

Lawrence Berkeley National Laboratory

Recent Work

Title

BELL'S THEOREM AND THE DIFFERENT CONCEPTS ON LOCALITY

Permalink

<https://escholarship.org/uc/item/3nt850mv>

Author

Eberhard, P.H.

Publication Date

1978-03-01

To be Submitted for Publication

LBL-7537
Preprint

c.2

BELL'S THEOREM AND THE
DIFFERENT CONCEPTS OF LOCALITY

P. H. Eberhard

RECEIVED
LAWRENCE
BERKELEY LABORATORY

MAY 9 1978

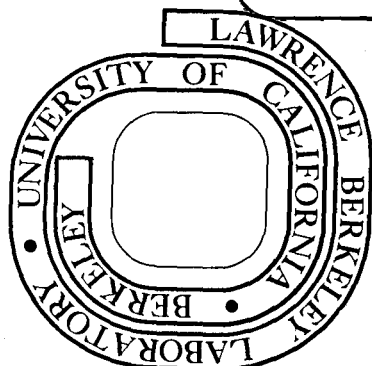
March 1978

LIBRARY
DOCUMENTS SECTION

Prepared for the U. S. Department of Energy
under Contract W-7405-ENG-48

TWO-WEEK LOAN COPY

*This is a Library Circulating Copy
which may be borrowed for two weeks.
For a personal retention copy, call
Tech. Info. Division, Ext. ~~5786~~ 6782*



LBL-7537

c.2

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

Bell's Theorem and the Different
Concepts of Locality.

P. H. EBERHARD

Lawrence Berkeley Laboratory, Berkeley, California 94720

Summary.— Four definitions of the principle of local causes, each of which, when applied to a theory, leads to a different mathematical property of the theoretical predictions, are considered and physical justifications given. The predictions of quantum theory are shown to contradict three of those four concepts of locality. Conclusions are drawn about the physical process and about the interpretations of quantum theory or any other theory that would provide the same predictions. Several interpretations are still possible.

Work supported by the U. S. Department of Energy.

1.--Mathematical properties related to locality.

Since Bell's theorem was first demonstrated for deterministic hidden variables theories,⁽¹⁾ generalizations of the theorem to other theories have been published.⁽²⁻⁷⁾ Recently, it was claimed that these generalizations were not valid and that the theorem applied only to hidden variables theories.⁽⁸⁾ It is the contention of this paper that much of the controversy is about the meaning of the words "principle of local causes." Quantum theory, or any other theory that would give the same probabilistic predictions, has certain mathematical properties that can be demonstrated. Whether these properties are in contradiction with one's concept of the principle of local causes depends on what interpretation is given to the theory. Essentially, the conflict occurs when one tries to describe the quantum system itself or when one considers the deterministic or random process by which the measurement results are produced. It can be ignored whenever one is concerned only about the possible action that one observer can have on another.

The principle of local causes, hereafter called locality, is derived from the principle of causality, which requires that, in time, the cause must precede the effect, and from the principle of relativity, which implies that all reference frames are equivalent. Let us consider two points in space, P_A and P_B . Suppose that at P_A [P_B] and at time t_A [t_B], an experimenter sets the knob of an apparatus to a position A [B], that at time t_α [t_β] the measurement M_A [M_B] is made, and that the measurement provides the result α [β]. Let $p(a,A,b,B)$ be the probability that the results of M_A and M_B turn

out to be $\alpha = a$ and $\beta = b$ when the knob settings are A and B. In the rest frame in which our study is made, the time sequence is: setting of A, measurement M_A , setting of B, measurement M_B . Therefore,

$$(1) \quad \begin{cases} t_A < t_\alpha < t_B < t_\beta, \\ t_\beta - t_A < \text{time of propagation of light from } P_A \text{ to } P_B. \end{cases}$$

Causality requires that the setting of the knob at P_B cannot affect M_A , that is, that α is somehow "independent" of B. In addition, according to the principles of special relativity, there can be observers for whom the knob setting at P_A occurs after the measurement M_B and causality should hold for every observer; therefore, the knob setting at P_A can have no effect on M_B so β is also "independent" of A. That is locality. However, the measurement result β at P_B may be correlated with the measurement result α at P_A because past events may have affected both α and β at the same time. The result $\alpha[\beta]$ may also depend on A[B]. Because of the possible correlations between β and α , between α and A and between β and B, it is not trivial to sort out an expression for the independence of β from A and of α from B.

The following mathematical properties could be attributed to the function $p(a,A,b,B)$ to express different concepts of causal independence, therefore of locality. Property 1 is convenient for expressing locality for theories in which there are no hidden variables or hidden quantities. Property 2 is suited for theories where some uncertainty about certain quantities explicitly contributes to the probabilistic character of the predictions. Property 3 is a generalization of

properties 1 and 2; it requires nothing but a gross estimation of $p(a,A,b,B)$ in particular circumstances and is suited for theories that may not be well defined and may not correspond to a completely explicit form for the probabilities. Property 4 is relevant to the question of message transmission, between experimenters, at speeds greater than that of light. Their mathematical definitions are discussed in the following paragraphs; physical justifications of the properties are presented at length in sect. 2.

1.1 Property 1. The probability distribution $p(a,A,b,B)$ is the product of two functions: $f(a,A)$ and $g(b,B)$ of a and A and of b and B , respectively:

$$(2) \quad p(a,A,b,B) = f(a,A) g(b,B) .$$

When the probability distribution has property 1, α and β are said to be "statistically independent." Property 1 insures the independence of α from B and of β from A by forcing the correlation between α and β to be zero.

1.2 Property 2. The probability distribution $p(a,A,b,B)$ involves a statistical mixture of cases λ with a statistical distribution $\rho(\lambda)$ and corresponding to conditional probabilities $f(\lambda,a,A)$ [$g(\lambda,b,B)$], independent of $B[A]$, that $\alpha = a[\beta = b]$ when the knob settings are A and B :

$$(3) \quad p(a,A,b,B) = \sum_{\lambda} \rho(\lambda) f(\lambda,a,A) g(\lambda,b,B) .$$

Property 2 allows for some statistical correlation between α and β . It is more general than property 1 and includes property 1 if

$\rho(\lambda) = 1$, for one particular case λ , and if $\rho(\lambda) = 0$ for all others. Therefore,

(4) property 1 implies property 2 .

1.3 The families of independent results. These families are defined to describe a more general property than property 2. They are defined in connection with two values A_1 and A_2 for the knob setting A and two values B_1 and B_2 for the knob setting B. A "family of independent results" is defined as a set of values $a_1, b_1, a_2,$ and b_2 such that

$$(5) \quad \left\{ \begin{array}{l} p(a_1, A_1, b_1, B_1) > 0 , \\ p(a_1, A_1, b_2, B_2) > 0 , \\ p(a_2, A_2, b_1, B_1) > 0 , \\ p(a_2, A_2, b_2, B_2) > 0 . \end{array} \right.$$

Consider four experiments $E_{1,1}, E_{1,2}, E_{2,1},$ and $E_{2,2}$ performed with AB set at the values $A_1B_1, A_1B_2, A_2B_1,$ and A_2B_2 , respectively. If $a_1, b_1, a_2,$ and b_2 form a family of independent results, the pairs of values $a_1b_1, a_1b_2, a_2b_1,$ and a_2b_2 are possible values of the results $\alpha\beta$ in the experiments $E_{1,1}, E_{1,2}, E_{2,1},$ and $E_{2,2}$, respectively, with nonzero probabilities. If the outcomes of these four experiments actually turned out this way, there would be the same result $\alpha[\beta]$ in the different experiments with different $B[A]$ but with the same $A[B]$.

1.4 Property 3. This property is defined relative to the four values, $A_1, A_2, B_1,$ and B_2 of the knob settings of A and B and

relative to a small number ϵ . Consider all the possible outcomes of the four experiments $E_{1,1}$, $E_{1,2}$, $E_{2,1}$, and $E_{2,2}$ defined above. Eliminate a certain number of possible results amounting to a fraction ϵ in probability in each experiment. Property 3 means that there is a family of independent results in those possible outcomes that are left after the elimination, no matter how the elimination is made. In other words, one can find four values a_1 , b_1 , a_2 , and b_2 such that a_1b_1 , a_1b_2 , a_2b_1 , and a_2b_2 are possible results of the experiments $E_{1,1}$, $E_{1,2}$, $E_{2,1}$, and $E_{2,2}$, performed with knob settings A_1B_1 , A_1B_2 , A_2B_1 , and A_2B_2 , even if we disregard a fraction ϵ of the possible results considered to be "pathological" in each experiment. Elimination of a small fraction ϵ of pathological results will be necessary to the demonstration of sect. 3 that the predictions of quantum theory do not have property 3 in general. At the limit of infinite statistics, ϵ can be considered to be zero; for full mathematical rigor, however, ϵ is just a small positive number.

Let $S(\epsilon, A, B)$ be the set of possible outcomes $\alpha\beta$ when the knob settings are A and B , and after a fraction ϵ of them in probability have been eliminated, using a given elimination process. Property 3 for the function $p(a, A, b, B)$ means that there are values a_1 , b_1 , a_2 , and b_2 such that

$$(6) \quad \left\{ \begin{array}{l} a_1b_1 \in S(\epsilon, A_1, B_1) , \\ a_1b_2 \in S(\epsilon, A_1, B_2) , \\ a_2b_1 \in S(\epsilon, A_2, B_1) , \\ a_2b_2 \in S(\epsilon, A_2, B_2) . \end{array} \right.$$

If two values, ϵ_1 and ϵ_2 , are considered for ϵ , it will be

easier to find a family of independent results if only a few possibilities are eliminated rather than if many are eliminated. Therefore,

$$(7) \quad \left\{ \begin{array}{l} \text{Property 3 for } \epsilon_1 \text{ implies property 3 for } \epsilon_2, \\ \text{if } \epsilon_1 > \epsilon_2. \end{array} \right.$$

Moreover, suppose that property 2 holds. The function $p(a,A,b,B)$ is of the form of eq. (3). It is possible to generate all the possible results of the four experiments $E_{1,1}$, $E_{1,2}$, $E_{2,1}$, and $E_{2,2}$ by the following Monte Carlo procedure.⁽⁹⁾ First, λ is generated according to the distribution $\rho(\lambda)$, then α_1 according to $f(\lambda,a,A_1)$, and then β_1, α_2 , and β_2 according to $g(\lambda,b,B_1)$, $f(\lambda,a,A_2)$, and $g(\lambda,b,B_2)$, respectively. The distribution of the results $\alpha_i \beta_k$ is proportional to the function $p(a,A_i,b,B_k)$ of eq. (3). It is representative of the experimental results of E_{ik} . However, for the same λ , we have generated possible results of the four experiments which have the same α for the same A , and the same β for the same B , that is, we have generated a family of independent results. If any fraction $\epsilon < \frac{1}{4}$ of the Monte Carlo results is disregarded in each of the four experiments, at least one family of independent results will survive intact. Property 3 is satisfied for any $\epsilon < \frac{1}{4}$. Therefore,

$$(8) \quad \text{property 2 implies property 3 for } \epsilon < \frac{1}{4}.$$

1.5 Property 4. The probability distribution of $\alpha[\beta]$ integrated over $b[a]$ is independent of $B[A]$:

$$(9) \quad \left\{ \begin{array}{l} \sum_b p(a,A,b,B) = F(a,A), \\ \sum_a p(a,A,b,B) = G(b,B), \end{array} \right.$$

Here any correlation between α and β is simply ignored because of the summations over $a[b]$ in eqs. (9). Property 4 is a property only of the probability distributions of $\alpha[\beta]$ regardless of $\beta[\alpha]$, that is, summed over the possible outcomes of the other measurement $M_B[M_A]$.

If the function $p(a,A,b,B)$ satisfies eq. (3), it is easily verifiable that it satisfies eqs. (9) by summing the expression (3) over $a[b]$ and taking into account the normalization property of the conditional probability $f(\lambda,a,A) [g(\lambda,b,B)]$:

$$(10) \quad \text{property 2 implies property 4}$$

From eqs. (10) and (4), we see that property 1 implies property 4 also. From eqs. (8) and (4), we see that property 1 implies property 3 as well for any $\epsilon < \frac{1}{4}$. However, property 3 does not necessarily imply property 4 as can be seen from the example shown in fig. 1. If only a fraction less than, let us say, 1/100 of the areas of the two rectangular domains OPQR and OPQ'R' is eliminated for each pair of values $A_i B_k$, it is still possible to find families of independent results. Therefore, the probability distribution has property 3 for values of $\epsilon < 1/100$. However, the projection of the probability distribution on the Oa axis is not the same for $A = A_1, B = B_1$ where it extends from O to R and for $A = A_1, B = B_2$ where it extends from O to R'. Therefore, the distribution does not have property 4 as defined by eqs. (9). And because of eqs. (10) and (4) it has neither property 1 nor property 2.

Conversely, property 4 does not necessarily imply property 3.

This can be demonstrated by using the example of fig. 2. The projection of the distribution on the Oa[Ob] axis is the same for the

same $B[A]$. Therefore, the distribution $p(a,A,b,B)$ satisfies eqs. (9) and thus has property 4. However, even for $\epsilon = 0$, it is impossible to find a family of independent results, that is, values $a_1, b_1, a_2,$ and b_2 , for which the combinations $a_1b_1, a_1b_2, a_2b_1,$ and a_2b_2 belong to $S(\epsilon, A_1B_1), S(\epsilon, A_1B_2), S(\epsilon, A_2B_1),$ and $S(\epsilon, A_2B_2)$, respectively. Therefore, because of eq. (7), the distribution does not have property 3 for any ϵ , and because of eqs. (8) and (4) it has neither property 1 nor property 2.⁽¹⁰⁾ Of course, property 4 implies that the projection of the ranges of possible results on the $a[b]$ axis be the same for $A = A_i [B = B_k]$ whatever $B[A]$ is. However, as in the case of fig. 2, it may happen that, in these ranges, no two particular values a_1 and $a_2 [b_1$ and $b_2]$ can be chosen for $A = A_1$ and $A_2 [B = B_1$ and $B_2]$ independently of $B[A]$ that still provide combined results $a_i b_k$ that would be physical in the four combinations of A and B .

The only implications that exist between the properties are summarized by the following statements, derived from eqs. (4), (7), (8), and (10).

$$(11) \quad \left\{ \begin{array}{l} \text{property 1 implies property 2,} \\ \text{property 2 implies property 3 for } \epsilon < \frac{1}{4} \text{ and property 4,} \\ \text{property 3 for any } \epsilon \text{ implies property 3 for a smaller } \epsilon. \end{array} \right.$$

In sect. 2, the relationship between these mathematical properties and different concepts of locality are discussed. In sect. 3, we will show that quantum theory, or any other theory that would provide the same predictions, does not have property 3, and therefore has neither property 1 nor 2, although it has property 4. This demonstration,

called Bell's theorem, shows that none of the physical considerations that justify properties 1, 2, and 3 in sect. 2 applies to quantum theory, although the physical justifications of property 4 do apply.

2.—Physical significance of the properties.

2.1 Property 1. There is at least one class of theories which may have different possible interpretations depending on whether property 1 does or does not apply. Theories in this class specify all the properties of a quantum system by a single state $|\lambda\rangle$, just as quantum theory does for pure cases. When the measurement M_A is made, the result α is obtained with a probability distribution that we define as $f(a,A)$. At the same time, the state is modified by the interaction with the measurement apparatus (wave function collapse). For $\alpha = a$ and the knob-setting A ,

$$(12) \quad |\lambda\rangle \rightarrow |\lambda'(\lambda,a,A)\rangle.$$

When the measurement M_B is made, the result β has a probability distribution that depends on the new state $|\lambda'(\lambda,a,A)\rangle$ and on B . Let $g(\lambda'(\lambda,a,A),b,B)$ be the probability that $\beta = b$. The combined probability for M_A to give $\alpha = a$ and for M_B to give $\beta = b$ is of the form

$$(13) \quad p(a,A,b,A) = f(a,A) g(\lambda'(\lambda,a,A), b, B).$$

If we suppose that the state describes only the physical system and not our knowledge of it, the locality concept requires that the probabilities at P_B be not affected by the transformation (12) due to M_A at P_A . The probability distribution for β should be the same

function $g(b,B)$ as if M_A has not been made. It is a function of b and B , not of a and A . Therefore,

$$(14) \quad g(\lambda'(\lambda,a,A), b, B) = g(b,B) .$$

Using eq. (14) for a substitution in eq. (13), we obtain eq. (2); that is, property 1. If property 1 does not apply, either the state $|\lambda\rangle$ does not represent only the physical system or the evolution of the state violates locality. Note that property 1 involves no hidden quantity at all.

2.2 Property 2. Even in classical physics there are quantities that evolve in a "nonlocal" manner. Suppose a probability distribution $p(x)$ results from our uncertainty about the position x of a particle. When that particle is observed at point x_0 , it is sensible to have $p(x)$ collapse instantaneously at any point other than x_0 . However, that probability distribution corresponds to our knowledge about the particle while x is the variable that describes the position of the particle itself. The constraints of locality are imposed only on quantities which physically describe the physical system itself,⁽¹¹⁾ not our knowledge of it.

In order to generalize that concept, we consider states $|\lambda\rangle$ that physically describe the quantum system but do not necessarily have the deterministic behavior of classical physics. Moreover, these states can be affected by a measurement as shown in eq. (12). In addition, there is some uncertainty as to which state the system is in. For this reason, the state can be called a "hidden state." Let $\rho(\lambda)$ be the probability that the system is in the state $|\lambda\rangle$ and let

$f(\lambda, a, A) [g(\lambda, a, B)]$ be the conditional probability that $\alpha = a$ [$\beta = b$] when the state is $|\lambda\rangle$. When the measurement M_A is made, the state collapses as in eq. (12). The probability that $\alpha = a$ and $\beta = b$ is

$$(15) \quad p(a, A, b, B) = \sum_{\lambda} \rho(\lambda) f(\lambda, a, A) g(\lambda'(\lambda, a, A), b, B).$$

An example of a "hidden states" theory is quantum theory for statistical mixtures, identifying $\rho(\lambda)$ with the eigenvalues and $|\lambda\rangle$ with the eigenvectors of the density matrix. Other examples are the deterministic hidden variables theories where, given one value for λ and $A[a, A$ and $B]$, the function $f(\lambda, a, A) [g(\lambda'(\lambda, a, A), b, B)]$ is equal to 1 for a particular value of $a[b]$ and to 0 for all the others.

Since the state $|\lambda\rangle$ is supposed to describe the system physically, a natural interpretation of the concept of locality is that the probability distribution of β for a given λ is the same whether or not M_A is performed. It is the same argument as for property 1. Therefore,

$$(16) \quad g(\lambda'(\lambda, a, A), b, B) = g(\lambda, b, B).$$

Using eq. (16) to substitute $g(\lambda'(\lambda, a, A), b, B)$ in eq. (15), one obtains eq. (3), that is, property 2.

Property 2 can also be arrived at, following a different line of arguments.⁽³⁾ Let λ designate all the conditions that can affect the measurements M_A and M_B prior to t_A and let $\rho(\lambda)$ designate the probability of their occurrence. To satisfy locality, it seems that the probability of getting $\alpha = a$ [$\beta = b$] must be a function $f(\lambda, a, A) [g(\lambda, b, B)]$ of $\lambda, a[b]$ and $A[B]$ only. Equation (3) follows, that is, property 2. Property 2 means that the correlation between α

and β is due to the circumstances λ prior to t_A and that λ is independent of A and B .

To describe a deterministic or random process that generates the results α and β ,⁽¹²⁾ and to satisfy the requirement that $\alpha[\beta]$ be independent of $B[A]$, it can be shown that one has to impose property 2 to $p(a,A,b,B)$. This can be seen when an attempt is made to reproduce the statistical distributions of α and β by Monte Carlo simulation.⁽⁹⁾ Let us define λ as a label for the conditions before measurements. These conditions can be generated by the Monte Carlo simulation according to the probability of their occurrence, which we call $\rho(\lambda)$. Then α can be generated according to a conditional probability distribution defined as $f(\lambda,a,A)$, because λ and A are the only variables at our disposal at time t_α . The generation of α may require the use of λ , of A , and of a few random numbers. A logical interpretation of locality is that the generation of β should not depend on A . If β involves one or more of the random numbers involved also in the generation of α , those random numbers can be incorporated in λ and generated at the same time as λ . Therefore, the generation of β is then made independent of α as well as of A . The β distribution is of the form $g(\lambda,b,B)$. It follows that the distribution of α and β is a function $p(a,A,b,B)$ that satisfies eq. (3). Therefore, the function $p(a,A,b,B)$ has property 2 as long as the process that generates $\alpha[\beta]$ does not require the knowledge of $B[A]$. This statement applies also when we consider the real process by which nature generates α and β if, of course, we are interested in the way nature operates.

These justifications of property 2 for $p(a,A,b,B)$ are related to the description of the system by a state or to mechanisms by which the probability distributions can be generated.⁽¹²⁾ Of course, none of these justifications implies that the theory which leads to $p(a,A,b,B)$ is deterministic. However, it should be noted that any probabilistic theory, whether it has property 2 or not, can be replaced by a deterministic theory yielding the same predictions. This statement is demonstrated in appendix A.⁽¹³⁾

2.3 Property 3. Because of eqs. (11), any justification of either property 1 or property 2 is a justification of property 3. In particular, if one attempts to describe the mechanism that produces the results α and β in a way that is consistent with locality, one is led to impose property 2, therefore property 3, on $p(a,A,b,B)$. However, property 3 is more general than properties 1 and 2. For instance, the probability distribution defined in fig. 1 has property 3 for small values of ϵ but not properties 1 or 2, as has been seen in sect. 1.5. Moreover, let us consider the following probability distribution for α and β :

$$(17) \quad p(a,A,b,B) = \eta \sum_{\lambda} \rho(\lambda) f(\lambda,a,A) g(\lambda,b,B) + (1-\eta) H(a,A,b,B) ,$$

where $\rho(\lambda)$, $f(\lambda,a,A)$, $g(\lambda,b,B)$, and $H(a,A,b,B)$ are any positive functions normalized to 1 in λ , a , b , or a and b , respectively, and where η is a positive number less than 1. At least a fraction η of the time, the results α and β will have a distribution of the type of eq. (3) and then form families of independent results.

The function $p(a,A,b,B)$ has property 3 for $\epsilon < \frac{\eta}{4}$ although, for $\eta < 1$, if $H(a,A,b,B)$ does not have property 2, neither does $p(a,A,b,B)$. Property 3 for $\epsilon < \frac{\eta}{4}$ may be justified using the same arguments as for properties 1 and 2, but only in a fraction η of the cases.

Suppose the function $p(a,A,b,B)$ does not have property 3 for $\epsilon = 0$, as in the case of fig. 2. It is clear that it is never possible to make a deterministic or random choice of $\alpha[\beta]$ unless we know $B[A]$ and then obtain a possible result for each of the four experiments defined in sect. 1.3. Suppose we make any choice α_1 and α_2 [β_1 and β_2] for $\alpha[\beta]$ when $A = A_1$ and $A = A_2$ [$B = B_1$ and $B = B_2$], independently of $B[A]$. There is a combination $\alpha_1\beta_1$, $\alpha_1\beta_2$, $\alpha_2\beta_1$, or $\alpha_2\beta_2$ which is not physical in at least one of the four experiments considered. If $p(a,A,b,B)$ does not have property 3 for a finite ϵ , the choice of $\alpha[\beta]$ without knowing $B[A]$ is possible in no more than a fraction 4ϵ of the cases. Property 3 is associated with the possibility in a fraction of the cases, of generating $\alpha[\beta]$ independently of $B[A]$ and of still getting results allowed by the theory.

For $\epsilon = 0$, property 3 is a property of the domain where the function $p(a,A,b,B)$ is not zero. For small ϵ 's, property 3 concerns the domain of possible results after a small fraction of pathological results is eliminated. Therefore, whereas properties 1 and 2 are suited to theories that determine the function $p(a,A,b,B)$ exactly, property 3 is meaningful even for less well-defined theories, ones that would even grossly reproduce the predictions of quantum theory.

Another justification of this property 3 relies on a concept which was called "contrafactual definiteness,"^(2,6) a concept used in daily

life whenever we must make a choice. Contrafactual definiteness means that, in a given situation, the consequence of each of the possible courses of action can be considered even though the only sequence of events that can be known for certain is the one produced by the final single choice. That is, we can hypothesize about the event sequences following the courses of action that will not be chosen. Three different causal relationships can be conceived when considering different choices and subsequent events:

1) For some events, we assume there would be no difference in the sequences following different choices.

2) For other events we do not know if the sequences following the various choices would be the same or if they would be different, within the domain of their probability distribution.

3) There are events that would surely be different because the probability distributions of the events for the different choices do not overlap.

For deterministic theories, only concepts (1) and (3) make sense. For probabilistic theories, concept (2) may be used but it is customary to give "causal independence" a definition still founded on concept (1), even for events that are not completely predicted in a deterministic way. (14) If now such a causal independence is assumed to exist between α and B and between β and A , α would be equal to the same value $a_1[a_2]$ when $A = A_1[A_2]$ for the two choices $B = B_1$ and B_2 , and $\beta = b_1[b_2]$ when $B = B_1[B_2]$ for the two choices $A = A_1$ and A_2 . (2,6,7) Therefore, whatever the result a_1b_1 is for the experiment with A_1B_1 , there are a_2 and b_2 which form a family of independent results with

a_1 and b_1 . Moreover, if we consider all possibilities and disregard a certain number of possible pairs of results $\alpha\beta$ amounting to a fraction ϵ in probability in each one of the experiments with A_1B_1 , A_1B_2 , A_2B_1 , and A_2B_2 , there is at most a fraction 4ϵ of families with a missing member. For any $\epsilon < \frac{1}{4}$, some are left intact. Therefore, one at least still remains and property 3 is valid for $\epsilon < \frac{1}{4}$. Conversely, if property 3 does not hold, one is obliged to think that the result α would be different for a different choice of B , or β would be different for a different choice of A .

Anyway, to justify property 3 for any $\epsilon < \frac{1}{4}$, it is possible not to invoke the four hypothetical experiments of which only one can be performed.⁽⁷⁾ It is possible to define a concept of locality based on the requirement that, in four different experiments performed at different times with the four different combinations of knob settings, it is possible to group all the possible results to form families of independent results.

2.4 Property 4. For a theory that agrees with the principles of relativity, causal independence according to property 4 is necessary to exclude mechanisms by which information could be transmitted backward in time. If property 4, expressed by the second eq. (9), were violated, the probability distribution for β regardless of α , $\sum_a p(a,A,b,B)$, would depend on A . By repeating the experiment enough times, the statistical distribution of the results β at P_B would reproduce the probability $\sum_a p(a,A,b,B)$. By looking at this A -dependent distribution of β , the parameter A , which was set at P_A at time t_A , could be

known at P_B at time t_β , outside of the light cone. Consider the observer for whom, according to relativity, the measurement M_B occurs before the setting of A . For him, the knowledge of A is transmitted backwards in time and, according to the principles of relativity, what this observer can do can be done by any other observer. With a series of similar mechanisms it would be possible to send information to any point of the past. Therefore, the violation of property 4 would lead to severe difficulties in our concept of the universe.

Conversely, property 4 forbids transmission of information, at speeds greater than the speed of light, between the two experimenters at P_A and P_B using the measurements M_A and M_B . Consider the information carried by $\alpha[\beta]$ when $\beta[\alpha]$ is not known. The corresponding probability distribution is $\sum_b p(a,A,b,B) [\sum_a p(a,A,b,B)]$. If that distribution obtained after summation over $b[a]$ does not depend on $B[A]$, the knowledge of $\alpha[\beta]$ does not supply the experimenter at $P_A[P_B]$ with any information about the knob setting at $P_B[P_A]$, that is, the parameter set by the other experimenter.

Property 4 concerns only causality between observers, that is, between the one who sets A and the one who reads β or the one who sets B and the one who reads α .

3. Properties of the quantum theoretical predictions.

Properties 1 and 2 lead to inequalities called locality inequalities which are violated by quantum theory.⁽¹⁵⁾ One of them is demonstrated in appendix B. By itself, property 3 leads to another locality inequality,

shown here, that is also violated by quantum theory. Because of eqs. (11) the latter inequality alone shows that the quantum theory predictions do not have properties 1, 2, or 3.

3.1 The locality inequality. Suppose the results α and β have many different possibilities that can be labeled using a binary system. The result $\alpha[\beta]$ is represented by a set of N numbers $\alpha_j[\beta_j]$ that are either 0 or 1. We define

$$(18) \quad C(\alpha, \beta) = \frac{1}{N} \sum_j |\alpha_j - \beta_j|.$$

For a given j in the list, the term $|\alpha_j - \beta_j|$ is 0 with $\alpha_j = \beta_j$ and 1 when $\alpha_j \neq \beta_j$. Given four lists $a_1, b_1, a_2,$ and b_2 of terms $a_{1j}, b_{1j}, a_{2j},$ and b_{2j}

$$(19) \quad \begin{aligned} |a_{1j} - b_{1j}| &= |(a_{1j} - b_{2j}) - (a_{2j} - b_{2j}) + (a_{2j} - b_{1j})| \leq \\ &\leq |a_{1j} - b_{2j}| + |a_{2j} - b_{2j}| + |a_{2j} - b_{1j}|. \end{aligned}$$

Summing inequality (19) over j and using eq. (18), we got

$$(20) \quad C(a_1, b_1) \leq C(a_1, b_2) + C(a_2, b_2) + C(a_2, b_1).$$

Inequality (20) applies to any four lists of numbers. In particular it applies to lists $a_1, b_1, a_2,$ and b_2 that are possible results of measurements at P_A and P_B . If the function $p(a, A, b, B)$ has property 3 for a value ϵ , there is a family of independent results $a_1, b_1, a_2,$ and b_2 such that eq. (6) is verified. In each set $S(\epsilon, A, B)$, there is a minimum $C_{\min}(\epsilon, A, B)$ and a maximum $C_{\max}(\epsilon, A, B)$ for the quantity $C(a, b)$ defined by eq. (18) and applied to the elements of the set. Therefore, if ab belongs to $S(\epsilon, A, B)$

$$(21) \quad C_{\min}(\epsilon, A, B) \leq C(a, b) \leq C_{\max}(\epsilon, A, B).$$

Since property 3 implies the existence of a family of independent results satisfying eqs. (6), eqs. (18) and (21) can both be applied to the elements of that family, a_1, b_1, a_2 , and b_2 . Therefore, we derive the locality inequality:

$$(22) \quad \begin{aligned} C_{\min}(\epsilon, A_1, B_1) &\leq C(a_1, b_1) \leq C(a_1, b_2) + C(a_2, b_2) + C(a_2, b_1) \leq \\ &\leq C_{\max}(\epsilon, A_1, B_2) + C_{\max}(\epsilon, A_2, B_2) + C_{\max}(\epsilon, A_2, B_1). \end{aligned}$$

3.2 The two-photon experiment. Inequality (22) can be used to test the predictions of quantum theory in the case of an idealized two-photon experiment with N events. One event consists of the detection at P_A and P_B of photons γ_A and γ_B in a total 0^+ spin-parity state emitted in opposite directions from a point P_C located between P_A and P_B (fig. 3). There is 100% correlation between the planes of polarization of the photons and complete symmetry of this common polarization plane about the axis of propagation. At P_A [P_B], the detection is made behind a polarization analyzer making an angle A [B] with respect to an x-axis normal to the z-axis that is the propagation axis $P_B P_C P_A$. The efficiencies are 100%. The result α [β] of the measurement M_A [M_B] is constituted of N numbers α_j [β_j] such that α_j [β_j] = +1 if the γ_A [γ_B] photon of the j th event passes the polarizer and α_j [β_j] = 0 if it does not. The quantity $C(\alpha, \beta)$ of eq. (18) can be known by measuring the single rates in P_A and P_B . In appendix C, it is shown that, whatever N is, the expectation value $\langle C(\alpha, \beta) \rangle$ of $C(\alpha, \beta)$ is

$$(23) \quad \langle C(\alpha, \beta) \rangle = \sin^2 (A-B).$$

For a large enough number of events N , most values $C(\alpha, \beta)$ are close to the expectation value. Given two arbitrary small numbers ϵ and σ , it is possible to find an N large enough so that the difference between $C(\alpha, \beta)$ and $\langle C(\alpha, \beta) \rangle$ is larger than σ only in a fraction ϵ of the time. Therefore, if we eliminate that fraction ϵ of the possible results, we constitute sets $S(\epsilon, A, B)$ such that

$$(24) \quad \begin{cases} C_{\min}(\epsilon, A, B) = \sin^2 (A-B) - \sigma, \\ C_{\max}(\epsilon, A, B) = \sin^2 (A-B) + \sigma. \end{cases}$$

No matter how small ϵ is, we will choose N large enough so that

$$(25) \quad \sigma < \frac{\sqrt{2}-1}{4} \cong 0.103,$$

and the angles will be chosen as

$$(26) \quad \begin{cases} A_1 = 0^\circ, \\ B_1 = 67.5^\circ, \\ A_2 = 45^\circ, \\ B_2 = 22.5^\circ. \end{cases}$$

Then, using eq. (24), inequality (25), and eq. (26), we get

$$(27) \quad C_{\min}(\epsilon, A_1, B_1) = \sin^2 (67.5^\circ) - \sigma = \frac{2+\sqrt{2}}{4} - \sigma > \frac{3}{4}$$

$$(28) \quad \begin{aligned} & C_{\max}(\epsilon, A_1, B_2) + C_{\max}(\epsilon, A_2, B_2) + C_{\max}(\epsilon, A_2, B_1) = \\ & = 3\{ \sin^2(22.5^\circ) + \sigma \} = 3\{ \frac{2-\sqrt{2}}{4} + \sigma \} < \frac{3}{4}. \end{aligned}$$

Inequalities (27) and (28) show that locality inequality (22) does

not hold for these quantum theoretical predictions. Therefore, these predictions of quantum theory do not have property 3 for any $\epsilon > 0$. It follows that they do not have properties 1 or 2 either, because of eqs. (11). The contradiction between quantum theory and properties 1 and 2 can also be demonstrated directly, as is done in appendix B.

As to property 4, it should be noted that it is always a property of the quantum theory predictions. This is insured by the fact that the measurement M_A at P_A at t_α and the measurement M_B at P_B at t_β correspond to operators that commute with one another in the conditions of eq. (1). This statement is demonstrated in detail in appendix D.

4.—Conclusions

The predictions of quantum theory do not have properties 1, 2, or 3 for any value of $\epsilon > 0$. Any of the arguments spelled out in sect. 2 to justify them do not apply to quantum theory or to any other theory that would give the same predictions.

Whenever the computation of the probabilities involves a single state $|\lambda\rangle$, that state has to collapse, when the measurement M_A is performed, as it does in conventional quantum theory and in eq. (12). Because property 1 does not apply, the probability distribution of β is affected by that collapse due to M_A , in contradiction to an intuitive concept of locality if the state describes the quantum system itself instead of our knowledge of it. Even if, on the contrary, there is a statistical mixture of several hidden states $|\lambda\rangle$, the physical states $|\lambda\rangle$ still have to collapse similarly during

M_A because property 2 does not apply. When introduced in a theory, hidden quantities do not resolve the conflict with locality, but neither do they create it. It is just impossible to describe the physical system in a way which is compatible with locality as expressed by eq. (16).

Since the probability distribution does not have property 2, it is also impossible to explain the correlation between α and β only by conditions existing in their common past light cone.⁽³⁾ The result $\alpha[\beta]$ at $P_A[P_B]$ is correlated directly with the knob setting $B[A]$ at $P_B[P_A]$. Moreover, the contradiction to properties 2 and 3 makes it impossible to imagine a process⁽¹²⁾ that can generate α without the knowledge of B and, at the same time, of β without the knowledge of A .⁽¹²⁾ In this respect, quantum theory probabilities are like the ones illustrated by fig. 2, but where the domains S 's would represent the $S(\epsilon, A, B)$ for a small nonzero ϵ . It is impossible to choose α independently of B and β independently of A and still obtain physical results that do not fall into that small category of results we call pathological, that is, corresponding to a small probability in one or the other of the four experiments.

If we consider the different sequences of events that would follow the different choices of A and B we may make, (i.e., if we use the concept called "contrafactual definiteness"^(2,6)), it is impossible to imagine, at the same time, that there is a choice of A and B for a given experiment performed at a given time; that the predictions of quantum theory apply to all the possible combinations of A and B ; and, in addition, that $\alpha[\beta]$ would be the same for the same choice of $A[B]$.

This consideration further illustrates how impossible it is for any process to produce a value for $\alpha[\beta]$ that would be independent of $B[A]$.

Determinism has not been assumed in our demonstrations, although it could be restored for local and nonlocal theories (see appendix A). Historically, the first attempts to describe the quantum system in itself were invoking some determinism.⁽¹⁶⁾ This is why Bell's theorem was first demonstrated for theories with deterministic hidden variables⁽¹⁾ and why the contradiction of the quantum theory predictions with locality are sometimes linked to hidden variables theories⁽⁸⁾ (see appendix E). However, Bell's theorem now has been proved for probabilistic theories as well.⁽²⁻⁷⁾ To make quantum theory predictions and locality compatible, more than just determinism has to be given up among the concepts we inherited from classical physics. What is of most relevance is the possibility of describing the real random or deterministic process⁽¹²⁾ that produces the measurement results.

Since quantum theory has property 4, it predicts that the probability distributions of $\beta[\alpha]$, integrated over the variables $a[b]$ at $P_A[P_B]$, are independent of the knob setting $A[B]$. By changing $A[B]$ it is not possible to transmit a message from $P_A[P_B]$ at $t_A[t_B]$, to an observer at $P_B[P_A]$ at $t_B[t_\alpha]$, that is, to have an effect on him, before light has time to reach the point $P_B[P_A]$. It follows that quantum theory and any other deterministic or probabilistic equivalent theory are compatible with locality if the only causal dependence that counts is an effect of one observer on another.

From these considerations, we conclude that several interpretations

of causality lead to different interpretations of the theorem. The conflict with locality arises when the question is asked, "How does nature do it?" According to the Copenhagen interpretation, quantum theory is not supposed to provide a description of the physical process, but only a set of rules by which predictions can be made. Bell's theorem demonstrates that such a description of the process will never be possible unless it violates that independence of $\alpha[\beta]$ from $B[A]$ which we call locality of the physical process. (12)

There are essentially four possibilities remaining; they are as follows. (17)

4.1 Quantum theory breaks down for the cases where Bell's theorem applies. After all, neither inequality (22) nor any other locality inequality has ever been shown to be incorrect, experimentally, without additional assumptions. Quantum theory may be invalid in the two-photon experiment of sect. 3.2. In order to fully check these quantum theoretical predictions, in the conditions expressed by eq. (1), nearly 100% efficiencies and very fast changes in polarizer angles A and B are needed. (18) These combined requirements seem beyond present technology.

However, experiments that approximate the two-photon experiment described in sect. 3.2 were performed and the quantum theoretical prediction was upheld. (19-22) More generally, in a wide domain of physics, small failures of quantum theory have been looked for (23) but not found. (24)

4.2 The principles of relativity break down. If so, then all rest frames may not be equivalent. (25,26) One of them, R_0 , is fundamental and, in this rest frame R_0 only, causality applies. Causal effects can

propagate faster than the velocity of light as long as the cause precedes the effect in R_0 . No causal loop can be made then. In any other rest frame R , the time sequence between events with a time-like separation is the same as in R_0 . Therefore, the usual causal chains in the light cone are the same as expected from relativity. For events with a space-like separation, the cause may seem to occur after the effect in R if the time sequences in R and R_0 are opposite. However, this may have an interpretation: only the time in the rest frame R_0 is the real physical time, and the other rest frames seem to be equivalent to the fundamental one R_0 because the laws of nature just happen to have Lorentz invariance. Variations on this idea have been suggested,⁽²⁵⁾ in which R_0 is not one of the Lorentz rest frames but a rest frame obtained by a nonlinear transformation in space and time.

If this option is taken seriously, it may be interesting to investigate the possibility that the special rest frame R_0 may also have special properties in different domains of physics. The experimentally verified Lorentz invariance may be only an approximation. It has been suggested that the phenomenon of consciousness may obey different laws than the ordinary subjects of experiment in physics.⁽²⁷⁾

4.3 Very basic concepts of causality have to be reviewed. It has been speculated that the laws of nature are fully deterministic, including the hands of the experimenters, so that the knob setting $A[B]$ is really predetermined by conditions in the past that can be known at $P_B[P_A]$ at time $t_\beta[t_\alpha]$.⁽¹⁷⁾

A second speculation⁽²⁸⁾ is that the reality exists in many universes simultaneously. Then, the decision about which measurement result

coexists with which observer could be made long after the measurements M_A and M_B have been made. At that instant, the information about A and B would have time to be communicated at speeds less than the velocity of light to the point where the observer is dragged into one of these universes only.

A third speculation⁽²⁹⁾ has it that in special circumstances, the effect may occur before the cause, in all rest frames and in the light cone. This speculation would explain how A[B] may be known at the point P_C at time of emission of the two photons. Then, this information could be transmitted downstream, in time, with the photon $\gamma_B[\gamma_A]$ to the point $P_B[P_A]$.

None of these speculations leads to a theory that is practical in the sense of its being used instead of conventional quantum theory.

4.4 One should not care about "how nature does it." In conformity with the Copenhagen point of view, the goal of a theory is to prescribe mathematical procedures and it is not worth looking for a description of what is going on.⁽²⁸⁾ Of course, there is a conflict between locality and any description of the quantum system or of the physical process, but it is not important. The only causality that counts concerns the causal action one observer can have on another one. The compatibility of quantum theory with property 4 insures that this concept of locality holds. In this case, Bell's theorem is less important because that different definition of locality is the important one.

The inconvenience of that fourth option is that the desire to understand how nature works has been historically a powerful motivation

for scientific research. Modifications of logics have been suggested to approach these questions. (28)

At the present time, all the four above options are being pursued, and it is difficult to ascertain which one will turn out to be the right one. Consequently, any attempt to discourage the work that is being performed in any one of those four directions is either futile or counterproductive.

* * *

The author is indebted to Dr. T. G. TRIPPE and Professor R. R. ROSS for in-depth discussions and for reviewing the manuscript; to Professor J. S. BELL, Drs. H. P. STAPP and J. F. CLAUSER for many discussions about several topics treated in this paper; to Drs. P. A. MOLDAUER, A. BERTHELOT, and N. HERBERT for a useful exchange of letters; and to Mrs. J. BARRERA for help with the manuscript.

APPENDIX A

Restoration of determinism in a theory.

Determinism is defined here as a property, of certain theories, according to which the initial conditions of a system completely determine its future evolution and observations and where the probabilistic character of some predictions is entirely due to some uncertainty we have about these initial conditions. Given any probability distribution $p(a_1, a_2, \dots, a_j, \dots, a_n)$ of n measurement results $\alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_n$, it is possible to construct a deterministic theory with hidden variables that predicts that probability distribution.⁽¹³⁾ This can be demonstrated considering the Monte Carlo generation⁽⁹⁾ of the α_j 's according to $p(a_1, a_2, \dots, a_j, \dots, a_n)$. It first requires the generation of α_1 according to the probability distribution

$$(A.1) \quad h_1(a_1) = \sum_{a_2, \dots, a_j, \dots, a_n} p(a_1, a_2, \dots, a_j, \dots, a_n),$$

then of all the subsequent α_j 's, according to the conditional probability

$$(A.2) \quad h_j(a_1, a_2, \dots, a_j) = \frac{\sum p(a_1, a_2, \dots, a_j, \dots, a_n)}{h_{j-1}(a_1, a_2, \dots, a_{j-1})}.$$

Each of these Monte Carlo generations requires the generation of a random number ζ_j of uniform probability between 0 and 1 and the determination of α_j from the values of $\alpha_1, \alpha_2, \dots, \alpha_{j-1}$ and ζ_j :

$$(A.3) \quad \zeta_j = H_j(\alpha_1, \alpha_2, \dots, \alpha_{j-1}, \alpha_j),$$

where

$$(A.4) \quad H_j(a_1, a_2, \dots, a_{j-1}, a_j) = \sum_{a'_j \leq a_j} h_j(a_1, a_2, \dots, a_{j-1}, a'_j).$$

Consider a theory in which the system is described by a state with hidden variables $\zeta_1, \zeta_2, \dots, \zeta_j, \dots, \zeta_n$ distributed according to a uniform probability between 0 and 1 and where the α_j 's are bound to the ζ_j 's by eq. (A.3). Then that hidden variable theory is deterministic and predicts probability distributions identical to $p(a_1, a_2, \dots, a_j, \dots, a_n)$. (13)

If in a particular Lorentz rest frame the results $\alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_n$ have been ordered according to time of measurement, if the corresponding measurements depend on external parameters $A_1, A_2, \dots, A_j, \dots, A_n$ that are set by some experimenter, and if $p(a_1, a_2, \dots, a_j, \dots, a_n)$ has been computed according to quantum theory, it can be proved that each function $h(a_1, a_2, \dots, a_j)$, (thus the generation of α_j) is independent of A_{j+1}, \dots, A_n , in agreement with causality. Let D be the initial density matrix and $Q_j(a_j, A_j)$ the projection operator associated with the result $\alpha_j = a_j$ when the external parameter is A_j , then

$$(A.5) \quad \left\{ \begin{array}{l} Q_j^2(a_j, A_j) = Q_j(a_j, A_j), \\ \sum_{a_j} Q_j(a_j, A_j) = I. \end{array} \right.$$

The procedure of quantum theory permits the computation of the probability:

$$(A.6) \left\{ \begin{array}{l} p(a_1, a_2, \dots, a_j, \dots, a_n) \\ = \text{Tr} \{ Q_{n-1}(a_{n-1}, A_{n-1}) \dots Q_j(a_j, A_j) \dots Q_1(a_1, A_1) DQ_1(a_1, A_1) \dots Q_j(a_j, A_j) \\ Q_{n-1}(a_{n-1}, A_{n-1}) Q_n(a_n, A_n) \} \\ = \text{Tr} \{ DQ_1(a_1, A_1) \dots Q_j(a_j, A_j) \dots Q_{n-1}(a_{n-1}, A_{n-1}) Q_n(a_n, A_n) \\ Q_{n-1}(a_{n-1}, A_{n-1}) \dots Q_j(a_j, A_j) \dots Q_1(a_1, A_1) \}. \end{array} \right.$$

Using eqs. (A.1) and (A.5), we obtain:

$$(A.7) \quad h_1(a_1) = \text{Tr} \{ DQ(a_1, A_1) \},$$

which does not depend on $A_2, \dots, A_j, \dots, A_n$ and, using eqs. (A.2) and (A.5), we obtain

$$(A.8) \left\{ \begin{array}{l} h_j(a_1, a_2, \dots, a_j) \\ \text{Tr} \{ DQ_1(a_1, A_1) \dots Q_{j-1}(a_{j-1}, A_{j-1}) Q_j(a_j, A_j) Q_{j-1}(a_{j-1}, A_{j-1}) \dots Q_1(a_1, A_1) \} \\ = \frac{\quad}{h_{j-1}(a_1, a_2, \dots, a_{j-1})} \end{array} \right.$$

which does not depend on A_{j+1}, \dots, A_n if $h_{j-1}(a_1, a_2, \dots, a_{j-1})$ does not. This way, by induction we prove that the generation of α_j can be made without the knowledge of $\alpha_{j+1} \dots \alpha_n$ and of the parameters $A_{j+1} \dots A_n$ that are set after α is measured. Such generation is then compatible with the concept of causality.

If two measurements of α_j and α_{j+1} are performed outside of the light cone with two measurement apparatus with the knob settings A_j and A_{j+1} , respectively, it follows that it is possible to generate α_j [or α_{j+1}] without the knowledge of A_{j+1} [or A_j]. However, it is not possible to generate both α_j without the knowledge

-31-

of A_{j+1} and α_{j+1} without the knowledge of A_j unless $p(a_j, A_j, a_{j+1}, A_{j+1})$ has property 2 (see sect. 2.2), a property that the quantum theoretical predictions do not always have (see sect. 3).

Demonstrations that deterministic hidden variables theories cannot have the same predictions as quantum theory have been attempted.^(30,31) Those demonstrations have been shown to rely on unnecessary assumptions^(32,33) that this demonstration does not make.

APPENDIX B

Inequality generated by properties 1 and 2.

Following is a proof of one of the possible locality inequalities that can be demonstrated. Let us define the probability that $\alpha = a$ and $\beta \neq b$ as

$$(B.1) \quad p(a, A, \bar{b}, B) = \sum_b p(a, A, b, B) - p(a, A, b, B),$$

and the probability that $\beta = b$ and $\alpha \neq a$ as

$$(B.2) \quad p(\bar{a}, A, b, B) = \sum_a p(a, A, b, B) - p(a, A, b, B)$$

for the knob settings of A and B. Let us consider two values A_1 and A_2 for knob setting A, two values B_1 and B_2 for knob setting B, and four values $a_1, b_1, a_2,$ and b_2 .

Property 1 is defined by eq. (2) where $f(a, A)$ and $g(b, B)$ can be taken as normalized to 1. Therefore,

$$(B.3) \quad p(a, A, \bar{b}, B) = f(a, A) (1 - g(b, B)),$$

$$(B.4) \quad p(\bar{a}, A, b, B) = (1 - f(a, A)) g(b, B).$$

Since $f(a, A)$ and $g(b, B)$ are positive and ≤ 1 ,

$$(B.5) \quad \begin{aligned} f(a_1, A_1) (1 - g(b_2, B_2)) &\geq \\ &\geq (1 - f(a_2, A_2)) f(a_1, A_1) (1 - g(b_2, B_2)) (1 - g(b_1, B_1)), \end{aligned}$$

$$(B.6) \quad \begin{aligned} (1 - f(a_2, A_2)) g(b_2, B_2) &\geq \\ &\geq f(a_1, A_1) (1 - f(a_2, A_2)) g(b_2, B_2) (1 - g(b_1, B_1)), \end{aligned}$$

$$(B.7) \quad f(a_2, A_2) (1 - g(b_1, B_1)) \geq f(a_1, A_1) f(a_2, A_2) (1 - g(b_1, B_1)).$$

Adding inequalities (B.5), (B.6), and then (B.7), the following result is obtained:

$$(B.8) \quad \begin{aligned} & f(a_1, A_1)(1-g(b_2, B_2)) + (1-f(a_2, A_2))g(b_2, B_2) + \\ & + f(a_2, A_2)(1-g(b_1, B_1)) \geq f(a_1, A_1)(1-g(b_1, B_1)). \end{aligned}$$

Therefore, using eqs. (B.3) and B.4),

$$(B.9) \quad \begin{aligned} & p(a_1, A_1, \bar{b}_2, B_2) + p(\bar{a}_2, A_2, b_2, B_2) + p(a_2, A_2, \bar{b}_1, B_1) \geq \\ & \geq p(a_1, A_1, \bar{b}_1, B_1). \end{aligned}$$

Inequality (B.9) is one of the locality inequalities that can be derived from property 1.

Property 2 is defined by eq. (3). Therefore, the functions defined by (B.1) and (B.2) can be written as

$$(B.10) \quad p(a, A, \bar{b}, B) = \sum_{\lambda} \rho(\lambda) f(\lambda, a, A) (1-g(\lambda, b, B)),$$

$$(B.11) \quad p(\bar{a}, A, b, B) = \sum_{\lambda} \rho(\lambda) (1-f(\lambda, a, A)) g(\lambda, b, B).$$

The functions $f(a, A)$ and $g(b, B)$ can be replaced by $f(\lambda, a, A)$ and $g(\lambda, b, B)$ in inequalities (B.5), (B.6), and (B.7), and therefore in inequality (B.8). Then, both sides of the inequality can be multiplied by $\rho(\lambda)$ and summed over λ . Taking eqs. (B.10) and (B.11) into account, we arrive again at inequality (B.9). Therefore, inequality (B.9) is valid also for theories that have property 2, which is a more general property than property 1.

It can be shown that locality inequality (B.9) is not verified by the quantum theoretical predictions for the two-photon experiment of sect. 3.2, if the number of events $N = 1$ and if the polarizer angles defined in eqs. (26) are used. The quantum theoretical predictions

are computed in detail in appendix C. From eqs. (C.6) and (C.9) of appendix C,

$$(B.12) \quad p(1, A, \bar{1}, B) = p(1, A, 0, B) = \frac{1}{2} \sin^2 (A-B),$$

$$(B.13) \quad p(\bar{1}, A, 1, B) = p(0, A, 1, B) = \frac{1}{2} \sin^2 (A-B).$$

Therefore, using the angles of eq. (26),

$$(B.14) \quad \begin{aligned} & p(1, A_1, \bar{1}, B_2) + p(\bar{1}, A_2, 1, B_2) + p(1, A_2, \bar{1}, B_1) = \\ & = \frac{3}{2} \sin^2 (22.5^\circ) = \frac{3}{8} (2-\sqrt{2}) \cong 0.220. \end{aligned}$$

$$(B.15) \quad p(1, A_1, \bar{1}, B_1) = \frac{1}{2} \sin^2 (67.5^\circ) = \frac{2+\sqrt{2}}{8} \cong 0.427.$$

Equations (B.14) and (B.15) are incompatible with inequality (B.9), showing that quantum theory has neither property 1 nor property 2.

APPENDIX C

Quantum theoretical computation.

For an experiment such as that of sect. 3.2 but with a number of events $N = 1$, the procedure of quantum theory calls for the definition of an initial state vector for two photons in a 0^+ spin-parity state

$$(C.1) \quad |\lambda\rangle = \frac{1}{\sqrt{2}} (|\gamma_A, x\rangle |\gamma_B, x\rangle + |\gamma_A, y\rangle |\gamma_B, y\rangle),$$

where

$$(C.2) \quad \begin{cases} |\gamma_A, x\rangle = \text{state vector with } \gamma_A \text{ photon polarized along } x, \\ |\gamma_A, y\rangle = \text{state vector with } \gamma_A \text{ photon polarized along } y, \\ |\gamma_B, x\rangle = \text{state vector with } \gamma_B \text{ photon polarized along } x, \\ |\gamma_B, y\rangle = \text{state vector with } \gamma_B \text{ photon polarized along } y. \end{cases}$$

The computation of quantum theory follows the procedure described in sect. 2.1 with eqs. (12) and (13). The result is noted $\alpha = 1$ [$\beta = 1$] if the photon γ_A [γ_B] passes the polarizer and $\alpha = 0$ [$\beta = 0$] if it does not. The probability that the photon γ_A passes [is absorbed in] the polarizer at P_A is $f(1,A)$ [$f(0,A)$]:

$$(C.3) \quad f(1,A) = f(0,A) = \frac{1}{2}.$$

If the photon γ_A passes the polarizer and is detected, the state for the photon γ_B instantaneously collapses as in eq. (12),

$$(C.4) \quad |\lambda'(\lambda,1,A)\rangle = \cos A |\gamma_B, x\rangle + \sin A |\gamma_B, y\rangle,$$

and the probability that γ_B will not be detected in B is

$$(C.5) \quad g(\lambda'(\lambda,1,A),0,B) = 1 - \cos^2(A-B) = \sin^2(A-B).$$

Therefore, the probability that $\alpha = 1$ and $\beta = 0$ is derived from eq. (13):

$$(C.6) \quad p(1,A,0,B) = \frac{1}{2} \sin^2(A-B).$$

If the photon γ_A is absorbed in the polarizer in A, the state of the photon γ_B instantaneously collapses into

$$(C.7) \quad |\lambda'(\lambda,0,A)\rangle = -\sin A |\gamma_{B,x}\rangle + \cos A |\gamma_{B,y}\rangle.$$

Therefore,

$$(C.8) \quad g(\lambda'(\lambda,0,A),1,B) = \sin^2(A-B),$$

$$(C.9) \quad p(0,A,1,B) = \frac{1}{2} \sin^2(A-B).$$

The quantity $|\alpha-\beta|$ is 1 [0] if $\alpha = 1$ and $\beta = 0$ [$\beta = 1$] or if $\alpha = 0$ and $\beta = 1$ [$\beta = 0$]. Therefore, its expectation value is

$$(C.10) \quad \langle |\alpha-\beta| \rangle = p(0,A,1,B) + p(1,A,0,B) = \sin^2(A-B).$$

For N events produced in the same condition, the statistical average $C(\alpha,\beta)$ of eq. (18) has the same expectation value as

$|\alpha_j - \beta_j|$ for each event. Therefore

$$(C.11) \quad \langle C(\alpha,\beta) \rangle = \sin^2(A-B).$$

APPENDIX D

Proof that quantum theory has property 4.

Here we will not consider that the density matrix D represents the statistical mixture of hidden states as we did in sect. 2.2. For simplicity, we will consider D as a state itself to be treated like the states of sect. 2.1. The procedures of quantum theory for computing the function $p(a,A,b,B)$ calls for the definition of a set of projection operators $Q(a,A)$ [$R(b,B)$] associated with the measurements M_A [M_B] when the knob settings are A [B] and the result is $\alpha = a$ [$\beta = b$]:

$$(D.1) \quad Q^2(a,A) = Q(a,A),$$

$$(D.2) \quad R^2(b,B) = R(b,B).$$

$$(D.3) \quad \sum_a Q(a,A) = \sum_b R(b,B) = I.$$

The probability $f(a,A)$ that $\alpha = a$ is expressed by

$$(D.4) \quad f(a,A) = \text{Tr}\{Q(a,A) D\},$$

and, at the same time, as in eq. (12) the wave function collapse produces a change of the density matrix that becomes

$$(D.5) \quad D' = \frac{Q(a,A) D Q(a,A)}{f(a,A)}.$$

The same procedure is repeated for M_B using the $R(b,B)$'s and D' as the density matrix. It leads to the computation of the probability that $\alpha = a$ and $\beta = b$, using eq. (13):

$$(D.6) \quad p(a,A,b,B) = \text{Tr}\{Q(a,A) R(b,B) Q(a,A) D\}.$$

In the formalism of quantum theory, locality is expressed by the requirement that the measurement operators outside of the light cone commute. Then $Q(a,A)$ and $R(b,B)$ commute for any $a, A, b,$ and B . Using eqs. (D.1) and (D.3),

$$(D.7) \quad p(a,A,b,B) = \text{Tr}\{Q(a,A) R(b,B) D\},$$

$$(D.8) \quad \begin{aligned} \sum_a p(a,A,b,B) &= \text{Tr}\{ \sum_a Q(a,A) R(b,B) D\} = \\ &= \text{Tr}\{ R(b,B) D\} = G(b,B) \text{ independently of } A. \end{aligned}$$

Similarly,

$$(D.9) \quad \sum_b p(a,A,b,B) = F(a,A) \text{ independently of } B.$$

Equations (D.8) and (D.9) are equivalent to eqs. (9). Therefore, property 4 is a property of quantum theory.

Note that the commutation relations have another consequence. The probabilities can be computed either by assuming the wave function collapse to occur during M_A , and using the collapsed wave function to compute the probability distribution of β , or vice versa, by collapsing in M_B and using the collapsed wave function for α . The two modes of computation give the same result $p(a,A,b,B)$. This property is another necessary property which prevents us from seeing a fundamental difference between Lorentz rest frames. We cannot establish which one of the two measurements M_A or M_B has modified the probability distribution of the other and, therefore, which one actually occurred first.

APPENDIX E

A broad definition of hidden variables.

As in sect. 3, the contradiction between property 3 and the predictions of quantum theory is usually demonstrated using the example of the experiment of sect. 3.2 with N events produced in the same conditions.^(2,6,7) The measurement index j of eq. (18) is equated to the event number and used in the demonstration. It has been suggested that this event number j be called a "hidden variable," thus restricting the effect of Bell's theorem to "hidden variables" theories.⁽⁸⁾

Actually, in the experiment, the event number j is known before the event occurs; therefore, it is not "hidden." Moreover, though it is known, the event number does not provide any more information about what the future measurement result will be; therefore, it is not a system variable either. To call the event number j a "hidden variable" supposes an extension of the usually accepted definition of these words in general and of the meaning given them by BELL⁽¹⁾ in particular. There are experiments for which the event number j may provide more information about the upcoming measurement result. If the experimental conditions are changing even slightly from one event to the next, then the probability distributions depend on j and that dependence can be used to refine the predictions in any theory. In this sense, j may be considered as a system variable though, of course, it is still not a "hidden" quantity in the common sense.

In order to justify property 1 in sect. 2.1, we have considered

theories where the quantum theory state corresponds to one physical state only. Nothing is "hidden" in these theories. Calling j a hidden variable would cause these theories, and quantum theory for pure cases, to be "hidden variable theories" whenever the event number j is used to compute probability distributions. For this reason, this extension of the definition has been judged misleading and beside the point. It is not adopted in this paper. Here, as in BELL,⁽¹⁾ the words "hidden variables" are attributed only to quantities which, according to some theories, would specify the state of the quantum system so completely that its future behavior would be determined for a specific set of hidden variables. Hidden variables would not be known before measurement and our uncertainty about them would be responsible for the probabilistic character of quantum theory.

Even though this restricted definition of hidden variables is adopted, the question can be raised whether Bell's theorem applies only to deterministic theories.⁽⁸⁾ If the question implies that there may be probabilistic theories which do not need a wave function collapse or another nonlocal evolution that violated eqs. (14) or (16), the answer is no, as has been shown in this paper and elsewhere.⁽²⁻⁷⁾ If the question means that all the theories that are affected by Bell's theorem can be replaced by a deterministic one giving the same predictions, the answer is yes, because all probabilistic theories can be replaced by a deterministic one (see appendix A) whether or not they abide with properties 1, 2, 3, or 4.⁽¹³⁾ The question of determinism consequently seems orthogonal to the points raised here.

To demonstrate the contradiction between quantum theory and

property 3, a comparison is made, as in inequality (19), between possible results of the different experiments $E_{1,1}$, $E_{1,2}$, $E_{2,1}$, and $E_{2,2}$. Since only one of these experiments can be performed at any one time, it has been suggested that, in the context of quantum theory, such demonstrations be called invalid on the grounds that the actual results cannot be obtained coincidentally. ⁽⁸⁾ If such a point of view were adopted, it would provide some philosophical background to be associated with the fact that the function $p(a,A,b,B)$ derived from quantum theory does not have property 3. However, the function $p(a,A,b,B)$ still does not have property 3 for any $\epsilon > 0$ and, therefore, no deterministic or random process ⁽¹²⁾ can possibly generate a value for $\alpha[\beta]$ without the knowledge of $B[A]$, as has been shown in this paper.

REFERENCES

- (1) J. S. BELL: Physics, 1, 195 (1964).
- (2) H. P. STAPP: Phys. Rev. D, 3, 1303 (1971).
- (3) J. F. CLAUSER and M. A. HORNE: Phys. Rev. D, 10, 526 (1974).
- (4) J. S. BELL: Res Th 2053 CERN and communication at the VI Gift Conference, Jaca, June 2-7, 1975.
- (5) N. HERBERT: Am. J. Phys., 43, 315 (1975).
- (6) H. P. STAPP: Nuovo Cimento, 40B, 191 (1977).
- (7) P. H. EBERHARD: Nuovo Cimento, 38B, 75 (1977).
- (8) P. A. MOLDAUER: Bell's Theorem Without Hidden Variables?, submitted to Nuovo Cimento (1977).
- (9) The Monte Carlo simulation is a technique to generate values of experimental results, statistically distributed like the theoretical probability distribution, using random numbers.
- (10) Property 3 would be satisfied if any one of the four domains $S(0,A,B)$ on fig. 2 were rotated by 90° about its center of gravity located between the two ellipses.
- (11) As in the time-dependent Boltzmann equations.
- (12) The description of the "process" or "mechanism" is meant to be an algorithm that could generate values of α and β according to the theoretical probability distribution. Of course, if the theory is not deterministic, the algorithm must use random numbers.
- (13) I owe to J. S. BELL the idea that Monte Carlo simulation can easily demonstrate that determinism can be restored in quantum theory.

J. S. BELL, private communication, July 1977. A different approach

- is taken by D. BOHM, Phys. Rev. 85, 166 and 180 (1952).
- (14) For instance, in day-to-day life, if it is assumed that somebody's decision would not depend on a particular set of circumstances, chosen by somebody else, it means that the first person's decision would be the same whatever choice is made by the second one. However, there is no deterministic theory to predict the first person's decision.
- (15) J. F. CLAUSER, M. A. HORNE, A. SHIMONY and R. A. HOLT: Phys. Rev. Letters, 23, 80 (1969).
- (16) A. EINSTEIN, B. PODOLSKY and W. ROSEN: Phys. Rev., 47, 777 (1935).
- (17) I owe to H. P. STAPP the idea of presenting the remaining possible interpretations of quantum theory in this fashion. H. P. STAPP, private communication (1974).
- (18) A. ASPECT: Phys. Rev. D., 14 1944 (1976).
- (19) S. J. FREEDMAN and J. F. CLAUSER: Phys. Rev. Letters, 28, 938 (1972).
- (20) E. S. FRY and R. C. THOMPSON: Phys. Rev. Letters, 37, 465 (1976).
- (21) A. R. WILSON, J. LOWE, and D. K. BUTT: J. Phys. G: Nucl. Phys., 2 613 (1976).
- (22) For a review of such experiments, see J. F. CLAUSER and A. SHIMONY, LLL Preprint UCRL-80745, Lawrence Livermore Laboratory, Livermore, Calif., submitted to Reports on Progress in Physics.
- (23) For instance, P. H. EBERHARD: CERN 72-1 (1972) unpublished.
- (24) For instance, P. H. EBERHARD et al.: Phys. Letters, 53B, 121 (1974); W. C. CARITHERS et al.: Phys. Rev. D, 14, 290 (1976).
- (25) H. P. STAPP: LBL 3837, Lawrence Berkeley Laboratory, Berkeley, California (1975).

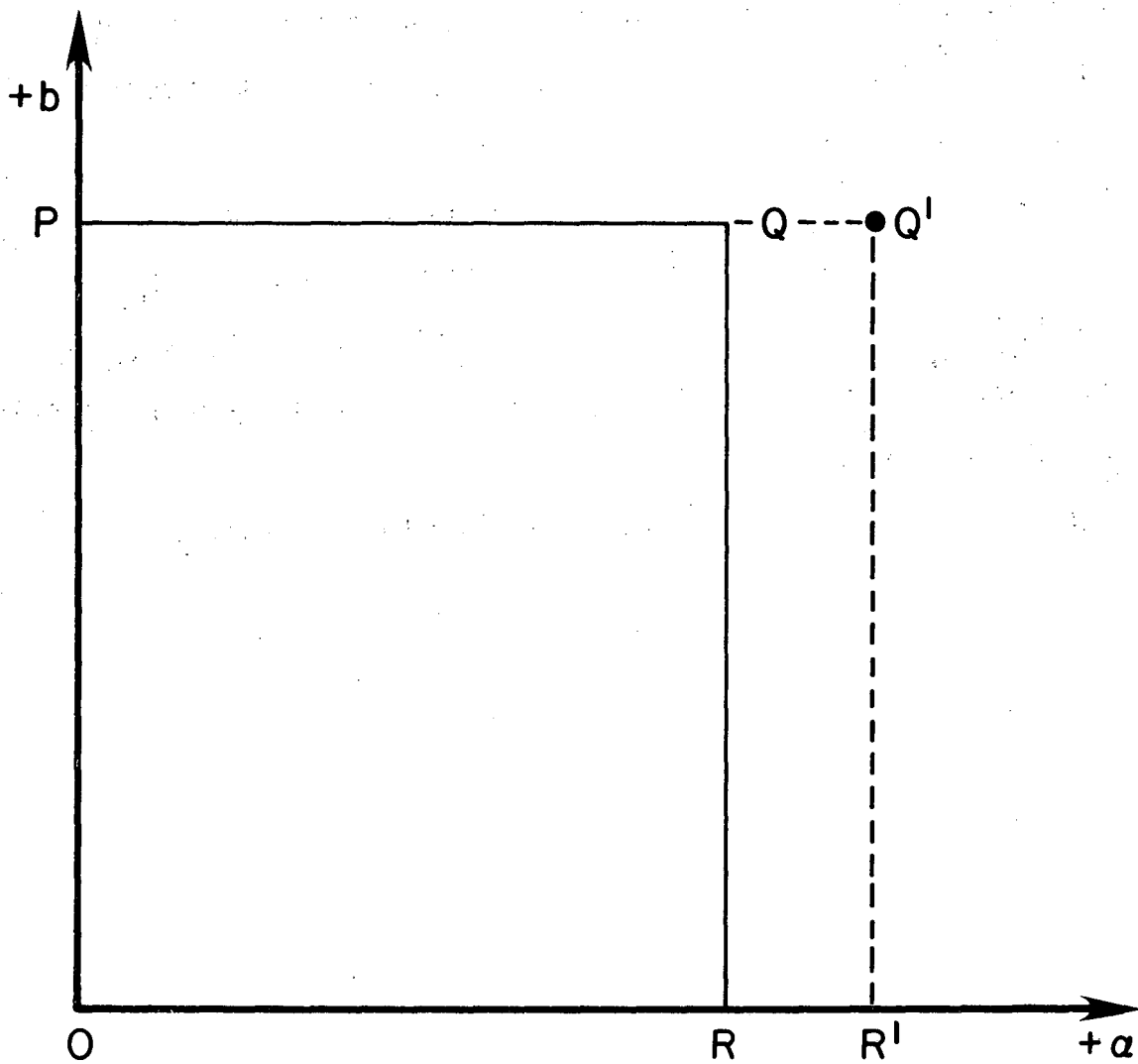
- (26) P. H. EBERHARD remark at the Thinkshop on Physics, Experimental Quantum Mechanics, Erice, Sicily, (1976); made in response to the paper presented by O. COSTA de BEAUREGARD.
- (27) E. P. WIGNER: Symmetries and Reflections (Indiana Univ. Press, Bloomington, Ind.) (1967).
- (28) For an extensive list of references, see M. JAMMER: The Philosophy of Quantum Mechanics (John Wiley & Sons, New York, 1974).
- (29) O. COSTA de BEAUREGARD: Dialectica, 19, 280 (1965).
- (30) J. VON NEUMANN: Mathematical Foundations of Quantum Mechanics (Princeton Univ. Press, Princeton, N.J.) (1955).
- (31) A. M. GLEASON: J. Math. & Mech., 6, 885 (1957).
- (32) J. S. BELL: Rev. Mod. Phys., 38, 447 (1966).
- (33) For a general review, see F. J. BELINFANTE: A Survey of Hidden Variables Theories (Pergamon Press, N.Y.) (1973).

Figure Captions

Fig. 1.— Domains $S(O,A,B)$ of the variables a and b for which a function $p(a,A,b,B)$ is not zero. The function is uniform in a and b in the rectangle $OPQR$ and zero outside for $A = A_1$ and $B = B_1$. For A and B equal to A_1 and B_2 , A_2 and B_1 , or A_2 or B_2 , the distribution is uniform in the rectangle $OPQ'R'$ and zero outside.

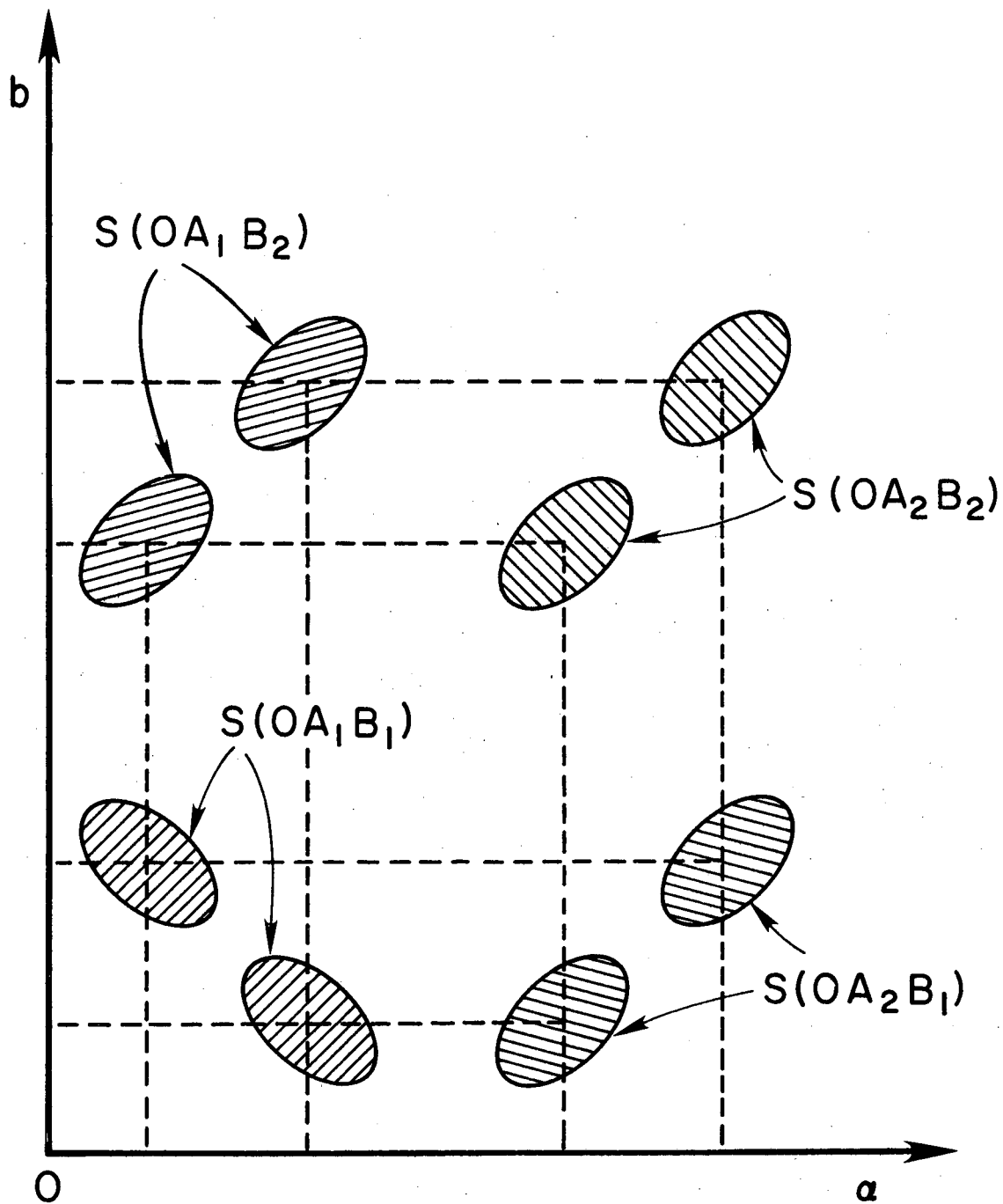
Fig. 2.— Domains $S(O,A_1,B_1)$, $S(O,A_1,B_2)$, $S(O,A_2,B_1)$, and $S(O,A_2,B_2)$ of the variables a and b for which distributions $p(a,A_1,b,B_1)$, $p(a,A_1,b,B_2)$, $p(a,A_2,b,B_1)$, and $p(a,A_2,b,B_2)$ are not zero. Each of them consists of the inside of two ellipses inside of which the distribution $p(a,A,b,B)$ is uniform in a and b .

Fig. 3.— Schematic of the experiment described in sect. 3.2



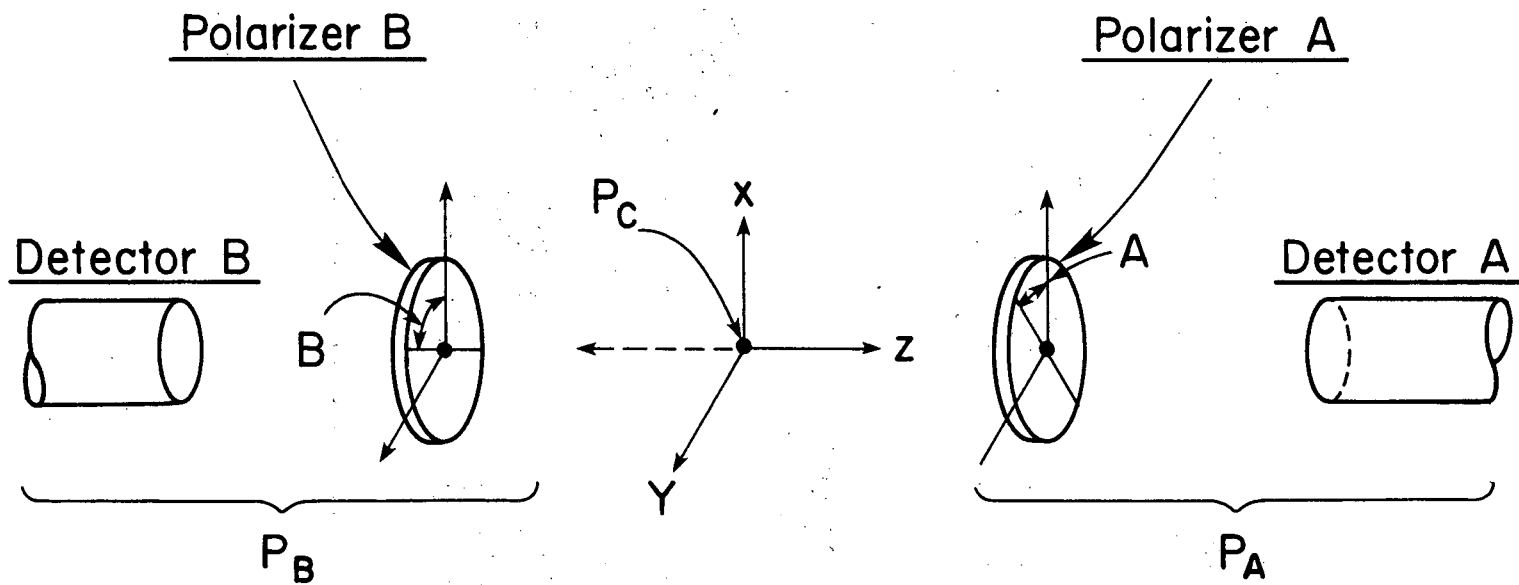
XBL784-552

Fig. 1



XBL 784 - 550

Fig. 2



XBL784 - 551

Fig. 3

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720