of negative-energy radiation because it has been nullified by the absorption process.

The above, in a simplified one-dimensional form that will be expanded below, is the emitter-absorber transaction. The emitter can be considered to produce an "offer" wave F_1 which travels to the absorber. The absorber then returns a "confirmation" wave G_2 to the emitter, and the transaction is completed with a "handshake" across space-time. To an observer who has not viewed the process in the pseudotime sequence[14] employed in the above discussion, there is no radiation before T_1 or after T_2, but a wave has traveled from emitter to absorber. This wave can be reinterpreted as a purely retarded wave because its advanced component G_2, a negative-energy wave traveling backwards in time from absorber to emitter, can be *reinterpreted* as a positive-energy wave traveling forward in time from emitter to absorber, in one-to-one correspondence with the usual description.[15]

Thus the WF time-symmetric description of electrodynamic processes is completely equivalent in all observables to the conventional electrodynamic description. Time-symmetric electrodynamics, in both classical and quantum-mechanical forms, leads to predictions identical with those of conventional electrodynamics. For this reason it is not possible to devise experimental tests that will distinguish between time-symmetric and conventional electrodynamics. The intrinsic untestability of time-symmetric electrodynamics reveals that it should be considered an alternative *interpretation* of the electrodynamic formalism rather than an alternative formulation.

It is this alternative interpretation of the electrodynamic formalism which we have generalized (Cramer, 1980) to include all quantum-mechanical processes and which leads to the alternative interpretation of quantum mechanics presented here. The fundamental element of this interpretation is the emitter-absorber transaction, a simple plane-wave version of which was described above. The transaction is a "handshake" between the emitter and absorber participants of a quantum event, occurring through the medium of an exchange of advanced and retarded waves. The description just presented is basically one dimensional (in space) and is not fully applicable to the case of three space dimensions with quantization boundary conditions. Before discussing the applications of the interpretation, we shall generalize the transaction model from one to three spatial dimensions.

There are two problems with the one-dimensional plane-wave description employed above: (1) it does not explicitly deal with the attenuation and modification of wave amplitude due to propagation through space or to passage through attenuating media; and (2) it does not explicitly include the quantum conditions on the transfer of energy, angular momentum, charge, etc., which are an important aspect of all quantum-mechanical processes. In the case of quantum electrodynamics, the photon energy quantization condition $E = \hbar\omega$ places an extra constraint on the electromagnetic wave equation, requiring that an integer number of quanta be exchanged between emitter and absorber despite the action of intervening space, filters, mirrors, slits, wave plates, etc., in reducing or modifying the amplitudes of the advanced and retarded waves exchanged between emitter and absorber.

For this reason, the two-step pseudotime sequence (see

[14]The account of an emitter-absorber transaction presented here employs the semantic device of describing a process extending across a lightlike or timelike interval of space-time as if it occurred in a time sequence external to the process. The reader is reminded that this is only a pedagogical convention for the purposes of description. The process is atemporal and the only observables come from the superposition of all "steps" to form the final transaction. (See also the Appendix on this point.)

[15]The statement that an advanced wave may be reinterpreted as a retarded wave (or vice versa), which is propagating in the opposite direction with sign-reversed energy and frequency, is actually a classical oversimplification of the quantum-mechanical formalism. Advanced waves have the characteristic time-dependent phase $\exp(i\omega t)$, while retarded waves have phase $\exp(-i\omega t)$. These functions are orthogonal and in principle distinguishable. Reinterpretation is permissible because it is consistent with the observed transfer of energy, momentum, etc., and because the time phases are no observed.

footnote 14) of Fig. 3 and the associated plane-wave description must be replaced by a multistep sequence allowing for spherical and more complicated wave forms and proceeding until all relevant conditions are met. In particular, we must view the transaction as occurring in pseudosequential form, which includes an "offer," a "confirmation," and a completed transaction.

Figure 4 illustrates this more general form of transaction. In the first pseudosequential step (1) the emitter, located at $(\mathbf{R}_1,T_1)$, sends out "offer" waves $F_1(\mathbf{r},t\geq T_1)$ and $G_1(\mathbf{r},t\leq T)$ (which can be of spherical or more complicated forms) in all possible spatial directions. In step (2) the absorber located at $(\mathbf{R}_2,T_2)$, receives the attenuated retarded-wave front $F_1(\mathbf{R}_2,T_2)$ and is stimulated to produce a response wave $G_2(\mathbf{r},t)$, which has an initial amplitude proportional to the local amplitude of the incident wave that stimulated it:

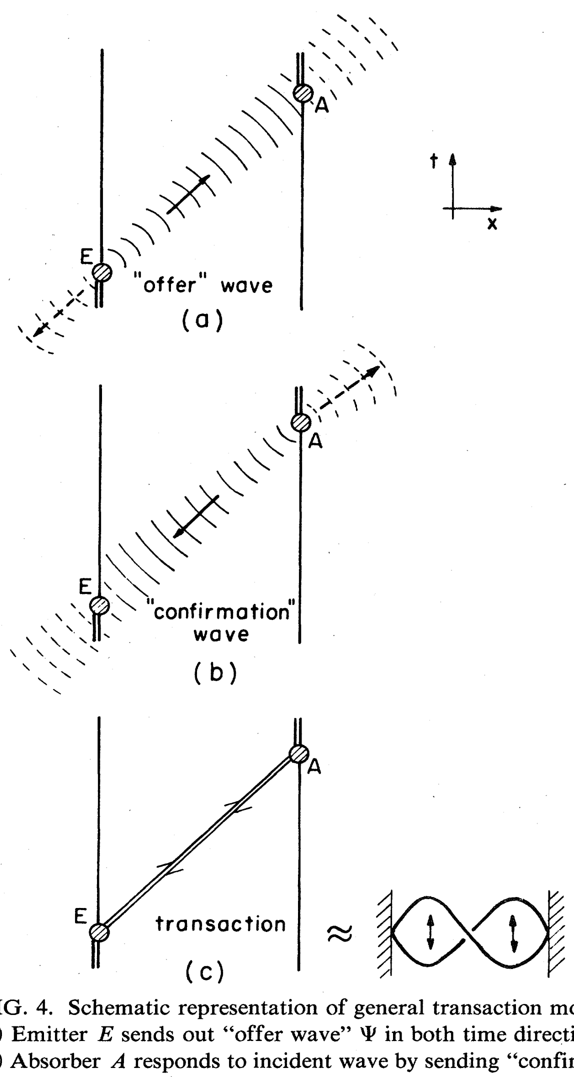


FIG. 4. Schematic representation of general transaction model. (a) Emitter E sends out "offer wave" Ψ in both time directions. (b) Absorber A responds to incident wave by sending "confirmation wave" echo Ψ^* back to emitter. Echo amplitude at emitter locus has value $\Psi\Psi^*=P$ (probability of transaction). (c) Process continues to completion with satisfaction of quantum boundary conditions at emitter and absorber loci, in analogy with a standing wave between terminating walls.

$$G_2(\mathbf{r},t)\propto F_1(\mathbf{R}_2,T_2)g_2(\mathbf{r},t)\ . \tag{6}$$

Here $g_2(\mathbf{r},t)$ is a unit advanced wave, i.e., the advanced equivalent of the retarded wave $F_1(\mathbf{r},t)$ in that $g_2(\mathbf{r},t-T_2)=[F_1(\mathbf{r},t-T_1)]^*$.

In step (3) the advanced wave G_2 propagates back to the locus of emission, at which it has an amplitude proportional to its initial amplitude $F_1(\mathbf{R}_2,T_2)$ multiplied by the attenuation it has received in propagating from the absorption locus to the emission locus. The advanced wave G_2 travels across the same spatial interval and through the same attenuating media encountered by F_1, but in reverse. For this reason, the unit amplitude wave $g_2(\mathbf{R}_1,T_1)$ arriving back at the emitter has an amplitude proportional to $F_1^*(\mathbf{R}_2,T_1)$, the time reverse of the retarded wave that reached the absorber. Thus at the emission locus the advanced-wave amplitude G_2 is

$$G_2(R_1,T_1)\propto F_1(\mathbf{R}_2,T_2)F_1^*(\mathbf{R}_2,T_2)=|F_1(\mathbf{R}_2,T_2)|^2\ . \tag{7}$$

This means that the advanced "confirmation" or "echo" wave that the emitter receives from the absorber as the first exchange step of the incipient transaction is just the absolute square of the initial "offer" wave, as evaluated at the absorber locus. The significance of this $\Psi\Psi^*$ echo and its relation to Born's probability law will be discussed in Sec. III.H.

In step (4) the emitter responds to the "echo" and the cycle repeats until the response of the emitter and absorber is sufficient to satisfy all of the quantum boundary conditions ($E=h\nu$ and various conservation laws), at which point the transaction is completed. Even if many such echoes return to the emitter from potential absorbers, the quantum boundary conditions can usually permit only a single transaction to form. The transaction formation can be considered as analogous to the establishment of a four-vector standing wave across the interval bounded by $(\mathbf{R}_1,T_1)$ and $(\mathbf{R}_2,T_2)$, the two loci forming terminating "walls" outside which the wave amplitude must make no contribution to the process. Note that at the completion of step (4) the local fields in the vicinity of both the emitter and the absorber are *real* (as opposed to complex) because they are a superposition of an advanced and a retarded wave of equal amplitude and the same phase. The significance of this for the problem of complexity is discussed in Sec. III.F.

To summarize the transaction model, the emitter produces a retarded offer wave (OW), which travels to the absorber, causing the absorber to produce an advanced confirmation wave (CW), which travels back down the track of the OW to the emitter. There the amplitude is $\mathrm{CW}_1\propto\mathrm{OW}_2^2$, where CW_1 is evaluated at the emitter locus and OW_2 is evaluated at the absorber locus. The exchange then cyclically repeats until the net exchange of energy and other conserved quantities satisfies the quantum boundary conditions of the system, at which point the transaction is complete. Of course the pseudotime sequence (see footnote 14) of the above discussion is only a semantic convenience for describing the onset of the transaction. An observer, as in the simpler plane-wave case, would perceive only the completed transaction, which he could reinterpret as the passage of a single retarded (i.e., positive-energy) photon traveling at the speed of light from emitter to absorber (see footnote 15).

An equally valid interpretation of the process is that a four-vector standing wave has been established between emitter and absorber. As a familiar three-space standing wave is a superposition of waves traveling to the right and left, this four-vector standing wave is the superposition of advanced and retarded components. It has been established between the terminating boundaries of the emitter, which blocks passage of the advanced wave further down the time stream, and the absorber, which blocks passage of the retarded wave further up the time stream. This space-time standing wave is the transaction we shall use as a basis for the discussion that follows.

It should be emphasized that the transactional interpretation is an *interpretation* of the existing formalism of quantum mechanics rather than a new theory or revision of the quantum-mechanical formalism. As such, it makes no predictions that differ from those of conventional quantum mechanics. It is not testable except on the basis of its value in dealing with interpretational problems. We have found it to be more useful as a guide for deciding which quantum-mechanical calculations to perform than as an aid in the performance of such calculations. As will be demonstrated in Sec. IV, the main utility of the transactional interpretation is a conceptual model which provides the user with a way of clearly visualizing complicated quantum processes and of quickly analyzing seemingly "paradoxical" situations (e.g., Wheeler's delayed-choice experiments, Herbert's paradox, the Hanbury-Brown-Twiss effect, and the Albert-Aharonov-D'Amato prediction), which would otherwise require elaborate mathematical analysis. It is a way of thinking rather than a way of calculating. It may have value as a pedagogical tool for the teaching of quantum mechanics to students. It also seems to have considerable value in the development of intuitions and insights into quantum phenomena that up to now have remained mysterious.

C. The transaction model and relativistic quantum mechanics

The transaction model discussed in the preceding section deals with the emission and absorption of photons arising from electromagnetic interactions. The model uses advanced- and retarded-wave functions that are solutions of the electromagnetic wave equation:

$$(\hbar c)^2\nabla^2\psi=\hbar^2\partial^2\psi/\partial t^2 \ . \qquad (8)$$

Note that this differential equation is second order in the time variable. As was shown in a previous publication (Cramer, 1980), the same transaction model can be applied to the emission and absorption of massive particles, either neutral or electrically charged, e.g., electrons.[16] The only requirement for this application is that the wave equations describing the particles of interest, like the electromagnetic wave equation, have both advanced and retarded solutions.

This requirement would seem to present a problem. The wave equation that has been the focus of most of the discussion surrounding the interpretation of quantum mechanics is the Schrödinger equation,

$$-(\hbar^2/2m)\nabla^2\psi=i\hbar\partial\psi/\partial t \ , \qquad (9)$$

where m is the mass of the particle described by the equation. This equation is first order in the time variable and for this reason does *not* have advanced solutions. Therefore, if $\psi=F(r,t)$ is a solution of the Schrödinger equation, then $\psi^*=G(r,t)$ is not a solution, nor is a linear combination of F and G as used in the transactional model.

We must bear in mind, however, that the Schrödinger equation is ultimately not physically correct because it is not relativistically invariant.[17] It should properly be considered as the limiting case, in a restricted nonrelativistic domain, of some more physically reasonable relativistical-

[16]A transaction involving charged particles, particularly in the presence of external electric and magnetic fields, must be more carefully described. To obtain the path of the advanced wave, the operation of time reversal must be performed not only on the charged-particle state vector but also on the external fields that it "sees" as acting along its time-reversal path back to the emitter. The time-reversal operation on such electric and magnetic fields must be applied to the proper Lorentz six-vector describing the electromagnetic field. This has the effect of reversing the signs of (real) magnetic fields while leaving the sign of (real) electric fields unchanged. Since the particle momentum also changes sign under time reversal, the Lorentz forces act in the same way as on the "forward" particle, and the time-reversed particle retraces the path of the former. Clearly there is nothing about the time-reversal properties of a magnetic field that causes problems for the transactional model because any system with a uniform external magnetic field can be transformed into a Lorentz frame in which the external field is purely electric.

[17]The time-dependent Schrödinger equation is an operator equation version of the nonrelativistic kinematics equation $p^2/2m=E$, where p is the momentum and E is the kinetic energy of a particle of mass m. The momentum operator is $(\hbar/i)\nabla$, and the energy operator is $(i\hbar)\partial/\partial t$, in that the indicated operations when performed on a quantum-mechanical wave function, e.g., $\psi=A_0\exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)]$, yield the unaltered wave function ψ multiplied by its momentum eigenvalue ($\hbar\mathbf{k}$) or energy eigenvalue ($\hbar\omega$). If these operators are substituted in the above kinematic equation, the Schrödinger equation results. The equivalent expression for relativistic kinematics is $(pc)^2+(mc^2)^2=W^2$, where W is the total mass energy. The operator equation version of this expression is the Klein-Gordon equation (Bjorken and Drell, 1964).

ly invariant wave equation, e.g., the Dirac equation or the Klein-Gordon equation. These relativistic equations, like the electromagnetic wave equation, have both advanced and retarded solutions.[18]

Considering the Schrödinger equation as a limiting case, we can resolve the apparent problem created by its lack of advanced solutions. When a suitable relativistic wave equation is reduced to the Schrödinger equation by taking a nonrelativistic limit (Bjorken and Drell, 1964), the reduction procedure leads to *two* distinct equations, the Schrödinger equation and another equation of the form

$$-(\hbar^2/2m)\nabla^2\psi = -i\hbar\partial\psi/\partial t \ , \tag{10}$$

which is the complex conjugate or time reverse of the Schrödinger equation. This equation has *only* advanced solutions. Equations (9) and (10) are equally valid nonrelativistic reductions of relativistic dynamics, but Eq. (10) is usually dropped because it has negative-energy eigenvalues. From this it should be clear that $F(r,t)$ and $G(r,t)$ (or ψ and ψ^*) are equally valid solutions of the dynamics which underlies the Schrödinger equation. It is therefore valid to use advanced solutions in the transactional model in the nonrelativistic limit as if they were solutions of the Schrödinger equation.

We can also look at the need for relativistic invariance in another way. The interpretational problem of locality (see Sec. II.D), which the recent tests of Bell's inequality have brought into sharp focus, is essentially a relativistic problem. If the velocity of light were infinite, the locality problem would not exist: there would be no difference between local and nonlocal descriptions. The Schrödinger equation can be considered as the limiting case of a relativistically invariant wave equation when the velocity of light goes to infinity. Therefore it is not particularly surprising that an explicitly nonlocal description such as the transactional model may have intrinsic inconsistencies with the Schrödinger equation and may require certain properties of relativistically invariant wave equations. This is a subtle link between relativity and quantum mechanics which has not, perhaps, been previously appreciated.

There is another implication of relativistic quantum mechanics which should also be discussed here. In the relativistic domain the uncertainty principle of Heisenberg (C-1) must be reconsidered because of the added restrictions of special relativity (Landau and Peierls, 1931; Berestetskii *et al.* 1971). In particular, introduction of the limiting velocity c imposes a new uncertainty relation on the precision with which momentum p can be measured: $\Delta p = \hbar/(c\Delta t)$. This relation can be considered to arise from the fact that the position localization Δq (assumed to be small initially) cannot spread at a rate greater than c. The distance $c\Delta t$ is therefore the maximum possible amount by which the position localization can be broadened in a time interval Δt to permit a smaller localization of momentum. This places a limit on the precision Δp with which momentum p can be measured in a time interval Δt.

There is an analogous limitation of the determination of position q, which arises from another characteristic of relativistic quantum field theories. As mentioned above, the solutions to the relativistically invariant wave equations for massive particles include advanced or negative-frequency solutions. When a particle is localized to a sufficiently small region of space, these negative-frequency functions appear explicitly in the expansion of its position wave packet. Landau and Peierls (1931) have suggested that in order to avoid the inclusion of "physically meaningless" negative-frequency solutions it is reasonable to confine position determinations to a domain that does not include such processes. This corresponds to a limiting position uncertainty to $\Delta q = \hbar c/W$, where W is the total mass-energy of the particle. This position uncertainty limit for a particle of mass m with an initially small momentum is just $\Delta q = \hbar/mc$, the de Broglie wavelength of the particle.

Berestetskii *et al.* (1971) justify this limit on position localization in a related but slightly different way. They interpret the negative-frequency or negative-energy components of the wave function as indicating the onset of particle-antiparticle production when the momentum becomes large enough to correspond to a free energy greater than $2mc^2$. They argue that when this threshold is reached in a measurement, for example, in determining the position of the electron, "the formation of new particles in a way which cannot be detected by the process itself clearly renders meaningless the measurement of the electron coordinates." Thus the broadening in momentum is cut off at this limit, leading to the $\hbar c/W$ limit on the position localization.

Landau and Peierls (1931) have argued that these relativistic limits on determinations of position and momentum irretrievably compromise the utility of these dynamical variables for measurement in the sense of nonrelativistic quantum mechanics. Neither the position nor the momentum of a particle can, even in principle, be determined to arbitrary accuracy in a finite time interval, nor can either be considered to have a particular value at a particular time. This would seem, in effect, to invalidate Born's statistical interpretation of quantum mechanics (C-2), in that the description of the state vector as a mathematical representation of the probability of finding a definite value of a particular observable as a result of a measurement made at a given instant is untenable. However, Bohr and Rosenfeld (1933; see also Rosenfeld, 1955) defused this problem by demonstrating that in the relativistic formalism of quantum electrodynamics, in which the field quantities are not represented by point functions but

[18]See, for example, Messiah (1961), p. 61, for a clear statement of the widely held supposition that proper quantum-mechanical wave equations should be first order with respect to time so as to uniquely specify the state and the time evolution of the system. See also the discussion of "pathology" of the Klein-Gordon equation in this context on p. 65.

by functions of space-time regions, there is no discrepancy between limits imposed by the relativistic uncertainty principle and the physical possibilities of measurement.

To state it slightly differently, this is a nonproblem. The relativistic limits on the precision with which dynamic variables can be measured do indeed make these variables less directly relevant to the coordinates of relativistic particles. However, this invalidates neither their use nor their usefulness. In a related nonrelativistic case, the angle of rotation θ remains a valid and sometimes useful dynamical variable, although its measurable value is rendered completely uncertain by quantization of the conjugate angular-momentum variable. The specification of a dynamical variable "at a given instant" as considered by Landau and Peierls (1931) is neither needed nor desirable. In the relativistic formalism it is integrals over space-time regions rather than point values which lead to the predictions of observables. In this context, the atemporal and nonlocal character of the transactional interpretation, as discussed below, provides a natural way of describing the atemporal collapse of the state vector to some localized value of a dynamic variable. In fact, the notion that the state vector collapses to a particular value of a variable "at a given instant" is inconsistent with the transactional description. See Sec. IV.C for a discussion of this point.

Another problem that raises some concern about the statistical interpretation of quantum mechanics in the relativistic domain is the observation that the solutions of the field equations cannot always be used to construct a relativistically invariant or positive definite probability density (Bjorken and Drell, 1964). Thus, while differential field equations and their solutions remain an indispensable feature of relativistic quantum field theory (Bjorken and Drell, 1965), the problem of the identity of these solutions (see Sec. II.A) is made more severe because of the statistical interpretation, at least in its simplest form, has proved inadequate. The state vector cannot be identified as a simple carrier of probability in the relativistic domain.

The naive statement of the statistical interpretation is clearly insufficient in the relativistic domain, particularly when applied to space-time regions where no measurement is actually made. Furthermore, some of the formal procedures of nonrelativistic quantum mechanics, e.g., the integration of wave-function products over a large volume of space at a fixed time, are manifestly inconsistent with special relativity. However, the development of relativistic quantum theory has resulted in a formalism with calculational procedures that are appropriate to the relativistic domain. A generalized form of the statistical interpretation is implicit in these procedures for calculation of observables and matrix elements. It is therefore our view that the statistical interpretation should be (and has been) generalized for the relativistic domain rather than discarded.

The scope of this paper is limited to the interpretation of quantum mechanics in the low-velocity nonrelativistic limit, the arena where almost all of the previous discussion on the interpretation of quantum mechanics has been focused. For this reason we shall not discuss further the interpretation of relativistic quantum mechanics. However, we are not aware of any new interpretational problems that are added be a fully relativistic quantum field theory beyond those associated with particle creation and space-time delocalization, as just discussed. The transactional interpretation of quantum mechanics presented in the next section is based on solutions of relativistically invariant differential field equations, is fully consistent with special relativity, and seems to accommodate these additional features of a relativistic quantum theory in a very natural way. We are therefore confident that the interpretation presented here, perhaps with minor embellishments, is appropriate for the interpretation of a fully relativistic theory of quantum mechanics.

D. The transactional interpretation

Now we are prepared to specify the premises of the transactional interpretation. In doing this we shall use a framework as similar as possible to the description of the Copenhagen interpretation given in Sec. II. We shall similarly use five principle elements, which we enumerate here.

(T-1) The uncertainty principle is as in (C-1). It is a consequence of the fact that a transaction in going to completion can project out and localize only one of a pair of conjugate variables from the other wave. This will be discussed further in Sec. III.H.

(T-2) The statistical interpretation is unchanged from (C-2). It is a consequence of the fact that the "echo" received by the emitter in initiating the transaction follows the Born probability law $P=\Psi\Psi^*$. This will also be discussed further in Sec. III.H, where the character of randomness in quantum mechanics is examined.

(T-3) All physical processes have equal status. The observer, intelligent or otherwise, has no special status. Measurement and measuring apparatus have no special status, except that they happen to be processes that connect to observers. The "wholeness" of (C-3) exists, but is not related to any special character of measurements but rather to the connection between emitter and absorber through the transaction. The "complementary" concept of (C-3) likewise exists, but like the uncertainty principle is just a manifestation of the requirement that a given transaction going to completion can project out only one of a pair of conjugate variables.

(T-4) The fundamental quantum-mechanical interaction is taken to be the transaction, as defined in the preceding section. The state vector of the quantum-mechanical formalism is a real physical wave with spatial extent and is identical with the initial "offer wave" of the transaction. The particle (photon, electron, etc.) and the collapsed state vector are identical with the completed transaction. The transaction may involve a single emitter and absorber or multiple emitters and absorbers, but is only complete when appropriate quantum boundary conditions are satisfied at all loci of emission and absorption. Particles transferred have no separate identity independent from the satisfaction of these boundary conditions. The correspondence of the state vector with "knowledge of the system" of (C-4) is a fortuitous but deceptive consequence of the transaction, in that such knowledge must follow and describe the transaction.

(T-5) A distinction is made between observable and inferred quantities. The former are firm predictions of the overall theory and may be subjected to experimental verification. The latter, particularly those that are complex quantities, are not verifiable and are useful only for interpretational and pedagogical purposes. It is assumed that both kinds of quantities must obey conservation laws, macroscopic causality conditions, relativistic invariance, etc. Resorting to the positivism of (C-5) is unnecessary and undesirable.

In summary, the transactional interpretation adopts the first two elements of the Copenhagen interpretation and is also able to accommodate aspects of elements (C-3). It drops the assertion of (C-4) that the solutions of a simple second-order differential equation relating mass, energy, and momentum are somehow related to "knowledge," and instead employs a usually neglected solution of that equation to construct the emitter-absorber transaction of (T-4). The transactional interpretation drops the positivism of (C-5) because the positivist curtain is no longer needed to hide the nonlocal backstage machinery.

It should also be pointed out that the substitution of element (T-4) for (C-4), in giving objective reality to the state vector, colors all of the other elements of the interpretation. Although the uncertainty principle (T-1) and the statistical interpretation (T-2) are formally the same as in the Copenhagen interpretation, their philosophical implications, about which so much has been written from the viewpoint of the Copenhagen interpretation, may be rather different.

E. The transactional interpretation and the quantum-mechanical formalism

Comparison of the formal notation used in quantum wave mechanics with the description of the transaction model as discussed in Sec. III.B shows an excellent correspondence between the two. Once the transactional interpretation is in firm conceptual grasp, description of quantum processes can be perceived in the mathematical procedures and notations used in the calculation of observables. In particular, the quantum-mechanical formalism makes extensive use of the operation of complex conjugation. For simple systems this operation is equivalent to the operation of time reversal (Wigner, 1950), which transforms retarded waves into advanced waves (see footnote 16). Thus, Ψ^* is the advanced confirmation wave equivalent to Ψ which is the retarded offer wave. As mentioned above, $\Psi\Psi^*$ is the offer-confirmation wave echo, which the emitter receives from a particular direction. Similarly,

$$\int \Psi\Psi^* dv \tag{11a}$$

is the sum of all such OW-CW echoes from all possible locations in space.

We consider other examples. A quantum-mechanical "overlap integral" of the form

$$\int \Psi_1\Psi_2^* dv \tag{11b}$$

can be interpreted in the transaction model as representing an average over all space of the "echoes" that an emitter sending out OW Ψ_1 receives from all possible absorbers sending back CW's which confirm transactions involving a final state described by Ψ_2. Moreover, the calculation of an expectation value of the variable x, which is deduced from a given wave function Ψ by the operator $\underline{X}$, so that $\underline{X}\Psi = x\Psi$, has the form

$$\langle x \rangle = \int \Psi_2^* \underline{X} \Psi_1 dv \; . \tag{11c}$$

This can be viewed as an average over space of the possible values of x that the operator $\underline{X}$ projects from the components of the OW which appear in the completed transaction. This interpretational approach can also be applied to other aspects of the quantum-mechanical formalism.

From one point of view; the transactional interpretation is so apparent in the Schrödinger-Dirac form of the quantum-mechanical formalism, which its combinations of normal and time-reversed waves, that one might fairly ask why this obvious interpretation of the formalism had not been made previously. No one can, of course, explain why something did *not* occur in the history of the development of quantum physics, but several relevant observations can be made. (1) The Schrödinger equation, which has been the focus of interpretational investigations from the 1920s to the present, does not have advanced solutions because it involves only the first time derivative (as discussed above in Sec. III.C), and so the association of Ψ^* with advanced solutions of the wave equation is by no means obvious. (2) After Heisenberg devised (C-4), concern about the nonlocality of quantum mechanics was effectively quenched for a long period, and considerations of interpretational problems were directed elsewhere, so it is only recently that concern about the nonlocality of the formalism has reemerged because of the Bell inequality test. (3) The relativistic wave equations that do have advanced as well as retarded solutions have been treated with suspicion because they do not uniquely describe the state of a given system (see footnote 18). Furthermore, the "purpose" of the advanced solutions was thought to have been found with the discrepancy of the antimatter counterparts of "normal" fermions (e.g., positrons, antiprotons, etc.). Therefore other "purposes" for the advanced-wave solutions were not sought, and they were not associated with the Ψ^* of the formalism. The Appendix discussed the few previous attempts to used advanced waves in quantum-mechanical interpretation.

F. Identity and complexity in the transactional interpretation

In dealing with the problem of identity, (T-2) gives the state vector the same meaning as did (C-2), as the medium for describing the probability of various possible quantum events. (T-4) deals more directly with the problem of identity and deals with it in a way quite different from that of (C-4). It asserts that the state vector is a real physical wave generated by the emitter, and travels through space to the final absorber as well as to many other space-time loci and many other potential absorbers. The state vector is the "offer wave" that initiates the transaction. Because of the quantum-mechanical boundary conditions, the transaction is only completed between a single emitter and absorber in a one-quantum effect. The particle itself (photon, electron, etc.) is not identical with the state vector but with the completed transaction, of which the state vector is only the initial phase.

Schrödinger's (see footnote 11) original attempt to interpret the state vector in a way similar to this was unsuccessful because of two problems: (1) it was found that quantum-mechanical waves exhibit nonlocal or action-at-a-distance behavior when they are interpreted as real waves physically present in space, and (2) it was found not to be possible to describe a particle as a "wave packet" which remained in a tight group envelope as it propagated. The transactional interpretation deal with problem (1) directly because it is explicitly nonlocal, and moreover the components of the SV which travel in directions other than that of the eventual absorber do not have to "disappear"; they are only virtual in the sense that they transfer no energy or momentum and participate in no transaction. Problem (2) is not a problem for the transactional interpretation because it is the formation of the transaction which localizes the energy and momentum transfer. The SV itself is therefore not required to stay in a tight packet in order to account for the particlelike behavior of the quantum event.

Since the transactional interpretation considers the SV to be physically present in space, it must deal directly with the problem of complexity. However, the transaction model is able to do this because at each point of the completed transaction at which there is a physical interaction, e.g., at the emitter and absorber loci, there is also a superposition of an advanced and a retarded wave of equal amplitude. Since $\Psi+\Psi^*=2\,\mathrm{Re}(\Psi)$, the collapsed SV becomes real, and there is no residual imaginary part of the SV to require explanation. The reader is referred to Sec. IV.D for a detailed example showing the real net amplitude from a transaction. In regions of space where there was no interaction, the SV is permitted to be complex, but these components are "virtual" and noninteractive since they produce no transfer of energy or momentum. Thus the transactional interpretation has restored algebraic reality to the description of the microcosm.

G. Collapse and nonlocality in the transactional interpretation

In the transactional interpretation the collapse of the state vector is interpreted as the completion of the transaction started by the OW and the CW exchanged between emitter and absorber. The emergence of the transaction from the SV does not occur at some particular location in space or at some particular instant of time, but rather forms along the entire four-vector that connects the emission locus with the absorption locus (or loci in the case of multiple correlated particles). The transaction employs both retarded and advanced waves, which propagate, respectively, along positive and negative lightlike (or timelike) four-vectors. Since the sum of these four-vectors can span spacelike and negative timelike and lightlike intervals, the "influence" of the transaction in enforcing the correlations of the quantum event is explicitly both nonlocal and atemporal.

Figure 5 shows an example of such combinations of four-vector for a two-photon transaction corresponding to an event in the Freedman-Clauser experiment. Note that, although all of the waves in the transaction lie along lightlike world lines, the "influence" that enforces the correlations between the two polarization measurements spans a spacelike interval and is therefore nonlocal. This nonlocality is an explicit feature of the transactional interpretation arising from the use of advanced waves.

Schrödinger (1935), in analyzing the EPR paradox, concluded that at least part of the problem lies in the way that time is used in quantum mechanics (in the context of the Copenhagen interpretation). The Copenhagen interpretation treats time in an essentially classical nonrelativistic way, and as we have seen in Sec. III.E, this leads to inconsistencies with relativity or causality in any nonsubjective Copenhagen description of collapse in, for example, the Freedman-Clauser experiment. The root of the inconsistencies lies in the implicit assumption of the Copenhagen interpretation that the SV collapse occurs at

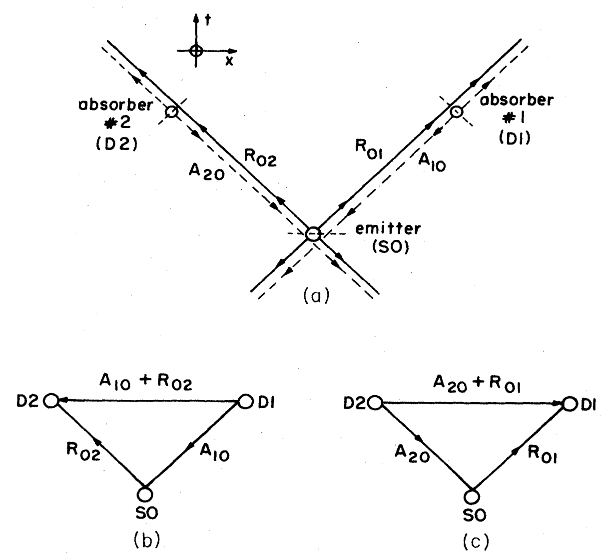


FIG. 5. (a) Minkowski diagram of Freedman-Clauser experiment showing lightlike world lines of advanced and retarded waves from emitter (solid line) and absorbers (dashed line), providing nonlocal enforcement of polarization correlations between absorption loci D_1 and D_2. (b) Four-vector sum of advanced and retarded waves connecting D_1 to D_2. (c) Four-vector sum of advanced and retarded waves connecting D_2 to D_1.

a particular instant at which a particular measurement is made and "knowledge" is gained, that before this instant the SV is in its full uncollapsed state, and that there can be a well-defined "before" and "after" in the collapse description. In the transactional interpretation the collapse, i.e., the development of the transaction, is atemporal and thus avoids the contradictions and inconsistencies implicit in any time-localized SV collapse.

Furthermore, the transactional description does not need to invoke arbitrary collapse triggers, such as consciousness, etc., because it is the *absorber* rather than the *observer* which precipitates the collapse of the SV, and this can occur atemporally and nonlocally across any sort of interval between elements of the measuring apparatus. This will be discussed further in the context of *Gedankenexperimente* in Sec. IV.

H. Completeness and predictivity in the transactional interpretation

As was shown in Sec. II.E, the solution of the problem of completeness posed in the EPR paper is implicit in the formalism of quantum mechanics, provided the SV is interpreted as a real physical quantity. Since the transactional interpretation treats the SV as a real physical quantity, quantum mechanics as interpreted with the transaction model is a complete theory. In particular, the SV brings to each potential absorber the full range of possible outcomes, and all have "simultaneous reality" in the EPR sense. The absorber interacts so as to cause one of these outcomes to emerge in the transaction, so that the collapsed SV manifests only one of these outcomes. The quantum-mechanical formalism ensures that if one of a pair of canonically conjugate quantities is localized in such a transaction, the other quantity is correspondingly delocalized, as required by the uncertainty principle.

The transactional interpretation also clarifies, but does not solve, the problem of predictivity. As was discussed in Sec. III.B, the beginning of a transaction can be viewed as the emitter sending out a retarded "offer" wave in various directions and receiving an "echo" back from the absorber in the form of an advanced confirmation wave, which has an amplitude proportional to Ψ^* (where Ψ is the complex OW evaluated at the absorber locus). In the usual circumstances there are a very large number of potential future absorbers, and if all provide such echoes, the emitter, at the instant of emission, has a large number of possible transaction possibilities from which to choose. In a single quantum event the boundary conditions will permit only one event to occur.

Born's probability law is therefore a statement that the probability of occurrence of a given transaction is proportional to the magnitude of the echo corresponding to that transaction which the emitter receives. This would seem to be a very plausible assumption. The quantum event, from this point of view, is a solution to a differential equation (the appropriate wave equation) for which a definite set of boundary conditions restrict the solutions but do not uniquely specify the solution. In this situation, the probability of a given solution is proportional to the "connectedness" of the participants, as indicated by the size of the echo that the absorber sends back to the emitter. The emitter is presented with echoes from potential absorbers which form a weighted list of possible transactions, from which only one may be chosen. The future absorbers can influence the past emission event only through the strength of their echo entry on this list, but cannot influence which entry is actually chosen for the transaction.

We note that there are analogous classical situations in which a system is specified by a set of differential equations with incompletely specified boundary conditions, for example, in the fluid dynamics of turbulent flow. Interestingly enough, there has been significant recent progress in these fields through the application of new mathematical techniques, such as the catastrophe theory, the theory of strange attractors, etc. It seems possible that similar techniques might some day be applied to the statistical processes of quantum mechanics.

Therefore, while the transactional interpretation does not alter the essentially statistical character of quantum mechanics, it has provided a glimpse of the "dice" that are at work in the statistical processes. The "dice" work to ensure an outcome consistent with the quantum boundary conditions of a transaction and are "loaded" in proportion to the magnitude of the echo that an emitter receives from potential absorbers.

I. Relativity and causality in the transactional interpretation

Several times we have mentioned the related constraints of nonlocality, relativistic invariance, and causality. As was previously mentioned, it would seem that the nonlocality of the transaction as defined above would give severe problems with both of the latter constraints by permitting both simultaneity tests across spacelike intervals and backward-in-time communication. However, this is not the case, as we shall show here.

The emitter-absorber transaction, although it has the effect of enforcing nonlocal correlations between separated parts of the system, cannot be used for nonlocal communication between observers. There are no residual advanced effects when the transaction is complete, and the reinterpretation of the advanced waves ensures that the result is observationally the same as if only retarded waves were present. Furthermore, as shown mathematically (Eberhard, 1977,1978; Ghirardi and Weber, 1979; Ghirardi *et al.*, 1980; Mittelstaedt, 1983), the nature of the correlations enforced between the separated parts of a FC experiment is such as to preclude the possibility of nonlocal communication between observers.

Since the transaction is atemporal, forming along the entire interval separating emission locus from absorption locus "at once," it makes no difference to the outcome or the transactional description if separated experiments occur "simultaneously" or in any time sequence. There is likewise no issue of which of the separated measurements occurs first and precipitates the SV collapse, since in the transactional interpretation both measurements participate equally and symmetrically in the formation of the transaction. Furthermore, the paths across which the correlation enforcing exchange takes place are lightlike four-vectors and remain so under any Lorentz transformation. Therefore the outcome and the transactional description of any correlation experiment is the same independent of the inertial reference frame from which it is viewed, as it must be if quantum mechanics and relativity are to be compatible theories.

The obvious "backwards-in-time" character of the transaction model warrants careful consideration of whether causality is preserved. In a sense, the transactional interpretation tells us that absorber "causes" the transaction that precedes it in time sequence, in violation of cause before effect. To come to terms with this aspect of the transactional interpretation it is necessary to consider carefully the nature of causality and the physical evidence that supports it. In a previous paper (Cramer, 1980) we have made the distinction between the strong principle of causality, which asserts that a cause must always precede its effect in any reference frame, and the weak principle of causality, which asserts the same thing, but only as it applies to macroscopic observations and observer-to-observer communication. There is no present experimental evidence in support of any causal principle stronger than the weak principle.

The transactional interpretation is completely consistent with the weak principle of causality. As discussed previously, the completion of the transaction removes all interacting advanced fields except the one connecting emitter with absorber, and the remaining advanced plus retarded superposition can be reinterpreted as purely retarded. Thus there are no "advanced effects," no evident acausal behavior, even at the microscopic level. Dispersion relations, etc., are completely consistent with microcausality at it is conventionally interpreted.

Nature, in a very subtle way, may be engaging in backwards-in-time handshaking. But the use of this mechanism is not available to experimental investigators even at the microscopic level. The completed transaction erases all advanced effects, to that no advanced-wave signaling is possible. The future can affect the past only very indirectly, by offering possibilities for transactions.

J. The arrow of time in the transactional interpretation

The formalism of quantum mechanics, at least in its relativistically invariant formulation, is completely evenhanded in dealing with the "arrow" of time, the distinction between future and past time directions. Even the apparently asymmetric action of an ideal macroscopic measurement in "preparing" a system in a definite quantum-mechanical state can be formally described in the context of the probability interpretation in a completely time-symmetric way (Aharonov *et al.*, 1964).

In the discussion of the Copenhagen interpretation in Sec. II.C, the point was made that the description of collapse in the Copenhagen interpretation is intrinsically unsymmetric in time. The transaction model of Sec. III.B gives the appearance of being more symmetrical, in that it treats past emitter and future absorber as equal terminators of the transaction that develops between them. However, the careful reader will perceive that there is a more subtle time asymmetry in the transactional description of the quantum event which is implicit in (T-2). There, the probability of a quantum event with emission from $(\mathbf{R}_1,T_1)$ to an absorber at $(\mathbf{R}_2,T_2)$ is assumed to be

$$P_{12}=|\Psi_1(\mathbf{R}_2,T_2)|^2 \tag{12}$$

rather than

$$P_{12}=|\Psi_2(\mathbf{R}_1,T_1)|^2 , \tag{13}$$

i.e., in the transaction model the emitter is given a privileged role because it is the echo received by the emitter, rather than that received by the absorber, which precipitates the transaction. Thus the past determines the future (in a statistical way) rather than the future determining the past.

The assumption of Eq. (12) is consistent with the usual formulation of quantum mechanics, the "post" formulation, which employs this rule in the evaluation of event probabilities. The alternative "prior" formulation, which employs Eq. (13) to evaluate probabilities, is rarely used but in the absence of violations of time-reversal invariance must give the same result for an exact calculation (DeVries *et al.*, 1974).

This symmetry between the post and prior formulations and the equivalent symmetry between post and prior versions of the transactional interpretation based on Eqs. (12) and (13) might be taken as a sufficient evenhandedness in the handling microreversibility, except for one problem. Nature has exhibited a clear violation of time-reversal invariance at the microscopic level in the decay of the K_L^0 meson. From experimental investigations of the *CP*-violating decay modes of the K_L^0 system, it is inferred that the time direction of the reaction $K_L^0+e^+\rightarrow\pi^++\bar{\nu}_e$ will be apparent in its cross section, i.e., the inverse reaction will have a qualitatively different reduced cross section from that of the forward reaction. It is not possible to provide a fixed target of any of the particles participating in this reaction, and therefore it is not experimentally feasible to observe directly either of these reaction modes. Therefore this should be considered as a *Gedankenexperiment.* Even so, it implies that the post and prior formalisms and interpretations are in principle distinguishable and therefore not equivalent.

The work of Aharonov *et al.* (1964) mentioned above showed that for ideal systems and measurements a plausible time-symmetric probability interpretation could be formulated to replace the usual one of Born (C-2). However, they found that in order to use this rule in a way that gave the same quantum-mechanical predictions as those of the conventional probability law of Eq. (12), they were forced to employ a time-asymmetric side condition on its use. Belinfante (1975) expanded this analysis to include nonideal measurements and systems and found that the importance of time asymmetries was even more apparent in the more general case. This body of work leads to conclusions similar to those stated above concerning the existence and inevitability of a quantum-mechanical arrow of time.

This microscopic quantum-mechanical arrow of time must be accounted for. Fortunately, a justification of such a time asymmetry for the case of WF electrodynamics has already been presented by the author in a previous publication (Cramer, 1983). A boundary condition model of the $T=0$ big bang was used to relate the electromagnetic arrow of time (the macroscopic dominance of retarded electromagnetic radiation) to the cosmological arrow of time (the direction of expansion of the universe). The arguments presented in that paper apply equally to the transaction model presented here and justify the use of probability law (12) rather than (13).

IV. EXAMPLES OF APPLICATION OF THE TRANSACTIONAL INTERPRETATION

The previous discussion contrasting the interpretations of quantum mechanics has been fairly abstract. Now we conclude with a more concrete elucidation of the transactional interpretation, demonstrating the power of its imagery by presenting examples of its application. The transactional interpretation is a conceptual model which provides a way of clearly visualizing complicated quantum processes. It is a way of thinking rather than a way of calculating. It may have considerable pedagogical potential as a more intuitive way of teaching quantum mechanics. It can provide insight and intuition that have been unavailable up to now in considering quantum phenomena. Here, we shall use the transaction model to examine and illuminate some of the "paradoxes," *Gedankenexperimente,* and real experiments that have been collected in the quantum-mechanics museum of curiosities.

We have elected to describe most of these *Gedankenexperimente* using photons of visible light passing through polarizing filters or beam splitters before detection by quantum-sensitive photomultiplier tubes. We have used circular and linear polarization of light as archetypes of noncommuting observables. Other authors (e.g., Feynman *et al.*, 1965) prefer to describe the equivalent experiments using electrons instead of photons, electron-spin orientation instead of photon polarization, Stern-Gerlach apparatus instead of polarizing filters and splitters, and electron counters instead of photomultipliers. While there may be differences in detail in such descriptions, there should be no fundamental change in the quantum effects illustrated. We believe that on balance the polarized-photon description is better connected to direct experience and that once embarked on a particular mode of description it is desirable to use consistent apparatus. However, the reader is assured that these *Gedankenexperimente* can be "performed" with electrons, protons, or even neutrinos with only minor descriptional differences. Analogous experiments could, in most cases, have been constructed using as noncommuting observables electron-spin states along the x axis and the y axis or even position and momentum observables. In most cases a real experiment could most easily be actually performed using photons of visible light and polaration observables.

A. Renninger's negative-result *Gedankenexperiment*

This is a *Gedankenexperiment* focusing on the collapse of the SV produced by *the absence* of an interaction of the system measured (a photon) with measurement apparatus. It was suggested by Renninger (1953) and was featured by de Broglie (1964) in his book on the interpretation of quantum mechanics. Dicke (1981) has recently stimulated renewed interest in this kind of "interaction-free measurement." The experimental arrangement is shown in Fig. 6.

Source S is located at the center of a spherical shell E_2 of radius R_2. The interior of E_2 is lined with a scintillating material that will produce a detectable flash of light, which will be seen by the observer, if E_2 is struck by a charged particle, e.g., an α particle. Inside E_2 is a partial concentric sphere E_1 of radius R_1, also lined with scintillator viewed by the observer. Partial sphere E_1 subtends solid angle Ω_1 as viewed from the position of source S. The portion of E_2 that is not shadowed by E_1 therefore subtends a solid angle $\Omega_2=4\pi-\Omega_1$. The source S is arranged so that on command it will emit exactly one α particle with an angular dependence that is completely isotropic, and with a velocity equal to V.

Now we consider the state vector $|S(t)\rangle$ as a function of time t, where t is the time that has elapsed since the source S has been commanded to emit an α particle. Before the α particle has traversed the distance R, i.e., for

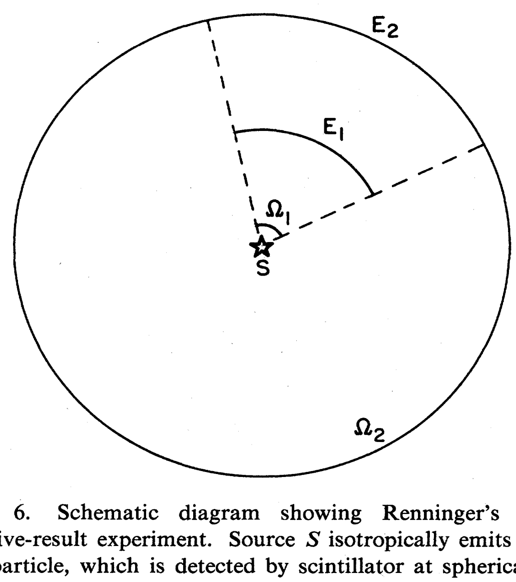


FIG. 6. Schematic diagram showing Renninger's (1953) negative-result experiment. Source S isotropically emits an alpha particle, which is detected by scintillator at spherical surfaces E_1 and E_2 depending on respective solid angles Ω_1 and Ω_2.

$0 < t < (R_1/V)$, the probability that the particle will produce a scintillation at E_1 is $P_1 = \Omega_1/4\pi$, and the probability that it will produce a scintillation at E_2 is $P_2 = \Omega_2/4\pi$. Thus the state vector might be written as

$$|S(t)\rangle = p_1|E_1\rangle + p_2|E_2\rangle ,$$

where (14)

$$|p_1|^2 = P_1 \text{ and } |p_2|^2 = P_2 .$$

Now suppose that time t becomes greater than R_1/V and that the observer *does not observe a scintillation from* E_1. Then the state vector must collapse, with the result that the probabilities become $P_1 = 0$ and $P_2 = 1$ and the state vector becomes $|S(t)\rangle = |E_2\rangle$ for $t > (R_1/V)$. The interpretational problem, as stated by Renninger and de Broglie, is that the state vector has collapsed abruptly and nonlinearly, and yet "the observer sees nothing at all on screen E_1, where nothing has happened." Thus the *absence* of an interaction with the measurement apparatus leading to the *absence* of an observation can collapse the SV as readily as a positive and definite observation.

This *Gedankenexperiment* helps us to understand why von Neumann (1932) and Wigner (1962) stressed the need for a conscious and intelligent observer as the triggering agent for the collapse of the SV. The change in "knowledge" when no scintillation is observed at E_1 at $t = R_1/V$, requires a *deduction* on the part of the observer as to what *should have happened* if the α particle had been aimed at E_1. It correspondingly casts some doubt on Schrödinger's (1935) principle of state distinction and on Heisenberg's (1960) irreversibility criterion, since no state-distinguishing record is made at $t = R_1/V$ and no irreversible process is initiated. Furthermore, one could imagine a more elaborate version of this experiment with a very large number of partial spheres inside E_2, so complicated that no human observer could possibly keep track of all the times and expectations of flashes that would signal the occurrence or elimination of various possible outcomes. One could speculate on how the SV collapse might occur in that situation.

The transactional interpretation avoids the conceptual problems implicit in this experiment by eliminating any SV collapse that occurs at some definite instant, such as $t = R_1/V$. Instead it employs an atemporal four-space description implicit in the transaction model: the state vector is emitted from the source at $t = 0$ as a retarded OW, which grows as a spherical wave front, part of which encounters E_1 at $t = R_1/V$ and the remainder of which encounters E_2 at $t = R_2/V$. The boundary condition of S that only a single α particle is emitted permits one and only one transaction to occur between S and E_1 or E_2. The transaction will occur with a probability proportional to the CW echoes that S receives from the two possible absorbers. These echoes will be proportional to the solid angles subtended by the two possible absorbers, i.e., Ω_1 and Ω_2, as expected. A single transaction forms in accordance with these probabilities through the exchange of advanced and retarded waves characterizing the transition of an α particle from S to E.

B. Wheeler's delayed-choice experiment

The previous *Gedankenexperiment* illustrated how the *absence* of an observation could collapse the SV. Now we consider a *Gedankenexperiment* in which the SV must *avoid* collapsing after interacting with the apparatus while the experimenter decides what experiment he wishes to perform. It is an example of the "delayed-choice" experiments proposed by Wheeler (1978). It is shown schematically in Fig. 7.

Here we have the usual Young's two-slit interference apparatus, illuminated by an ideal source S, which is a distance L from the slits. The source S emits one and only one photon in the general direction of the slits, on command from the observer who is operating the apparatus. Downstream of the slits are two different measuring devices. One of these devices is E, a photographic emulsion which, when placed in the path of the photon, will record photons at the position where they strike the emulsion. After many such events are recorded they will form a two-slit interference pattern characteristic of the photon's wavelength λ and *momentum* h/λ. The other measuring device consists of T_1 and T_2, a pair of tightly collimated telescopes with single-quantum-sensitive photomultiplier tubes at their image foci, each of which is focused on one of the two slits. A photon registered by T_1 or T_2 means that the photon has passed through slit 1 or 2, respectively. Therefore, T_1 and T_2 constitute a determination of photon *position*.

Such an apparatus is often used to illustrate the wave-

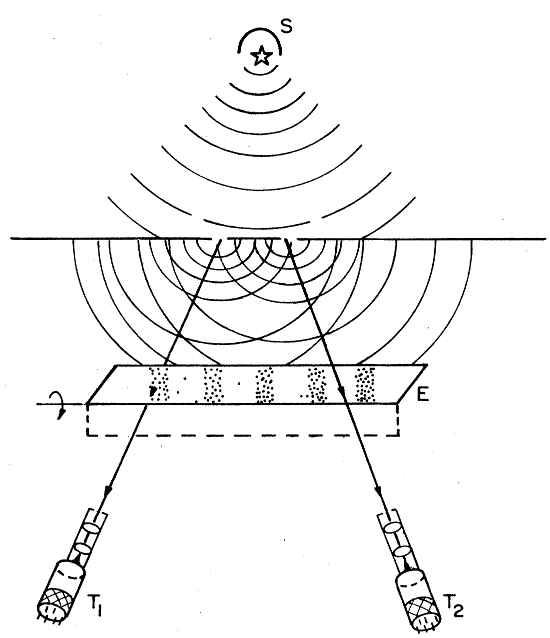


FIG. 7. Schematic diagram showing Wheeler's (1978) delayed-choice experiment. Source S emits a single photon in direction of the two-slit apparatus. On passing slits, the photon is absorbed by emulsion E if it is up. If emulsion is down, the photon continues on to detection in either telescope T_1 or telescope T_2, which are focused on slits. Operator decides position of emulsion after photon has passed through slit system.

particle duality of light. The light waves that form the interference pattern on the emulsion must have passed through *both* slits of the apparatus in order to interfere at the emulsion, while the photon particles that strike the photomultiplier surfaces can have passed through only the *one* slit at which the telescope was aimed. The emulsion measures momentum and the telescopes measure position, i.e., conjugate variables. Thus the two experimental measurements are "complementary" in Bohr's sense. The uncertainty principle is not violated, however, because only one of the two experiments can be performed on a given photon.

The emulsion E is mounted on a fast-acting pivot mechanism so that it can on command either be raised into position to intercept the photon from S or alternatively dropped out of the way so that the photon can proceed to T_1 or T_2. Thus when E is up we make an interference measurement requiring the photon to pass through both slits. When E is down we make a position measurement requiring that the photon pass through only one slit.

Wheeler's innovative modification of this old *Gedankenexperiment* is the following: The time $t > L/c$ at which the photon has safely passed the slits but not yet reached the apparatus is known to the experimenter, and he refrains from deciding which experiment to do, i.e., whether to place E up or down, until a time t when the photon must have already passed through the slit or slits. Therefore, the photon *has already emerged from the slit system* when the experimenter decides whether it should be caused to pass through one slit (E down) or both slits (E up). In a sense then, the cause (emulsion down or up) has come after the effect (passage through one or two slits).

This *Gedankenexperiment* demonstrates that the physical interaction of the photon with the slit system has not collapsed the SV, which must remain uncollapsed at least until the experimenter decides which experiment to do. Since the experimenter, after having made the decision, knows whether the photon will pass through one or both slits, it can be argued from (C-4) that it is his mental process of deciding which has precipitated the collapse of the SV rather than its subsequent interaction with E, T_1, or T_2, since after that decision is made he has the unambiguous knowledge of how many slits the photon has passed through.

Wheeler has explored the physical and philosophical implications of this and similar experiments, and has been led to assert the often quoted paradigm: "No phenomenon is a phenomenon until it is an observed phenomenon." In this statement he is emphasizing the role of the observer in precipitating an underlying indefinite reality into a definite observed state by the act of deciding on a measurement and then performing it. Again it would seem that observers, and indeed intelligent and decisive observers, are required to interpret this class of *Gedankenexperimente* using reasonable variants of the Copenhagen interpretation.

The transactional interpretation, however, is able to give an account of the delayed-choice experiment without resort to such observers as the triggers of collapse. In the transactional description the source S emits the retarded OW, which propagates through both slits and reaches the locus of E, where (a) it finds that the emulsion E is up and is absorbed by it, as illustrated in Fig. 8(a), or (b) it finds E down and proceeds to T_1 where it is absorbed, as illustrated in Fig. 8(b), or (c) it finds E down and proceeds to T_2 where it is absorbed, as illustrated in Fig. 8(c). For case (a), in which the photon is absorbed by E, the advanced CW retraces the path of the OW, traveling in the negative time direction through both slits and back to source S. Therefore the final transaction, as shown in Fig. 8(a), forms along paths that pass through both slits in connecting the source S with the emulsion E. The transaction is therefore a "two-slit" quantum event. The photon can be said to have passed through both slits to reach the emulsion.

For cases (b) and (c) the OW also passes through both slits on its way to the photomultiplier telescopes T_1 and T_2. However, when the absorption takes place at *one* of the telescopes (not both because of the single quantum boundary condition), the collimation system of that telescope prevents the CW from passing through more than one of the slits, since the collimation only permits passage through the slit at which the telescope is aimed. Thus the CW passes through only one slit in passing from T_1 (or T_2) to S, and the transaction that forms is characteristic

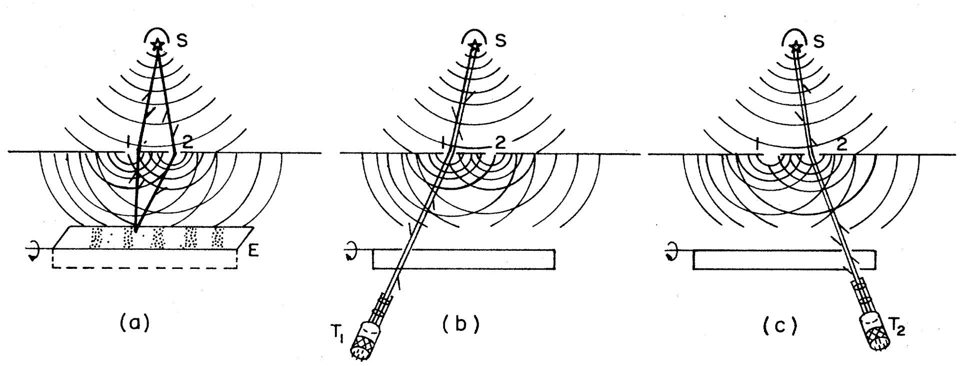


FIG. 8. Schematic diagram of possible transactions in Wheeler's delayed-choice experiment (see the text). (a) Photon is detected at emulsion E. (b) Photon is detected by photomultiplier telescope T_1. (c) Photon is detected by photomultiplier telescope T_2.

of a "one-slit" quantum event. The photon can be said to have passed through only one slit to reach the telescope.

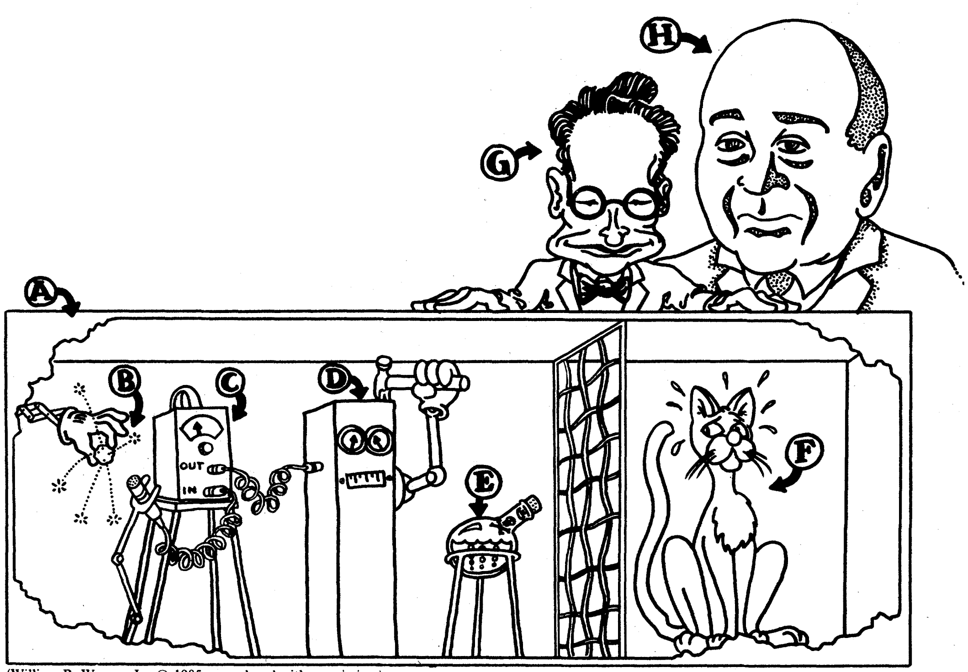
Since in the transactional description the transaction forms atemporally, the issue of when the observer decides which experiment to perform is no longer significant. The observer determined the experimental configuration and boundary conditions, and the transaction formed accordingly. Furthermore, the fact that the detection event involves a measurement (as opposed to any other interaction) is no longer significant, and so the observer has no special role in the process. To paraphrase Wheeler's paradigm, we might say, "No offer is transaction until it is a confirmed transaction."

C. Schrödinger's cat and Wigner's friend

Perhaps the most famous *Gedankenexperiment* demonstrating an interpretational "paradox" of quantum mechanics is the Schrödinger's cat paradox (1935) illustrated in Fig. 9. Schrödinger describes a "höllenmaschine" which dramatizes the interpretational problem. An ideally isolated system (a sealed, soundproof, and well-insulated box with an adequate oxygen supply) is prepared so that it contains a Geiger counter placed near a radioactive source that emits γ rays. The source of the γ rays is adjusted in strength so that in a period of 1 h it has a probability of exactly 50% of causing the Geiger counter to record one count. The counter mechanism is connected to a solenoid device which, if a count occurs, will shatter a flask of prussic acid, thereby filling the box with lethal fumes. Of course there is also a probability of 50% that no count will occur and the flask will remain intact.

The experimenter places a cat inside the box, seals it, and leaves the system undisturbed for 1 h. At the end of the hour the experimenter deactivates the counter, opens the box, and observes the state of the system. Two states are possible: a state $|\mathscr{A}\rangle$ (alive cat) in which the flask is unbroken and the cat remains alive and a state $|\mathscr{D}\rangle$ (dead cat) in which the flask has shattered and the cat has been killed. Schrödinger's question is: "What is the quantum-mechanical state vector of the system immediately before the box is opened and the observation is made?"

Quantum mechanics, as interpreted by the Copenhagen interpretation, would seem to tell us that the SV was $[\alpha\,|\mathscr{A}\rangle+\beta\,|\mathscr{D}\rangle]$, where $\alpha\alpha^*=\beta\beta^*=\frac{1}{2}$. In other words, the SV of the system consists of equal components of the live cat wave function $|\mathscr{A}\rangle$ and the dead cat wave function $|\mathscr{D}\rangle$ until such time as the observer collapses the SV into one or the other of these states by making an observation, since it is the change in the observer's knowledge that precipitates the SV collapse. In the period just before the observation is made the SV describes the cat as 50% alive and 50% dead. This description, which may seem plausible enough when applied to a microscopic system (or even to a statistically large ensemble of



(William R. Warren, Jr., © 1985, reproduced with permission.)

FIG. 9. Schematic diagram illustrating the Schrödinger cat *Gedankenexperiment* (Schrödinger, 1935). A sealed and insulated box (A) contains a radioactive source (B), which has a 50% chance during the course of the "experiment" of triggering Geiger counter (C), which activates a mechanism (D) causing a hammer to smash a flask of prussic acid (E) and kill the cat (F). An observer (G) must open the box in order to collapse the state vector of the system into one of the two possible states. A second observer (H) may be needed to collapse the state vector of the larger system containing the first observer (G) and the apparatus (A)—(F). And so on.

Schrödinger's cat experiments), appears rather absurd when applied to an individual complex organism like a cat.

Wigner (1962) further heightened the weirdness implicit in the Copenhagen description by replacing the cat with a "friend," i.e., an intelligent observer, and at the same time replacing the prussic acid mechanism with a less lethal piece of apparatus, e.g., a light bulb that is switched on when a count is recorded. The experimenter then performs the experiment, which can be considered as two experiments: (a) treating friend + plus as a system, the experimenter makes an observation and (b) treating the counter mechanism as a system, the friend makes an observation that is subsequently reported to the experimenter.

We shall not reproduce Wigner's detailed analysis of this *Gedankenexperiment* here, but will state his conclusion: consciousness must have a special role in the collapse of the SV, for otherwise one must deal (at least on the philosophical level) with uncollapsed SV's containing conscious observers in a multiplicity of alternative states. Several others have suggested alternative ways of avoiding uncollapsed SV's describing conscious observers. Heisenberg (1960) has suggested that the SV collapses when the system enters the domain of thermodynamic irreversibility, e.g., as soon as a piece of macroscopic apparatus becomes involved. Schrödinger (1935) suggested that as soon as a permanent record of the system's state is made, e.g., by smashing the flask, the SV is collapsed. Everett (1957) has dispatched the interpretational problem posed by these *Gedankenexperimente* by suggesting that the SV *never* collapses. Instead, in the Everett-Wheeler interpretation of quantum mechanics (see the Appendix), the universe "splits" with each quantum event into alternate universes, each characterized by one of the possible outcomes of the event. None of these modifications of the basic Copenhagen interpretation has gained wide acceptance, and, as discussed in Sec. II and the Appendix, none is without its own interpretational problems.

The central focus of the problems posed by Schrödinger's cat and Wigner's friend is the question of *when* the SV actually collapses. The transactional interpretation avoids the implicit dilemma because in the transaction model the SV collapse, i.e., the formation of the transaction, is atemporal. During the entire 1-h period that the box is closed the radioactive source S of Schrödinger's apparatus sends out a very weak OW. This OW may or may not, with equal 50% probabilities, be confirmed by a CW from the Geiger counter so that a completed transaction is formed. If a transaction is formed, then the count is recorded, the flask shattered, the cat killed. If such a transaction is not formed, then the cat remains alive. The SV (or OW) does indeed have implicit in it both live and dead cat possibilities, but the completed transaction allows only one of these possibilities to become real. Because the collapse does not have to await the arrival of the observer, there is never a time when "the cat is 50% alive and 50% dead." And the need for consciousness, permanent records, thermodynamics, or alternate universes never arises. The "buck stops" at the absorber, in this case the Geiger counter, and the uncollapsed SV need not be tracked any further.

To state this another way, Schrödinger's question is: "When can a quantum event be considered finished?" Is it when the γ ray leaves the radioactive nucleus? Is it when it interacts with the Geiger counter? When the flask is smashed? When the cat dies? When the observer looks in the box? When he tells a colleague what he observed? When he publishes his observation? When ...? A billiard shot is over when the billiard balls stop colliding and come to rest. The atomic "billiard balls" of a quantum billiard shot continue to collide forever, never coming to rest so that the shot can be considered finished.

The source of confusion here is that the wrong question is being asked. The Copenhagen interpretation has led us to ask *when* the SV collapses instead of *how* it collapses. There is not a "when," not a point in time at which the quantum event is finished. The event is finished when the transaction forms, which happens along a set of world lines that include all of the events listed above, treating none of them as the special conclusion of the event. If there is one particular link in this event chain that is special, it is not the one that ends the chain. It is the link at the beginning of the chain when the emitter, having received various CW's from its OW, reinforces one of them in such a way that it brings that particular CW into reality as a completed transaction. The atemporal transaction does not have a "when" at the end.

D. Transmission of photons through noncommuting polarizing filters

The behavior of quantum systems in response to measurements of noncommuting variables is often cited as one of the interpretational problems of quantum mechanics and has been used as a justification for the development of quantum logics. However, one can usually find excellent classical analogs of such measurements, e.g., the Fourier time-frequency complementarity of electric pulse wave forms and the transmission of light through successive polarizing filters.

Therefore, it is instructive to consider the quantum-mechanical treatment of the transmission of light through polarizing filters as an illustration of the application of the transactional interpretation. We shall specifically select a case where the handling of complex amplitudes is required so that this aspect of the transactional interpretation can be shown. Figure 10(a) shows the system to be considered. A single photon of light in the form of a plane wave is emitted by source S and travels along an optical bench to the single-quantum-sensitive photomultiplier detector D. In traversing this path it passes through three polarizing filters, which we shall call H, R, and V, to indicate that they transmit with 100% efficiency light in a pure state of horizontal linear polarization, right circular polarization, and vertical linear polarization, respectively, while completely absorbing light that has the orthogonal polarization.

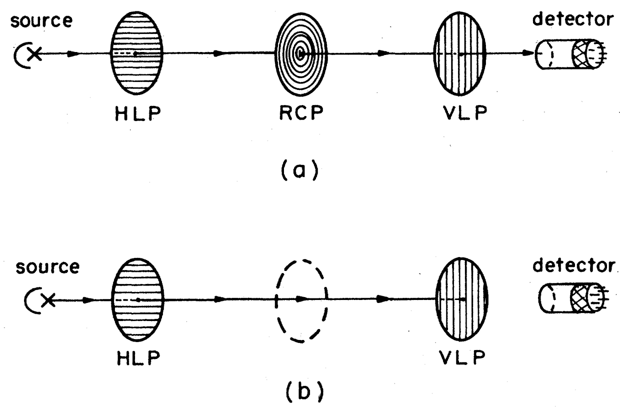


FIG. 10. (a) Schematic diagram showing the passage of a single photon through successive noncommuting polarizing filters F_1, F_2, and F_3 (see the text). (b) Same diagram with filter F_2 removed.

This example is chosen because the operators characterizing linear polarization eigenstates do not commute with the operators characterizing circular polarization eigenstates, and so linear and circular polarization are noncommuting variables. The two descriptions linear and circular, represent two related bases. In particular, if $|H\rangle$, $|V\rangle$, $|R\rangle$, and $|L\rangle$ represent pure states, respectively, of horizontal linear, vertical linear, right circular, and left circular polarization, then they are related by the transformation equations

$$|R\rangle=\alpha(|H\rangle+i|V\rangle), \tag{15a}$$

$$|L\rangle=-\alpha(|H\rangle-i|V\rangle), \tag{15b}$$

$$|H\rangle=\alpha(|R\rangle+|L\rangle), \tag{15c}$$

$$|V\rangle=-i\alpha(|R\rangle-|L\rangle), \tag{15d}$$

where $\alpha=(2)^{-1/2}$ and $i=(-1)^{1/2}$.

The transactional interpretation provides the following description of the transmission of a photon from S to D: The source S produces a retarded OW in the form of a general SV including all possible states of polarization. This wave then passes through filter H. The filter transmits[19] only $|H\rangle$, i.e., that component of the SV which corresponds to a state of pure horizontal linear polarization (HLP). This wave then travels to filter R, which transmits only that component of $|H\rangle$ in a pure state of right circular polarization (RCP). From Eq. (15c) this is $\alpha|R\rangle$. This RCP wave then travels to filter V, which transmits only that component in a pure state of vertical linear polarization (VLP). From Eq. (15a), this will be $\alpha(|\alpha|V\rangle)=(i/2)|V\rangle$. This VLP wave then strikes the photocathode of D and is absorbed and detected.

According to the transaction model this is only half of the story. In absorbing the incident retarded wave, the photocathode must produce a "time-mirrored" advanced wave or CW. This wave will be the complex conjugate of the incident OW and has the form

$$\text{CW}=\text{OW}^*=[(i/2)|V\rangle]^*=(-i/2)\langle V| \ . \tag{16}$$

The advanced CW travels back along the track of the incident OW until it encounters filter V, where it is perfectly transmitted since it is already in a state of pure VLP.

The CW then proceeds along the track of the OW until it reaches filter R, where only its RCP component is transmitted. We can use Eqs. (15) for changing the basis of advanced waves by taking the complex conjugates (i.e., the time reverse) of both sides of the equations to obtain a new set of transformation equations, which we shall denote as Eqs. (15′). Employing that procedure, Eq. (15d′) shows us that the transmitted CW will have the form

$$(-i/2)(i\alpha\langle R|)=\tfrac{1}{2}\alpha\langle R| \ . \tag{17}$$

The CW then proceeds until it reaches filter H, where only its HLP component is transmitted. Equation (15a′) shows us that the transmitted wave will be

$$\text{CW}=\tfrac{1}{2}\alpha(\alpha\langle H|)=\tfrac{1}{4}\langle H| \ . \tag{18}$$

Thus the source has sent out an OW of unit amplitude and has received back a CW in state $\langle H|$ with an amplitude of $\frac{1}{4}$. The amplitude factor is the product of the OW amplitude at D, i.e., $(i/2)$, and the complex conjugate of that amplitude or $(-i/2)$, the complex conjugate arising from the time-reversed character of the CW. Note that the final amplitude of the echo CW is a real number,[20] even though complex amplitudes were found at many intermediate points of the transaction. This is a general feature of the transactional interpretation.

This is a concrete example of the assertion (T-2) that the probability of a transaction is proportional to the amplitude of the CW echo from a potential absorber and is also an illustration of the operation of the Born probability law $P=\Psi\Psi^*$. The transaction will be confirmed and the photon transmitted from S to D with a probability of $\frac{1}{4}$ and will arrive at D in a state of pure vertical polarization. There will also be a probability of $\frac{3}{4}$ that the photon will not be transmitted to D but instead will be absorbed by one of the filters. These are the same transmis-

[19]Here "transmits" means that the state vector (or offer wave) is split into components that stop at the polarizer and components that continue on the original path. The "transmitted" OW is in the latter class. Thus a transaction involving components that are not transmitted cannot form on the path beyond the polarizer.

[20]The net wave at the absorber locus is also real. Assume that in the vicinity of the absorber $|V\rangle$ depends on x and t as $|V\rangle=\exp[i(kx-\omega t)]$. Then $\langle V|=[|V\rangle]^*=\exp[-i(kx-\omega t)]$. Thus at the absorber locus the net wave is the superposition $(i/2)|V\rangle+(-i/2)\langle V|=-\sin(kx-\omega t)$, which is a real quantity.

sion and absorption probabilities as are given by classical optics[21] for the transmission of an initially HLP beam of light from S to D.

Now consider the modification of the apparatus shown in Fig. 10(b), in which the second filter R has been removed. Now the OW is placed in a pure state of HLP by filter H, so that when it travels to filter V it cannot be transmitted. Therefore, no OW reaches the detector D and no transaction from S to D takes place. With filter R removed, the transmission of the apparatus drops from 25% to 0%.

The transactional description of other experiments involving noncommuting variables can be constructed by employing the same procedures used above (see Sec. IV.G, for example). In each case it will be found that the probability of the quantum event under consideration is just the real and positive amplitude of the echo CW response to the OW from the emitter.

E. The Freedman-Clauser experiment and Herbert's paradox

As discussed in Sec. I, the Freedman-Clauser experiment (1972) and later tests of the Bell Inequality (Clauser and Shimony, 1978; Aspect *et al.*, 1982a, 1982b) have demonstrated that quantum mechanics (and nature) cannot simultaneously have the properties of CFD and locality. In the present work we have advocated the view that the solution to this dichotomy is that quantum-mechanical description of nature is intrinsically nonlocal. Quantum mechanics as viewed by the transactional interpretation is an explicitly nonlocal description of quantum processes. It is therefore informative to apply this description to the multiparticle quantum events of the Freedman-Clauser experiment.

We refer to Fig. 1(a) describing the FC experiment in Sec. II.D. Source S produces a pair of photons that are only allowed to leave the source in directions 180° apart and that are constrained by angular-momentum conservation to be in the same state of polarization—which may be either a helicity (circular polarization) state or a linear polarization state. As discussed in Sec. IV.D, we can use any circular or linear basis to describe all such states.

The FC apparatus employs polarimeters consisting of linear polarizing filters placed in front of single-quantum-sensitive photomultiplier detectors. The linear polarizing axes of the filters are rotated about the axis defined by the line of flight of the photons. The coincidence counting rate between photon events detected in the two photomultiplier tubes is recorded as a function of the angle settings of the two filters. This coincidence rate is expected by symmetry to depend only on the relative angle θ between the angle settings of the two filters.

As discussed in Sec. II.D, quantum mechanics predicts, if the filters are ideal, that the coincidence rate will be proportional to $\cos^2\theta$, i.e., maximum when the polarimeter axes are aligned and zero when they are at right angles. It is this deceptively innocent result which stands in violation of Bell's inequality and which implies some nonlocality in the enforcement of this correlation between separated measurements.

The analysis of this experiment with the transactional interpretation will be similar to that of Sec. IV.D. The source S generates two correlated OW's $|A_1\rangle$ and $|B_1\rangle$, which are constrained by angular-momentum conservation to be in the same state of polarization, but which are in an indefinite and uncollapsed state. These OW's propagate down the two arms of the apparatus until they encounter the polarizing filters F_A and F_B. Each filter transmits only $|A_2\rangle$ or $|B_2\rangle$, the component of the OW that matches its orientation angle. These components then reach the photomultiplier tubes A and B and are absorbed. The absorption process produces CW's $\langle A_2|$ and $\langle B_2|$, the time reverses of the incident OW's, and these propagate in the negative time direction back through the apparatus to the filters. Since each CW is in a state that matches the filter orientation, they are transmitted unmodified to the source.

Thus, the source receives CW echoes $\langle A_2|$ and $\langle B_2|$ in response to the two emitted OW's. However, these responses will not, for arbitrary settings of the polarizing filters, satisfy the boundary condition that the two photons be in the same polarization state, and therefore they cannot participate fully in the formation of the transaction, which is a double quantum event involving the emission and detection of both photons. Rather, as required by (T-4), the transaction will project from one of these CW's only the component that matches the other and will ignore the nonmatching or orthogonal component. Thus, if the polarizing filters are aligned, the CW's will match and there will be a maximum coincidence rate, while if the polarizing filters are at right angles, the CW's will be in orthogonal states, no transaction can be formed, and the coincidence rate will be zero.

As discussed in Sec. III.F and illustrated in Fig. 5, the combination of four-vectors in such a two-photon transaction provides a "bridge" across the spacelike interval between the two detection events to enforce their correlation, even though the structural members of the

---

[21]The transmission of the initial photon through the first polarizing filter has not been considered in this calculation. There is a 50% probability that a photon of indefinite polarization would be absorbed by the first filter, giving a net transmission of a beam of photons from the source of the detector of $\frac{1}{8}$ rather than $\frac{1}{4}$, which is the transmission from the first filter on. It should be emphasized that there is a difference in definition between "unpolarized" in the classical sense and "of indefinite polarization" in the quantum mechanical sense. This represents a fundamental difference between the quantum-mechanical and classical treatments of polarization and the lack thereof. It is related to the Furry modification of quantum mechanics discussed in Sec. II.D. This difference is not related to the interpretation, but to the quantum-mechanical formalism. The latter accommodates randomly polarized light in the classical sense by averaging over polarization orientations.

bridge are lightlike four-vectors which transform properly under Lorentz transformations. Thus the transactional description of the FC experiment is explicitly nonlocal. As demonstrated by several authors (Eberhard), 1977, 1978; Ghiradi and Weber, 1979; Ghirardi *et al.*, 1980), it is not possible to exploit this nonlocality for the purposes of nonlocal observer-to-observer communication.

Recently, however, Herbert (1982) has suggested a *Gedankenexperiment* that is a modification of the basic FC apparatus intended for communication. Herbert's arguments suggested that nonlocal observer-to-observer communication might be experimentally possible. A simplified version of Herbert's apparatus is illustrated in Fig. 11. Essentially he has added a laser "gain tube" amplifier to arm B of the FC apparatus so that the process of stimulated emission is used to "clone" the photon from the source, making multiple copies of the $|B_1\rangle$ state vector so that its state of polarization can be definitively determined by using beam splitters to route duplicate photons to multiple measurements with polarimeters aligned along many axes. Such measurements, it is asserted, will reveal the kind of polarization measurement that was performed on $|A_1\rangle$, e.g., a circular polarization measurement or a linear polarization measurement. This, it is asserted, will permit an observer at arm A to "telegraph" a message to a second observer at arm B across a spacelike interval by encoding the message in the time structure of the measurements performed (e.g., as 1's and 0's in a binary code).

To illustrate the paradox more explicitly, we consider a simplified version of Herbert's apparatus in which (in the rest system of the apparatus) the single photon arriving in arm A is measured slightly earlier than whatever measurements occur in arm B. Moreover, we assume that the photon in arm B is duplicated *once* and that each photon separately is subjected to a linear polarization (LP) measurement which determines whether it is in a horizontal (H) or a vertical (V) state of linear polarization. Herbert argues that if the "sender" at arm A has chosen the LP basis for his measurement, i.e., the same basis as that of the pair of measurements in arm B, then the photons reaching arm B will be in a pure state of linear polarization and the measurements there will always match, i.e., the measurements will always show $H+H$ or $V+V$ and never $H+V$ or $V+H$. On the other hand, it is argued, if the sender at A uses the circular polarization (CP) basis for his measurement then the photons reaching arm B will be in a mixed state of linear polarization and the measurements there will show matching ($H+H$ or $V+V$) and nonmatching ($H+V$ or $V+H$) responses with equal probability. Thus a repeated sending of the message at arm A would permit its reception in arm B, even in the presence of strong spontaneous emission noise in the laser amplifier.

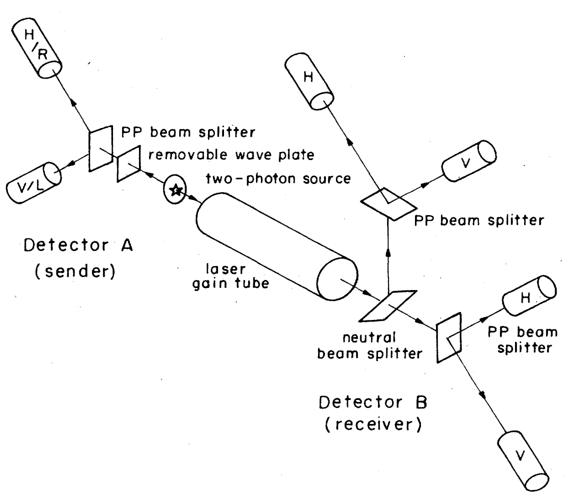


FIG. 11. Schematic diagram of Herbert's (1982) suggested experiment. The Freedman-Clauser apparatus is modified by placing a laser gain tube in the path of photon 2, so that it is "duplicated." Beam splitters then permit separate polarization measurements of the resulting photons.

Herbert's arguments are implicitly based on the Copenhagen interpretation's account of state vector collapse. It is assumed that once the measurement at arm A has occurred, the SV for the photons reaching arm B will have collapsed into a definite state of polarization, which can then be determined to discover what A measurement had been made. This description is consistent with the Copenhagen interpretation, but it does give a special and rather unsymmetric role to one of the three measurements that are performed on the same system. This is perhaps justified by the time sequence of the measurements if the arm A measurement occurs first.

Even before Herbert's paper appeared in print there were several "refutations" published which concerned themselves with spontaneous emission noise in the stimulated emission process (Wooters and Zurek, 1982; Milonni and Hardies, 1982) or with the description of the quantum state that such an amplification process would produce. The former analyses are probably irrelevant to the heart of the problem because the envisioned communication could in principle persist even in the presence of such noise. On the other hand, the analysis of Dieks (1982) seems to provide a satisfactory (if rather formal) resolution of the "paradox."

Analysis of the *Gedankenexperiment* with the transactional interpretation is relatively straightforward. The transaction corresponding to each detection event must involve the absorption of all three photons and must connect loci at the source, three of the detectors, and the laser amplifier. The transaction can be verified only if all of the CW's are in nonorthogonal states. Thus the polarimeters in arm B can *never* record events that are $H+V$ or $V+H$, no matter what measurement the sender performs, because this would generate orthogonal CW's, which would not verify the overall multiphoton transaction. Therefore, even under ideal circumstances (which may not be realizable experimentally), the responses in arm B must always match, and no observer-to-observer message can be transmitted. Dieks used a more detailed analysis employing the quantum-mechanical formalism to arrive at the same conclusion: multiple polarimeters in arm B

of the apparatus will always give matching responses, no matter what is measured in arm A.

The basic fallacy in Herbert's argument, from the viewpoint of the transactional interpretation is that he was led by the Copenhagen interpretation to assume that the measurement in arm A collapses the SV, so that the apparatus in arm B is presented with a photon in a definite state, and that the correlations between arms A and B are enforced by this collapse while the correlations between the duplicate photons in arm B are not. The transactional interpretation requires the enforcement of *all correlations equally.* There is no single measurement that collapses the SV presented to the others, but rather a transaction enforcing correlation of all measurements, independent of the time sequence in which they occur. This enforcement can only be accomplished if the measurements in arm B match not only the measurement in arm A but also match each other.

F. The Hanbury-Brown-Twiss effect

The Hanbury-Brown-Twiss (HBT) effect is an example of the interference of radiation sources that are incoherent (Klauder and Sudarshan, 1968). It has been applied to the measurement of the diameters of nearby stars with radio interferometry and to investigation of the "hot spot" developed in a relativistic heavy-ion collision in which pions are produced (Gyulassy *et al.*, 1979). The effect applies equally well to classical waves and to particlelike quanta.

A simplified version of a HBT interference measurement is illustrated in Fig. 12. Sources 1 and 2 are separated by a distance d_{12}. Both sources emit photons of the same energy $h\nu$ but are completely incoherent. The radiation from the two sources is detected by detectors A and B, which are separated by a distance d_{AB}. The line of centers of the sources is parallel to the line of centers of the detectors, and the two lines are separated by a distance L.

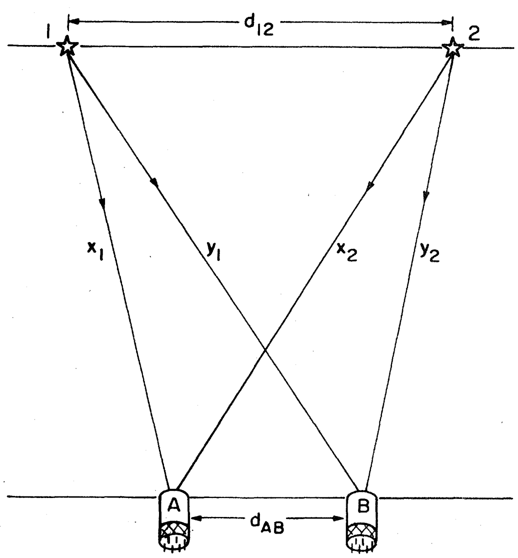


FIG. 12. Schematic diagram of the Hanbury-Brown—Twiss experiment (Klauder and Sudarshan, 1968) demonstrating coherent interference in light from incoherent sources. Sources 1 and 2 are separated by distance d_{12} and emit photons of identical wavelengths, which are detected by detectors A and B located a distance d_{AB} apart. The distance between source plane and detector plane is L. A product of coincidence between detector outputs results in a composition signal exhibiting an interference effect depending on both d_{12} and d_{AB}.

It will not be demonstrated here, but a signal that is a product of (or coincidence between) the signals received at A and B (indicating that photons have simultaneously triggered both detectors) reflects the *coherent* interference of the two sources and depends on the source separation d_{12} as well as the detector separation d_{AB}. Measurements made at a number of values of d_{AB} can therefore be used to determine d_{12}, in a manner analogous to moving a single detector in an interference pattern to determine the separation of a pair of coherent sources. This is the HBT interference effect.

There is a lesson for applications of the transactional interpretation in this kind of interference phenomenon: particles like photons and electrons cannot consistently be described as blobs that travel from point 1 or 2 to point A or B. In the HBT effect a whole photon is assembled at the detector out of half-photons contributed by each of the two sources. Consider a transaction in which photons are emitted by 1 and 2 and detected by A and B so that their product signal exhibits HBT interference. In the transactional description of such an HBT event, retarded OW's $|x_1\rangle$ and $|y_1\rangle$ are emitted by the source 1 and travel to detectors A and B, respectively. Similarly, OW's $|x_2\rangle$ and $|y_2\rangle$ are emitted by the source 2. Detector A receives a composite OW $|A_{12}\rangle$, which is a linear superposition of $|x_1\rangle$ and $|x_2\rangle$ and seeks to absorb the "offered" photon by producing advanced CW $\langle A_{12}|$, the time reverse of that superposition. Detector B similarly responds to composite OW $|B_{12}\rangle$. These advanced waves then travel back to the two sources, each of which receives a different linear superposition of $\langle A_{12}|$ and $\langle B_{12}|$.

A HBT transaction is formed which removes one energy quantum $h\nu$ from each of the two sources 1 and 2 and delivers one energy quantum $h\nu$ to each of the two detectors A and B. For many combinations of source and detector separation distances, the superimposed OW's and/or CW's are nearly equal and opposite, so that the composite wave is very weak and the transaction is very improbable. For a few ideal combinations of source and detector separation distances, all of the composite waves are strong because their components coherently reinforce, and in this case the transaction is much more probable. The transaction probability depends on the separation distances in just the way predicted by quantum mechanics. Thus the HBT effect is completely consistent with the transactional interpretation.

However, there is an interesting point here: neither of the photons detected by A or B can be said to have originated uniquely in one of the two sources. Each detected photon originated partly in each of the two sources. It might be said that each source produced two fractional photons and that fractions from two sources combined at a detector to make a full-size photon. This is what (T-4) emphasized in stating that "particles transferred have no separate identity independent from the satisfaction of these (the quantum-mechanical) boundary conditions." The boundary conditions here are those imposed by the HBT geometry and detection criteria.

This two-photon event may be viewed as a simple case of more general multiphoton (or multiparticle) events, which may involve many sources and many detectors. Such transactions can be viewed as assembling particles at a detector from contributions derived from a number of sources, with no one-to-one correspondence between particles emitted and particles detected except in overall number. One way of stating this is to emphasize that the spatial localization of the emitter (or the absorber) may be very unclear and indefinite, as long as all boundary conditions are satisfied. Likewise the time localization of the emission event (or absorption event) can be made very indefinite by a choice of experimental conditions, e.g., very low emission probability as in the Pflegor-Mandel experiment (1967).

G. The Albert-Aharonov-D'Amato predictions

The predictions of Albert, Aharonov, and D'Amato (1985) clarify an old problem, the question of retrospective knowledge of a quantum state following successive measurements of noncommuting variables (Aharonov *et al.*, 1964). The assumption of contrafactual definiteness mentioned in the Introduction plays an important role in the Albert, Aharonov, and D'Amato (AAD) predictions because these concern the retrospective knowledge of the observer about the outcome of experiments which *might have been* performed on the system in the time interval between one of the measurements and the other. We need the CFD assumptions (a) that the various alternative possible measurements which might have been performed on the system would each have produced a definite (although unknown and possibly random) observational result and (b) that we are permitted to discuss these results. Under the assumption of CFD, the AAD predictions provide a challenging quantum-mechanical interpretational problem.

As a simple example of the AAD predictions, consider the experiment illustrated in Fig. 13(a). A photon is emitted from source S and is transmitted through a filter H, which passes only horizontal linearly polarized light. It then travels a distance L and is transmitted through a second filter R, which passes only right circularly polarized light. The photon is then detected by a quantum-sensitive photomultiplier tube D, which generates an electrical signal registering the arrival of the photon. The questions that are addressed by AAD are (1) what is the quantum state of the photon in the region L, which lies in the region between H and R, and (2) what would have been the outcome of measurements on the photon that might have been performed in that region?

AAD use the formalism of quantum mechanics as ap-

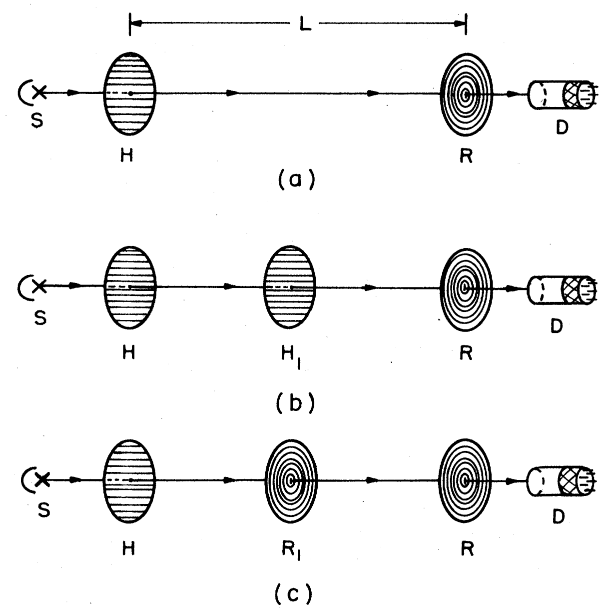


FIG. 13. Schematic diagram showing the three experimental situations considered in the Albert-Aharonov-D'Amato (1985) predictions: (a) The photon emerges from the source S, passes through a horizontal linear polarizing filter H, and then through a right circular polarizing filter R, before being detected by a photomultiplier tube D. (b) An intermediate horizontal linear polarizing filter H_1 is inserted. (c) An intermediate right circular polarizing filter R_1 is inserted. These additional measurements [(b) and (c)] are said (Albert *et al.*, 1985) to demonstrate that the photon is simultaneously in a state of linear and circular polarization in the intermediate region.

plied to the joint probability of a series of measurements (Aharonov *et al.* 1964) to demonstrate a remarkable pair of predictions (here applied to the present example): (1) if a linear polarization measurement had been performed [Fig. 13(b)] in region L the photon would have been found to be in a HLP state, and (2) if a circular polarization measurement [Fig. 13(c)] had been performed in region L the photon would have been found to be in a RCP state. In other words, the intermediate measurement of polarization appears to be equally influenced by the past linear polarization measurement which was performed at H and by the future circular polarization measurement which will be performed at R, in that both seem equally to prepare the system in a definite state which "forces" the outcome of the intermediate measurement.

This completely valid application of the quantum-mechanical formalism appears to be in at least *interpretational* conflict with the uncertainty principle (C-1) and with complementarity (C-3), which assert that, since RCP and HLP states are eigenstates of noncommuting variables, a photon cannot be in both of these eigenstates simultaneously. AAD on the other hand, interpret their result as indicating that "without violating the statistical predictions of quantum mechanics, it can be consistently supposed . . . that noncommuting observables can simultaneously be well defined" and that indeed, "given those statistical predictions, ... it is inconsistent to suppose anything else." The AAD result has been summarized in a popular account ("*Science* and *the Citizen*," 1985) as indicating that "the measurement on Friday caused, in some sense of the word cause, the smeared-out values of spin on Wednesday to collapse into some definite configuration. The logical puzzle about time and causality that this development engenders has not yet been fully explored."

It is therefore of considerable interest to apply the transactional interpretation to this new interpretational puzzle, both as a means of gaining insight into the problem and as a test of the utility of the transactional interpretation for resolving the interpretational paradoxes of quantum mechanics. The transactional analysis of this problem follows that of Sec. IV.D, which also dealt with the transmission of a photon through polarizing filters. The three experimental configurations considered are illustrated in Figs. 13(a)—13(c). Figures 14(a)—14(c) show diagrammatically the corresponding SV descriptions that will be discussed. These experimental configurations must be treated as separate (but related) quantum-mechanical systems, and each must be analyzed separately with the transactional interpretation. We first consider Fig. 13(b).

The transactional interpretation provides the following description of the transmission of the photon from S to D with an intermediate HLP measurement: The source S produces a retarded OW in the form of a general SV including all possible states of polarization. This wave then passes through filter H. The filter transmits only $|H\rangle$, i.e., that component of the SV which corresponds to a state of pure horizontal linear polarization. This wave

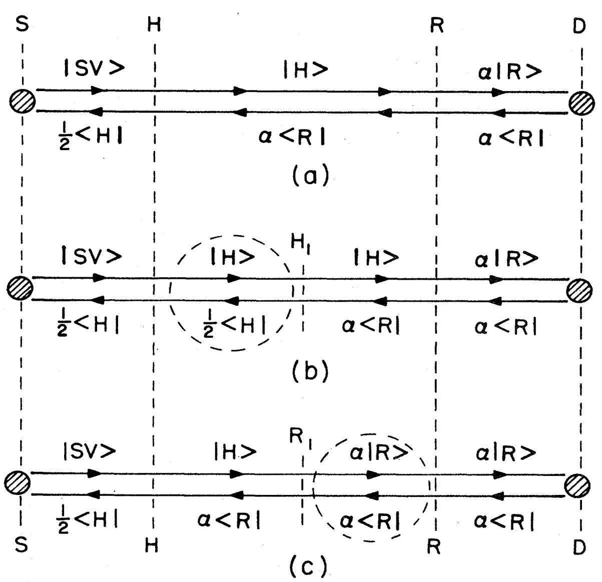


FIG. 14. Schematic diagram showing the transactional interpretation descriptions of the three Albert-Aharonov-D'Amato experiments. Note the changes in the quantum state in the intermediate region indicated by the dashed circles in case (b) and case (c).

then travels to filter H_1, which transmits $|H\rangle$ unchanged. This HLP wave then travels to filter R, which transmits only that component in a pure state of right circular polarization. From Eq. (15c), this will be $\alpha|R\rangle$. This RCP wave then strikes the photocathode of D and is absorbed and detected, producing the advanced wave $\alpha\langle R|$, the CW which travels back along the track of the incident OW to confirm the transaction. When the CW reaches R, it is transmitted without modification because it is already in a state of RCP. However, when it reaches H_1, only its HLP component is transmitted, so from Eq. (15a) it becomes $\alpha(\alpha\langle H|)=\frac{1}{2}\langle H|$. It retains this form as it passes through the filter H and back to the source S.

The description of the transmission of the photon from S to D with an intermediate RCP measurement illustrated in Fig. 13(c) is very similar: The source S produces a retarded OW in the form of a general SV including all possible states of polarization. This wave then passes through filter H, which transmits only $|H\rangle$. This wave then travels to filter R_1, which transmits only that component in a pure state of right circular polarization. From Eq. (15c), this will be $\alpha|R\rangle$. This RCP wave then travels to filter R, which transmits $\alpha|R\rangle$ unchanged. It strikes the photocathode of D and is absorbed and detected, producing the advanced wave $\alpha\langle R|$, the CW which travels back along the track of the incident OW to confirm the transaction. When the CW reaches R and R_1 it is transmitted without modification because it is already in a state of RCP. However, when it reaches H, only its HLP component is transmitted, so from Eq. (15a) it becomes $\alpha(\alpha\langle H|)=\frac{1}{2}\langle H|$. It retains this form as it passes back to the source S.

In cases (b) and (c) of Fig. 13 the insertion of the intermediate polarizing filter does not alter the statistical aspects of the measurement from that of case (a) of Fig. 13, where there is no intermediate measurement, and so the three cases are equivalent in the observational sense. However, the transactional interpretation gives us the opportunity to examine the intermediate quantum states in each case, and when this is done we find that the transaction that is confirmed is quite different in each of the three cases. This is illustrated in Fig. 14. In case (a) where there is no intermediate measurement the state in the intermediate region between H and R is in an indeterminate quantum state, in that the OW is $|H\rangle$ while the CW is $\alpha\langle R|$. This is also the case for the region between H_1 and R for case (b) and for the region between H and R_1 for case (c). However, we see that for case (b) the photon in the region between H and H_1 is in a state of pure HLP (circled), while for case (c) the photon between R_1 and R is in a state of pure RCP (circled).

The transactional resolution of the riddle posed by the AAD predictions is that the uncertainty principle is *not* compromised, nor can noncommuting observables simultaneously be well defined, as AAD have suggested. However, as was suggested above in another context, the circular polarization measurement that occurs later at R does cause, in some sense of the word cause, the smeared-out values of circular polarization between R and H to earlier "collapse into some definite configuration." The transactions that form in the three cases are not identical, even though they have the same observables, because each transaction is a separate self-consistent solution to the wave equation. Each satisfies a different set of boundary conditions. The insertion of the intermediate filter, while not altering the statistics of the measurement, brings into being a different transaction, which has different characteristic eigenstates in the intermediate region between H and R. Thus the two predictions of the AAD calculation concern intrinsically different quantum systems and cannot be construed as implying the presence "simultaneously" of the eigenstates of non-commuting variables.

V. CONCLUSION

In Sec. I we set out the ground rules by which an interpretation of quantum mechanics could be judged, assuming that no experimental tests are possible. The suggested criteria were economy, compatibility, plausibility, and insightfulness. In the succeeding sections we examined the Copenhagen interpretation and the transactional interpretation in the ways that they deal with the problems of quantum mechanics and can be applied to experiments and *Gedankenexperimente.*

In the context of economy we have shown that the Copenhagen interpretation is somewhat less compact than the transactional interpretation because (C-1), (C-2), and (C-4) are essentially independent and seemingly unrelated elemental postulates of the interpretation. On the other hand, the transactional model which is the essential content of (T-4) brings with it the uncertainty principle (T-1) and the statistical interpretation (T-2) as consequences of the model.

In the context of compatibility we have shown that the SV collapse model implicit in (C-4), if the SV is taken to be a mathematical description of the state of an extended system common to all observers of the system (C-4a), is descriptionally inconsistent with relativistic invariance and/or causality, and also appears manifestly to violate time-reversal invariance at the descriptional level. The transactional interpretation applied to the same system is fully consistent with relativistic invariance and macroscopic causality. While the transactional interpretation implicitly involves an arrow of time in its account of the statistical interpretation (T-2), this arrow has been accounted for elsewhere in terms of a boundary condition model (Cramer, 1983).

The criterion of plausibility is admittedly more subjective than the other criteria. However, we feel that the assertion of the transactional interpretation that the advanced solutions of the relativistic quantum-mechanical wave equations play a role in quantum events (T-4) is intrinsically more plausible than the assertion of the Copenhagen interpretation that the solutions of these equations are somehow mathematical descriptions of knowledge (C-4). This is reinforced by the evidence of the deep-rooted philosophical and epistemological problems implicit in the "knowledge" assertion.

Finally, in the context of insightfulness we have shown that by providing an explicit nonlocal mechanism for describing quantum events the transaction model has provided new insights into the reality behind the quantum-mechanical formalism. In doing so, it has also shown that a synthesis is possible between the ideas of Heisenberg and Bohr, who emphasized the intrinsic uncertainty and complementarity of quantum processes, and the ideas of Einstein, who emphasized the need to view the reality behind the formalism with an interpretation that is compatible with our understanding in other areas of physics.

Despite these advantages, there are many who will find the use of advanced waves in the transaction model difficult to accept because there is no experimental evidence for the existence of such a phenomenon. Certainly there are others who will feel that the whole notion of considering alternative interpretations of quantum mechanics when there is no possible experimental means of deciding between one interpretation and another is a kind of philosphizing that has no place in physics.

We answer such criticisms by reiterating that interpretation does have an important role on our understanding of physical theory, even if it is untestable at the experimental level, and that as far as possible, an interpretation should interpret (rather than decline to do so). What is presented here is a paradigm shift in the mode of quantum-mechanical interpretation. As Kuhn (1962) has made clear, such paradigm shifts are always difficult because of the intellectual gear changing they require. A shift away from the Cophenhagen interpretation may be particularly difficult because of its traditional role in the teaching of quantum mechanics over 5 decades.

The value of new interpretational insights into physical processes, however, should not be underestimated. Experience in many fields of physics has shown that progress and new ideas and approaches are stimulated by the ability to visualize clearly physical phenomena. Schrödinger described quantum mechanics as "a formal theory of frightening, indeed repulsive, abstractness and lack of visualizability" (Heisenberg, 1927). The visualization of quantum phenomena has been denied us for half a century, not by the abstract quantum-mechanical formalism but by the Copenhagen interpretation. The transactional interpretation of the same formalism now makes this long-sought visualization possible.

ACKNOWLEDGMENTS

This work was supported in part by the Division of Nuclear Sciences of the U.S. Department of Energy. Teaching University of Washington undergraduates about quantum phenomena provided much of the stimulation and original inspiration for the ideas presented here. The opportunity for much of the contemplation and writing was provided by a sabbatical leave supported jointly by the University of Washington and the Hahn-Meitner Institute of Berlin. The author is indebted to many individuals for valuable discussions and correspondence, helpful suggestions, and good advice (sometimes ignored), and to many institutions for the opportunity to present these ideas at various stages of evolutionary development to varied audiences. Among the individuals who deserve particular mention are David Boulware, Wil Braithwaite, John Clauser, Kevin Coakley, Jorrit deBoer, Dieter Gross, Ernest Henley, Nick Herbert, Max Jammer, Hans Krappe, Dan Larsen, Albert Lazzarini, Chris Morris, Rudolph Moessbauer, Riley Newman, Rick Norman, Pierre Noyes, Rudi Peierls, Dennis Sciama, Saul-Paul Sirag, Derek Storm, Dan Tieger, John Wheeler, and Larry Wilets. Among the institutions that have made possible the presentation of these ideas in colloquia, seminars, and discussions are Carnegie-Mellon University, Michigan State University, the University of Washington, Los Alamos National Laboratory, the Esalen Institute, the University of Munich, the Hahn-Meitner Institute—Berlin, the Demokritos Laboratory for Nuclear Research—Athens, the Ruhr University—Bochum, and the University of Cologne. Thanks are due to Bill Warren for the Schrödinger's cat illustration. Finally, the author is deeply indebted to his wife Pauline B. Cramer, a pragmatic mechanical engineer, for many discussions, for the injection of much common sense, and for a great deal of encouragement.

APPENDIX: ALTERNATIVE INTERPRETATIONS OF QUANTUM MECHANICS

In the discussion in the main body of this paper we have focused on the orthodox formalism of quantum mechanics and on its interpretation by the Copenhagen and transactional approaches. In doing this we have not discussed the efforts of individuals and groups who have investigated other ways of dealing with the interpretational problems of the quantum-mechanical formalism. For a comprehensive review of such efforts the reader is referred to the excellent survey provided by Jammer (1974). Here we briefly summarize some of the alternative interpretations and theories that have had a significant impact on the field and consider these in the framework of the discussion of interpretational problems in Sec. II of this paper.

1. Hidden-variable theories

The hidden-variable alternatives to the formalism of quantum mechanics (Belinfante, 1973; Jammer, 1974) have been aimed primarily at the problems of completeness and predictivity (see Secs. II.E and II.F) and have conventionally started from the assumption of locality. By asserting the existence of unobserved variables, which would eliminate the indeterminacy of quantum mechanics if their values were known, the proponents of these theories have attempted to demonstrate that quantum mechanics is an incomplete theory.

Hidden-variable theories are able to deal at some level with some of the interpretational problems of Sec. II by avoiding SV collapse through the use of deterministic hidden variables. The SV is treated as an average and incomplete description of the system, and so the "knowledge" issue does not arise. The Bell inequality experiments have at a stroke invalidated all hidden-variable theories based on the locality assumption. While a hidden-variable theory could, in principle, be constructed that was nonlocal and compatible with the Bell inequality experimental results, such an approach would lose much of its intrinsic classical appeal and would run the risk of conflicts with relativity and causality. It remains to be seen whether any new hidden-variable theory can successfully come to terms with nonlocality, the experimental results, and achieve compatibility with relativity and causality.

2. Semiclassical interpretations

Perhaps the most widely "accepted" semiclassical interpretation is the "disturbance model" (Herbert, 1985). This is the notion, often introduced as a pedagogical tool in elementary textbooks on modern physics, that canonically conjugate variables of a particular system under study, e.g., position and momentum, can "actually" have simultaneous well-defined values, but that the act of making a measurement of one of these variables "disturbs" the other so that no knowledge of it can be obtained. Heisenberg reportedly used this model in his early thinking (Rosenfeld, 1971a) but later discarded it when he realized its inadequacy. The disturbance model has been refuted again and again, most recently by the experimental tests of Bell's inequality, but remains as a widely held interpretation of quantum mechanics used by a sizable segment of the community of practicing physicists. A second early and unsuccessful semiclassical interpretation, that of Schrödinger, has already been discussed in Sec. II.

The "guide wave" interpretation of de Broglie, which is also a semiclassical interpretation, was invented very early in the development of quantum mechanics (about 1925) and is said to be responsible for stimulating the development of the Schrödinger equation and for the emphasis on the wave aspects of quantum mechanics that were so important to the early development.

The guide wave interpretation suggests a specific underlying mechanism for the interplay of waves and particles in a quantum event. It has been described in a number of publications of de Broglie (1926,1927a, 1927b,1960,1964,1968). For example, de Broglie (1964) gave the following summary: ". . . a particle is a very small object which is constantly localized in space, and a wave is a physical process which is propagated in space in the course of time according to a given equation of propagation. . . . the wave has a very low amplitude and does not carry energy, at least not in a noticeable manner. The particle is a very small zone of highly concentrated energy incorporated in the wave, in which it constitutes a sort of generally mobile singularity. By reason of this incorporation of the particle in the wave, the particle possesses an internal vibration which, as it moves, remains constantly in phase with the vibration of the wave. . . . the mean path of the particle is determined according to the shape of the wave by a certain 'guidance law,' but this motion has superimposed on it continual fluctuations corresponding to a hidden variable behavior of the particles."

From this summary it should be apparent that the guide wave interpretation presents a very different view of quantum events from that of the Copenhagen interpretation. The "wave" in the above description is the SV itself, which has a definite but limited reality in that it can physically travel through space but cannot carry energy, momentum, etc. The collapse does not occur, but is replaced by the action of the particle, which "rides" the SV and arrives with the largest probability at the locations where the SV has the largest amplitude, the general properties of the SV being separated from those of the specific particle that tracks it.

The problem of complexity is not addressed, but is not serious because the SV is given only limited reality. The predictivity of the picture is a problem, since the particle does not follow the path of greatest amplitude of the SV, as might be expected, but rather follows the wave in a random "thermodynamic" way.

The most serious problem of the guide wave interpretation is that it makes no provision of nonlocalities of the second kind and is implicitly local. It is therefore inconsistent with the Bell inequality experiments. There are also grounds for believing that it may be inconsistent with the formalism of quantum mechanics. Recent papers (Garuccio *et al.*, 1981,1982) have asserted that there are experimental tests that can distinguish the predictions of the guide wave interpretation from those of orthodox mechanics. This is because in certain situations involving the interference of incoherent sources the guide wave interpretation would predict interference effects that are absent in orthodox calculations. (See, also, Costa de Beauregard, 1982.)

From a certain point of view, the guide wave interpretation can be taken as a kind of preliminary version of the transactional interpretation presented here. It is completely consistent with the transactional interpretation in most of its aspects, and its principal shortcomings are its lack of a nonlocal mechanism that can account for correlations in separated measurements and the *ad hoc* way that the particle and its properties are introduced. But if de Broglie's particle is identified with the transaction of the transaction model, then the picture presented is very close to that presented in Sec. III above. Thus the penetrating intuitive insight of de Broglie was not only crucial to the early development of quantum mechanics, but it also came very close to the nonlocal interpretation presented here.

3. "Collapse" interpretations

As discussed in Sec. II.C, the abrupt and discontinuous collapse of the SV implied by the formalism of quantum mechanics and its treatment by the Copenhagen interpretation have been the source of many of the interpretational problems. Therefore, there have been a number of attempts to provide a more plausible account of the SV collapse process.

One of the early alternative collapse models was suggested first by Darwin (1929), but was more widely publicized through the work of von Neumann (1932), London and Bauer (1939), and Wigner (1962). Wigner in particular popularized the model, and we quote him to describe it. "... the result of an observation modifies the wave function of a system. The modified wave function is, futhermore, in general unpredictable before the impression gained at the interaction has entered our consciousness: it is the entering of an impression into our consciousness which alters the wave function because it modifies our appraisal of the possibilities of different impressions which we expect to receive in the future. It is at this point that the consciousness enters the theory unavoidably and unalterably." Wigner goes on to introduce what has become known as the Wigner's friend paradox, which was discussed in Sec. IV.F above. He uses this *Gedankenexperiment* to illustrate the plausibility of his model and the implausibility of several alternatives.

This "consciousness" interpretation, while it is a reasonable working hypothesis for an observer who does not wish to find himself dissolved into the state vector of the system he is measuring, does beg a number of questions. Did the SV of the universe remain uncollapsed until the first consciousness evolved? Where is the borderline between consciousness and unconsciousness? Will "smart" measuring instruments eventually achieve the ability to collapse SV's, and how will one know when they do? And so on.

Schrödinger (1935) suggested an alternative to the consciousness interpretation, which he called the principle of state distinction and which asserts, "states of a microscopic system which could be told apart by macroscopic observation are distinct from each other whether observed or not." In other words, the SV collapses as soon as some macroscopic record of the result of a measurement is made, whether a conscious observer looks at that record or not. Heisenberg (1960) and others have suggested a variant of this position which asserts that as soon as the quantum measurement passes from the domain of reversible processes into the domain of thermodynamic irreversibility the SV collapses.

The latter two "collapse triggers" are more appealing to most physicists than the former because they avoid giving some special significance to consciousness and because, as pointed out by Weisskopf (1959,1980), they correspond more closely to the operating assumptions that practicing physicists use in thinking about how quantum measurements are done. However, these models also beg the question of borders: Where precisely is the border between macrophysics and microphysics and the border at which irreversibility begins? This point seems particularly troublesome when one realizes that present experimental techniques permit the result of a quantum measurement to be "recorded" in the spin orientation of a single electron in a Penning trap or in the trapping of a single magnetic flux quantum in a split superconducting ring.

Indeed, in the context of the latter apparatus this point has been made more quantitative by recent work. Leggett (1980) has carefully considered the question of macroscopic quantum effects. He has introduced a semiquantitative measure called "disconnectivity" for characterizing the degree to which a quantum system is isolated from effects that would average away any coherent quantum interference phenomena. He has shown that this is a useful criterion for separating the macrocosm from the microcosm, i.e., classical from quantum phenomena. He finds that while many experiments, and in particular the Schrödinger's cat *Gedankenexperiment,* do indeed have a very small disconnectivity which places them in the classical domain, there remains a class of possible macroscopic experiments that satisfy the criterion of high disconnectivity and that should exhibit quantum interference effects at the macroscopic level. Of particular interest are experiments involving macroscopic quantum tunneling of magnetic flux in a superconducting Josephson junction, because this collective behavior involves many degrees of freedom as well as macroscopic dissipation. Recent experiments (Voss, and Webb, 1981; Ouboter *et al.*, 1982) have confirmed some of Leggett's predictions for macroscopic quantum tunneling in this macroscopic junction phase of Josephson junctions. The border between macrophysics and microphysics seems to have become less sharp with the improvement of such experimental techniques, making the issue of collapse less clear.

Another problem with these collapse models is that they are not full interpretations of quantum mechanics. They have focused only on the cause of collapse and have provided no insights into the related interpretational problems listed in Sec. II. In particular, the nonlocal aspects of the SV collapse, which have become the focus of the recent Bell inequality tests, are not clarified by these interpretations.

4. The many-worlds interpretation

At a conference on gravitation, Everett (1957) and his thesis supervisor at Princeton, Wheeler (1957), presented related papers on Everett's thesis research, which has come to be known (DeWitt, 1970; DeWitt and Graham, 1973) as the Everett-Wheeler, Everett-Wheeler-Graham, or many-worlds interpretation of quantum mechanics. Like the interpretations discussed in the Appendix, the many-worlds interpretation addresses the problem of collapse. Unlike the other interpretations the many-worlds interpretation asserts that collapse never occurs. Instead each component of the SV of a quantum event represents a separate and equal physically real reality. In other words, with each quantum event the universe splits into a number of branch universes, each containing a different possible outcome for the event. What we perceive as collapse is, in the many-worlds interpretation, simply a result of the fact that our consciousness took a particular path through these branches and therefore observed one set of results instead of another of the myriad possibilities. Presumably other copies of our consciousness are observing all of the other possible outcomes in other branch universes.

This is perhaps the most "heroic" of the efforts to deal with the problem of collapse. It addresses identity by giving the SV the status of objective reality, for the electron in each branch universe is identical with its component in the original SV. Complexity is not addressed, and is presumably more troublesome for the many-worlds interpretation than for the Copenhagen interpretation because the SV is a physical entity. Locality is not specifically addressed in Everett's paper, but he labels the locality problem stated by Einstein, Podolsky, and Rosen (1935) as a "fictitious paradox" and asserts that it can be easily investigated and clarified with the many-worlds interpretation.

From one point of view this is perhaps true, for in a situation with two separated measurements the "earlier" of the two measurements will, in the many-worlds interpretation, split the universe containing the second measurement such that its outcome is always correlated properly with that of the first. From this point of view the many-worlds interpretation is compatible with nonlocality. However, from another point of view the many-worlds interpretation would appear to have severe problems in this area. With each splitting of the universe, spatial regions megaparsecs distant from an event locus are instantaneously split into alternate realities due to the distant quantum event. It would seem that both the propagation speed of the splitting and its simultaneity are manifestly inconsistent with relativistic invariance.

The many-worlds interpretation is interesting from another point of view. It represents an interpretation of quantum mechanics to which the assumption of contrafactual definiteness, as discussed in Sec. I, does not apply. The many-worlds interpretation characterizes our world as one of many equally real alternatives, and so some alternative experiment that might have been performed would not have a single definite outcome as CDF asserts, but rather would have had *all possible* outcomes, one for each branch universe split off by the measurement. Thus the CFD assumption is not applicable to the many-worlds model. Therefore, in the context of the many-worlds interpretation, the Bell inequality experimental results can be taken as an experimental demonstration of the invalidity of CFD rather than of locality.

The many-worlds interpretation also has an intrinsic time asymmetry. In its description, the universe splits only in the future time direction, and never in the past time direction. Thus there is an intrinsic arrow of time built into the interpretation that is inconsistent, as has been said of the Copenhagen interpretation, with the evenhandedness with which microphysics deals with the flow of time. Perhaps this kind of approach is just what William of Occam had in mind in warning that hypotheses should not be multiplied beyond necessity.

5. Advanced-action interpretations

In addition to the present work, two other approaches to the interpretation of quantum mechanics have appeared in the literature which have suggested the use of advanced waves. The first of these is the "advanced-action" interpretation of Costa de Beauregard (1953, 1965,1976,1977a,1977b,1978,1979a,1979b,1982,1985), as discussed in a series of papers. He points out that the timelike symmetry of electrons and positrons in the Feynman picture can, in principle, account for the nonlocal structure of quantum mechanics as applied to electrons and positrons in a creation-annihilation event.

The approach, however, may have deficiencies. In a recent paper, Garuccio and his co-workers (1980) have pointed out that the advanced-action interpretation employs negative-energy solutions of the Dirac equation that are assumed to propagate in the positive time direction, an impossibility due to the complementarity of the time and energy variables (see Cramer, 1980, for a discussion of this point). They also argue that the advanced-action interpretation violates causality and energy conservation. Indeed there may be causality problems with the advanced-action approach, for Costa de Beauregard (1979a,1985) has suggested this interpretation as an explanation for parapsychological phenomena. Selleri and Vigier (1980) have also argued that the advanced-action interpretation is inconsistent with quantum electrodynamics because it implies a rejection of Feynman's D_c propagator.

A second interpretation using advanced waves was suggested by Davidon (1976), who proposed that "an operator which factors into a tensor product of advanced and retarded solutions of the time-dependent Schrödinger equation" could lead to "a local and objective description ... for each of the remote parts in an Einstein-Podolsky-Rosen situation." Since the time-dependent Schrödinger equation (see Sec. III.C), being first order in its time derivative, does not *have* advanced solutions, it is not clear what is the actual content of Davidon's model.

Both of these advanced-wave approaches, from the point of view of the present work, were on the right track, but missed the crucial role of the transaction in mediating the transfer of energy, momentum, etc., and in erasing all residual traces of the advanced waves. Without this key concept, these interpretations lead inevitably to causality paradoxes and inconsistency with quantum-mechanical predictions and experimental observations.

6. Stapp's nonlocal model

Stapp (1976,1977,1980) has proposed a very general nonlocal model, which we regard as a precursor of the present work. It is not an interpretation because it does not propose any specific mechanism for quantum events. Rather, it is a world model similar to that originally proposed by A. N. Whitehead, but suitably modified to deal with the manifest nonlocality demonstrated by the tests of the Bell inequality.

A particular feature of this work is that although the elemental events considered connect loci of space-time across spacelike intervals, the events retain a sequentiality that is independent of the reference frame and that proceeds from the past to the future. Thus, the model is atemporal (like the transaction model), but it preserves a sequentiality consistent with causality and with our perception of the causal arrow of time.

The transactional interpretation is fully consistent with Stapp's model, but goes beyond it in providing a specific and plausible mechanism through which Stapp's nonlocal sequentiality can operate. In particular, if one equates the transaction concept defined above and Stapp's elemental quantum event concept, then the transactional interpretation provides a meaningful context in which Stapp's very general but rather nonspecific world model can be seen in operation.

REFERENCES

Aharonov, Y., P. G. Bergmann, and J. L. Lebowitz, 1964, Phys. Rev. **134**, B1410.
Albert, D. Z., Y. Aharonov, and S. D'Amato, 1985, Phys. Rev. Lett. **54**, 5.
Aspect, A., J. Dalibard, and G. Roger, 1982a, Phys. Rev. Lett. **49**, 91.
Aspect, A., J. Dalibard, and G. Roger, 1982b, Phys. Rev. Lett. **49**, 1804.
Audi, M., 1973, *The Interpretation of Quantum Mechanics* (University of Chicago, Chicago, Illinois).
Ballentine, L. E., 1970, Rev. Mod. Phys. **42**, 358.
Belinfante, F. J., 1973, *A Survey of Hidden Variable Theories* (Pergamon, Oxford).
Belinfante, F. J., 1975, *Measurements and Time Reversal in Objective Quantum Theory* (Pergamon, Oxford).
Bell, J. S., 1965, Physics **1**, 195.
Bell, J. S., 1966, Rev. Mod. Phys. **38**, 447.
Berestetskii, V. B., E. M. Lifschitz, and L. P. Pitaevskii, 1971, *Relativistic Quantum Theory* (Pergamon, Oxford).
Bjorken, J. D., and S. D. Drell, 1965, *Relativistic Quantum Mechanics* (McGraw-Hill, New York).
Bjorken, J. D., and S. D. Drell, 1965, *Relativistic Quantum Fields* (McGraw-Hill, New York).
Bohr, N., 1928, in *Atti del Congresso Internazionale dei Fisici Como* (Zanchelli, Bologna), Vol. 2, pp. 565—588.
Bohr, N., 1935a, Nature (London) **136**, 65.
Bohr, N., 1935b, Phys. Rev. **48**, 696.
Bohr, N., and L. Rosenfeld, 1933, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **12**, 8; translated in Wheeler and Zurek (1983), pp. 479—534.
Born, M., 1926a, Z. Phys. **37**, 863; translated in Wheeler and Zurek (1983), pp. 52—55.
Born, M., 1926b, Z. Phys. **38**, 803.
Born, M., 1927, Nature (London) **119**, 354.
Born, M., and A. Einstein, 1979, *The Born-Einstein Letters* (Walker, New York), with comments by M. Born.
Clauser, J. F., and A. Shimony, 1978, Rep. Prog. Phys. **41**, 1881.
Costa de Beauregard, O., 1953, C. R. Acad. Sci. **236**, 1632.
Costa de Beauregard, O., 1965, Dialectica **19**, 280.
Costa de Beauregard, O., 1976, C. R. Acad. Sci. **282**, 1251.
Costa de Beauregard, O., 1977a, Phys. Lett. A **60**, 93.
Costa de Beauregard, O., 1977b, Nuovo Cimento B **42**, 41.
Costa de Beauregard, O., 1978, Phys. Lett. A **67**, 171.
Costa de Beauregard, O., 1979a, Lett. Nuovo Cimento **26**, 135.

Costa de Beauregard, O., 1979b, Nuovo Cimento B **51**, 267.
Costa de Beauregard, O., 1982, Phys. Lett. A **89**, 171.
Costa de Beauregard, O., 1985, Found. Phys. **15**, 871.
Cramer, J. G., 1980, Phys. Rev. D **22**, 362.
Cramer, J. G., 1983, Found. Phys. **13**, 887.
Darwin, C. G., 1929, Proc. R. Soc. London, Ser. A **124**, 375; see also the discussion in Rosenfeld (1971).
Davidon, W. C., 1976, Nuovo Cimento B **36**, 34.
Davies, P. C. W., 1970, Proc. Cambridge Philos. Soc. **68**, 751.
Davies, P. C. W., 1971, J. Phys. A **4**, 836.
Davies, P. C. W., 1972, J. Phys. A **5**, 1025.
de Broglie, L., 1926, C. R. Acad. Sci. **183**, 447.
de Broglie, L., 1927a, C. R. Acad. Sci. **184**, 273.
de Broglie, L., 1927b, J. Phys. Radium **8**, 225.
de Broglie, L., 1960, *Nonlinear Wave Mechanics, A Causal Description* (Elsevier, Amsterdam).
de Broglie, L., 1964, *The Current Interpretation of Wave Mechanics* (Elsevier, Amsterdam).
d'Espagnat, B., 1976, *Conceptual Foundations of Quantum Mechanics,* 2nd ed. (Benjamin, Menlo Park, CA).
d'Espagnat, B., 1979, Sci. Am. **241**, 158.
DeVries, R. M., G. R. Satchler, and J. G. Cramer, 1974, Phys. Rev. Lett. **32**, 1377.
DeWitt, B. S., 1970, Phys. Today **23**(9), 30; see also the letter comments, Phys. Today **24**(4), 36, and **24**(11), 10.
DeWitt, B. S., and N. Graham, 1973, Eds., *The Many-Worlds Interpretation of Quantum Mechanics* (Princeton University, Princeton, N.J.).
Dicke, R. H., 1981, Am. J. Phys. **49**, 925.
Dieks, D., 1982, Phys. Lett. A **92**, 271.
Dirac, P. A. M., 1930, *The Principles of Quantum Mechanics* (Oxford University Press, London).
Dirac, P. A. M., 1938, Proc. R. Soc. London, Ser. A **167**, 148.
Eberhard, P. H., 1977, Nuovo Cimento B **38**, 75.
Eberhard, P. H., 1978, Nuovo Cimento B **46**, 392.
Eberhard, P. H., 1982, Phys. Rev. Lett. **49**, 1474.
Einstein, A., 1928, in *Electrons et Photons—Rapports et Discussions du Cinquième Conseil de Physique tenu à Bruxelles du 24 au 29 Octobre 1927 sous les Auspices de l'Institut International de Physique Solvay* (Gauthier-Villar, Paris).
Einstein, A., 1932, in a letter to M. Born, in Born (1979); see also his remark in Wheeler and Zurek (1983), p. 8.
Einstein, A., 1947, in a letter to M. Born (see Born, 1979).
Einstein, A., 1949, in *Albert Einstein, Philosopher Scientist,* edited by P. A. Schilpp (Library of Living Philosophers, Evanston, IL).
Einstein, A., B. Podolsky, and N. Rosen, 1935, Phys. Rev. **47**, 777.
Everett III, H., 1957, Rev. Mod. Phys. **29**, 454; see also Wheeler (1957).
Feinberg, G. 1967, Phys. Rev. **159**, 1089.
Feinberg, G., 1978, Phys. Rev. D **17**, 1651.
Feynman, R. P., 1949, Phys. Rev. **76**, 769.
Feynman, R. P., 1967a, *The Character of Physical Law* (MIT, Cambridge, MA), p. 129.
Feynman, R. P., 1967b; students of Wheeler-Feynman electrodynamics are directed to the comments of a certain "Mr. X" on this subject in Gold (1967).
Feynman, R. P., R. B. Leighton, and M. Sands, 1965, *The Feynman Lectures* (Addison-Wesley, Reading, MA), Vol. III.
Freedman, S. J., and J. F. Clauser, 1972, Phys. Rev. Lett. **28**, 938.
Furry, W. H., 1936a, Phys. Rev. **49**, 393.
Furry, W. H., 1936b, Phys. Rev. **49**, 476.
Garuccio, A., G. N. Maccarone, E. Recami, and J. P. Vigier, 1980, Lett. Nuovo Cimento **27**, 60.
Garuccio, A., K. R. Popper, and J. P. Vigier, 1981, Phys. Lett. A **86**, 397.
Garuccio, A., V. Rapisarda, and J. P. Vigier, 1982, Phys. Lett. A **90**, 17.
Ghirardi, G. C., A. Rimini, and T. Weber, 1980, Lett. Nuovo Cimento **27**, 293.
Ghirardi, G. C., and T. Weber, 1979, Lett. Nuovo Cimento **26**, 599.
Gold, T., 1967, Ed., *The Nature of Time* (Cornell University, Ithaca).
Gyulassy, M., S. K. Kauffmann, and L. W. Wilson, 1979, Phys. Rev. C **20**, 2267.
Hartle, J. B., 1969, Am. J. Phys. **36**, 704.
Heisenberg, W., 1925, Z. Phys. **33**, 879.
Heisenberg, W., 1927, Z. Phys. **43**, 172, translated in Wheeler and Zurek (1983), pp. 62—84.
Heisenberg, W., 1958, Daedalus (Stockholm) **87**, 95.
Heisenberg, W., 1960, *Physics and Beyond* (Harper and Row, New York), pp. 60—62.
Heisenberg, W., 1961, *Planck's Discovery and the Philosophical Problems of Atomic Physics,* in *On Modern Physics,* edited by C. N. Polter (Orion, New York), pp. 9 and 10.
Herbert, N., 1975, Am. J. Phys. **43**, 315.
Herbert, N., 1982, Found. Phys. **12**, 1171.
Herbert, N., 1985, *Quantum Reality: Beyond the New Physics* (Anchor-Doubleday, Garden City, N.Y.).
Herbert, N., and J. Karush, 1978, Found. Phys. **8**, 313.
Hiley, B., 1981, New Scientist **6**, quoting P. A. M. Dirac.
Hoyle, F., and J. V. Narlikar, 1969, Ann. Phys. (N.Y.) **54**, 207.
Hoyle, F., and J. V. Narlikar, 1971, Ann. Phys. (N.Y.) **62**, 44.
Jammer, M., 1966, *The Conceptual Development of Quantum Mechanics* (McGraw-Hill, New York).
Jammer, M., 1974, *The Philosophy of Quantum Mechanics* (Wiley, New York).
Klauder, J. R., and E. C. G. Sudarshan, 1968, *Fundamentals of Quantum Optics* (Benjamin, New York).
Konopinski, E. J., 1980, *Electromagnetic Fields and Relativistic Particles* (McGraw-Hill, New York), Chap. XIV.
Kuhn, T. S., 1962, *The Structure of Scientific Revolutions* (University of Chicago, Chicago, IL).
Landau, L. D., and R. Peierls, 1931, Z. Phys. **69**, 56; translated in Wheeler and Zurek (1983), pp. 465—476.
Leggett, A. J., 1980, Prog. Theor. Phys. Suppl. **69**, 80.
London, F., and E. Bauer, 1939, *La Theorie de L'Observation en Mécanique Quantique* (Hermann, Paris); translation in Wheeler and Zurek (1983), pp. 217—259.
Mermin, N. D., 1981, Am. J. Phys. **49**, 940.
Mermin, N. D., 1985, Phys. Today **38**(4), 38.
Messiah, A., 1961, *Quantum Mechanics* (North-Holland, Amsterdam), Vol. 1.
Milonni, P. W., and M. L. Hardies, 1982, Phys. Lett. A **92**, 321.
Mittelstaedt, P., 1983, Int. J. Theor. Phys. **22**, 293.
Ouboter, X., R. de Bruyn, and D. Bol, 1982, Physica **112B**, 15.
Partridge, R. B., 1973, Nature (London) **244**, 263.
Pegg, D. T., 1975, Rep. Prog. Phys. **38**, 1339.
Peierls, R., 1979, *Surprises in Theoretical Physics* (Princeton University, Princeton, N. J.).
Pflegor, R. L., and L. Mandel, 1967, Phys. Rev. **159**, 1084.
Popper, K. R., 1967, in *Quantum Theory and Reality,* edited by M. Bunge (Springer, Berlin).
Reichenbach, H., 1944, *Philosophical Foundations of Quantum Mechanics* (University of California, Berkeley, CA).

Renninger, M., 1953, Z. Phys. **136**, 251.
Rosenfeld, L., 1955, in *Quantum Theory and Measurement,* edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, N. J., 1983), p. 477.
Rosenfeld, L., 1971a, in *Quantum Theory and Measurement,* edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, N. J., 1983), p. 58.
Rosenfeld, L., 1971b, *Quantum Theory in 1929* (Nordita Publications No. 387, Copenhagen).
Schrödinger, E., 1926a, Ann. Phys. (Leipzig) **79**, 361.
Schrödinger, E., 1926b, Ann. Phys. (Leipzig) **79**, 486.
Schrödinger, E., 1926c, Ann. Phys. (Leipzig) **79**, 734.
Schrödinger, E., 1926d, Ann. Phys. (Leipzig) **81**, 109.
Schrödinger, E., 1927a, Ann. Phys. (Leipzig) **82**, 257.
Schrödinger, E., 1927b, Ann. Phys. (Leipzig) **82**, 265.
Schrödinger, E., 1927c, Ann. Phys. (Leipzig) **83**, 956.
Schrödinger, E., 1935, Proc. Cambridge Philos. Soc. **31**, 555.
Science and the Citizen, 1985, Sci. Am. **252**, 82.
Selieri, F., and J. P. Vigier, 1980, Lett. Nuovo Cimento **29**, 7.
Stapp, H. P., 1971, Phys. Rev. D **3**, 1303.
Stapp, H. P., 1975, Nuovo Cimento B **36**, 34.
Stapp, H. P., 1977, Found. Phys. **7**, 313.
Stapp, H. P. 1980, Found. Phys. **10**, 767.
Stapp, H. P., 1982, Phys. Rev. Lett. **49**, 1470.
von Neumann, J., 1932, *Mathematische Grundiagen der Quantenmechanik* (Springer, Berlin).
von Weizaecker, C. H., 1971, in *Quantum Theory and Beyond,* edited by T. Bastin (Cambridge University Press, London).
Voss, R. F., and R. A. Webb, 1981, Phys. Rev. Lett. **47**, 265.
Weisskopf, V., 1959, Daedalus **88**, 592.
Weisskopf, V., 1980, letter to the editor in Sci. Am. **242**, 8.
Wheeler, J. A., 1957, Rev. Mod. Phys. **29**, 463, see also Everett (1957).
Wheeler, J. A., 1978, in *The Mathematical Foundations of Quantum Mechanics,* edited by A. R. Marlow (Academic, New York).
Wheeler, J. A., and R. P. Feynman, 1945, Rev. Mod. Phys. **17**, 157.
Wheeler, J. A., and R. P. Feynman, 1949, Rev. Mod. Phys. **21**, 425.
Wheeler, J. A., and W. H. Zurek, 1983, Eds., *Quantum Theory and Measurement* (Princeton University, Princeton, N. J.).
Wigner, E. P., 1950, *Group Theory* (Academic, New York).
Wigner, E. P., 1962, in *The Scientist Speculates,* edited by I. J. Good (Heinemann, London).
Wooters, W. K., and W. H. Zurek, 1982, Nature (London) **299**, 802.