A note on information theoretic characterizations of physical theories

Hans Halvorson[*]
Department of Philosophy, Princeton University

December 15, 2003

Abstract

Clifton, Bub, and Halvorson [*Foundations of Physics* 33, 1561–1591, (2003)] have recently argued that quantum theory is characterized by its satisfaction of three information-theoretic axioms. However, it is not difficult to construct apparent counterexamples to the CBH characterization theorem. In this paper, we discuss the limits of the characterization theorem, and we provide some technical tools for checking whether a theory (specified in terms of the convex structure of its state space) falls within these limits.

1 Introduction

Some would like to argue that quantum information theory has revolutionary implications for the philosophical foundations of QM (see, e.g., Bub, 2004; Fuchs, 2003). Whether or not this claim is true, there is no doubt that quantum information theory presents us with new perspectives from which we can approach traditional questions about the interpretation of QM. One such question asks whether there are natural physical postulates that capture the essence of QM — postulates that tell us what sets QM apart from other physical theories, and in particular from its predecessor theories. The advent of quantum information theory suggests that we look for *information-theoretic* postulates that characterize (i.e., are equivalent to) QM.

A positive answer to this question has been supplied by Clifton, Bub, and Halvorson (2003). Clifton, Bub and Halvorson (CBH) show that, within

[*] hhalvors@princeton.edu. This is version 2.

the C^*-algebraic framework for physical theories, quantum theories are singled out by their satisfaction of three information-theoretic axioms: 1. no superluminal information transfer via measurement; 2. no broadcasting;[1] and 3. no unconditionally secure bit commitment. Nonetheless, the creative thinker will have little trouble concocting a "theory" that satisfies these three axioms, but which does not entail QM (see Spekkens, 2003; Smolin, 2003). Such toy theories might be thought to show that the three information-theoretic axioms are not sufficient to recover the full structure of QM.

Since the CBH characterization theorem is a valid mathematical result, there is a problem of application here — these apparent counterexamples must not satisfy the premises of the theorem. Besides the three information-theoretic axioms, the only other premise of the theorem is the C^* assumption (i.e., the assumption that a theory's observables be representable by the self-adjoint operators in a C^*-algebra). However, in specific cases, it may be difficult to ascertain whether or not a theory satisfies the C^* assumption. In particular, since the axioms for C^*-algebras are rather intricate, and some of these axioms have no direct physical interpretation (e.g., the C^*-algebraic product of non-commuting observables does not correspond to any physical operation on observables), there is a prima facie difficulty in relating the C^* assumption to specific features of a physical theory.

In this paper, we address the difficulty in determining whether a toy theory satisfies the C^* assumption. In particular, it is sometimes easier to ascertain the convex structure of the state space of a theory (i.e., which states are mixtures of which other states) than to ascertain the algebraic structure of the observables of that theory. Furthermore, due to the deep mathematical results of Alfsen et al. (detailed in Alfsen & Shultz, 2003), specifying the convex structure of the state space of a theory is sufficient to determine whether that theory can be formulated within the Jordan-Banach (JB) algebraic framework. Since a theory permits a C^*-algebraic formulation only if it permits a JB algebraic formulation, showing that a theory does not permit a JB algebraic formulation is sufficient to show that it falls outside of the range of validity of the CBH theorem.

The structure of this paper is as follows. In Section 2, we review the basics of the theory of JB algebras, and of the dual (but more general) theory of convex sets. We also prove a "Root Theorem," which forms the basis for our results in subsequent sections. In Section 3, we address the worry that the three information-theoretic axioms are not sufficient to entail QM.

[1] For the case of pure states, broadcasting reduces to cloning.

In particular, we look at a certain class of toy theories that satisfies the axioms, and we show that these toy theories do not permit a JB algebraic formulation. In Section 4, we consider a class of theories that are locally quantum-mechanical, but which, unlike QM, do not have nonlocally entangled states. We show that the simplest of these theories does not permit a JB algebraic formulation; and we adduce considerations which indicate that no such theory permits a JB algebraic formulation.

2 The JB algebraic framework for physical theories

2.1 Jordan-Banach algebras

The CBH theorem shows that among the theories within the C^*-algebraic framework, quantum theories are precisely those that satisfy the three information-theoretic axioms. One limitation of this result is that it excludes from consideration those theories that employ real or quaternionic Hilbert spaces (and so the result does not shed any light on the physical significance of the choice of the underlying field for a Hilbert space). We can get past this limitation by moving to the broader JB algebraic framework.

Let $M_n(\mathbb{K})$ be the set of $n \times n$ matrices over $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}, \mathbb{C}$, or $\mathbb{H}$ (the quaternions). The set $H_n(\mathbb{K})$ of Hermitian matrices in $M_n(\mathbb{K})$ is a vector space over $\mathbb{R}$. If we set $A \circ B = \frac{1}{2}(AB + BA)$, where AB is the usual matrix product of A and B, then it follows that

$$((A \circ A) \circ B) \circ A = (A \circ A) \circ (B \circ A). \tag{1}$$

The matrix algebra $H_n(\mathbb{K})$ with product $\circ$ is the prototype for the notion of a Jordan algebra: a Jordan algebra is any real vector space equipped with a commutative (not necessarily associative), bilinear product $\circ$ satisfying Eqn. 1.

For an element $A \in H_n(\mathbb{C})$, we define

$$\|A\| = \sup\{\|Ax\| : x \in \mathbb{C}^n,\ \|x\| = 1\}, \tag{2}$$

where the norm on the right is the vector norm on $\mathbb{C}^n$. This norm is complete in the sense that for any Cauchy sequence $\{A_i\}$ in $H_n(\mathbb{C})$, there is an $A \in H_n(\mathbb{C})$ such that $\lim_i \|A_i - A\| = 0$. That is, $H_n(\mathbb{C})$ is a Banach space. Furthermore, the norm satisfies the inequalities:

$$\|A \circ B\| \leq \|A\|\|B\|, \quad \|A \circ A\| = \|A\|^2, \quad \|A \circ A\| \leq \|(A \circ A) + (B \circ B)\|, \tag{3}$$

for all A, B in $H_n(\mathbb{C})$. In general, a Jordan-Banach (JB) algebra is a Jordan algebra that is complete relative to some norm satisfying Eqns. 3.

States on $H_n(\mathbb{C})$ are given, in the first place, by (equivalence classes of) unit vectors in $\mathbb{C}^n$. In particular, if $|\alpha\rangle$ is a unit vector in $\mathbb{C}^n$, then $\langle\alpha|A|\alpha\rangle$ gives the expectation value of A in the state $|\alpha\rangle$. Note that the map $A \mapsto \langle\alpha|A|\alpha\rangle$ on $H_n(\mathbb{C})$ is linear and continuous. Furthermore, $\langle\alpha|I|\alpha\rangle = 1$, where I is the identity matrix, and and $\langle\alpha|(A \circ A)|\alpha\rangle \geq 0$ for any A. Generally, we define a state of a JB algebra $\mathfrak{A}$ to be a linear and continuous mapping $\omega : \mathfrak{A} \to \mathbb{R}$ such that $\omega(I) = 1$ and $\omega(A \circ A) \geq 0$ for all A in $\mathfrak{A}$.

The state space K of a JB algebra $\mathfrak{A}$ is a convex set; that is, if x and y are states, and $\lambda \in (0,1)$, then $\lambda x + (1-\lambda)y$ defines a state in a natural way. The set K also carries two standard topologies. First, a net $\{\omega_a\}$ of states converges in the weak* topology to a state ω just in case the numbers $\{\omega_a(A) : A \in \mathfrak{A}\}$ converge pointwise to the numbers $\{\omega(A) : A \in \mathfrak{A}\}$. Since K is a weak* closed subset of the unit ball of the Banach space dual $\mathfrak{A}^*$ (all continuous linear functionals on $\mathfrak{A}$), the Alaoglu-Bourbaki theorem (Kadison & Ringrose, 1997, Thm. 1.6.5) entails that K is weak* compact. Second, K inherits the standard norm topology from $\mathfrak{A}^*$. A net $\{\omega_a\}$ converges in norm to ω just in case the numbers $\{\omega_a(A) : A \in \mathfrak{A}\}$ converge uniformly to the numbers $\{\omega(A) : A \in \mathfrak{A}\}$. Thus, the norm topology on K is always finer that the weak* topology. In the finite dimensional case, pointwise convergence entails uniform convergence, and so the weak* and norm topologies are equivalent. But in the infinite dimensional case, K will not typically be compact in the norm topology. (For example, the state space of the JB algebra $\mathcal{B}(\mathcal{H})_{\mathrm{sa}}$ of all self-adjoint operators on an infinite dimensional Hilbert space $\mathcal{H}$ is not compact in the norm topology.)

There is a canonical mapping from the category of C^*-algebras into the category of JB algebras. Indeed, if $\mathfrak{A}$ is a C^*-algebra, and $\mathfrak{A}_{\mathrm{sa}}$ is the real vector space of self-adjoint operators in $\mathfrak{A}$, then $\mathfrak{A}_{\mathrm{sa}}$ with the symmetric product is a JB algebra (Landsman, 1998, Thm. 1.1.9). Furthermore, the state space of $\mathfrak{A}$ is affinely isomorphic (see the definition below) to the state space of $\mathfrak{A}_{\mathrm{sa}}$. In contrast, the nonassociative JB algebra $H_2(\mathbb{R})$ has linear dimension 3, whereas there is no C^*-algebra $\mathfrak{A}$ such that $\mathfrak{A}_{\mathrm{sa}}$ is a 3-dimensional, nonassociative JB algebra. ($\mathfrak{A}_{\mathrm{sa}}$ is associative iff $\mathfrak{A}$ is abelian.) Therefore, $H_2(\mathbb{R})$ is not isomorphic to the self-adjoint part of C^*-algebra, and the JB algebraic framework is genuinely broader than the C^*-algebraic framework.

2.2 Convex sets

All JB algebra state spaces are convex sets. But the converse is not true — not all convex sets are JB algebra state spaces. We now briefly recall some of the main definitions in the theory of convex sets.

A point x in a convex set K is *extreme* just in case for any $y, z \in K$ and $\lambda \in (0,1)$, if $x = \lambda y + (1-\lambda)z$, then $x = y = z$. We let $\partial_e K$ denote the set of extreme points in K. If K is the state space of an algebra, we also call extreme points *pure states*. A subset F of a convex set K is said to be a *face* just in case F is convex, and for any $x \in F$, if $x = \lambda y + (1-\lambda)z$ with $\lambda \in (0,1)$, then $y \in F$. Clearly the intersection of an arbitrary family of faces is again a face. For $x, y \in K$, we let $\mathrm{face}(x, y)$ denote the intersection of all faces containing $\{x, y\}$. A pair of faces F, G in K is said to be *split* if every point in K can be expressed uniquely as a convex combination of points in F and G. A convex set K is said to be a *simplex* if mixed states have unique decompositions into pure states. More precisely, K is a simplex if for all $w, x, y, z \in \partial_e K$, when

$$\lambda w + (1-\lambda)x = \mu y + (1-\mu)z, \tag{4}$$

with $\lambda, \mu \in (0,1)$, then either $w = y$ or $w = z$. (This definition differs slightly from the standard definition; see (Alfsen & Shultz, 2001, p. 8).) If K and L are convex sets, a mapping $\phi : K \to L$ is an *affine isomorphism* just in case ϕ is bijective, and

$$\phi(\lambda x + (1-\lambda)y) = \lambda\phi(x) + (1-\lambda)\phi(y), \tag{5}$$

for all $x, y \in K$ and $\lambda \in (0,1)$. If there is an affine isomorphism ϕ from K onto L, then K and L are said to be *affinely isomorphic*.

2.3 The root theorem

Drawing on the results of Alfsen et al., we now derive some easily checked necessary conditions for a theory to admit a JB algebraic formulation. (In this theorem and subsequently, we let B^n denote the closed unit ball in $\mathbb{R}^n$.)

Root Theorem. *Let K be a convex set. If K is affinely isomorphic to the state space of a JB algebra, then:*

1. *For any distinct $x, y \in \partial_e K$,* $\mathrm{face}(x, y) = B^n$ *for some $n \geq 1$.*
2. *If K is not a simplex, then for any distinct $x, y \in \partial_e K$,* $\mathrm{face}(x, y) = B^n$ *for some $n \geq 2$.*

3. *If $x, y \in \partial_e K$ are connected by a norm-continuous path, then* $\mathrm{face}(x, y) = B^n$ *for some $n \geq 2$.*

The statement of (3) could use some clarification: since we have not made any assumptions about a topology on K, saying that $\partial_e K$ is connected does not really make sense. However, if K is affinely isomorphic to the state space K' of a JB algebra, then there is a map $\phi : K \to K'$. Thus, (3) should be understood as referring to the topology on K that is induced, via the mapping ϕ, by the norm topology on K'.

Proof. (1.) The first statement is a non-trivial result (Corollary 5.56 in Alfsen & Shultz, 2003) that depends on a number of lemmas. Due to space constraints, we just sketch the structure of the proof for the simple case where K is a subset of a finite-dimensional vector space.

Suppose that K is the state space of a JB algebra $\mathfrak{A}$. Since x, y are pure states, they correspond to minimal projection operators $P, Q \in \mathfrak{A}$, and $\mathrm{face}(x, y)$ is the state space of the "projected" algebra

$$\mathfrak{A}_{P \vee Q} = \{(P \vee Q)A(P \vee Q) : A \in \mathfrak{A}\}. \tag{6}$$

The identity of $\mathfrak{A}_{P \vee Q}$ (namely, $P \vee Q$) is the sum of two orthogonal projections P and $R = (P \vee Q) - P$. We now consider the two cases where this projected algebra is associative or nonassociative. If $\mathfrak{A}_{P \vee Q}$ is associative, then it is isomorphic to the algebra of real valued functions on a two-point set, and its state space consists of two pure states and their convex combinations. That is, the state space of $\mathfrak{A}_{P \vee Q}$, and therefore $\mathrm{face}(x, y)$, is isomorphic to B^1. If $\mathfrak{A}_{P \vee Q}$ is nonassociative, then in fact the center of $\mathfrak{A}_{P \vee Q}$ is trivial. By the comparison theorem for projections, there is a symmetry $U \in \mathfrak{A}_{P \vee Q}$ (that is, $U \circ U = I$) such that

$$2(U \circ (P \circ U)) - P = Q. \tag{7}$$

Thus, the identity in $\mathfrak{A}_{P \vee Q}$ is the sum of two "exchangeable" minimal projections. Finite dimensional JB algebras with this property have been completely classified (see Alfsen & Shultz, 2003, Prop. 3.37), and their state spaces are isomorphic to B^n, for some $n \geq 2$.

(2.) If K is not a simplex, then there are $w, x, y, z \in \partial_e K$ such that

$$\lambda w + (1 - \lambda)x = \mu y + (1 - \mu)z, \tag{8}$$

where $\lambda, \mu \in (0, 1)$, $w \neq y$ and $w \neq z$. But then w is an extreme point in $\mathrm{face}(y, z)$. We know from part 1 that $\mathrm{face}(y, z) = B^n$, for some $n \geq 1$. Since there are three distinct extreme points of $\mathrm{face}(y, z)$, it follows that $n \geq 2$.

(3.) We prove the contrapositive. Suppose that $x, y \in \partial_e K$, and $\text{face}(x, y) = B^1$. Then there are split faces F, G of K such that $x \in F$ and $y \in G$ (Alfsen & Shultz, 2003, Lemma 5.54). Let $U = F \cap \partial_e K$ and let $V = G \cap \partial_e K$. Since F and G are closed in the norm topology (Alfsen & Shultz, 2001, Prop. 1.29), U and V are closed in $\partial_e K$. Since $\partial_e K \subseteq F \cup G$, it follows that $\partial_e K = U \cup V$, and U and V are open in $\partial_e K$. Since $x \in U$ and $y \in V$, there is no continuous path in $\partial_e K$ connecting x and y. $\square$

3 Sufficiency of the axioms

The state space of a quantum system has *ambiguous mixtures* — i.e., mixed states with more than one decomposition into pure states — and this fact is responsible for some of the interesting information-theoretic features of QM. For example, the BB84 (Bennett & Brassard, 1984) bit commitment protocol is perfectly concealing because two distinct ensembles can be absolutely identical relative to a local observer (since these ensembles correspond to the same quantum state). Thus, in order to find a toy theory that simulates some of the information-theoretic features of QM, it would be natural to look for simple theories with ambiguous mixtures.

One such theory has been recently described by Spekkens (2003). (Smolin (2003) proposes a different sort of theory that satisfies the three axioms. We look at Smolin's theory in (Halvorson & Bub, 2003).) The state space S of Spekkens' theory (for a local system) has exactly seven points: the pure states correspond to the unit vectors $\{e_i, -e_i : i = 1, 2, 3\}$ in $\mathbb{R}^3$, and the mixed state corresponds to the origin $\mathbf{0}$ in $\mathbb{R}^3$. In order to equip S with partial binary operations corresponding to superposition and mixture, we identify S with a subset of the Bloch sphere. That is, $\mathbf{0}$ is an equal mixture of e_i and $-e_i$, for $i = 1, 2, 3$. However, S does not contain unequal mixtures of e_i and $-e_i$, nor does S contain mixtures of e_i and e_j when $i \neq j$. Similarly, e_i and $-e_i$ can be superposed with equal weights to obtain any of the states in $\{e_j, -e_j : j \neq i\}$. However, e_i and $-e_i$ cannot be superposed with unequal weights, nor can e_i be superposed with e_j when $i \neq j$.

Since S is not convex, it is obviously not the state space of a JB algebra. However, the failure of convexity can be easily remedied by passing to the modified theory that allows arbitrary mixtures of Spekkens' states — i.e., the theory whose state space is the convex hull $K = \text{co}(S)$ of S.[2] Clearly,

[2] Presumably, Spekkens would want to say that the "transition probability" between e_1 and e_2 is $\frac{1}{2}$, since e_2 is supposed to be an equally weighted superposition of the orthogonal states e_1 and $-e_1$. However, the natural geometric transition probability of e_1 and e_2,

K has ambiguous mixtures, and has exactly six pure states. We now show that convex sets of this sort are not state spaces of JB algebras.

Theorem 1. *Let K be a convex set, and suppose that K is affinely isomorphic to the state space of a JB algebra. If K is not a simplex, then $|\partial_e K| \geq |\mathbb{R}|$.*

Proof. Suppose that K is not a simplex. Then part 1 of the Root Theorem entails that there are $x, y \in \partial_e K$ such that $\text{face}(x, y) = B^n$, with $n \geq 2$. Since every extreme point in $\text{face}(x, y)$ is an extreme point in K, we have $|\partial_e K| \geq |\partial_e B^n| = |\mathbb{R}|$. □

Spekkens' theory does not permit a JB algebraic formulation, and *a fortiori*, does not not permit a C^*-algebraic formulation. So, this theory falls outside the range of validity of the CBH theorem. Nonetheless, since Spekkens' theory has no obvious physical pathologies, it would be a interesting test case for the claim that physical theories should permit, at the very least, a JB algebraic formulation.

4 Independence of the axioms

All parties agree that, in the presence of the C^* assumption (i.e., the assumption that theories permit a C^*-algebraic formulation), the three information-theoretic axioms entail QM. It seems, then, that the real question is whether the C^* assumption is true, warranted, reasonable, or something like that. Unfortunately, it seems that it would be extremely difficult to give a decisive answer to this question.

However, there is reason to think that the C^* assumption is doing too much work in the CBH theorem. In particular, given the C^* assumption, QM is a logical consequence of the first two axioms alone. In fact, given the C^* assumption, the no bit commitment axiom is a logical consequence of the no superluminal signaling and no cloning axioms. We prove this fact here with the one simplifying assumption that the relevant algebras are actually von Neumann algebras (i.e., the algebras act on some concrete Hilbert space $\mathcal{H}$, and are closed in the weak-operator topology).

relative to the convex set K, is 0. (See the definition of affine ratio below.) In particular, if every affine function from K into $\mathbb{R}$ corresponds to an observable (as is usually assumed in the convex sets approach), then there is a measurement that can distinguish with certainty between e_1 and e_2; for example, the measurement corresponding to the function $f(x) = \frac{1}{2}(1 + (e_1 - e_2 + e_3) \cdot x)$. Presumably, then, Spekkens would wish to impose some restriction on the set of observables.

Theorem 2. *Suppose that $\mathfrak{A}$ and $\mathfrak{B}$ are von Neumann algebras. If the composite system $(\mathfrak{A}, \mathfrak{B})$ satisfies the no superluminal signaling and no cloning axioms, then:*

1. *$(\mathfrak{A}, \mathfrak{B})$ has nonlocally entangled states; and*

2. *$(\mathfrak{A}, \mathfrak{B})$ satisfies the no bit commitment axiom.*

Proof. Suppose that the pair $(\mathfrak{A}, \mathfrak{B})$ satisfies the no superluminal signaling and no cloning axioms. On the one hand, the no superluminal signaling axiom entails that observables in $\mathfrak{A}$ commute with observables in $\mathfrak{B}$ (Clifton et al., 2003, Thm. 1). On the other hand, the no cloning axiom entails that $\mathfrak{A}$ and $\mathfrak{B}$ are nonabelian (Clifton et al., 2003, Thm. 2).

(1.) When $\mathfrak{A}$ and $\mathfrak{B}$ are nonabelian, a theorem by Landau (1987) shows that there are nonlocally entangled (indeed, Bell correlated) states across $(\mathfrak{A}, \mathfrak{B})$.

(2.) By the generalized HJW theorem (Halvorson, 2003), for any two equivalent measures μ, ν on the state space of $\mathfrak{B}$ (i.e., these measures correspond to the same quantum state), there is an entangled state ψ of $(\mathfrak{A}, \mathfrak{B})$ such that either μ or ν can be prepared from ψ by local operations on $\mathfrak{A}$. So, for any bit commitment protocol for $(\mathfrak{A}, \mathfrak{B})$, if the protocol is concealing, then it is not binding. □

4.1 The Schr*dinger theory

Theorem 2 is somewhat surprising. From an apparently local axiom (no cloning), it follows that there are nonlocally entangled states. In slogan form: *any locally quantum mechanical theory is nonlocal.* We may contrast this result with Schrödinger's claim that there should be a locally quantum mechanical theory without entangled states. Schrödinger says:

> Indubitably, the situation described here [in which there are nonlocally entangled states] is, in present QM, a necessary and indispensable feature. The question arises, whether it is so in Nature too. I am not satisfied about there being sufficient experimental evidence for that...
>
> It seems worth noticing that the paradox could be avoided by a very simple assumption, namely if the situation after [two systems] separating were described by the expansion
>
> $$c_1|01\rangle + c_2|10\rangle, \tag{9}$$

> but with the additional statement that the knowledge of the *phase relations* between the complex constants c_1 and c_2 has been entirely lost in consequence of the process of separation. This would mean that not only the parts, but the whole system, would be in the situation of a mixture, not of a pure state. ...it would utterly eliminate the experimenter's influence on the state of that system which he does not touch.
>
> This is a very incomplete description and I would not stand for its adequateness. But I would call it a possible one, until I am told, either why it is devoid of meaning or with which experiments it disagrees.
>
> (Schrödinger, 1936, pp. 451–452. Eqn. 9 has been adapted to the present discussion.)

When Schrödinger speaks of the state in Eqn. 9, but with "the knowledge of the phase relations" lost, he presumably means the mixed state

$$|c_1|^2|01\rangle\langle 01| + |c_2|^2|10\rangle\langle 10|. \tag{10}$$

Thus, Schrödinger's hope is that the true theory will turn out to be locally quantum mechanical, but with some sort of selection rule that prohibits superposition of product states for systems that are spacelike separated.

We now know — due to experimental verification of the violation of Bell's inequality — that Schrödinger's hoped-for theory disagrees with experiment. But, Theorem 2 shows that Schrödinger's hoped-for theory is "devoid of meaning" — well, at least if all meaningful theories admit a C^*-algebraic formulation. But can Schrödinger's hope be realized within the broader JB algebraic framework? While we do not currently know the answer to this question, we will proceed to show that the answer is negative in one particularly simple case.

Consider the simplest composite quantum system, a pair of qubits. Of course, we cannot simply throw away the nonlocally entangled states without doing violence to the linear structure of $\mathbb{C}^2 \otimes \mathbb{C}^2$. But since the complement of the set of nonlocally entangled states is a convex set, we can throw away the entangled states and still end up with a theory with a convex state space. More precisely, recall that a density operator $D \in \mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ is a *pure product state* just in case $D = E \otimes F$, where E, F are projections onto rays in $\mathbb{C}^2$. The set of pure product states is a closed subset of the pure state space of $\mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$; as such, it is a closed, bounded subset of $\mathbb{R}^{15}$. (The set of 4×4 Hermitian complex matrices has real dimension 16, and the subset of positive, trace-1 matrices has real dimension 15.) Let K denote the set of convex combinations of pure product states; in other words, K is the space of *separable states* of $\mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$, and corresponds to a compact convex subset of $\mathbb{R}^{15}$. (For some results on the geometry of K, see (Bertlmann, Narnhofer, & Thirring, 2002).)

Since K is a convex set, it gives a genuine theory in the convex sets approach; we call this theory the *Schr*dinger theory*. (This ad hoc construction is for conceptual purposes only; we do not think that Schrödinger really had this theory in mind when he expressed his hope for an alternative to QM.) The observables of the Schr*dinger theory are the elements of the real vector space $A(K)$ of affine functions from K to $\mathbb{R}$. The expectation value of observable $f \in A(K)$ in state $x \in K$ is $f(x)$. Clearly, each self-adjoint operator $A \in \mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ gives an observable for the Schr*dinger theory via the mapping $x \mapsto \mathrm{Tr}(xA)$.

In order to clarify the information-theoretic properties of the Schr*dinger theory, we need to define a notion of transition probability for arbitrary convex sets (see Mielnik, 1969; Landsman, 1998, Prop. 2.8.1).

Definition. Let K be a convex set, and let $A(K)$ be the set of affine functions from K into $\mathbb{R}$. If $x, y \in \partial_e K$, then the affine ratio (or transition probability) of x and y relative to K is given by

$$\mathbf{p}_K(x/y) \;=_{\text{def}}\; \inf\{f(y); f \in A(K),\ \text{range}(f) \subseteq [0,1],\ \text{and}\ f(x) = 1\}.$$

The affine ratio has a natural geometrical interpretation. In particular, if K is a contained in the vector space V, then each affine function $f : V \to \mathbb{R}$ foliates V into a family of hyperplanes $\{f^{-1}(t)\}_{t\in\mathbb{R}}$. Now consider those f's where K lies between the 0 and 1 hyperplanes, and where x lies in the intersection of K with the 1 hyperplane. Then any $y \in \partial_e K$ falls within a unique $t \in [0,1]$ hyperplane. Finally, consider all such foliations, and take the infimum of the t such that $f(y) = t$.

In some nice cases, there is a unique affine function f such that $f(x) = 1$ and $\min\{f(y) : y \in K\} = 0$; thus, $\mathbf{p}_K(x/y) = f(y)$. For example, when K is the unit sphere in $\mathbb{R}^3$, and x, y are points on the surface of the sphere, then $\mathbf{p}_K(x/y) = \frac{1}{2}(1 + x \cdot y)$. That is, the transition probability is given by the (normalized) tangent function to K at x.

In fact, the sphere is a familiar case from QM: it is affinely isomorphic to the set of density operators on $\mathbb{C}^2$, and the affine ratio corresponds to the standard quantum-mechanical transition probability. Indeed, the equivalence between the two notions holds quite generally.

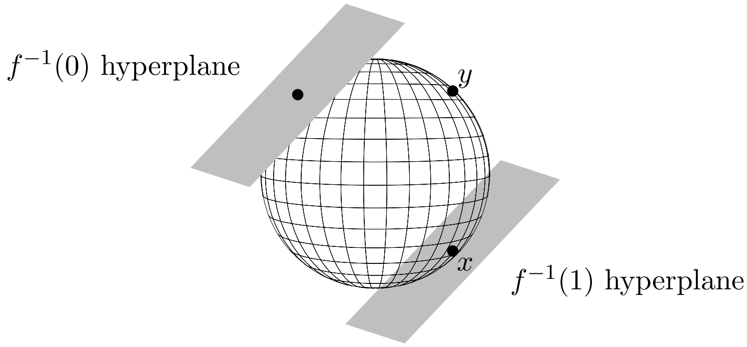

Figure 1: The Bloch sphere

Lemma 1. *Let $\mathcal{H}$ be a complex Hilbert space, and let K be the convex set of density operators on $\mathcal{H}$. Then for any projection operators $E, F \in \partial_e K$, $\mathbf{p}_K(E/F) = \mathrm{Tr}(EF)$.*

Proof. Consider the affine function $f : K \to [0,1]$ given by $f(D) = \mathrm{Tr}(ED)$, for all $D \in K$. We claim that $f(F) = \mathbf{p}_K(E/F)$, for all $F \in \partial_e K$. For this it will suffice to show that for any $g \in A(K)$, if range$(g) \subseteq [0,1]$ and $g(E) = 1$, then $g \geq f$. Let g be such a function. Since $A(K)$ is order-isomorphic to $\mathcal{B}(\mathcal{H})_{\mathrm{sa}}$, there is a self-adjoint operator A on $\mathcal{H}$ such that A has spectrum in $[0,1]$, and $g(F) = \mathrm{Tr}(AF)$, for all $F \in \partial_e K$. Let $|\alpha\rangle$ be a unit vector in the range of E, and let $|\beta\rangle$ be a unit vector in the range of F. Then $\langle\alpha|A|\alpha\rangle = g(E) = 1$, and it follows that $A|\alpha\rangle = |\alpha\rangle$. By the spectral theorem, $EA = AE = E$, and so $\langle\beta|(I-E)A|\beta\rangle \geq 0$. Therefore,

$$\langle\beta|A|\beta\rangle = \langle\beta|EA|\beta\rangle + \langle\beta|(I-E)A|\beta\rangle \geq \langle\beta|E|\beta\rangle. \tag{11}$$

That is, $g(F) \geq f(F)$, for all $F \in \partial_e K$. □

We can now show that the Schr*dinger theory satisfies the no cloning axiom, but violates the no bit commitment axiom. First, we claim that the permissible state transformations of a theory with state space L correspond to affine endomorphisms of L. (A transformation is reversible iff it is one-to-one.) If η is an affine endomorphism of L, and if $L' = \eta(L)$, then

$$\mathbf{p}_{L'}(\eta(x)/\eta(y)) \geq \mathbf{p}_L(x/y), \qquad \forall x, y \in L. \tag{12}$$

In the Schr*dinger theory, we have two subsystems A, B (each qubits) combined in a nonstandard way into a composite system AB. Suppose for reductio ad absurdum that states of system A can be cloned by using system

B as a cloning machine. That is, there is a ready state x_0 of B and a state transformation η on AB such that $\eta(x \otimes x_0) = x \otimes x$ for all pure states x of A. Let x and y be non-orthogonal pure states; that is, $0 < \mathbf{p}_A(x/y) < 1$. It then follows that

$$\mathbf{p}_A(x/y) \leq \mathbf{p}_{AB}(x \otimes x_0/y \otimes x_0) \leq \mathbf{p}_{AB}(x \otimes x/y \otimes y) \leq \mathbf{p}_A(x/y)^2, \quad (13)$$

in contradiction with the assumption x and y are non-orthogonal. (The first inequality follows from the fact that $x \mapsto x \otimes x_0$ is an affine embedding of A into AB. The second inequality follows from the fact that the cloning map cannot decrease transition probabilities. The final inequality follows from the fact that $\mathbf{p}_{AB}(x \otimes x/y \otimes y) \leq T$, where T is the transition probability of $x \otimes x$ and $y \otimes y$ relative to the full state space of $\mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$; and $T = |\langle x \otimes x | y \otimes y \rangle|^2 = |\langle x|y \rangle|^4 = \mathbf{p}_A(x/y)^2$.) Therefore, the cloning map η does not exist.

In order to see that the Schr*dinger theory allows an unconditionally secure bit commitment protocol, consider the direct analogue of the BB84 protocol (Bennett & Brassard, 1984). In this protocol, Alice encodes bit 0 into the mixed state $D_0 = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|)$, and she encodes bit 1 into the mixed state $D_1 = \frac{1}{2}(|\alpha\beta\rangle\langle\alpha\beta| + |\beta\alpha\rangle\langle\beta\alpha|)$, where $|\alpha\rangle = 2^{-1/2}(|0\rangle + |1\rangle)$ and $|\beta\rangle = 2^{-1/2}(|0\rangle - |1\rangle)$. Since $\mathrm{Tr}_A(D_0) = \mathrm{Tr}_A(D_1)$, this protocol is perfectly concealing. If Alice could prepare the EPR-Bohm state E, i.e., the projection onto the vector $2^{-1/2}(|01\rangle - |10\rangle)$, then this protocol would not be binding; because E can be transformed by local nonselective operations into either D_0 or D_1. However, in the Schr*dinger theory, there are no such entangled states; indeed, there is no state that Alice can transform into either D_0 or D_1. Therefore, this protocol is perfectly binding.

Since the Schr*dinger theory prohibits cloning, but allows unconditionally secure bit commitment, it follows from Theorem 2 that it does not admit a C^*-algebraic formulation. We devote the next section to establishing a stronger claim: the Schr*dinger theory does not admit a JB algebraic formulation.

4.2 Pathology of the Schr*dinger theory

By using the generalized definition of transition probability for arbitrary convex sets, we can see classical transition probabilities are always in $\{0, 1\}$, whereas quantum mechanical transition probabilities can lie anywhere in the unit interval. We will use the transition probability to define a generalized notion of the superposition of two pure states.

Definition. Pure states $x, y \in \partial_e K$ are said to be *orthogonal* just in case $\mathbf{p}_K(x/y) = 0$. Two orthogonal states $x, y \in \partial_e K$ are said to be *superposable in* K just in case there is a $z \in \partial_e K$ such that $\mathbf{p}_K(x/z) = \frac{1}{2} = \mathbf{p}_K(y/z)$.

This definition is motivated by the following considerations. If observables correspond to affine functions on K (as is usually assumed in the convex sets approach), a measurement designed to distinguish x from y can be represented by an affine function $f : K \to [0, 1]$, where $f(x) = 1$ and $f(y) = 0$. If there is such a function f, then $\mathbf{p}_K(x/z) = \frac{1}{2} = \mathbf{p}_K(y/z)$ iff $f(z) = \frac{1}{2}$. That is, when the system is in state z, the x and y outcomes of an f-measurement are equally likely.

For the case of JB algebra state spaces, if two pure states can be connected by a continuous path, then they can be coherently superposed.[3]

Lemma 2. *Let K be the state space of a JB algebra, and let x, y be orthogonal states in $\partial_e K$. If x and y are connected by a norm-continuous path in $\partial_e K$, then x and y are superposable in K.*

Proof. Suppose that $x, y \in \partial_e K$ are orthogonal, and that x and y are connected by a norm-continuous path in $\partial_e K$. Let $F = \text{face}(x, y)$. By part 3 of the Root Theorem, there is an affine isomorphism ϕ from F onto B^n, with $n \geq 2$. Let $\{e_1, e_2, \ldots, e_n\}$ be the canonical orthonormal basis for $\mathbb{R}^n$. Since there is an affine automorphism of B^n that maps $\phi(x)$ to e_1, we may suppose that $\phi(x) = e_1$. An exercise in elementary geometry shows that $\phi(y) = -e_1$. (ϕ preserves affine ratios, and $-e_1$ is the unique $r \in B^n$ such that $\mathbf{p}_{B^n}(e_1/r) = 0$.) Furthermore, $\mathbf{p}_{B^n}(e_1/e_2) = \mathbf{p}_{B^n}(-e_1/e_2) = \frac{1}{2}$. Thus, if we choose $z = \phi^{-1}(e_2)$, then $\mathbf{p}_K(x/z) \geq \mathbf{p}_F(x/z) = \frac{1}{2}$ and $\mathbf{p}_K(y/z) \geq \mathbf{p}_F(y/z) = \frac{1}{2}$. □

The Schr*dinger theory seems to have some non-trivial superselection rule, because it does not seem to allow coherent superpositions of, say, $|01\rangle$ and $|10\rangle$. We begin by confirming that such states are not superposable.

Lemma 3. *Let K be the set of separable states of $\mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$. Then $|01\rangle$ and $|10\rangle$ are not superposable in K.*

[3] According to Hardy (2001), QM is differentiated from classical probability theory by the assumption that there is a continuous transition between any two pure states. Hardy's claim is true in the JB algebraic framework (if we take the relevant topology to be the norm topology), if we think of "quantum" systems as those systems that have a single non-trivial superselection sector, and "classical" systems as those systems whose superselection sectors are singletons.

Proof. Let $x = |01\rangle\langle 01|$, let $y = |10\rangle\langle 10|$, and suppose for reductio ad absurdum that there is a $z = |\alpha\beta\rangle = |\alpha\rangle|\beta\rangle \in \partial_e K$ such that $\mathbf{p}_K(x/z) = \frac{1}{2} = \mathbf{p}_K(y/z)$. If L is the full state space of $\mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$, then $\mathbf{p}_K(v/w) \leq \mathbf{p}_L(v/w)$ for any $v, w \in \partial_e K$ (since $\partial_e K \subseteq \partial_e L$). Thus, Lemma 1 entails that

$$\begin{aligned} 1 &= \mathbf{p}_K(x/z) + \mathbf{p}_K(y/z) && (14) \\ &\leq \mathbf{p}_L(x/z) + \mathbf{p}_L(y/z) && (15) \\ &= |\langle 01|\alpha\beta\rangle|^2 + |\langle 10|\alpha\beta\rangle|^2 \ \leq \ 1. && (16) \end{aligned}$$

We now show that either $|\langle 01|\alpha\beta\rangle|^2 = 0$ or $|\langle 10|\alpha\beta\rangle|^2 = 0$. For this, let

$$a = |\langle 0|\alpha\rangle|^2, \quad b = |\langle 1|\beta\rangle|^2, \quad c = |\langle 0|\beta\rangle|^2, \quad d = |\langle 1|\alpha\rangle|^2. \tag{17}$$

Thus, Eqn. 16 becomes $ab + cd = 1$. Since $\{|0\rangle, |1\rangle\}$ is an orthonormal basis for $\mathbb{C}^2$, we also have $b = 1 - c$ and $d = 1 - a$. Hence, $a + c - 2ac = 1$. The functions $[0,1] \ni a \mapsto a + c - 2ac$ (for fixed $c \in [0,1]$) and $[0,1] \ni c \mapsto a + c - 2ac$ (for fixed $a \in [0,1]$) are affine. Thus, $a + c - 2ac$ achieves its maximum value only at extreme points of the convex set $[0,1] \times [0,1]$. Checking these points, we find that $a + c - 2ac \leq 1$, with equality achieved only when $(a,c) = (1,0)$ or $(a,c) = (0,1)$. If $c = 0$, then $|\langle 10|\alpha\beta\rangle|^2 = cd = 0$. Similarly, if $a = 0$, then $|\langle 01|\alpha\beta\rangle|^2 = ab = 0$. Applying Lemma 1 again, it follows that either

$$\mathbf{p}_K(x/z) \leq \mathbf{p}_L(x/z) = |\langle 01|\alpha\beta\rangle|^2 = 0, \tag{18}$$

or

$$\mathbf{p}_K(y/z) \leq \mathbf{p}_L(y/z) = |\langle 10|\alpha\beta\rangle|^2 = 0, \tag{19}$$

both of which contradict our assumption that z is an equally weighted superposition of x and y. Therefore, $|01\rangle$ and $|10\rangle$ are not superposable in K. □

The previous Lemma shows that if the Schr*dinger theory permits a JB algebraic formulation, then the states $|01\rangle$ and $|10\rangle$ must lie in different superselection sectors. However, $|01\rangle$ and $|10\rangle$ are connected by a continuous path of pure product states.[4]

Lemma 4. *Let K be the set of separable states of $\mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$. Then there is a continuous path in $\partial_e K$ between $|01\rangle$ and $|10\rangle$.*

[4] Since K is not a topological space, this statement doesn't really make sense. However, since K is a subset of $\mathbb{R}^n$, $(n < \infty)$, there is a unique topology τ on K that is compatible with its affine structure; namely, the relative topology from $\mathbb{R}^n$. So, if K were isomorphic to a JB algebra state space, the transported norm topology would be equivalent to τ.

Proof. By symmetry, and since path-connectedness of points is transitive, it will suffice to show that there is a norm-continuous path in $\partial_e K$ between $|01\rangle$ and $|11\rangle$. Let $\|\cdot\|_2$ denote the Hilbert-Schmidt norm on $\mathcal{B}(\mathbb{C}^n)$; i.e., $\|A\|_2 = \mathrm{Tr}(A^*A)^{1/2}$. There is a $\|\cdot\|_2$-continuous function f from $[0,1]$ into the set of one-dimensional projections on $\mathbb{C}^2$ such that $f(0) = |0\rangle\langle 0|$ and $f(1) = |1\rangle\langle 1|$. Define a function $g : [0,1] \to \partial_e K$ by setting $g(t) = f(t) \otimes |1\rangle\langle 1|$. Since $\|A \otimes B\|_2 = \|A\|_2\|B\|_2$ for all operators A, B on $\mathbb{C}^2$, it follows that

$$\|g(t) - g(t')\|_2 = \|(f(t) - f(t')) \otimes |1\rangle\langle 1|\,\|_2 = \|f(t) - f(t')\|_2, \tag{20}$$

for all $t, t' \in [0,1]$. Therefore, g is $\|\cdot\|_2$-continuous as a mapping into $\partial_e K$ with the relative topology inherited from the state space of $\mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$. □

The previous two Lemmas show that the topology and affine structure of the separable state space do not mesh in the way that these structures mesh in JB algebra state spaces.

Theorem 3. *The set of separable states of* $\mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ *is not affinely isomorphic to the state space of a JB algebra.*

Proof. Suppose for reductio ad absurdum that K is affinely isomorphic to the state space of a JB algebra. Let $x = |10\rangle\langle 10|$ and let $y = |01\rangle\langle 01|$. By Lemma 4, there is a continuous path between x and y, and so Lemma 2 entails that x and y are superposable in K. But this contradicts Lemma 3. Therefore, K is not affinely isomorphic to the state space of a JB algebra. □

This result shows that the simplest Schrödinger-like theory — viz., the Schr*dinger theory — does not admit a JB algebraic formulation. Thus, it provides some evidence for the claim that even within the JB algebraic framework, locally quantum mechanical theories have nonlocally entangled states; and it suggests that even within the JB algebraic framework, the no cloning axiom entails the no bit commitment axiom.

The upshot, then, of this section is to confirm worries that the C^* assumption is doing too much work in the CBH theorem; and, furthermore, it probably wouldn't help matters if we were to derive a generalized CBH theorem for JB algebras. But, of course, there is still hope that within a suitably broader mathematical framework (e.g., Segal-algebras (see Segal, 1947), and the dual theory of spectral convex sets), the three axioms are independent, and together entail QM.

5 Conclusion

This note attempts to clarify the limits of recent information-theoretic characterizations of QM. However, in doing so, it has raised a number of further questions, both of a technical and a philosophical nature.

First, we conjecture that the three information-theoretic axioms are independent in the Segal-algebraic framework, and that the conjunction of the axioms entails QM. This generalized version of the CBH theorem would not only address the worries raised in the previous section, but might also help shed light on traditional questions, such as physical reasons for using complex coefficients rather than reals or quaternions.

Second, the considerations in this paper suggest that we take a closer look at different ways of putting together composite systems, where all systems are assumed to have convex state spaces. It is known that there are several different notions of the "tensor product" of compact convex sets (see, e.g., Namioka & Phelps, 1969). Thus, it would be interesting to see which of these products preserve which information-theoretic properties of the component systems. More specifically, suppose that $\otimes$ is a tensor product of compact convex sets that preserves the defining properties of JB algebra state spaces. Then does it follow that $K \otimes L$ has nonlocally entangled states whenever K and L are not simplexes? Or does the JB algebraic framework permit the existence of a Schrödinger-like theory? If the JB algebraic framework does not, does the broader Segal-algebraic framework permit the existence of a Schrödinger-like theory?

Finally, our discussion has raised the question of the role of constraints (either *a priori* or operational) on theory construction. On the one hand, if there are no constraints on theory construction — i.e., if there is no minimum amount of mathematical structure shared by all theories, and if any fairy tale can count as a legitimate "toy theory" — then it would be hopeless to try to *derive* QM from information theoretic principles, or from any other sort of principles for that matter. (E.g., why assume that the results of measurements are real numbers? Why assume that measurements have single outcomes? Why assume that the laws of physics are the same from one moment to the next?) On the other hand, the idea that it is legitimate to assume a fixed background framework for physical theories seems to come into tension with the empiricist attitude that drove the two major revolutions in physics in the 20th century; and the last thing we want is to impede the search for a future theory that would generalize QM.

Acknowledgments

Many of the ideas in this paper originated from conversations with Jeff Bub. Thanks also to Rob Spekkens and to an anonymous referee for comments on an earlier draft.

References

Alfsen, E., & Shultz, F. (2001). *State spaces of operator algebras.* Boston: Birkhäuser.

Alfsen, E., & Shultz, F. (2003). *Geometry of state spaces of operator algebras.* Boston: Birkhäuser.

Bennett, C., & 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). IEEE.

Bertlmann, R., Narnhofer, H., & Thirring, W. (2002). A geometric picture of entanglement and Bell inequalities. *Physical Review A*, *66*, 032319.

Bub, J. (2004). Why the quantum? (This volume)

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

Fuchs, C. (2003). Quantum mechanics as quantum information, mostly. *Journal of Modern Optics*, *50*, 987–1023.

Halvorson, H. (2003). Generalization of the Hughston-Jozsa-Wootters theorem to hyperfinite von Neumann algebras. (arXiv.org: quant-ph/0310001)

Halvorson, H., & Bub, J. (2003). Can quantum cryptography imply quantum mechanics? Reply to Smolin. (arXiv.org: quant-ph/0311065)

Hardy, L. (2001). Quantum theory from five reasonable axioms. (arXiv.org: quant-ph/0101012)

Kadison, R., & Ringrose, J. (1997). *Fundamentals of the theory of operator algebras.* Providence, RI: American Mathematical Society.

Landau, L. (1987). On the violation of Bell's inequality in quantum theory. *Physics Letters A*, *120*, 54–56.

Landsman, N. (1998). *Mathematical topics between classical and quantum mechanics.* New York: Springer.

Mielnik, B. (1969). Theory of filters. *Communications in Mathematical Physics*, *15*, 1–46.

Namioka, I., & Phelps, R. (1969). Tensor products of compact convex sets. *Pacific Journal of Mathematics*, *31*, 469–480.

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

Segal, I. E. (1947). Postulates for general quantum mechanics. *Annals of Mathematics*, *48*, 930–948.

Smolin, J. (2003). Can quantum cryptography imply quantum mechanics? (arXiv.org: quant-ph/0310067)

Spekkens, R. (2003). In defense of the epistemic view of quantum states. (Unpublished talk, New Directions in the Foundations of Physics, University of Maryland)