Studies in History and Philosophy of Modern Physics 38 (2007) 863–878

Studies in History and Philosophy of Modern Physics

www.elsevier.com/locate/shpsb

# On local realism and commutativity

Allen Stairs*, Jeffrey Bub

*Department of Philosophy, University of Maryland, College Park, MD 20742, USA*

Received 31 August 2006; received in revised form 21 February 2007; accepted 27 February 2007

**Abstract**

James D. Malley and Arthur Fine have argued that interest in noncontextual hidden variable theories is misplaced because at bottom such theories are simply attempts to turn noncommuting quantities into commuting ones. We disagree and argue that the issues are richer and more complex.


*Keywords:* Quantum foundations; Hidden variables; Commutation; Entanglement

## 1. Introduction

In various papers (Fine, 1982; Malley, 2004, 2006; Malley & Fine, 2005), James D. Malley and Arthur Fine, both separately and together, have been critical of interest in noncontextual hidden variables in general and, more particularly, of attempts to construct experimental tests of Bell-type inequalities. They have claimed (Malley & Fine, 2005, p. 52) that purely local experiments on a single system can refute what they refer to as 'local realism,' and that 'trying to make sense of nonlocal or contextual phenomena for systems where the standard Bell (or KS) conditions fail is ... the ultimately unsatisfying program of trying to denature noncommuting objects into commuting ones.' More recently, Malley (2006) has argued that for Hilbert spaces of dimension three or greater, assuming the possibility of noncontextual hidden variables forces the conclusion that the Hilbert space collapses to a single one-dimensional projector. In Malley's view (2006, p. 125), the upshot is that 'it is difficult to see any useful outcome for physics in further exploration of this form of determinism.'

*Corresponding author.
*E-mail addresses:* stairs@umd.edu (A. Stairs), jbub@umd.edu (J. Bub).

1355-2198/$ - see front matter 
doi:10.1016/j.shpsb.2007.02.003

We do not dispute any of the mathematical claims that Malley and/or Fine make, but we see the issues about commutativity and the significance of the various no-go results rather differently. In what follows, we explain why.

## 2. Review of results

Malley and Fine state their conditions somewhat differently in different papers, but they often refer to the strictures on noncontextual hidden variables by the acronym BKS for 'Bell–Kochen–Specker.' For reasons that will emerge later, we are uncomfortable with tying the ideas behind Bell's result quite so closely to the Kochen–Specker theorem, and so we will avoid the label 'BKS.' However named, Malley and Fine offer essentially the following conditions defining noncontextual hidden variable theories. First, define a probability measure $\Pr[\cdot]$ on the lattice $\mathcal{L}(\mathcal{H})$ of projectors (alternatively, subspaces) of a Hilbert space in the usual way, as what Gleason would call a frame function with weight 1:

Pr(1): For each projector $A$, $0 \leqslant \Pr[A] \leqslant 1$.
Pr(2): $\Pr[I] = 1$; $\Pr[O] = 0$, where $I$ is the identity and $O$ is the null projector.
Pr(3): If $\{A_i\}$ is a set of mutually orthogonal projectors, then $\Pr[\sum_i A_i] = \sum_i \Pr[A_i]$.

A noncontextual hidden variable theory is a scheme satisfying the following conditions:

HV(1): Each projector $A$ is associated with a subset $f(A) = a$ in a phase space $\sigma(\Lambda)$, where $\Lambda$ is a set and $\sigma$ is a $\sigma$-field of subsets of $\Lambda$. If $A \neq B$ then $a \neq b$.
HV(2): The function $f$ satisfies the conditions:
(i) $f(I) = \Lambda$ and $f(O) = \emptyset$.
(ii) If $A$ and $B$ are commuting projectors, then $f(AB) = a \cap b$.
(iii) If $\{A_i\}$ is a set of mutually orthogonal projectors, then $\{f(A_i)\} = \{a_i\}$ is a set of mutually disjoint subsets of $\Lambda$ and $f(\sum_i A_i) = a_1 \cup a_2 \cup \ldots$.
HV(3): Given a probability measure $\Pr[\cdot]$ on $\mathcal{L}(\mathcal{H})$, there is a measure $\mu$ on $\sigma(\Lambda)$ such that:
(i) For all projectors $A$, $\mu(a) = \Pr[A]$.
(ii) If $A$ and $B$ are commuting projectors, then $\mu(a \cap b) = \Pr[AB]$.
(iii) If $\{A_i\}$ is a set of mutually orthogonal projectors, then
$\mu(a_1 \cup a_2 \cup \cdots) = \mu(a_1) + \mu(a_2) + \cdots = \Pr[\sum_i A_i] = \sum_i \Pr[A_i]$.

Let us say that $\mu$ is a measure that *corresponds* to Pr. (We need not assume that $\mu$ is unique.) Although the conditions are stated in terms of projectors, they straightforwardly determine conditions on all observables.

In Malley (2006), Malley characterizes noncontextual hidden variables purely in Hilbert space terms:

KS(1): There is a valuation function $v$ on $\mathcal{L}(\mathcal{H})$ such that $v(A) = 1$ or $v(A) = 0$ for any projector $A$.
KS(2): If $\{A_i\}$ is an orthogonal set of one-dimensional projectors that spans $\mathcal{H}$, then $v$ assigns the value 1 to exactly one member of $\{A_i\}$.

KS(1) and KS(2) amount to the claim that there is a (dispersion-free) probability measure on $\mathcal{L}(\mathcal{H})$ that takes values in $\{0, 1\}$.

With these definitions and requirements, Malley and Malley–Fine prove a number of results. Before we review them, we begin with a useful lemma. Suppose we are given a

quantum system for which HV(1)–HV(3) are satisfied. Then, for any probability measure $\mu$ on the phase space $\sigma(\Lambda)$, we can explicitly define a probability measure $\Pr[\cdot]$ on $\mathcal{L}(\mathcal{H})$, where $\mu$ is a measure corresponding to $\Pr[\cdot]$.

**The Probability Lemma.** *Let $\mu$ be any probability measure on $\sigma(\Lambda)$. Then there is a probability measure* $\Pr[\cdot]$ *on $\mathcal{L}(\mathcal{H})$ satisfying* $\Pr[A] = \mu(a)$ *for all projectors $A$.*

**Proof.** Consider the subsets of $\sigma(\Lambda)$ such that for some projector $A$, $f(A) = a$ and define the function $F[\cdot]$ by $F[A] = \mu(a)$. Since $\mu$ is a probability measure, $F[\cdot]$ satisfies Pr(1). Also, since $\mu(\Lambda) = 1$ and $\mu(\emptyset) = 0$, the fact that $f(I) = \Lambda$ and $f(O) = \emptyset$ ensures that $F[\cdot]$ satisfies Pr(2). Finally, if $\{A_i\}$ is a mutually orthogonal set of projectors, then $\{a_i\}$ is a mutually disjoint set of subsets in $\sigma(\Lambda)$. Since $\mu$ is a classical probability measure, $\mu(a_1 \cup a_2 \cup \cdots) = \mu(a_1) + \mu(a_2) + \cdots$. But, by HV(2), $f(\sum_i A_i) = a_1 \cup a_2 \cup \cdots$ and so the rule $F[A] = \mu(a)$ requires $F[\sum_i A_i] = \mu(a_1 \cup a_2 \cup \cdots) = \mu(a_1) + \mu(a_2) + \cdots = F[A_1] + F[A_2] + \cdots$, as required by PR(3). Thus $F$ is a probability measure on $\mathcal{L}(\mathcal{H})$. □

Now suppose that $\mathcal{H}$ is a Hilbert space of dimension three or greater, and that $\Pr[\cdot]$ is a probability measure on $\mathcal{L}(\mathcal{H})$. By Gleason's theorem, $\Pr[\cdot]$ will be determined in the usual way by a density operator $D$ on $\mathcal{H}$. Suppose $B$ is a projector such that $\Pr[B] \neq 0$, and let $\Pr[\cdot|B]$ be defined by the density operator $BDB/\mathrm{Tr}(BDB)$ so that $\Pr[A|B] = \mathrm{Tr}(BDBA)/\mathrm{Tr}(BDB) = \mathrm{Tr}(DBAB)/\mathrm{Tr}(DB)$. (The latter form will be useful later.) This is Lüders' rule, and we will refer to $\Pr[X|B]$ as the *quantum conditional probability* of $X$ on $B$. Lüders' rule has an important uniqueness property. If $\Pr[B] = \mathrm{Tr}(DB) \neq 0$, then $\Pr[\cdot|B]$ is the unique measure on $\mathcal{L}(\mathcal{H})$ with the property that for any projector $A \leqslant B$, $\Pr[A|B] = \Pr[A]/\Pr[B]$. (see Beltrametti & Cassinelli, 1981 or Hughes, 1995) Malley (2004) proves:

**The Conditional Probability Rule.** *Suppose* $\dim \mathcal{H} \geqslant 3$, HV(1)–HV(3) *hold, and that* $\Pr[\cdot]$ *is a measure on $\mathcal{L}(\mathcal{H})$ determined by a density operator $D$. Suppose* $\Pr[B] \neq 0$. *Then* $\Pr[\cdot|B] = \mu[\cdot|b]$, *where $\mu[\cdot|b]$ is the measure on $\sigma(\Lambda)$ that results from conditionalizing on $b$.*

**Proof.** Consider a measure $\Pr[\cdot]$ on $\mathcal{L}(\mathcal{H})$ such that $\Pr[B] \neq 0$ and let $\mu$ be a corresponding measure on $\sigma(\Lambda)$. Then $\mu(b) \neq 0$, and $\mu(\cdot|b)$ is well-defined. By the Probability Lemma, the measure $\mu(\cdot|b)$ on $\sigma(\Lambda)$ determines a measure $\Pr'[\cdot]$ on $\mathcal{L}(\mathcal{H})$ via the rule $\Pr'[X] = \mu(x|b)$. If $A \leqslant B$, and hence $a \subseteq b$, we have $\mu(a|b) = \mu(a)/\mu(b)$, and hence $\Pr'[A] = \Pr[A]/\Pr[B]$. As noted above, there is only one measure on $\mathcal{L}(\mathcal{H})$ that updates $\Pr[\cdot]$ in this way: the measure $\Pr[\cdot|B]$ determined by Lüders' rule. □

In Malley (2004), Malley uses the Conditional Probability Rule to derive this result:

**The Commutation Theorem.** *If* $\dim \mathcal{H} \geqslant 3$ *and* HV(1)–HV(3) *hold, then all projectors on $\mathcal{H}$* (*and hence all observables*) *commute.*

**Proof.** Let $A$ and $B$ be any two projectors. The conditions on a classical probability measure together with the Conditional Probability Rule require that:

(i) $\mu(a \cap b) = \mu(a|b)\mu(b) = \frac{\mathrm{Tr}(DBAB)}{\mathrm{Tr}(DB)}\mathrm{Tr}(DB) = \mathrm{Tr}(DBAB)$.

(ii) $\mu(b \cap a) = \mu(b|a)\mu(a) = \frac{\mathrm{Tr}(DABA)}{\mathrm{Tr}(DA)}\mathrm{Tr}(DA) = \mathrm{Tr}(DABA)$.

This yields $\mathrm{Tr}(DBAB) = \mathrm{Tr}(DABA)$, and since this holds for all $D$, we have $BAB = ABA$. Using this identity and the idempotence of projectors, a little algebra shows that $[AB - BA] = 0$. $\square$

The Commutation Theorem is a general result. Malley and Fine (2005) prove more specialized versions: (i) if $\dim \mathcal{H} \geqslant 3$, then for any pair of noncommuting projectors $A, B \in \mathcal{H}$ there is a state $\phi$ such that $[AB - BA]\phi \neq 0$, but HV(1)–HV(3) require $[AB - BA]\phi = 0$, and (ii) for any pure state $\phi$, there is a pair of noncommuting projectors $A$ and $B$ such that the same pathology arises.

In Malley (2006), Malley reasons from KS(1) and KS(2) to prove:

**The Collapse Theorem.** *Suppose* $\dim \mathcal{H} \geqslant 3$ *and that* KS(1) *and* KS(2) *hold. Let $A$ be a ray such that $v(A) = 1$. Then if $B$ is any ray in $\mathcal{L}(\mathcal{H})$, then $B = A$.*

Although, in one sense, Malley's proof of this result is elementary, it uses an ingenious but not altogether intuitive affine construction. We prefer to derive the result in a more accessible way. Richard Friedberg proved a version of the Kochen–Specker theorem by a wonderfully simple geometric argument in 1969 (see Jammer, 1974, pp. 324–326). He showed (see Appendix A) that if $A$ and $B$ are rays separated by an angle $\arccos \frac{1}{3}$, then $v(A) = 1$ requires $v(B) = 0$. From this it follows that if $A$ and $B$ are two rays separated by an angle of $\arcsin \frac{1}{3}$, or about $19.5°$, then $v(A) = 1$ implies $v(B) = 1$. But let $A$ and $X$ be distinct rays separated by any angle whatsoever. Then we can reach $X$ from $A$ by a finite number of rotations of angle $\arcsin \frac{1}{3}$ around appropriate axes. If the chain is $A \to A' \to A'' \to \cdots \to X$, then $v(A) = 1$ implies $v(A') = 1$, etc., until we arrive at $v(X) = 1$. Thus, for any ray $A$, $v(A) = 1$ implies that $v(X) = 1$ for all rays $X$.

Call this the $1 \to 1$ Lemma. Malley's proof invokes a different result, which we will call the $1 \to 0$ Lemma: if $A$ and $B$ are any two rays and $v(A) = 1$, then $v(B) = 0$. To see this, let $C$ be a third ray such that the angle between $C$ and $B$ is $\arccos \frac{1}{3}$. Then $v(A) = 1$ implies $v(C) = 1$ by the $1 \to 1$ Lemma. But the fact about rays separated by $\arccos \frac{1}{3}$ now implies that $v(B) = 0$. [Readers inclined to point out that we also have $v(B) = 1$ are quite right, of course. But bear with us.]

With that in mind, here is a proof that follows the overall logic of Malley's reasoning. Suppose $A$ is a ray such that $v(A) = 1$. Let $B$ be an arbitrary ray in $\mathcal{H}$. Then it is impossible that $B$ is skew to $A$—neither orthogonal to nor identical with $A$. For suppose it is possible. Then let $B'$ be in the same plane as $A$ and $B$ and orthogonal to $B$; in lattice terms, $A \leqslant (B \vee B')$. Since $v(A) = 1$, KS(1) and KS(2) imply that either $v(B) = 1$ or $v(B') = 1$. (Proof an elementary exercise for the reader.) But the $1 \to 0$ Lemma requires $v(B) = 0$ and $v(B') = 0$, which gives us a contradiction. Suppose, then, that $B \perp A$. Then let $C$ be a ray distinct from $A$ and $B$ and in their span and repeat the preceding argument to arrive at another contradiction. This leaves one possibility, namely $B = A$, and hence the whole Hilbert space collapses to a single ray.

Taken at face value, this result adds weight to the claim that noncommutation is what underlies the no-go results. A one-dimensional Hilbert space is the only nonzero Hilbert space on which all operators commute.

## 3. Preliminary discussion of results

While all of these results are mathematically correct, just what they mean is less clear. Everyone agrees to this: if $\dim \mathcal{H} \geqslant 3$, then HV(1)–HV(3) lead to contradiction, as do

KS(1) and KS(2). If we begin with contradictory premises we can, of course, derive a wide array of conclusions. The problem is to sort out which, if any, are especially revelatory.

The Commutation Theorem and the related results of Malley and Fine (2005) are meant to 'fully identify the logical engine driving the no-go theorems with the most basic, non-classical feature of the quantum theory: that not all observables commute' (p. 52) and, as noted, to expose the hidden variable program as an attempt to turn noncommuting objects into commuting ones. A skeptic about this interpretation might suggest that we could equally well look at the matter through the lens of this result:

**The Noncommutation Theorem.** *Assume* $\dim \mathcal{H} \geqslant 3$, *let* $A$ *and* $B$ *be one-dimensional projectors and let* $D$ *determine a probability measure* $\Pr[\cdot]$ *such that* $\Pr[A] \neq \Pr[B]$, $\Pr[A] \neq 0$, $\Pr[B] \neq 0$. *Then, if* HV(1)—HV(3) *hold,* $\mu(a \cap b) \neq \mu(b \cap a)$.

**Proof.** From the conditions on a classical probability measure and the Conditional Probability Rule, we have:

$$\begin{aligned} \mu(a \cap b) &= \mu(a|b)\mu(b) \\ &= \Pr[A|B]\Pr[B], \\ \mu(b \cap a) &= \mu(b|a)\mu(a) \\ &= \Pr[B|A]\Pr[A]. \end{aligned}$$

Since $A$ and $B$ are both one-dimensional and since neither has zero probability, $\Pr[A|B]$ and $\Pr[B|A]$ depend only on the angle between $A$ and $B$, and $\Pr[A|B] = \Pr[B|A]$. But, by assumption, $\Pr[A] \neq \Pr[B]$, and so $\Pr[A|B]\Pr[B] \neq \Pr[B|A]\Pr[A]$. It follows, therefore, that $\mu(a \cap b) \neq \mu(b \cap a)$. $\square$

Thus, our skeptic might insist, we could equally claim that the hidden variable program is the unsatisfying attempt to turn commuting objects (intersections of sets or products of characteristic functions) into noncommuting ones.

The point is not quite as silly as it may seem. The requirements HV(1) and HV(2) amount to requiring that $\sigma(\Lambda)$ reproduce the orthogonality and inclusion structure of $\mathcal{L}(\mathcal{H})$. In fact, HV(1) and HV(2) require that $\sigma(\Lambda)$ contains an isomorphic copy of the partial Boolean algebra of projectors $\mathcal{B}(\mathcal{H})$. (See Kochen & Specker, 1967 for relevant definitions.) Looked at in this way, the program of 'denaturing' quantum objects might appear to carry with it a commitment to leaving them significantly untamed.

Malley could urge that his more recent result tips the scale by showing that the hidden variable theorist must reject the quantum mechanical structure in a very drastic way: by reducing it to a single dimension. However, the skeptic will point to an interesting feature of the logic of Malley's proof. We begin by assuming that the ray $A$ satisfies $v(A) = 1$ and then we let $B$ be any ray. We show that $B$ cannot be skew to $A$ on pain of contradiction. The only way, then, for $B$ to be distinct from $A$ is for $B$ to be orthogonal to $A$. But then Malley's proof asks us to assume the existence of a third ray $C$, skew to $A$. This is curious because the preceding portion of the proof has shown that there can be no such rays. Thus, the skeptic might say: why not take the following to be the 'real' theorem?

**The Superselection Theorem.** *Suppose* $\dim \mathcal{H} \geqslant 3$ *and let* $A$ *be a ray such that* $v(A) = 1$. *Then every other ray in* $\mathcal{H}$ *is orthogonal to* $A$.

**Proof.** Let $v(A) = 1$ and let $B$ be an arbitrary ray distinct from $A$. Then $B$ cannot be skew to $A$ (see the proof of the Collapse Theorem) and, therefore, $H$ contains no rays skew to $A$. □

Of course, it is no more possible that $H$ reduces to this 'superselected space' than that it reduces to a single ray. If $H$ is a Hilbert space of dimension three or greater, then any linear combination of any two vectors is also a vector in space. However, the Superselection Theorem at least has this in its favor: there is a valuation function $v$ on this space, even though the algebra of self-adjoint operators contains noncommuting pairs.

There is another way to get Malley's result. By KS, some ray $A$ must satisfy $v(A) = 1$. Let $B$ be an arbitrary ray. If we suppose $B \neq A$, then by the $1 \to 1$ Lemma, $v(B) = 1$, and by the $1 \to 0$ Lemma, $v(B) = 0$, producing a contradiction. But instead of concluding that $\mathcal{H}$ has been reduced to a single dimension, we might equally conclude that it is impossible for any ray to be assigned the value 1 consistently with the conditions KS. Instead of taking the existence of a ray $A$ such that $v(A) = 1$ as a given, and then using the assumptions of a $B$ either skew or orthogonal to $A$ as assumptions for *reductio*, we could treat $v(A)$ itself as the source of the *reductio*. We could go on to prove that $v(A) = 0$ also leads to a contradiction. The proof would be simple. Suppose that $v(A) = 0$ and let $\{B_j\}$ be a set of projectors orthogonal to $A$ and to each other such that $\{A\} \cup \{B_j\}$ spans $\mathcal{H}$. Then by KS(2), one of the $B_j$ must be assigned 1 by $v$, and we arrive at a contradiction by the same route as before. A version of this proof is clearly available to Malley. We could conclude that no rays can exist at all, since each ray $R$ must satisfy either but not both of $v(R) = 1$ or $v(R) = 0$, and each assumption leads to a contradiction. Thus we would have the truly drastic conclusion that the space collapse to the null space!

## 4. Commutation, locality and entanglement

All this suggests that the approach adopted by Malley and Fine to identify the 'logical engine' behind no-go proofs is problematic. Starting with contradictions and going on to derive untoward results is, to paraphrase Fine (1996) paraphrasing Einstein, a shaky game. However, there are some interesting considerations that might seem initially to support Malley and Fine's instincts and might also seem to provide support for their claim that there are 'elementary tests for local realism, using single particles and without reference to entanglement, thus avoiding experimental loopholes and efficiency issues that continue to bedevil the Bell inequality related tests' (Malley & Fine, 2005, p. 51).

We begin with the results of Clifton, Bub, and Halvorson (2003) and Halvorson's related result (2004). We will refer to these results/authors collectively as CBH. CBH show that a handful of information-theoretic constraints are necessary and sufficient for the main structural features of quantum theories: specifically, (i) that the algebras of observables of distinct physical systems commute, (ii) that any individual system's algebra of observables is noncommutative, and (iii) that the physical world is nonlocal, in the sense that space-like separated systems can occupy entangled states that persist as the systems separate. These constraints, which can be precisely formulated in the general framework of $C^*$-algebras, are:

CBH(1): the impossibility of superluminal information transfer between two physical systems by performing measurements on one of them,

CBH(2): the impossibility of perfectly broadcasting the information contained in an unknown physical state (which, for pure states, amounts to 'no cloning'),

CBH(3): the impossibility of communicating information so as to implement a bit commitment protocol with unconditional security (so that cheating is in principle excluded by the theory).

CBH(1) corresponds to a sort of 'separability' constraint. CBH characterize two systems, $A$ and $B$, as *physically distinct* if $A$-states impose no constraints on $B$-states, and conversely; more precisely, $A$ and $B$ are physically distinct systems just in case, for any state $s_A$ and $s_B$ of systems $A$ and $B$ respectively, there is a state $s$ of the joint system such that $s_A$ is the reduced state (marginal state) for system $A$ and $s_B$ is the reduced state for system $B$. It is easy to see that if any $A$-observable commutes with any $B$-observable, then the state transformation induced by a local measurement operation performed on system $A$ cannot convey any information to a physically distinct system $B$, i.e., the reduced state of $B$, and hence the expectation values of all $B$-observables, will be unchanged (and similarly, of course, local measurements on $B$ cannot convey information to $A$). CBH(1) entails the converse: if $A$ and $B$ are physically distinct, then it follows from CBH(1) that any $A$-observable must commute with any $B$-observable.

In a cloning process, a ready state $\sigma$ of a system and the state to be cloned $\rho$ of a second system are transformed into two copies of $\rho$. In a more general broadcasting process, a ready state $\sigma$ and the state to be broadcast $\rho$ are transformed to a new state $\omega$ of the composite system, where the marginal state $\omega$ with respect to both component systems is $\rho$. CBH(2) entails and is entailed by the failure of commutativity. Specifically, in the commutative case there is a universal broadcasting map that clones any pair of input pure states and broadcasts any pair of input mixed states. Conversely, if any two states can be perfectly broadcast, then any two pure states can be cloned; and if two pure states of a $C^*$-algebra can be cloned, then they must be orthogonal. So, if any two states can be broadcast, then all pure states are orthogonal, from which it follows that the algebra is commutative. Thus, the CBH conditions put noncommutativity at the core of quantum theories, which would seem to be good news for the position that Malley and Fine want to defend.

CBH(3) is where the connection with nonlocality emerges. Bit commitment is a primitive cryptographic protocol. In a bit commitment protocol, one party, Alice, supplies an encrypted bit, 0 or 1, to a second party, Bob. The information available in the encrypted bit should be insufficient for Bob to ascertain the value of the bit, but sufficient, together with further information supplied by Alice at a subsequent 'opening' stage when she is supposed to reveal the value of the bit, for Bob to be convinced that the protocol does not allow Alice to cheat by encrypting the bit in a way that leaves her free to reveal either 0 or 1 at will.

Now, Alice can send encrypted classical information to Bob that guarantees the truth of an exclusive classical disjunction (equivalent to her commitment to a 0 or a 1) only if the information is biased towards one of the alternative disjuncts (because a classical exclusive disjunction is true if and only if one of the disjuncts is true and the other false). No principle of classical physics precludes Bob from extracting this information, so the security of a classical bit commitment protocol can only be a matter of computational complexity. The hope was that there exists a quantum analogue of this procedure that is *unconditionally secure*: provably secure as a matter of physical law (according to quantum theory) against cheating by either Alice or Bob.

The idea, first proposed by Bennett and Brassard (1984), was to associate the 0 and 1 commitments with two different mixtures represented by the same density operator. However, as the authors showed, Alice can cheat by exploiting entanglement and adopting an 'EPR attack' or cheating strategy: she prepares entangled pairs of qubits, keeps one of each pair (the ancilla) and sends the second qubit (the channel particle) to Bob. In this way she can fake sending one of two equivalent mixtures to Bob and reveal either bit at will at the opening stage by effectively steering Bob's particle into the desired mixture by an appropriate measurement. Bob cannot detect this cheating strategy.

It turns out that the existence of entangled states, and the persistence of entanglement as the systems held by Alice and Bob become space-like separated, ensures that this sort of attack is perfectly general, so unconditionally secure quantum bit commitment is impossible (see Lo & Chau, 1997, 1998; Mayers, 1996, 1997). It follows that for any nonclassical theory where the algebras of observables of distinct physical systems commute, and any individual system's algebra of observables is noncommutative, requiring in addition the impossibility of secure bit commitment ensures the existence of entangled states that persist as systems become space-like separated.

However, from a certain point of view, entangled states are an inevitable consequence of noncommutativity. Begin with Hilbert space itself. The tensor product is the natural construction for embodying the sort of separability between two systems required by CBH(1). But it is a familiar fact that so long as $\mathcal{H}_1$ and $\mathcal{H}_2$ are both at least two-dimensional—so long, therefore, as the Hilbert spaces of the two systems include noncommuting operators—the tensor product will contain vectors that cannot be reduced to the form $\phi \otimes \psi$. In other words, it will contain entangled states. An analogous result holds in the $C^*$-algebraic framework that CBH adopt. If the algebras associated with systems 1 and 2 satisfy CBH(2) and thus are noncommutative, then requiring CBH(1) for the algebra that represents the pair will necessarily introduce states that cannot be expressed as products of system-1 and system-2 states. See Summers (1990) and Bacciagaluppi (1994).

There is even a lattice-theoretic analogue to this point. Sticking with the more straightforward case of finite dimensions, if $\mathcal{L}(\mathcal{H}_1)$ and $\mathcal{L}(\mathcal{H}_2)$ are lattices of subspaces of Hilbert spaces of dimension two or greater, then there will be lattice elements $x$ and $y$ that fail to 'commute'—i.e., for which the identity $x = (x \wedge y) \vee (x \wedge y^{\perp})$ fails. (These will correspond exactly to projectors that do not commute.) We can impose 'separability' conditions on any lattice suitable for representing the pair of systems. Roughly, these conditions will require that the product lattice be generated by independent isomorphic images of the factor lattices, and that if $\mathcal{B}_1$ and $\mathcal{B}_2$ are maximal Boolean sublattices of the factor lattices, then the product lattice will contain a maximal Boolean sublattice isomorphic the product of these two Boolean sublattices. (see Stairs, 1983 for details) It can be shown that products exist, that the lattice $\mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ fits the definition, and more important, that any lattice fitting the separability requirements will contain atoms (minimal nonzero elements) that are not of the form $A_1 \wedge B_2$. Thus, even when we consider the 'logical' or lattice-theoretic analogue of noncommutation, we find that straightforward and minimal conditions on the representation of pairs carry entanglement along in their wake.

Looking at things this way may appear to vindicate Malley and Fine completely. Noncommutation is the *sine qua non* of quantum theories, and when we add a plausible separability or locality condition, the mathematics guarantees the existence of entangled

states. Thus, it might seem, if we can confirm noncommutation by local experiments, we already have everything we need to rule out local hidden variables.

## 5. Hidden variables

In fact, all of the above results notwithstanding, we do not see things this way. Perhaps a good place to begin unraveling our differences with Fine and Malley is by quoting the abstract of Malley's 2004 paper:

> Under a standard set of assumptions for a hidden-variable model for quantum events we show that all observables must commute simultaneously ... And, despite Bell's complaint that a key condition of von Neumann's was quite unrealistic, we show that these conditions, under which von Neumann produced the first no-go proof, are entirely equivalent to those introduced by Bell and Kochen and Specker. As these conditions are also equivalent to those under which the Bell–Clauser–Horne inequalities are derived, we see that the experimental violation of the inequalities demonstrates only that quantum observables do not commute.

If Malley's various results really do show that the KS and Bell conditions amount to imposing commutativity on the Hilbert space, then this would make sense. Von Neumann's conditions would hold only if all quantum observables commuted, but so would Kochen and Specker's and so would Bell's. However, this is not quite right.

Let us begin with the idea that von Neumann's conditions are 'entirely equivalent' to Bell's and to Kochen and Specker's. von Neumann's requirement was that if $Q$ and $R$ are two self-adjoint operators, then any 'hidden' dispersion-free states must respect the relation

$$\mathrm{Exp}(Q+R) = \mathrm{Exp}(Q) + \mathrm{Exp}(R).$$

As is well-known, this rules out hidden variables even in the case of the spin observables of a spin-$\frac{1}{2}$ system, but as is also well-known, both Bell and Kochen and Specker provided explicit examples of hidden variables for this case—hidden variables that satisfy their own conditions. Since there is a case in which von Neumann's requirements rule out hidden variables while Bell's and Kochen and Specker's do not, it can hardly be correct to say that all three sets of conditions are 'entirely equivalent' and Malley nowhere attempts to construct such a proof.[1]

More important for what follows, we see no reason to read Bell's conditions as amounting to Kochen and Specker's. The conclusion that this is so comes from treating the Bell inequalities or the CHSH inequality simply as abstract constraints on joint distributions for sets of bivalent observables. While there is some interest in reading the inequalities that way (more later on this point), it would not do as a reading of Bell's concerns. Bell's conditions are intended to constrain joint probabilities for quantities $Q_i$ and $R_j$ when $Q_i$ and $R_j$ belong to distinct, *space-like separated* systems. This seems

[1]Such a proof would require showing that Bell's conditions imply the Kochen–Specker conditions, that the Kochen–Specker conditions imply von Neumann's conditions, and that von Neumann's conditions imply Bell's conditions. On the face of it, the two-dimensional case shows that these conditions cannot be equivalent. In their joint paper, Malley and Fine appeal to a result of Busch (2003) to maintain that their results extend to the two-dimensional case. But as one of us (Stairs, 2007) argues, Busch's result, though mathematically correct, cannot be used to provide a proof of the impossibility of noncontextual hidden variables in two dimensions.

particularly clear from Bell's (1964) 'local beables' paper, filled as it is with references to space-like separated systems and variables in the shared past light cone. We take his condition to say that *in this case*, quantum probabilities must be reproducible as averages over factorizable phase space probabilities. That is, for a given 'hidden state' $\lambda$, the probability $\mu_\lambda((Q_1 = q_1)\&(R_2 = r_2)) = \mu_\lambda(Q_1 = q_1)\mu_\lambda(R_2 = r_2)$. But this means that there are examples where Bell's conditions, understood in Bell's way, allow hidden variables and Kochen and Specker's do not.

As an example, a single spin-1 system will do. There is nothing in Bell's conditions that precludes a contextual hidden variable theory for such a system. Bell's concern is with nonlocal contextualism. He expresses the stricture in robust physical language, as in this passage (Bell, 1964, pp. 64–65):

> ... now we add the hypothesis of *locality*, that the setting $\hat{b}$ of a particular instrument has no effect on what happens, $A$, in a remote region, and likewise that $\hat{a}$ has no effect on $B$.

This can be mirrored in the style appropriate for HV(1)–HV(3) by requiring that each *locally maximal* observable be represented by a single random variable.

The distinction at issue here never comes up directly in Bell's writings, for the simple reason that in two-dimensional Hilbert space there are no nontrivial locally degenerate observables. That makes the qualification 'locally maximal' otiose. But, however we settle the exegetical issue about Bell, the distinction between completely noncontextual theories and theories that are at most locally contextual is clearly a coherent one.

On our view, the strictures of von Neumann, of Kochen and Specker, and of Bell form a series with decreasing strength. von Neumann rules out more varieties of hidden variables than Kochen and Specker, who in turn rule out more kinds than Bell. The three sets of conditions are certainly not equivalent, and their motivations are not identical. Now of course, it is possible to find sets of observables and states that scuttle all three kinds of hidden variables at once. Put in a way that is more useful for present purposes, there are sets of states and observables such that each set of hidden variable requirements leads to a contradiction. We might say that the conditions are equivalent modulo such a set, but they are not equivalent *simpliciter*.

## 6. Entanglement

How does all this square with what was said earlier about the centrality of noncommutativity and the fact that noncommutative structures entail the existence of entangled states? A pair of papers by Schrödinger offers a good place to begin exploring that question. In Schrödinger (1935, p. 555) wrote:

> When two systems, of which we know the states ... enter into temporary physical interaction ... and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before, viz. by endowing each of them with a representative of its own. I would not call that *one* but rather *the* characteristic trait of quantum mechanics, the one that forces its entire departure from classical lines of thought. By the interaction the two representatives (or $\psi$-functions) have become entangled.

Schrödinger goes on in his 1936 paper to describe in detail what he sees as the 'sinister' consequences of this fact: that if true, it allows for an experimenter in one region to 'steer' the remote system, with nonvanishing probability, into any state in the support of the density operator describing the reduced state of the remote system. In particular, he explores the consequences of the fact that a mixture of states on one system may be decomposed into infinitely many distinct linearly independent sets of states. This creates a striking possibility: a measurement by the experimenter 'here' will select a state for the system 'there.' The experimenter 'here' can entangle the local system with an ancilla system by a suitable unitary transformation and perform a measurement in a particular basis on the local composite system to choose which set of linearly independent states will contain the state that the measurement will select for the distant system. Similar results were later proved independently by Hughston, Jozsa, and Wootters (1993), Jaynes (1957), and Gisin (1989). (See Halvorson, 2004 for a generalization to hyperfinite von Neumann algebras.)

Schrödinger found this disconcerting. He wrote (Schrödinger, 1936, p. 451):

> Indubitably, the situation here described is, in present quantum mechanics, a necessary and indispensable feature. The question arises, whether it is so in nature too. I am not convinced that there is sufficient experimental evidence for that.

Schrödinger speculated that when two entangled systems separate, the entanglement itself decays, leaving the pair in a mixture that still embodies some correlations, but not of the 'sinister' sort. They would be correlations that would only show up in a preferred basis, and they would not allow for 'remote steering.'

Whether Schrödinger's physical speculation was correct or not, we think his intuitions about the conceptual situation were sound. Entanglement of distant systems is a strikingly nonclassical phenomenon with significant consequences. Whether or not it exists is a genuine experimental question to which Nature could answer 'No.' We know now that quantum systems can exist in entangled states, but if the answer had turned out to be no, this would not be the same as rejecting quantum theory wholesale. The point is very simple. On the one hand, it is a mathematical fact that noncommutative structures entail the mathematical possibility of entangled states. That is what the familiar facts about tensor products as well as the less familiar facts about $C^*$-algebras and orthocomplemented modular lattices tell us. But the *mathematical* possibility of a certain kind of state is not the same as its *physical* possibility. It is at least an abstract possibility that even though the state space of compound systems includes entangled states, those states never arise in nature. And it is a much less abstract possibility that even though entangled states exist in nature, they decay as soon as the systems that possess them achieve a certain minimal degree of separation. Our confident expectation is that this will not turn out to be true. But fidelity to the experimental facts requires us to admit that no experiment to date has shown otherwise in a way that avoids the issues about loopholes and inefficiencies that Malley and Fine think their approach can avoid. It is true as a matter of mathematics: any reasonable way of combining the state spaces of two spin-$\frac{1}{2}$ particles yields a space that contains entangled states. But it is not a mere matter of mathematics to claim that space-like separated systems can yield statistics governed by such states that persist as the systems separate.

## 7. Inequalities

As we noted above, there is more than one way to understand inequalities such as the CHSH/Bell inequalities. As we take Bell to have understood them, they tell us something about how space-like separated systems would behave if their 'deepest' states are factorizable. But understood in this way, *quantum systems* could be bound by the CHSH inequality. If space-like separated *quantum* systems only admit states of the form $\phi \otimes \psi$ and mixtures thereof, then the statistics of those systems will be governed by CHSH. Of course, quantum systems might, as a matter of empirical fact, be bound by CHSH understood as a constraint on space-like separated systems and exhibit purely local correlations that violate CHSH. This will be true, for example, if the full range of states is open to a pair of spin-$\frac{1}{2}$ systems that are not well-separated—whatever may be the case as the systems move apart. But as a sheer mathematical possibility, if the only actual quantum states have high enough entropy, then quantum systems, in spite of possessing noncommuting algebras of observables, would obey CHSH. It is simply that if this were so, it would be a fact that did not flow from the intrinsic structure of the systems themselves.

Understood in a more general way, then, CHSH and related inequalities tell us something about the strength of correlations that certain sorts of systems permit 'intrinsically.' For systems representable on a phase space, CHSH represents an upper bound on the strength of correlations compatible with this sort of property structure, though of course, a sufficiently disordered and random classical world might never reach the CHSH bound. The sorts of property structures embodied by quantum systems have a higher intrinsic limit—permit stronger correlations consistently with the property structure. And in fact we know that for quantum correlations, the Cirel'son bound (Cirel'son, 1980) is the maximum bound consistent with the property structure. But whether and in what conditions nature allows that bound to be approached is an empirical matter.

We also know that there are interesting possible systems, embodied by the Popescu–Rohrlich 'nonlocal box' models (Barrett & Pironio, 2005; Barrett et al., 2005; Popescu & Rohrlich, 1994), that exceed the Cirel'son bound. As far as we know, nature does not permit any physical systems to display this sort of behavior, but this is an empirical claim and could turn out to be mistaken.

Clearly if we look at these issues through the lens of quantum information theory, a rich set of questions opens up. We can ask what sorts of theories permit what sorts of behavior in principle (for example: do nonlocal box theories permit secure bit commitment?) And with respect to systems that we believe to be representable in certain sorts of structures (for example, quantum systems and $C^*$-algebras or Hilbert spaces), we can ask how close nature actually lets us come to the theoretical bounds imposed by relations such as Cirel'son's inequality. We miss all of this if we reduce everything to questions of commutativity.

## 8. The quantum world

However interesting nonlocal boxes or classical probability may be, we seem to live in a quantum mechanical world. Even taking that as a given, however, it should be clear that

we are not prepared to see no-go results as ineffectual wrestling matches with noncommutativity.

One reason is that even though it may now seem trivial that quantum statistics cannot be captured by representing observables as random variables on a phase space, this was not obvious at all a few decades ago. It is true: we can rule out Kochen–Specker style hidden variables by means of arguments that a high school student can follow; a bit of reasoning about a small handful of vectors will do. Much the same goes for attempts to understand entangled states by means of local hidden variables. (As explained above, we do not think that the issues here are simply the Kochen and Specker questions re-warmed.) There are utterly elementary proofs that local hidden variables will not work, at least for some simple, idealized cases. But we do not think it follows from the fact that these proofs are elementary that what we can learn from them is trivial. Simple though the proofs may be, the idea that nature might admit structures of properties that cannot be embedded in a Boolean algebra is surprising and, we think, profound. And the possibility that separated systems might admit correlations that cannot be captured by a classical model is central to widespread interest in quantum information, and underlies such unexpected phenomena as quantum teleportation.

And this brings us to our second and final point. The question of whether the world might admit unconditionally secure bit commitment strikes us as a very interesting one. Even though Schrödinger did not think of it in those terms, it clearly interested him. The sort of 'remote steering' that disturbed him so much is exactly what would thwart secure bit commitment if it really is possible. Remote steering depends on entanglement, as Schrödinger made abundantly clear. But even though the possibility of the required entangled states is a mathematical consequence of CBH(1) and CBH(2), it is not a *physical* consequence of those principles. That is why, in spite of Malley and Fine's skepticism, we think that there is an important issue behind attempts to find loophole-free evidence that Bell-type inequalities are actually violated in nature. If they are not, then we may live in a world where commutation fails but locality does not. That would be interesting to know if true, but it is not the sort of thing we could find out by experiments performed on a single qubit.

**Acknowledgments**

Allen Stairs would like to thank the General Research Board of the University of Maryland for supporting this research. Jeffrey Bub acknowledges support for research leading to this paper from the National Science Foundation under Grant No. 0522398, and from the Perimeter Institute for Theoretical Physics in Waterloo, Canada, during a stay as a long-term visiting researcher in 2006.

## Appendix A. On the $1 \to 1$ and $1 \to 0$ Lemmas

The $1 \to 1$ and $1 \to 0$ Lemmas depend on two simple geometrical results. In what follows, let $r, s, t \ldots$ denote unit vectors, and let $R, S, T \ldots$ denote the corresponding rays. Suppose $v$ satisfies KS. We can show that:

(a) If $v(R) = 1$, and $R$ and $S$ are separated by an angle $\arccos(\frac{1}{3})$, then $v(S) = 0$.
(b) If $v(R) = 1$, and $R$ and $T$ are separated by an angle $\arcsin(\frac{1}{3})$, then $v(T) = 1$.

The proof goes more smoothly if we restate it in terms of inner products. Suppose $|(r,s)| = \frac{1}{3}$ and $|(r,t)| = \sqrt{8}/3$. By a slight modification of Friedburg's argument in Jammer (1974), noted earlier, we can show that (a) if $v(R) = 1$ then $v(S) = 0$, and (b) if $v(R) = 1$ then $v(T) = 1$. The reasoning proceeds in $\mathcal{H}^3$, but it should be clear that it holds in higher dimensions as well.

Consider the following three orthonormal bases in $\mathcal{H}^3$:

$$x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$
$$a = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \quad b = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{-1}{\sqrt{2}} \end{pmatrix}, \quad y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$
$$c = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad d = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{1}$$

Now introduce two unit vectors:

$$f = \begin{pmatrix} \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \quad g = \begin{pmatrix} \frac{e^{i\theta}}{\sqrt{3}} \\ \frac{e^{i\theta}}{\sqrt{3}} \\ \frac{e^{i\theta}}{\sqrt{3}} \end{pmatrix}, \tag{2}$$

where $\theta$ is an arbitrary real number. Notice that the inner product $(f,g) = e^{i\theta}/3$, and so $|(f,g)| = \frac{1}{3}$. It is easy to see that $F \perp A$ and $F \perp C$, and hence that if $v(F) = 1$, $v(A) = 0 = v(C)$. Likewise, it is easily checked that $G \perp B$ and $G \perp D$, and hence that if $v(G) = 1$, $v(B) = 0 = v(D)$. This forces $v(Y) = 1 = v(Z)$, which violates KS. Therefore, if $v(F) = 1$, $v(G) = 0$.

Although we chose particular coordinates, the result depends only on the norm of the inner product between $f$ and $g$. To see why, let $r$ and $g$ be any two unit vectors in $\mathcal{H}^3$ such that $(r,g) = e^{i\theta}/3$. Then there is a unitary map $U$ such that $U(f) = r$, $U(g) = s$, and the argument can be repeated for $\{U(x), U(y), U(z)\}$, $\{U(a), U(b), U(y)\}$, $\{U(c), U(d), U(z)\}$, $r$ and $s$. Thus (a) holds for any two unit vectors $r, s$ such that $|(r,s)| = \frac{1}{3}$.

To prove (b), return to our original two vectors $f$ and $g$. Consider the unit vector

$$h = \begin{pmatrix} \frac{-2e^{i\theta}}{\sqrt{6}} \\ \frac{e^{i\theta}}{\sqrt{6}} \\ \frac{e^{i\theta}}{\sqrt{6}} \end{pmatrix}. \tag{3}$$

It is easy to check that $g \perp h$ and that $(f, h) = e^{i\theta}\sqrt{8}/3$, hence $|(f, h)| = \sqrt{8}/3$. Suppose $v(F) = 1$. Since $f$ is in the $g$–$h$ plane, KS requires that exactly one of $v(G)$ or $v(H)$ equals 1. To see this, note that $G$ and $H$ can be extended to a complete frame $\{G, H, J\}$. Since $F$ is orthogonal to $J$, $v(F) = 1$ requires $v(J) = 0$, and hence, by KS(2), either $v(G) = 1$ or $v(H) = 1$ but not both. But as we saw above, if $v(F) = 1$ then $v(G) = 0$. It therefore follows that $v(H) = 1$. By invoking the same unitary map $U$ that we appealed to above, it follows that if $r$ and $t$ are any two unit vectors in $\mathcal{H}_3$ such that $|(r, t)| = \sqrt{8}/3$, then $v(R) = 1$ requires that $v(T) = 1$, which proves (b). This provides us with the two pieces we need for the $1 \to 1$ and $1 \to 0$ Lemmas, whose proofs now proceed as in the body of the paper.

## References


Bacciagaluppi, G. (1994). Separation theorems and Bell inequalities in algebraic quantum mechanics. In P. Busch, P. Lahti, & P. Mittelstaedt (Eds.), *Symposium on the foundations of modern physics 1993: Quantum measurement, irreversibility and the physics of information* (pp. 29–37). Singapore: World Scientific.

Barrett, J., Linden, N., Massar, S., Pironio, S., Popescu, S., & Roberts, D. (2005). Nonlocal correlations as an information-theoretic resource. *Physical Review A*, *71*, 022101–022111.

Barrett, J., & Pironio, S. (2005). Popescu–Rohrlich correlations as a unit of nonlocality. *Physical Review Letters*, *95*, 2019–2039.

Bell, J. S. (1964). The theory of local beables. In *Speakable and unspeakable in quantum mechanics* (pp. 52–62). Cambridge: Cambridge University Press (1987).

Beltrametti, E. G., & Cassinelli, G. (1981). *The logic of quantum mechanics*. Reading, MA: Addison-Wesley.

Bennett, C. H., & Brassard, G. (1984). Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE international conference on computers, systems, and signal processing* (pp. 175–179). New York: IEEE.

Busch, P. (2003). Quantum states and generalized observables: A simple proof of Gleason's theorem. *Physical Review Letters*, *91*, 120403.

Cirel'son, B. S. (1980). Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, *4*, 379–385.

Clifton, R., Bub, J., & Halvorson, H. (2003). Characterizing quantum theory in terms of information-theoretic constraints. *Foundations of Physics*, *33*, 1561–1591.

Fine, A. (1982). Joint distributions, quantum correlations, and commuting observables. *Journal of Mathematical Physics*, *23*, 1306–1310.

Fine, A. (1996). *The shaky game*. Chicago: University of Chicago Press.

Gisin, N. (1989). Stochastic quantum dynamics and relativity. *Helvetica Physica Acta*, *62*, 363–371.

Halvorson, H. (2004). Remote preparation of arbitrary ensembles and quantum bit commitment. *Journal of Mathematical Physics*, *45*, 4920–4931.

Hughes, R. I. G. (1995). *The structure and interpretation of quantum mechanics*. Cambridge: Cambridge University Press.

Hughston, L. P., Jozsa, R., & Wootters, W. K. (1993). A complete classification of quantum ensembles having a given density matrix. *Physics Letters A*, *183*, 14–18.

Jammer, M. (1974). *The philosophy of quantum mechanics: The interpretation of quantum mechanics in historical perspective*. New York: Wiley Interscience.

Jaynes, E. T. (1957). Information theory and statistical mechanics ii. *Physical Review*, *108*, 171–190.

Kochen, S., & Specker, E. P. (1967). On the problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, *17*, 59–87.

Lo, H.-K., & Chau, H. F. (1997). Is quantum bit commitment really possible? *Physical Review Letters*, *78*, 3410–3413.

Lo, H.-K., & Chau, H. F. (1998). Why quantum bit commitment and ideal coin tossing are impossible. *Physica D*, *120*, 177–187.

Malley, J. D. (2004). All quantum observables in a hidden-variables model must commute simultaneously. *Physical Review A*, *69*, 022118–022120.

Malley, J. D. (2006). The collapse of Bell determinism. *Physics Letters A*, *359*, 122–125.

Malley, J. D., & Fine, A. (2005). Noncommuting observables and local realism. *Physics Letters A*, *347*, 51–55.

Mayers, D. (1996). Unconditionally secure quantum bit commitment is impossible. In *Proceedings of the fourth workshop on physics and computation* (pp. 224–228). Boston: New England Complex System Institute.

Mayers, D. (1997). Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, *78*, 3414–3417.

Popescu, S., & Rohrlich, D. (1994). Quantum nonlocality as an axiom. *Foundations of Physics*, *24*, 379–385.

Schrödinger, E. (1935). Discussion of probability relations between separated systems. *Proceedings of the Cambridge Philosophical Society*, *31*, 555–563.

Schrödinger, E. (1936). Probability relations between separated systems. *Proceedings of the Cambridge Philosophical Society*, *32*, 446–452.

Stairs, A. (1983). Quantum logic, realism, and value-definiteness. *Philosophy of Science*, *50*, 578–602.

Stairs, A. (1983). POVMs and hidden variables. *Physics Letters A*, *365*, 268–272.

Summers, S. (1990). On the independence of local algebras in quantum field theory. *Reviews in Mathematical Physics*, *2*, 201–247.