PHYSICAL REVIEW A VOLUME 49, NUMBER 5 MAY 1994

Stapp's algebraic argument for nonlocality

Michael Dickson
Department of Philosophy, University of Notre Dame, Notre Dame, Indiana 46556

Rob Clifton
Department of Philosophy, University of Western Ontario, London, Ontario, Canada N6A 3K7
(Received 31 March 1993; revised manuscript received 27 September 1993)

Stapp has recently presented a revised algebraic version of his (and others') earlier nonlocality arguments. Stapp's argument would seem to be an improvement over traditional arguments because he assumes only the so-called parameter independence condition, but neither determinism nor the so-called outcome independence condition. We show that, in fact, Stapp's argument rests on a logical fallacy.

PACS number(s): 03.65.Bz

I. INTRODUCTION

Nearly all attempts at deriving an inequality [1] or an algebraic contradiction [2] from a locality assumption coupled with quantum predictions have relied on one of two sets of assumptions, often described [3] as (a) determinism plus parameter independence (PI) (also called "locality"), and (b) parameter independence plus outcome independence (OI) (also called "completeness"). PI says that outcomes, or the probabilities for outcomes, at one detector [in the standard Einstein-Podolsky-Rosen (EPR)-Bohm experiment [1], the Greenberger-Horne-Zeilinger (GHZ) experiment [2], or, most recently, the Hardy experiment [4]] do not directly depend on parameter settings at the other, e.g., the orientation of the Stern-Gerlach magnet. Similarly, OI is the assumption that outcomes (or probabilities for outcomes) at one detector do not depend on outcomes at the other. Stapp and others have claimed in the past [5] that one can deny both determinism and OI, yet still derive a Bell inequality. However, his claim has been much disputed [6]. Recently [7–9] Stapp has taken a slightly new strategy, namely, to seek an algebraic contradiction (rather than a Bell inequality). In these new arguments, he reaffirms his commitment to a denial of determinism. (See especially Ref. [7].) And in [8], he makes it clear that his intention remains to avoid any assumption such as OI:

> Some proofs introduce *hidden variables*, and require a factorization property that entails that, for any fixed values of these hidden variables, the result of a measurement in one of two spacelike-separated regions must be independent of the *result* of the measurement in the other region.

Stapp claims that such proofs are "none too surprising." Two very well known proofs along these lines (and ones whose conclusions, we believe, *are* surprising) are those of Bell, and Clauser and Horne [1]. (We emphasize, however, that the introduction of hidden variables is inessential to OI, and that Bell's inequality or an algebraic contradiction can be derived without explicit reference to them, using just OI and PI— see ref. [10].) We claim that Bell and others are correct to require OI along with PI, given that they avoid determinism. We support this claim by showing that Stapp's argument is logically invalid. We shall then suggest that *given* the assumption of OI (or determinism, or, of course, both), the proof is valid.

II. STRICT CONDITIONALS AND COUNTERFACTUAL CONDITIONALS

The two essential features of Stapp's latest argument, apart from the quantum predictions, are his locality ('no-faster-than-light-influence") conditions—roughly, PI—and a rule of inference which elsewhere [9] he calls "elimination of eliminated conditions" (EEC). For reasons discussed at length elsewhere [6], we doubt whether Stapp's locality conditions are valid in a genuinely indeterministic framework. However, here we shall grant Stapp's locality conditions as unproblematic so that we may focus on the (perhaps more serious) difficulty arising from the supposed rule of inference EEC. We concentrate on the argument in [8], that of [7] being the same in logical structure.

Notation: We adopt Stapp's notation in [8] and his use of Hardy's nonlocality experiment [4], in which there are two spacelike-separated regions, A and B, in each of which there are two possible measurements, 1_A or 2_A and 1_B or 2_B, with possible results y (yes) or n (no) in each region. Letting $R = A$ or B and $i = 1$ or 2, we define: R_i means "the measurement i_R is performed in region R." Further, letting $\rho = y$ or n, we define: ρ_R means "the result ρ occurs in region R." Note that because only one measurement can be performed in a given region at a time, and only one result can occur in a given region at a time, the following inferences are true, with "$\rightarrow$" the strict conditional connective and "$\neg$" the logical negation operator:

1050-2947/94/49(5)/4251(6)/$06.00

$$R_i \to \neg R_j \quad (\text{for } j \neq i),$$
$$\rho_R \to \neg \sigma_R \quad (\text{for } \rho \neq \sigma). \tag{1}$$

Stapp [8] says explicitly that he, too, uses $\to$ as the strict conditional. In doing so, he correctly derives from the quantum predictions and his locality conditions the following nested conditional (where "$\wedge$" is the logical "and" connective):

$$(A_1 \wedge B_1 \wedge y_A) \to [(A_1 \wedge B_2) \to [(A_2 \wedge B_2) \to [(A_2 \wedge B_1) \to [(A_1 \wedge B_1) \to (A_1 \wedge B_1 \wedge n_B)]]]]. \tag{2}$$

Stapp says that (2) implies

$$(A_1 \wedge B_1 \wedge y_A) \to (A_1 \wedge B_1 \wedge n_B). \tag{3}$$

And (3) directly contradicts the quantum predictions for the Hardy [4] experiment. Hence, if Stapp has legitimately derived (3) from his locality conditions plus the quantum predictions, then he has indeed shown that they lead to a contradiction. He claims to justify the inference from (2) to (3) by use of the following rule (EEC):

$$\{(A_1 \wedge B_1 \wedge y_A) \to [(A_1 \wedge B_2) \to [(A_2 \wedge B_2) \to [(A_2 \wedge B_1) \to [(A_1 \wedge B_1 \to (A_1 \wedge B_1 \wedge n_B)]]]]\} \to [(A_1 \wedge B_1 \wedge y_A) \to (A_1 \wedge B_1 \wedge n_B)]. \tag{4}$$

The first thing to note about this argument is that Stapp cannot mean what he says when he says that $\to$ is a strict conditional, for straightforward truth-functional analysis reveals that EEC is then invalid. The strict conditional $p \to q$ is false if and only if there is a possible valuation: $p \equiv \top$ ("true") and $q \equiv \bot$ ("false") [11]. Using the valuation $(A_1 \equiv \top, A_2 \equiv \bot, B_1 \equiv \top, B_2 \equiv \bot, y_A \equiv \top, n_B \equiv \bot)$, EEC comes out false. Hence the rule thus interpreted is invalid. (Keep in mind that the possibility of a falsifying valuation is all that is required to invalidate a strict conditional.)

It is very important here to note that this valuation is permitted by the quantum predictions themselves. Indeed, it is exactly a prediction of the Hardy experiment [4]—in particular, his Eq. (17d)—that makes the falsifying valuation possible. Therefore, (17d) entails the invalidity of EEC. Furthermore, (17d) is explicitly used by Stapp in his derivation of an algebraic contradiction. And once one sees that (17d) and EEC are inconsistent, it is no surprise that the argument results in a contradiction. But, of course, deriving a contradiction from a contradiction does not constitute a nonlocality result.

On the other hand, the content of Stapp's argument suggests that he did not mean to say that $\to$ in the antecedent and consequent of EEC is the strict conditional, but rather the counterfactual conditional, usually written $\phi \,\Box\!\!\to \psi$ and read: "If ϕ had been the case, then ψ would have been the case." (Indeed, Stapp himself uses the counterfactual conditional in [9].) So, for example, the conditional

$$(A_1 \wedge B_1 \wedge y_A) \,\Box\!\!\to [(A_1 \wedge B_2) \,\Box\!\!\to (A_1 \wedge B_2 \wedge y_A)] \tag{5}$$

can be read: "If 1_A and 1_B had been measured with the result y in region A, then it would have been the case that, if 1_A and 2_B had been measured, it would have been the case that 1_A and 2_B were measured with the result y in region A." (5) is in fact one of Stapp's locality (PI) conditions, with "$\to$" interpreted as "$\Box\!\!\to$." (The first two conjuncts in the last consequent are obviously redundant, but we keep them in order to make it clear that our argument exactly follows Stapp's.) Under this necessary change of conditionals, EEC reads

$$\{(A_1 \wedge B_1 \wedge y_A) \,\Box\!\!\to [(A_1 \wedge B_2) \,\Box\!\!\to [(A_2 \wedge B_2) \,\Box\!\!\to [(A_2 \wedge B_1) \,\Box\!\!\to [(A_1 \wedge B_1) \,\Box\!\!\to (A_1 \wedge B_1 \wedge n_B)]]]]\} \to [(A_1 \wedge B_1 \wedge y_A) \,\Box\!\!\to (A_1 \wedge B_1 \wedge n_B)]. \tag{6}$$

We do not change the second-to-last conditional to a counterfactual because it expresses the entailment of the consequent of EEC by its antecedent.

An important question arises here, namely, whether *all* of the conditionals in (4) must be reinterpreted as in (6). The answer is no. On the one hand, it is clear that, for example, the second conditional must be a counterfactual because, given $A_1 \wedge B_1$, one can at best say what *would* have happened had $1_A \wedge 2_B$ been performed. [One cannot say what does or will happen, given that $1_A \wedge 2_B$ *is* performed because quite clearly it is not—given the assumption $A_1 \wedge B_1$, the concurrent assumption that $A_1 \wedge B_2$ would amount to the claim that the measuring device in region B simultaneously makes two different measurements, a contradiction of (1), above.] On the other hand, this problem does not plague the first and last conditionals in (6); hence, there is nothing preventing them from being strict conditionals. [A similar argument would make the first, but not the second, conditional in (5) a strict conditional.] Changing the first and last conditionals back to strict conditionals, EEC becomes

$$\{(A_1 \wedge B_1 \wedge y_A) \to [(A_1 \wedge B_2) \,\Box\!\!\to [(A_2 \wedge B_2) \,\Box\!\!\to [(A_2 \wedge B_1) \,\Box\!\!\to [(A_1 \wedge B_1 \,\Box\!\!\to (A_1 \wedge B_1 \wedge n_B)]]]]\} \to [(A_1 \wedge B_1 \wedge y_A) \to (A_1 \wedge B_1 \wedge n_B)]. \tag{7}$$

However, in the next section we shall show that if (6) is invalid, then so is (7). Indeed, of (4), (6), and (7), (6) is the logically weakest formulation of EEC, so that its failure entails the failure of (4) and (7). Therefore, we shall concentrate on (6) and demonstrate its invalidity. We emphasize that the move from the strict conditional to the counterfactual conditional is absolutely necessary if Stapp's argument has even a chance of validity, because EEC is clearly invalid when all its conditionals are read as strict.

III. TRUTH CONDITIONS FOR THE COUNTERFACTUAL CONDITIONAL

The counterfactual conditional is, of course, subject to different truth conditions from those of the strict conditional. The conditions most widely used—and the ones endorsed explicitly by Bedford and Stapp [9]—are those given in Lewis' [12] analysis in terms of possible worlds. Lewis proposes:

> "$\phi \Box\!\!\rightarrow \psi$" is true at a world w iff: (i) There are no ϕ-worlds (i.e., possible worlds where ϕ is true) or (ii) *every* ϕ-world among those worlds closest to w is also a ψ-world.

[One must say "among those closest to w" to allow for ties—two ϕ worlds equally close to w. Also, for simplicity we ignore the case where there are *no* ϕ worlds closest to w—just as there are no positive real numbers closest to zero. Cf. [12] (1976), Sec. 1.4.] Example: Let ϕ denote "a photon hits the photographic plate" and let ψ denote "the plate registers a hit." Presumably, "$\phi \Box\!\!\rightarrow \psi$" is then true in our world (i.e., the actual world, which, hereafter, we call α). And a remote (highly dissimilar) possible world where the laws of nature are different (so that photographic plates are unaffected by photons) is irrelevant, because such a world is farther from α than are possible worlds where the laws are such that in fact photons do affect photographic plates (as they do in α).

For present purposes, one need not worry unduly about a general definition of closeness of possible worlds. We shall take only the rather flexible position that the closest worlds are ones where no "indiscriminate" changes are made—indiscriminate here meaning "not forced by the antecedent of the counterfactual conditional." So, for example, because the counterfactual supposition ϕ about the photon did not require a change in the laws of nature, we presumed above that worlds where such changes obtain are farther from α than those which share our laws. In what follows, nothing we say turns on controversial assumptions about the precise criteria to be used for world similarity.

It is clear from the truth conditions for counterfactuals that (6) is weaker than (7). Recall that the invalidity of a strict conditional $p \rightarrow q$ is assured by the possibility of a falsifying valuation. In other words, if $p \rightarrow q$ is true, then it is the case that either $p \equiv \bot$ or $q \equiv \top$ (or both) in every possible world. (It is best in this case, however, to restrict ourselves to possible worlds where quantum predictions hold.) On the other hand, if $p \Box\!\!\rightarrow q$ is true, then it could still be that $p \equiv \top$ and $q \equiv \bot$ in *some* worlds. The truth of $p \Box\!\!\rightarrow q$ requires only that $p \equiv \bot$ or $q \equiv \top$ (or both) in the *nearest* p worlds. Therefore, the strict conditional is *easier* to falsify than is the counterfactual conditional. And conversely, if a counterfactual conditional is false, then *a fortiori* the corresponding strict conditional is false. Hence, if we want to show the invalidity of EEC in its weakest formulation, we must show that (6) is invalid. It will follow that (7) is invalid.

IV. INVALIDITY OF EEC

The consequent of a strict conditional is often called a "necessary condition" for the antecedent. In other words, a consequent, q, has a certain *restriction* placed upon it merely by virtue of its being the consequent of some antecedent, p, namely, that it must be true if p is. The consequent of a counterfactual conditional is similarly restricted— though now *counterfactually* restricted—by its antecedent, for which it may be called a 'counterfactually necessary condition." Hence, even if for some reason one supposes that the antecedent of some counterfactual $p \Box\!\!\rightarrow q$ is not of interest, still it cannot be ignored, because it imposes a certain restriction on q. Stapp's fundamental error in endorsing EEC is to ignore certain counterfactual antecedents which he considers to have been "countermanded." Even if he is correct, their having been countermanded is *not* license to discard these antecedents. The following rigorous analysis of EEC makes it clear that Stapp's brief and nonrigorous argument does not justify the EEC "collapse" of a nested counterfactual into a single one.

The way to think about (6) is to imagine moving from one set of possible worlds to another. Here we begin by putting ourselves into the nearest $(A_1 \wedge B_1 \wedge y_A)$ worlds to α and asking, what would have been the case (given that we are in that world) if we had performed $(1_A \wedge 2_B)$? To answer *that* question we jump into the nearest $(A_1 \wedge B_2)$ worlds—that is, the $(A_1 \wedge B_2)$ worlds nearest to the $(A_1 \wedge B_1 \wedge y_A)$ worlds where we were a moment ago. There we ask, what would have been the case if we had performed $(2_A \wedge 2_B)$? The completed process is illustrated in Fig. 1.

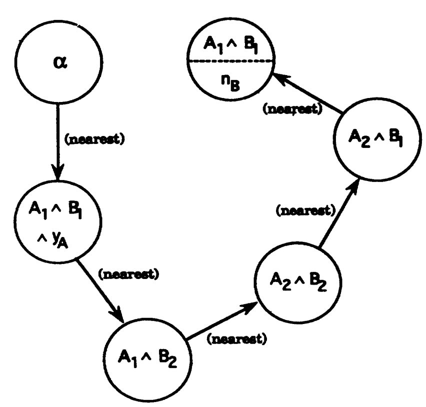

FIG. 1. Possible worlds analysis of EEC.

The process ends at the $(A_1 \wedge B_1)$ worlds nearest to the $(A_2 \wedge B_1)$ worlds nearest to the $(A_2 \wedge B_2)$ worlds nearest to the $(A_1 \wedge B_2)$ worlds nearest to the $(A_1 \wedge B_1 \wedge y_A)$ worlds nearest to α. Are those worlds at the end of the process exactly the same worlds as the $(A_1 \wedge B_1 \wedge y_A)$ worlds nearest to α? If they were, then all of the $(A_1 \wedge B_1 \wedge y_A)$ worlds nearest to α would also have to be n_B worlds, and (3) (interpreted with counterfactual conditionals) would come out true under Lewis' conditions. This result would mean that the diagram in Fig. 1 "closes." But there is no reason whatsoever to suppose that the diagram does close [13]. The worlds at the end of the process are n_B worlds *by virtue of the restrictions placed upon them.* But these restrictions are *not* also placed on the $(A_1 \wedge B_1 \wedge y_A)$ worlds nearest to α. (2) (interpreted with counterfactual conditionals) and Fig. 1 tell us *nothing whatever* about the direct relationship between α and the $(A_1 \wedge B_1 \wedge n_B)$ worlds at the end of the process. Indeed, (2) does not imply (3) for at least the following reasons:

(a) Nothing in (2) or Fig. 1 precludes the possibility that the $(A_1 \wedge B_1 \wedge y_A)$ worlds nearest to α are all y_B worlds. In this case, the diagram would not close. This possibility, however, is perhaps contrary to the quantum predictions for the Hardy experiment—whether it is might depend on difficult issues concerning the connection between probabilitistic and counterfactual statements.

(b) Nothing in (2) or Fig. 1 precludes the possibility that the $(A_1 \wedge B_1 \wedge n_B)$ worlds at the end of the process are all n_A worlds. Again, in this case the diagram would not close. But this possibility is subject to the same potential objection as (a).

(c) Nothing in (2) or Fig. 1 precludes the possibility that the $(A_1 \wedge B_1 \wedge y_A)$ worlds nearest to α are split, some being n_B worlds, others y_B worlds. This possibility amounts to the claim that the $(A_1 \wedge B_1 \wedge n_B)$ worlds at the end of the process are a subset of the $(A_1 \wedge B_1 \wedge y_A)$ worlds nearest to α, so that the diagram partially closes. In such a case, there is a tie in nearness to α between $(A_1 \wedge B_1 \wedge y_A \wedge y_B)$ worlds and $(A_1 \wedge B_1 \wedge y_A \wedge n_B)$ worlds. Hence, (3) is false under Lewis' conditions while (2) could still be true. The quantum predictions do *not* tell against this possibility, but indeed suggest quite strongly that it is the case, given that the outcome in region B is truly indeterministic, i.e., it might have gone either way.

(d) Because the outcome of the 1_B measurement in any $(A_1 \wedge B_1 \wedge y_A)$ world nearest to α is restricted only by α (i.e., by its being a counterfactually necessary condition for α), whereas the 1_B outcome in any $(A_1 \wedge B_1)$ world at the end of the process is subject to a long string of restrictions, there is no reason to suppose that these worlds are the same. The first is hardly restricted at all. The latter is highly restricted. Stapp's move from (2) to (3) is, in effect, a retention of the *effects* of the restrictions while dropping the restrictions themselves. But, of course, once one drops the restrictions, there is no reason to suppose that their effects continue to obtain.

Furthermore, it does not help matters to note, as Stapp sometimes does [5], that any given measurement event yields a "unique result," meaning that it yields only a single possible result out of the range of possible results. It does not help because the measurement event in $(A_1 \wedge B_1)$ worlds nearest to α is not the same as the measurement event in $(A_1 \wedge B_1)$ worlds at the end of the process, even though the settings are the same. These measurement events are not the same for the simple reason that they exist in different possible worlds. But Stapp cannot claim that measurement events in different possible worlds, when the settings are the same, must yield the *same* unique result—such a claim amounts to determinism.

Point (c) is sufficient to dispose of EEC, but point (d) is perhaps more relevant to Stapp's own attempt to justify EEC. With (d) in mind, one can see that his claim that the intermediate conditions are countermanded by the most recent condition $(A_1 \wedge B_1)$ in (2) is irrelevant to the truth or falsity of EEC. Proposition (2), and likewise the antecedent of (6), does *not* consist of the repeated replacement of one condition for another. It proceeds, rather, by piling new conditions on to the old ones. The point is easily seen in a simpler case:

$$p \mathrel{\Box\!\!\rightarrow} (q \mathrel{\Box\!\!\rightarrow} r). \tag{8}$$

The consequent of (8) is not a replacement of the original condition p by the condition q. It is, rather, itself a counterfactual conditional. So, (8) says that if p had been the case, then $q \mathrel{\Box\!\!\rightarrow} r$ would have been true. (8) is not equivalent to, nor does it imply that, $p \mathrel{\Box\!\!\rightarrow} r$, even if (indeed, *especially* if) $q \rightarrow \neg p$, i.e., even if q countermands p. The further complexity of (2) adds no means to avoid these fundamental facts of logic.

V. WHY RULES OF CLOSENESS DO NOT MATTER

It is important to realize that the argument of Sec. IV does not, as we said before, depend on any controversial claims about rules for closeness of worlds. To see why, consider again point (c) of Sec. IV. Could *any* acceptable rule for determining closeness of worlds *prevent* the possibility of the situation described in (c)? No. Recall that the problem in (c) arose because there is no guarantee that the $(A_1 \wedge B_1 \wedge y_A)$ worlds closest to α are not split between $(A_1 \wedge B_1 \wedge y_A \wedge y_B)$ worlds and $(A_1 \wedge B_1 \wedge y_A \wedge n_B)$ worlds. To prevent this possibility, one would have to adopt a rule of closeness which *guarantees* at least that the $(A_1 \wedge B_1 \wedge y_A)$ worlds nearest to α are all n_B worlds. The problem is that any such rule of closeness would, without any further argument, make true the following counterfactual:

$$(A_1 \wedge B_1 \wedge y_A) \mathrel{\Box\!\!\rightarrow} (A_1 \wedge B_1 \wedge n_B), \tag{9}$$

i.e., if one measures 1_A and 1_B and gets result y_A, then the result n_B is *guaranteed.* There are two points to make here. First, any rule of closeness which by itself implies (9) is clearly unacceptable. Indeed, (9) violates the quantum predictions for the Hardy experiment. [See his Eq. (17d).] Second, it is of course true that (9) is in fact the claim which Stapp wants to derive in order to get a contradiction, but (9) must be derived from a locality condition plus the quantum predictions, not guaranteed directly (quite apart from any locality considerations) by one's choice of the rules of closeness. Any rule which guarantees (9) already violates quantum mechanics and therefore cannot be adopted without presupposing from the start that quantum mechanics is false.

Therefore, in order to avoid the problem described in (c) by adopting some special rule of closeness, one would have to assume the falsity of the quantum predictions. But deriving a contradiction with quantum mechanics having already implicitly assumed the falsity of quantum mechanics does not constitute a nonlocality result. Note also that this argument is completely general—it applies to *any* rule of closeness which might be adopted to meet the objections of Sec. IV.

This independence from rules of closeness is a special case of a more general problem, namely, that any new assumption (i.e., an assumption derived from neither locality nor quantum mechanics) which might be used to justify EEC will necessarily also entail the falsity of quantum mechanics, the reason being that, as we have shown above, EEC is incompatible with the quantum predictions for the Hardy experiment. Hence, the proper response to any such argument is to deny not locality or quantum mechanics but the new assumption used to justify EEC. For the same reason, it is clear that one cannot simply adopt EEC as a postulate. The only option left is to show that EEC is a logical truth.

VI. CONCLUSIONS

We have shown that even in its weakest formulation, EEC is logically invalid. Hence, Stapp's algebraic argument, which relies essentially on EEC, is invalid. Note, however, that adopting OI will restore the argument. To see this point, recall that the $(A_1 \wedge B_1)$ worlds at the end of the process are restricted by (i.e., are counterfactually necessary conditions for) all of the antecedent families of worlds indicated in Fig. 1. For the present case, the important restrictions are imposed on outcomes at one detector by the parameter settings and outcomes in those antecedent families of worlds. But Stapp's PI guarantees that parameter settings at one detector cannot restrict outcomes at the other. (5), for example, says that if $(A_1 \wedge B_1 \wedge y_A)$ then the closest $(A_1 \wedge B_i)$ worlds must be $(A_1 \wedge B_i \wedge y_A)$ worlds, no matter what the value of i. (It is, of course, trivial for $i = 1$.) Thus far, however, there remains the possibility that introducing an outcome at region B will change matters, i.e., there remains the possibility that given some ρ, if $(A_1 \wedge B_1 \wedge y_A \wedge \rho_B)$, then it *need not* be the case that the closest $(A_1 \wedge B_1 \wedge \rho'_B)$ worlds are $(A_1 \wedge B_1 \wedge y_A \wedge \rho'_B)$ worlds, no matter what the value of ρ'. Leaving this possibility open amounts to allowing counterfactual restriction of outcomes by outcomes at the opposite detector. But one can eliminate such restrictions if, on analogy with Stapp's formulation of PI, one adopts the following form of OI:

$$(A_i \wedge B_j \wedge \rho_A \wedge \rho_B) \;\Box\!\!\rightarrow [(A_i \wedge B_j \wedge \rho'_B) \;\Box\!\!\rightarrow (A_i \wedge B_j \wedge \rho_A \wedge \rho'_B)]. \quad (10)$$

Indeed, if we add (10) to Stapp's assumptions, then a contradiction *can* be derived. But obtaining a contradiction with quantum predictions from PI *plus* OI is nothing new; hence, we shall not follow through the derivation of that contradiction here.

ACKNOWLEDGMENTS

M.D. thanks the University of Notre Dame, the Mellon Foundation, the International Center for Theoretical Physics, and S.I.S.S.A for generous financial support while this work was being done. R.C. similarly thanks Christ's College of Cambridge University, and the Social Sciences and Humanities Research Council of Canada.

[1] J. Bell, Physics **1**, 195 (1964); in *Foundations of Quantum Mechanics: Proceedings of the Enrico Fermi International School of Physics, Course II*, edited by B. d'Espagnat (Academic Press, New York, 1971), pp. 171–181; J. Clauser and M. Horne, Phys. Rev. D **10**, 526 (1974). See also the following extensive review: J. Clauser and A. Shimony, Rep. Prog. Phys. **41**, 1881 (1978).

[2] D. Greenberger, M. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Reidel, Dordrecht, 1989), pp. 69–72; D. Greenberger, M. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys. **58**, 1131 (1990); N. D. Mermin, *ibid.* **58**, 731 (1990); Phys. Today **43**, 9 (1990); Phys. Rev. Lett. **65**, 3373 (1990); R. Clifton, M. Redhead, and J. Butterfield, Found. Phys. **21**, 149 (1991).

[3] The observation that the factorizability assumption (used by the approaches in [1]) is decomposable into PI plus OI was anticipated in J. Clauser and M. Horne, Ref. [1]; P. Suppes and M. Zanotti, *Logic and Probability in Quantum Mechanics*, edited by P. Suppes (Reidel, Dordrecht, 1976), pp. 445–455; and B. van Fraassen, Synthese **52**, 25 (1982), but not fully and clearly demonstrated until 1984 by J. Jarrett, Noûs **18**, 569 (1984). Cf. J. Jarrett, in *Philosophical Consequences of Quantum Theory*, edited by E. McMullin and J. Cushing (Notre Dame University Press, Notre Dame, 1989), pp. 60–79; and L. Ballentine and J. Jarrett, Am. J. Phys. **55**, 696 (1987). The terms PI and OI were introduced by A. Shimony, in *Quantum Concepts in Space and Time*, edited by R. Penrose and C. Isham (Oxford University Press, Oxford, 1986), pp. 182–203. (PI and OI are not strictly equivalent to Jarrett's two principles, called "locality" and "completeness." Nor is PI strictly equivalent to Stapp's "no-faster-than-light-influence" condition. But these differences are not important for the argument at hand.)

[4] L. Hardy, Phys. Rev. Lett **68**, 2981 (1992). Cf. R. Clifton and P. Niemann, Phys. Lett. A **166**, 177 (1992); L. Hardy and E. Squires, *ibid.* **168**, 169 (1992).

[5] H. Stapp, Il Nuovo Cimento **40B**, 191 (1977); Found. Phys. **10**, 767 (1980); Am. J. Phys. **53**, 306 (1985); Found. Phys. **18**, 427 (1988); *Philosophical Consequences of Quantum Theory*, edited by E. McMullin and J. Cushing (Notre Dame University Press, Notre Dame, 1989), pp. 155–174. See also P. Eberhard, **35B**, 75 (1977); **46B**, 392 (1978).

[6] The most extensive critique of Stapp's argument is R. Clifton, J. Butterfield, and M. Redhead, Brit. J. Phil. Sci. **41**, 5 (1990). Other critiques (all agreeing that the argument is invalid) are G. Hellman, Synthese **53**, 461 (1982); M. Redhead, *Incompleteness, Nonlocality, and Realism* (Clarendon Press, Oxford, 1987); H. Krips, *The Metaphysics of Quantum Theory* (Clarendon Press, Oxford, 1987); J. Philos. Logic **18**, 155 (1989); R. Clifton, Found. Phys. Lett. **2**, 347 (1989); **2**, 531 (1989); A. Fine, in *Philosophical Consequences of Quantum Theory*, edited by E. McMullin and J. Cushing (Notre Dame University Press, Notre Dame, 1989), pp. 175–194; M. Dickson Stud. Hist. Philos. Mod. Phys. (to be published).

[7] H. Stapp, Phys. Rev. A **47**, 847 (1993).

[8] H. Stapp, Phys. Rev. A **46**, 6860 (1992).

[9] D. Bedford and H. Stapp, Lawrence Berkeley Laboratory Report No. LBL-29836 (unpublished).

[10] R. Clifton, C. Pagonis, and I. Pitowsky, in *Philosophy of Science 1992*, edited by D. Hull, M. Forbes, and K. Okruhlik (Philosophy of Science Association, East Lansing, MI, 1992), Vol. I, pp. 114–128.

[11] The material conditional, often denoted "$p \supset q$," is true iff $p \equiv \perp$ or $q \equiv \top$. The strict conditional (sometimes called "entailment") is the *necessity* of a material conditional, i.e., $\Box(p \supset q)$ where $\Box$ is the standard necessity operator in modal logic. It is true if and only if it is not possible that $p \equiv \top$ and $q \equiv \perp$. Stapp's EEC fails whether $\rightarrow$ is the material *or* the strict conditional.

[12] D. Lewis, *Counterfactuals* (Blackwell, Oxford, 1976); *Philosophical Papers* (Oxford University Press, Oxford, 1986), Vol. 2.

[13] This difficulty is analogous to the "broken square problem" discussed by Clifton, Butterfield, and Redhead *et al.* in Ref. [6].