Quantum Locality?

Henry P. Stapp
Lawrence Berkeley National Laboratory
University of California Berkeley, California 94720

Abstract. Robert Griffiths has recently addressed, within the framework of a 'consistent quantum theory' that he has developed, the issue of whether, as is often claimed, quantum mechanics entails a need for faster-than-light transfers of information over long distances. He argues that the putative proofs of this property that involve hidden variables include in their premises some essentially classical-physics-type assumptions that could are not entailed by the precepts of quantum mechanics. Thus whatever is proved is not a feature of quantum mechanics, but is a property of a theory that tries to combine quantum theory with quasi-classical features that go beyond what is entailed by quantum theory itself. One cannot logically prove properties of a system by establishing, instead, properties of a system modified by adding properties alien to the original system. Hence Griffiths' rejection of hidden-variable-based proofs is logically warranted. Griffiths mentions the existence of a certain alternative proof that does not involve hidden variables, and that uses only macroscopically described observable properties. He notes that he had examined in his book proofs of this general kind, and concluded that they provide no evidence for nonlocal influences. But he did not examine the particular proof that he cites. An examination of that particular proof by the method specified by his 'consistent quantum theory' shows that the cited proof is valid within that restrictive version of quantum theory. An added section responds to Griffiths' reply, which cites general possibilities of ambiguities that might make what is to be proved ill-defined, and hence render the pertinent 'consistent framework' ill defined. But the vagaries that he cites do not upset the proof in question, which, both by its physical formulation and by explicit identification, specify the framework to be used. Griffiths confirms the validity of the proof insofar as that pertinent framework is used. The section also shows, in response to Griffiths' challenge, why a putative proof of locality that he has described is flawed.

Keywords: Nonlocality, Consistent Quantum Theory, Counterfactual.

PACS: 01.70 03.65.bz

INTRODUCTION

Robert Griffiths begins his recent paper *Quantum Locality* [1] with the observation that "The opinion is widespread that quantum mechanics is nonlocal in the sense that it implies the existence of long range influences which act instantaneously over long distances, in apparent contradiction to special relativity". He says that the purpose of his paper "is to move beyond previous discussions by employing a fully consistent quantum mechanical approach" to "argue that the supposed nonlocal influences do *not* exist" and to "establish on the basis of quantum principles a strong statement of quantum *locality*: the objective properties of an isolated individual (quantum) system do not change when something is done to another non-interacting system."

Griffiths' claims, if valid, would constitute an extremely important achievement: it is difficult to find an issue as central to our understanding of nature as the question of whether or not far-flung parts of the universe are tied together by long-range faster-than-light transfers of information .

Almost all of Griffiths' paper is directed against arguments for nonlocality that are based on the concept of hidden variables: the paper is directed primarily against arguments that have stemmed directly from the works of John Bell pertaining to local deterministic and local stochastic *hidden-variable* theories. However, the local *stochastic* hidden-variable theories have been shown by Stapp [2], and also by Fine[3], to be essentially equivalent to local *deterministic* hidden-variable theories. But these latter theories are theories of an essentially classical-physics type, with statistically distributed unobservable hidden variables. Such theories could include Bohm's pilot-wave model if it were stripped of its nonlocal-interaction feature, which is, however, essential to its structure and its success, particularly in applications to the EPR-type correlation experiments that are the basis of the arguments for nonlocal influences.

In view of this basically *classical* character of the hidden-variable theories, it is obviously going to be extremely difficult to deduce, in any logically sound way, the properties of a quantum-mechanical world from the properties of hidden-variable models: How can one pass, logically, from fact that one needs to add nonlocal influences to any essentially classical model, in order to fit the quantum predictions, to conclusions about the quantum mechanical universe itself? The logical difficulty in deriving such a conclusion is that the hidden-variable premises contain classical reality assumptions that are incompatible with basic quantum concepts. In view of this basic logical problem, it is clear that a search for a strictly rational proof of the existence within the quantum universe of nonlocal influences should focus on arguments that do not use hidden variables; arguments that are not based on the failure of local hidden-variable theories! Griffiths nevertheless confines his attention mainly to arguments for nonlocality based on the failure of local hidden-variable theories.

Commenting upon this severe curtailment of the scope of his arguments Griffiths laments that "In an argument of modest length it is impossible to deal with all the published arguments that quantum theory is beset with nonlocal influences… In particular we do not deal with …Stapp's counterfactual arguments. …the problems associated with importing counterfactual reasoning into the quantum domain are treated in some detail in Ch. 19 of [4], and the conclusion is the same: there is no evidence for them."

In this paper I shall show that the methods that Griffiths developed lead, rather, to the opposite conclusion. His "fully consistent quantum approach" *validates* the counterfactual argument that he cites, but does not analyze. The validated nonlocal influence required by the assumed validity of certain predictions of quantum theory is fully concordant with the basic principles of relativistic quantum field theory, which ensure that the phenomena covered by the theory can neither reveal a preferred frame associated with these influences, nor allow "signals" (sender-controlled information) to propagate faster than the speed of light.

COUNTERFACTUALS IN PHYSICS

The word "counterfactual" engenders in the minds of minds of most physicists a feeling of deep suspicion. This wariness is appropriate because counterfactuals, misused, can lead to all sorts of nonsense. On the other hand, all arguments for the need, in a universe in which the predictions of quantum mechanics hold, for some faster-than-light transfer of information requires considering in a single logical analysis the predictions of quantum theory associated with (at least) four *alternative* possible measurements. Probably the only logically sound way to do this, without bringing in hidden-variables, is to use counterfactuals. This can be done in a completely logical and rational way. Indeed, Griffiths takes pains to show how valid counterfactual reasoning is to be pursued and validated within his "consistent quantum theory". His conclusion pertaining to the validation of counterfactual reasoning is the basis of the present work.

Griffiths begins his discussion of counterfactuals [4, p. 262] by noting that "Unfortunately, philosophers and logicians have yet to reach agreement about what constitutes valid counterfactual reasoning in the classical domain." It is certainly true that philosophers fall into disputes when trying to formulate general rules that cover all of the conceivable counterfactual situations that they can imagine, in a classical-physics, and hence deterministic, setting. But such a setting is strictly incompatible with the notion of "free choices" that underlies the idea of alternative possibilities. But what will be examined here is only a very simple special case, one in which the quantum mechanical laws (predictions) *themselves* specify all that we need to know about the outcomes of the contemplated measurements, and in which alternatives arising from alternative possible choices become theoretically possible because of the allowed entry of elements of chance into the dynamics of the choices of which measurements will be performed.

As a brief introduction to the subject of counterfactual statements, consider the following simple classical example: Suppose an electron that is moving in some fixed direction with definite but unknown speed is shot into a region in which there is an electric field E that is known to be uniform at one or the other of two known values, E1 or E2, with E2 twice E1. And suppose two detectors, D1 and D2, are placed so that one can assert, on the basis of the known laws of classical electromagnetism, that "If E is E1 and detector D1 clicks, then if, *instead*, E is E2, the detector D2 would have clicked." Under the appropriate physical conditions this can be a valid theoretical assertion, even though it cannot be empirically verified, since one can not actually perform both of the contemplated alternative possible experiments. But the postulated physical laws allow one to infer from knowledge of what happens in a certain performed experiment what would have happened if, instead, an alternative possible measurement had been performed, all else being the same. The concept "if, *instead*," becomes pertinent in a quantum context in which this choice between E1 and E2 is controlled by whether a certain quantum detection device "clicks" or not. This choice of which measurement is performed is then not determined by the quantum mechanical laws, but enters as a "random" variable.

Consider in this light the following formulation of a putative argument for the need for faster-than-light transmission of information.

Suppose in each of two space-like separated regions, L and R , with L earlier than R (in some frame) there will be performed one or the other of two alternative possible measurements, with each measurement having two alternative possible outcomes. The choices between alternative possible measurements are to be specified in way that can be considered, within the quantum framework, to be "free choices": they are not specified by any known law or rule. The question at issue is whether, under these conditions, it is possible to satisfy the orthodox predictions of quantum mechanics in the four alternative possible measurement situations, without allowing information about the free choice made in either region to be present in the other region.

Notice that the only things that enter the argument are the random choices of which macroscopically described measurement is performed in each region, and the predictions of the theory about which macroscopically described outcomes then appear. No microscopic quantities or properties enter into the argument.

GRIFFITHS' CONSISTENT QUANTUM THEORY

The proof in question of the need for faster-than-light transfer of information was given in [5], and repeated in the last two pages of [6]. But the purpose of this paper is not to recall old results. It is rather to comment upon Griffiths' "consistent quantum theory" approach, which has attracted interest due to references to it by Murray Gell-Mann and Jim Hartle (who, in contrast to Griffiths, use it in a "Many-Worlds" context), and in particular to show that the counterfactual argument cited but not analyzed by Griffiths is, contrary to Griffiths' implicit claim, *validated* within his "consistent quantum theory" framework, as currently defined. This validation of the need for faster-than-light transmission of information within the "consistent quantum theory" framework constitutes a serious failing of that approach, insofar as it claims to be superior to the von Neumann approach because it does not lead to nonlocal influences.

I begin by describing Griffiths' general theory and its relationship to the orthodox quantum theory of von Neumann, to which it is contrasted.

,
"Measurements" play a very important role in orthodox quantum mechanics. But they are not generated by the quantum evolution in accordance with the Schroedinger equation. The physical act of performing a measurement on a quantum system and getting a positive empirical outcome is represented in the orthodox quantum mathematics by the action a corresponding *projection operator* on the prior quantum state.

Generalizing from the concept of a set alternative possible measurement *outcomes* at one single time one arrives at the concept of a "framework", which involving a sequence times $\{t_0, t_1, t_2, \ldots, t_f\}$, with $t_{i+1} > t_i$ and for each of these times t_i a set of orthogonal projection operators that sum to unity .

A "history" is a time-ordered set of (Heisenberg Picture) projection operators (all operating in the usual Hilbert space of the full physical system) with one projection operator selected from the set at each time t_i . The different alternative possible "histories" labeled by index k (which runs over the set of possible histories) are mapped (by Griffiths' chain operator) into operators represented by the symbols F_k. For each F_k the Hermitian conjugate of F_k is represented by G_k. Let "rho" represent the initial density matrix. Then the set of histories is called a "consistent" if and only if Trace (G_g rho F_k) is zero when g is different from k. This condition is automatically satisfied if, as in the case to be examined here, all of the occurring projection operators, in context, commute. In our case, every nonzero F_k can be identified by a trajectory that moves from left to right on a temporal tree graph that starts from a single line on the far left, and ends at one of sixteen possible lines on the far right, with each non-final segment of the tree graph having a binary branching into two lines at its right-hand endpoint, which occurs at one of the four times t_i at which a choice (of a measurement or an outcome) is made. This leads to sixteen possible lines on the far right of the tree graph. Purely for simplicity, one can take the evolution between measurements to be represented by the unit operator. In order to allow an easy graphical check on Griffiths' rules for validating counterfactual arguments one can, and should, prune away any branches that are required to have zero amplitude for the Hardy initial state.

Griffiths' procedure for checking the validity of counterfactual reasoning is to draw a tree graph that starts at the far left with a single horizontal line that represents the original (in our case, Hardy) state. In our case this line bifurcates at time t_1 into an upper branch labeled by ML1, and a lower branch labeled by ML2. These two branches represent the two alternative possible observer-selected settings of the device in the earlier region L. Then at time t_2 the line ML1 bifurcates into an upper branch labeled by ML1+, and a lower branch labeled by ML1-, and the branch ML2 bifurcates in similar way into ML2+ and ML2-. These branches represent the two alternative possible states of the outcome indicator (pointer) on device ML set at state of readiness ML1, and, alternatively, on the device ML set at state of readiness ML2. At time $t_{3,}$ each of these four branches bifurcates into an upper branch MR1 and a lower branch MR2, and then at time t_4 each of the eight branches bifurcates into a plus and a minus branch, giving one branch for each of the sixteen orthogonal states of the pair of apparatuses together with their respective pointers. This graph represents one single framework, within which the entire argument can be carried out, thereby satisfying Griffiths' crucial "single framework rule". Due to the orthogonality of the states representing the alternative possible device settings and of the alternative possible pointer locations in each region, and the orthogonality of the apparatus-pointer "outcome" states in the two regions L and R, Griffiths' condition of "consistent histories" is satisfied. Thus we can proceed to check Griffiths' condition for valid counterfactual reasoning.

The pertinent counterfactual statement has the form:

SR: "If MR1 is performed and the outcome MR1+ appears, then if, instead of MR1, rather MR2 is performed then the outcome MR2+ must appear."

If the initial state is the Hardy state, then Hardy [7] gives four pertinent predictions of quantum theory:

S1: If ML1 and MR1+, then ML1+. [Hardy's (14.a)]
S2: If ML1+ and MR2, then MR2+ [Hardy's (14.c)]
S3: If ML2+ and MR1, then MR1+. [Hardy's (14.b)]

S4: If ML2+ and MR2, then sometimes MR2-." [Hardy's (14.d)]

[Connection to Hardy's notation:

Hardy's		Stapp's
	$U_1 = 0$	ML1+
	$U_1 = 1$	ML1-
	$D_1 = 0$	ML2-
	$D_1 = 1$	ML2+
	$U_2 = 0$	MR1-
	$U_2 = 1$	MR1+
	$D_2 = 0$	MR2+
	$D_2 = 1$	MR2-

Statement S1 follows from Hardy's (14.a), which entails that, in the Hardy state, if ML1 and MR1 are performed and outcome MR1+ ($U_2 = 1$) appears, then outcome ML1+($U_1 = 0$) must appear---since ML1- ($U_1 = 1$) cannot appear. Statement S2 follows from (14.c), [If MR2 and ML1 are performed and MR2 has outcome -, then ML1 must have outcome -: Use the fact that A→B is equivalent to NotB→NotA. Statement S3 is a direct translation of Hardy's (14.b), and S4 follows from Hardy's (14.d), which asserts that the probabability that both ML2+ ($D_1 = 1$) and MR2- ($D_2 = 1$) appear is (with nonzero A and B) nonzero.]

It is a straightforward exercise to show that if the initial state is the Hardy initial state, and if it is assumed that an outcome that occurs and is recorded in the *earlier* region L is left unchanged if instead of MR1 rather MR2 is performed *later* in R, then the statement SR is true if ML1 is performed in L but is false if ML2 is performed in L: The truth of the statement SR about possible happenings in R depends upon which experiment is "freely chosen" in the region L, which is spacelike separated from region R

Griffiths' validation of SR in the ML1 case follows from the fact that if the choice in L is ML1 then starting on branch MR1+, the quantum prediction S1 justifies the move back to the "pivot point" where ML1+ branches into MR1 and MR2. Then S2 justifies the move forward to MR2+.

But if the choice of measurement in L had been ML2 then sometimes the outcome ML2+ appears. But under that condition, if MR1 is chosen on the right, then S3 implies that the outcome on the right must be MR1+. But in this case where MR1+ must appear, if, instead, MR2 is chosen in R then, virtue of S4, MR2+ sometimes does not appear, and we have a counter example to what was proved true in the case that ML1 was chosen in L. All parts of the argument are represented in the tree graph that corresponds to a "single framework", in accordance with Griffiths very restrictive "single framework rule".

RESPONSE TO GRIFFITHS REPLY

Griffiths [8] points out, quite correctly, that multiple frameworks often exist for analyzing counterfactual arguments. They arise, in the examples he cites, from ambiguities in the detailed meanings of the counterfactual statements, and these ambiguities are, in his examples, often resolved by making precise the intended meaning of the statements in a way that allows the framework to be fixed.

The purpose of my paper is to validate within the CQT formalism my previously published proof of the nonlocality property of quantum theory. To achieve this goal it necessary to express within the CQT formalism the intended meanings of the statements in my previously published proof. In order to resolve any possible ambiguities, I extract, directly from the statements of my proof (which involve only macroscopic properties, and include prominently the key experimental settings in the earlier region that are the variables upon which the counterfactual statement SR is to be proved to depend) crucial conditions that specify unambiguously the specific Griffiths framework that captures the intended meanings of the statements in my proof. While I regard this specification as merely the capturing within the Griffiths' formalism of the normal meaning of the statements in my proof, within the particular context under consideration, one can regard this specification of the Griffiths framework as resolving, within the CQT formalism, any possible ambiguities pertaining to the intended logical meaning of the words in my proof. In either case, I am entitled to use the framework that represents the intended meaning of the statements in my proof, and Griffiths confirms that within that pertinent framework my proof is valid: if the statements in my proof mean what they are intended to mean, as fixed by the framework that I specify, then my proof is validated within

Griffiths' consistent quantum theory. Of course, nothing can be proved insofar the statements in the proof lack unambiguous meanings.

One very nice thing about Griffiths' reply to my paper is that he not only confirms that my proof is valid in the pertinent framework, but he also shows two diagrams that immediately display almost all of what needs to be proved. In particular, one can see immediately from his Fig.(1a), which pertains to the case in which the alternative possibilities in L are $[0]_a$(ML1+) and $[1]_a$(ML1-), that if one performs Z_b (MR1) and gets Z^{--}_b (MR1+) then if instead of Z_b (MR1) one performs X_b (MR2) then, by tracing back from the *unique* Z^{--}_b (MR1+), to the pivot point where the choice between Z_b (MR1) and X_b (MR2) is made, and them forward along the MR2 branch, one arrives unambiguously at X^+_b(MR2+). This validates, within the Griffiths framework that specifies the intended meanings of the words, the first half of what needs to be proved. It confirms, within CQT, the same conclusion that, by using the natural logical meanings of the occurring words, follows unambiguously from the orthodox predictions of quantum theory, in a Hardy-type experiment.

Most of the other half is then proved by looking at Figure (1b), which pertains to the case in which the two alternative possible macroscopic outcomes in L are the states that in my proof I call ML2+ and ML2-. Figure 1b shows that in this case the CQM rules for validating the counter-factual statement SR fail: the final outcome in this second case need not be X^+_b (MR2+).

Thus the structural properties of Griffiths' two diagrams, Fig. (1a) and Fig. (1b), display graphically the logical connections in Hardy-type experiments that are the basis of the counterfactuals-based proof of quantum nonlocality.

Griffiths claims that "Neither Stapp nor anyone else has yet found a defect in the relatively straightforward (no counterfactuals) demonstration of the principle of Einstein locality given in [2], a principle that directly contradicts Stapp's claim of nonlocality."

In regard to Griffiths' argument for locality, as described in his final note, one point needs to be stressed. Griffiths' two diagrams, Figs. (1a) and (1b), reveal certain differences pertaining to the later region R – which contains particle b – that depend upon conditions pertaining to the earlier region L, which contains particle a, upon which is performed an action, either L1 or

L2, controlled by the free choice of the state |c>. However, this difference in region R is not revealed by simple averages: the *probabilities* for outcomes in R do not depend upon which state, |c>=|0> or |c>=|1> is chosen in L. Yet my proof of nonlocality claims, essentially on the basis of the differences represented in Figures (1a) and (1b) (but with the parts pertaining to the earlier region L represented by *macroscopic* conditions, as demanded by the statement of my theorem) that the information about the choice between L1 and L2 (controlled by the free choice between |c>=|0> and |c>=|1>) must nevertheless be present in the later region R, which is spacelike displaced from the region L in which the free choice between |c>=|0> and |c>=|1> was made. Hence my proof must somehow access the properties that are represented in the two figures, but that are not revealed by simple averages!

The differences represented in Figures (1a) and (1b) pertain to different "histories". The point here is that a "history", as normally conceived, is backward-looking; it takes what it now known, and constructs, on the basis of some presumed causal process, what can be said about the course of events that led to the present state of affairs. Correspondingly, the two diagrams, (1a) and (1b), represent, for different conditions in L, the different possible courses of events in R that can lead to different specified outcomes in R. This graphical representation of possible histories, provides a rational basis for defining within CQT the meaning of the counterfactual concept of "instead of": given the present facts, the theory specifies, graphically, the possible histories that could have led to these facts. One can then trace back along the graphically specified history that leads to a specified final outcome, (back) to the earlier choice specified by the "instead of" condition, and then forward along the alternative path to see what can be said in case the identified earlier binary choice had been reversed. The gives the concept "instead of" a formal mathematical meaning that accords with its intuitive meaning. Exploiting this capacity of the CQT "histories" approach to provide this mathematical foundation for counterfactual reasoning allows my proof to access the information about region R (and particle b) represented in Figures (1a) and (1b). It allows one to prove by means of counterfactual reasoning something that is not revealed by looking merely at probabilities alone.

Griffiths mentions that his proof of locality is "relatively straightforward (it does not involve counterfactuals)". But the fact that the proof does not involve counterfactuals is not an asset. The CQT formalism is supposed to encompass counterfactual arguments, and hence a valid proof of locality must involve counterfactuals, in the sense of ruling out the possibility that a counterfactual argument could lead to a contradictions with its locality conclusion. No such demonstration is included. That this omission is a serious deficiency of the proof is demonstrated by fact that the CQT formalism, duly applied to represent my counterfactuals-based proof of nonlocality, entails a violation of the locality condition. The mere fact that this contradiction with Griffiths' proposed proof of locality can occur, within CQT, regardless of whether some alternative proposed formulations do not lead to contradictions with locality, means that Griffiths' proof is logically flawed: it does not rule out the possibility that a counterfactual argument could contradict the asserted conclusion.

ACKNOWLEDGEMENT

I thank Robert Griffiths for a long cordial correspondence that allowed us to propound our views in ways that clarify the basic issues, rather than obscuring them. This work was supported by the Director, Office of Science, Office of High Energy and Nuclear Physics, of the U.S. Department of Energy under contract DE-AC02-05CH11231

REFERENCES

1. Robert Griffiths, Quantum Locality, *Found. Phys*. 41, 705-733 (2011)
2. Henry P. Stapp, *Epistemological Letters*, June 1979, Colloquium on Bell's Theorem, Assoc. F. Gonseth, Case Postale 1081, Bienne, Switzerland. (1979) Reprinted in http://www-physics.lbl.gov/~stapp/AnEquivalenceTheorem.pdf
3. Arthur Fine, Hidden variables, Joint Probabilities, and Bell Inequalities, *Physical Review Letters*, 48, 291-295, (1982)
4. Robert Griffiths, *Consistent Quantum Theory*, Cambridge Univ. Press, Cambridge UK (2002)
5. Henry P. Stapp, A Bell-type theorem without hidden variables. *American Journal **of** Physics*, 72, 30-33, (2004)
6. Henry P. Stapp, *Mindful Universe: Quantum Mechanics and the Participating Observer*, Springer, Berlin, Heidelberg, (2007).
7. Lucian Hardy, Nonlocality for Two Particles Without Inequalities for Almost All Entangled States, *Physical Review Letters*, 71, 1665-1668 (1993)
8. Robert B. Griffiths, *Quantum Counterfactuals and Locality.* Following paper.