From Physics to Information Theory and Back

Wayne C. Myrvold
Department of Philosophy
University of Western Ontario

Abstract

Quantum information theory has given rise to a renewed interest in, and a new perspective on, the old issue of understanding the ways in which quantum mechanics differs from classical mechanics. The task of distinguishing between quantum and classical theory is facilitated by neutral frameworks that embrace both classical and quantum theory. In this paper, I discuss two approaches to this endeavour, the algebraic approach, and the convex set approach, with an eye to the strengths of each, and the relations between the two. I end with a discussion of one particular model, the toy theory devised by Rob Spekkens, which, with minor modifications, fits neatly within the convex sets framework, and which displays in an elegant manner some of the similarities and differences between classical and quantum theories. The conclusion suggested by this investigation is that Schrödinger was right to find the essential difference between classical and quantum theory in their handling of composite systems, though Schrödinger's contention that it is entanglement that is the distinctive feature of quantum mechanics needs to be modified.

1   Introduction

Quantum information theory is the study of how the peculiar features of quantum mechanics can be exploited for the purposes of information processing and transmission. A central theme of such a study is the ways in which quantum mechanics opens up possibilities that go beyond what can be achieved classically. This has in turn led to a renewed interest in, and a new perspective on, the differences between the classical and the quantum. Although much of the work along these lines has been motivated by quantum information theory— and some of it has been motivated by the conviction that quantum theory is essentially

about possibilities of information processing and transmission—the results obtained, and the frameworks developed, have interest even for those of us who are not of that conviction. Indeed, much of the recent work echoes, and builds upon, work that predates the inception of quantum information theory. The significance of such work extends beyond the setting of quantum information theory; the work done on distinguishing the quantum from the classical in the context of frameworks that embrace both is something worthy of the attention of anyone interested in the foundational issues surrounding quantum theory.

One of the striking features of quantum mechanics lies in its probabilistic character. A quantum state yields, not a definite prediction of the outcome of an experiment, but a probability measures on the space of possible outcomes. Of course, probabilities occur also in a classical context. In this context they have to do with situations in which the experimenter does not have complete control over the classical state to be prepared, and, as a consequence, we do not have complete knowledge of the classical state of the system subjected to the preparation procedure. The question arises, therefore, whether quantum probabilities can be construed as being like this. One way of framing this question is in terms of hidden variables: ought we to think of quantum-mechanical pure states as being probabilistic mixtures of states of a more encompassing theory, whose pure states would ascribe definite values to all variables? It is interesting to ask which of the peculiar features of quantum mechanics are traceable to ineliminable statistical dispersion in its states. Such features will be reproducible in an essentially classical theory with suitable restrictions on possibilities of state preparation.

Some of the recent work in quantum information theory has shown that some of the features of quantum mechanics that one might be inclined to think of as peculiarly quantum can, indeed, be recovered from a theory in which the preparable states are probabilistic mixtures of such classical states. The no-cloning theorem is a case in point. Arbitrary pairs of quantum states cannot be cloned; those pairs that can be cloned are orthogonal pairs. Though originally formulated in the context of quantum mechanics, it admits of a formulation applicable to classical mixed states, which are probability measures over a classical phase space. A pair of classical states is orthogonal iff the probability measures have disjoint support, and it can be shown that cloneable pairs are orthogonal in this sense. The similarity between these two theorems suggests that they are special cases of a more general theorem, and indeed, this is the case. Implicit in the proof of Lemma 3 of Clifton, Bub, and Halvorson (2003) is a proof that a pair of pure states of a C^*-algebra are cloneable iff they are orthogonal. Barnum, Barrett, Leifer, and Wilce (2006) prove a theorem of much greater generality. Within the convex sets framework (see section 4, below), they prove that a finite set of states on a compact, finite-dimensional state space is cloneable if and only if it is a jointly distinguishable set of states.

Some, notably Fuchs (2002) Spekkens (2001, 2007), have found in this fact—that some phenomena that might be thought to be distinctively quantum can be reproduced in a classical theory by imposing restrictions on state preparation—encouragement for the view that quantum probabilities are just like classical probabilities, epistemic probabilities bound

up with limitations on state preparation. A natural alternative is to conclude that those features that can be reproduced in an essentially classical setting ought not to have been considered distinctively quantum in the first place. This, is, I think, the right lesson to draw. If this is right, then we must seek deep distinctions between the classical and the quantum elsewhere. To anticipate a conclusion to be drawn below, a case can be made that Schrödinger (1936) was right to locate the essential difference between the classical and the quantum in its treatment of combined systems. However, Schrödinger's conclusion that it is entanglement that distinguishes the quantum from the classical requires qualification—as we shall see.

In this paper, I will compare and contrast two approaches to the construction of neutral frameworks in which theories can be compared. The first is the algebraic framework, which begins with an algebra, among the elements of which are included the observables of the theory. The second is the operational approach, which motivates the introduction of the convex sets framework. I will end with a discussion of one particular model, the toy theory devised by Rob Spekkens (2001, 2007), which, with minor modifications, fits neatly within the convex sets framework and which displays, in an elegant manner some of the similarities and differences between classical and quantum theories.

2   Algebraic frameworks

Clifton, Bub, and Halvorson (2003) (henceforth CBH) undertook the task of characterizing quantum mechanics in terms of information-theoretical constraints. They adopted a framework in which a physical theory is associated with a C^*-algebra, the self-adjoint elements of which represent the bounded observables of the theory. For the definition of a C^*-algebra, see the appendix; for our purposes it suffices to know that the set of all bounded operators on a Hilbert space is a C^*-algebra, as is any subalgebra of these that is closed under the operation of taking adjoints, and is complete in the operator norm. Moreover, the set of all bounded, continuous complex-valued functions on a classical phase space is a C^*-algebra, as is the set of all bounded, measurable complex-valued functions on a classical phase space. Thus, classical mechanics also admits of a C^*-algebra representation, as was shown by Koopman (1931) . The difference between the quantum case and the classical case is that, in the classical case, the algebra is abelian.

A *state* on a C^*-algebra $\mathfrak{A}$ is a positive linear functional $\rho : \mathfrak{A} \to \mathbb{C}$, normalized so that $\omega(I) = 1$. For self-adjoint A, the number $\omega(A)$ is to be interpreted as the expectation value of the observable corresponding to A, in state ω. The set of states is a convex set: for any states ρ, σ, and any real $\lambda \in (0,1)$, the functional defined by

$$\omega(A) = \lambda \, \rho(A) + (1 - \lambda) \, \sigma(A)$$

is also a state, a mixture of ρ and σ. A state that is not a mixture of any two distinct states is called *pure*. General state evolution is represented by completely positive norm-preserving linear maps, also known as non-selective operations.

A state is *dispersion-free* iff $\rho(A^2) = \rho(A)^2$ for all self-adjoint A. Any dispersion-free state is pure. It can be shown that a C^*-algebra is abelian iff all its pure states are dispersion-free. It can be also shown that a theory involving an abelian C^*-algebra admits of an essentially classical representation, in which the states are probability distributions on the set of its pure states (see Kadison and Ringrose 1983, Thm. 4.4.3). Thus within the C^*-algebraic framework, the classical theories are those whose algebras are abelian.

Within this framework, CBH characterize quantum theory via three properties of the algebra:

i). Algebras associated with distinct physical systems commute.

ii). Any individual system's algebra of observables is noncommutative.

iii). Spacelike separated systems at least sometimes occupy entangled states.

As CBH (2003, 1570) point out, the first two conditions entail that the state space of a composite systems contains nonlocally entangled states. What the third requirement is meant to do is to guarantee that these states are physically accessible. Thus, CBH allow for theories in which the set of preparable states is a proper subset of the full state space of the algebra of observables. If the set of preparable states is taken to be the full set of states of a C^*-algebra, then the third condition is redundant.

Though CBH describe their conditions as "definitive of what it means to be a quantum theory in the most general sense" (1563), it should be noted that these three properties do not suffice to characterize quantum mechanics. In a quantum theory, there are no states that are dispersion-free in all observables. This is not entailed by CBH's conditions, as can be shown by the following simple example. Let A, B be two separated systems, and associate with each of these the algebra $M(\mathbb{C}) \oplus M(\mathbb{C}^2)$— that is, the algebra of 3×3 complex matrices of the form:

$$\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & \gamma \\ 0 & \delta & \epsilon \end{pmatrix}$$

Associate with the composite system the algebra that is the tensor product of the algebras associated with A and B.

All three of CBH's conditions are satisfied. However, unlike either quantum mechanics or classical mechanics, some of the pure states are dispersion-free, and some are not. The state corresponding to the vector,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

is an eigenvector of every observable. To ensure that the state space of our theory is quantum mechanical, some additional condition is needed. Plausible candidates are symmetry conditions; one might impose, for example, the condition that, for any pair of pure states ρ, σ on

4

$\mathfrak{A}$ (or $\mathfrak{B}$), there is an automorphism of the algebra $\mathfrak{A}$ taking ρ into σ. This condition would entail that either all pure states are dispersion-free (in which case the algebra is abelian, and the state space is a classical simplex), or that none are (as in the quantum case).

The virtues of the C^*-algebraic formulation are that there is a rich and well-worked out theory of C^*-algebras, and that both classical and quantum theories—including quantum mechanics and quantum field theories—are readily formulated in such terms. This rich theory comes at a price, however. The assumption that the set of observables be the self-adjoint part of a C^*-algebra requires that sums and products of observables be well-defined even when the observables are incompatible ones. When self-adjoint operators A, B fail to commute, the product AB will not correspond to any observable; nevertheless, our definition of state requires that states assign numbers to such products, numbers that are not interpretable as expectation values of the results of measurement. The embedding of our observables into an algebra imposes non-trivial algebraic relations between expectation values assigned to observables.

It would be difficult to argue—or at least, at this point nobody knows how to argue—that any plausible physical theory would admit of a C^*-algebraic formulation. If we further require that the set of preparable states be the full state space of some C^*-algebra, then the C^*-algebraic framework becomes decidedly too restrictive. Halvorson (2004) discusses the case of a theory that he calls the Schr*dinger theory, in which elementary systems are like quantum systems, but in which entangled states decay into mixtures when the systems are separated. Such a theory is locally quantum, but admits of no Bell-inequality violating correlations. Such a theory was suggested by Schrödinger (1936), who pointed out that at the time there was little in the way of experimental evidence of nonlocally entangled states. Though there is now abundant evidence of nonlocal entanglement, it does not seem that the Schr*dinger theory is one that could, or should, have been ruled out in advance of experiment. Moreover, we would like a framework in which we can consider some admittedly artificial constructions, such as Spekkens' toy theory, discussed in section 5, below. As Halvorson (2004) has shown, the state space of the Spekkens theory is not the state space of a C^*-algebra.

Within the algebraic approach, one can also consider weakening of the algebraic assumptions. One can, for example, consider Jordan-Banach (JB) algebras, or Segal algebras (see Halvorson (2004) for definition and discussion), both of which contain C^*-algebras as special cases. Seeking a further widening of the algebraic framework, with constraints limited to those that, arguably, any reasonable physical theory must share, one is led naturally to effect algebras.[1]

An effect algebra is meant to represent the set of yes-no tests that can be performed on a physical system. It contains distinguished elements 0 and u (the unit element), representing

---

[1]The notion of an effect algebra has occurred in the writings of many authors working on the foundations of quantum mechanics. The presentation here is based on that of Beltrametti and Bugajski (1997). It should be pointed out that, despite the name, an effect algebra is not an algebra: mutiplication is not defined, and addition is a partial operation.

the two trivial tests whose outcomes, independent of the state, are 'no' and 'yes,' respectively. There is a *partial* operation $\oplus$. $a \oplus b$ is defined when a and b can represent alternative outcomes of a single experiment. The partial operation $\oplus$ is assumed to be symmetric and associative. Moreover, it is assumed that, for each effect a, there is a complementary effect $\bar{a}$ such that $a \oplus \bar{a} = u$. As a final condition: $a \oplus u$ is defined only when $a = 0$. A *resolution of the identity* is a set $\{a_i, ..., a_n\}$ such that

$$\sum_{i=1}^{n} a_i = u.$$

Associated with an n-outcome experiment is a resolution of the identity representing its possible outcomes.

An element a of a C^*-algebra $\mathfrak{A}$ is said to be *positive* iff $a = b^*b$ for some $b \in \mathfrak{A}$. We can define a partial order $\leq$ on $\mathfrak{A}$ by: $a \leq b$ iff $a + c = b$ for some positive c. The set of positive elements a of $\mathfrak{A}$ such that $0 \leq a \leq I$ forms an effect algebra, with $a \oplus b$ defined to be $a + b$ when $0 \leq a + b \leq I$, and undefined otherwise.

The conditions defining an effect algebra, unlike those defining C^*-algebras or JB-algebras, have clear significance in terms of their intended association with experiments. This gives us a broad framework within which we add conditions to delimit interesting classes of physical theories.

Though restricting ourselves to the C^*-algebraic framework excludes some theories that we might want to consider, such a framework nonetheless has a role to play. As mentioned, the C^* framework is broad enough to include both classical and quantum theory. Moreover, it lends itself readily to hybrid theories, as it permits hybrid systems, composite systems having both classical and quantum subsystems. This is useful for considering what can be done by a combination of classical and quantum information processing. Moreover, a deep theorem by Alfsen, Hanche-Olsen, and Shultz, discussed in section 4, below, shows that the C^*-assumption entails that the simplest systems are equivalent either to a classical bit or a qubit. This is an indication that the theory is too restrictive to be regarded as a *general* framework for physical theories, and suggests, rather that it is the minimally general framework that is broad enough to include both the classical and the quantum.

3   The operational approach

Lucien Hardy (2001, 2002) sets out to characterize quantum mechanics within a framework whose basic concepts are framed in terms of performable operations. Hardy invites the reader to imagine concrete devices: state preparation devices, transformation devices, and measurement devices, and characterizes quantum mechanics in terms of relations between possibilities of state preparation, transformation, and measurement. States, on this approach, are associated with preparation devices, and are regarded as compendia of probabil-

ities regarding the outcomes of any measurement that can be performed. Similar discussions can be found in Holevo (1982), Ludwig (1983), and D'Ariano (2006).

To take such concepts as basic may seem to smack of an operationalism that eschews on principle any talk of a reality not directly accessible to observation and manipulation. This is not, however, a necessary concomitant of this approach. If one wishes to construct a framework capable of embracing as wide an array as possible of physical theories, it can be a useful strategy to restrict one's attention to the features that *any* physical theory might be expected to have. Whatever else it might do, a physical theory should specify a set of possible states of a physical system, and say something about operations that can be performed on systems and experiments that can be done. The states, however (or whether) they may be conceived ontologically, ought, in the context of the physical theory, to yield probabilities for outcomes of any experiment that might be performed. The operationalism associated with this approach can be thought of as a *methodological operationalism*. Someone who takes as the goal of theorizing an account of a reality existing independently of us and our experimental manipulations might nonetheless wish to adopt a framework that begged no questions about what that reality is like. The framework itself is neutral between an attitude on which operational concepts are the only meaningful ones, and one that seeks to use them as a springboard for theorizing about the nature of physical reality.

It is also neutral between an attitude that rests content with taking the concepts of preparation, transformation, and experiment as primitives and does not seek to explain them within the theory to be constructed, and an attitude that seeks to "close the circle," to borrow a phrase from Abner Shimony (1993), by construing the experimental apparatus as among the physical systems to be dealt with by the theory, and the processes by which states are prepared and experimental results obtained to be among the dynamical evolutions allowed by the theory. It is this latter, of course, that has proved problematic in connection with quantum mechanics!

States are associated with preparation procedures, and yield probabilities over the results of experiments, given the preparation procedure. These probabilities might or might not represent the epistemic state of some human agent. An agent might be in doubt about what probabilities are most appropriate to associate with a given preparation procedure. She might, for example, be uncertain whether a given combination of a preparation and a two-outcome experiment yielded equal probabilities for both outcomes. The conjecture that the probabilities are equal can be subjected to test; by repetition our agent can satisfy herself, to any degree of precision and confidence required, about what probabilities she ought to associate with a given combination of preparation and experiment.

Proponents of a subjectivist interpretation of probability in quantum mechanics sometimes assert that it is nonsensical to speak of an unknown quantum state (see, *e.g.* Fuchs and Schack 2004). And indeed, if the probabilities associated with a state had to be epistemic probabilities, then a state could be unknown only if an agent were ignorant of her own state of mind. However, it is surely not nonsensical to say that a system has been subjected to some state preparation procedure, though no-one knows what procedure was

applied. In quantum information theory situations are routinely invoked in which one agent has prepared a state, and another has to guess what state it is. This situation would not be fundamentally different if it were automated; we can imagine a machine subjecting a system to some preparation procedure, making a record of which procedure was applied.

The operational approach gives us the following picture: we have a set of states, associated with preparation procedures, a set of possible transformations of the states, and a set of possible experiments that can be performed. The mathematical framework that lends itself naturally to this picture is the convex sets framework, to be discussed in the next section.

4   The convex set approach

Given any two preparation procedures ρ, σ and any real number $\lambda \in [0, 1]$, there will be a third, according to which one utilizes either procedure ρ or σ, with probabilities λ and $1 - \lambda$, respectively. Thus, we expect the set of states of any reasonable physical theory to be convex: it will contain all mixtures of states that it contains.

The convex set approach starts with a convex state space Ω.[2] Extremal points of this state space — that is, states that cannot be expressed as a nontrivial mixture of other states — are called *pure*. An *affine linear functional* on Ω is a mapping $a : \Omega \to \mathbb{R}$ that respects mixtures; that is,

$$a(\lambda \, \rho + (1 - \lambda) \, \sigma) = \lambda \, a(\rho) + (1 - \lambda) a(\sigma).$$

for all $\rho, \sigma \in \Omega$, $\lambda \in (0, 1)$. Among affine linear functionals are the constant functionals, taking on the same value on every state; we single out the unit functional u and the zero functional 0. There is a natural partial ordering on $\mathcal{A}(\Omega)$, the set of affine linear functionals: $a \leq b$ iff $a(\omega) \leq b(\omega)$ for all $\omega \in \Omega$. If $0 \leq a \leq u$, a is called an *effect*. The set of effects on a state space Ω will be denoted by $E(\Omega)$. The notion of an effect generalizes the notion, familiar from the quantum context, of an effect as a positive operator with spectrum in $[0, 1]$. A resolution of the identity u in terms of effects generalizes the notion of a positive operator-valued measure (POVM).

The set of effects has a natural structure as an effect algebra in the sense of section (2) , and, conversely, the set of all probability measures on an effect algebra forms a convex state space. Relations between these frameworks are further discussed in Beltrametti and Bugajski (1997).

There is a natural affine structure on the set of effects: for any effects a, b, and any $\lambda \in (0, 1)$, the functional c defined by

$$c(\omega) = \lambda \, a(\omega) + (1 - \lambda) \, b(\omega)$$

[2]The presentation of the convex set framework presented here is heavily indebted to the presentations in Beltrametti and Bugajski (1997), and in Barnum *et al.* 2006.

is also an effect. Thus, we can distinguish between effects that can be written as a nontrivial mixture of other effects, and those that cannot. The former are called *mixed* effects, the latter, *pure*. A resolution of the identity in terms of pure effects is what corresponds, in the context, to a quantum projection-valued measure (PVM).

Effects map states to probabilities. Equivalently, we can think of states as mapping effects to probabilities. For any state $\omega \in \Omega$, there is an affine linear mapping $\hat{\omega} : E(\Omega) \to \mathbb{R}$ defined by

$$\hat{\omega}(a) = a(\omega).$$

It will be convenient, when defining the tensor product space, to work with these dual states, which map effects to probabilities.

To specify an observable A, one specifies a measure space $\langle \mathrm{sp}(A), \mathcal{M}_A \rangle$, and an affine mapping α_A from states to probability measures on $\langle \mathrm{sp}(A), \mathcal{M}_A \rangle$. The set $\mathrm{sp}(A)$ is the outcome space of A, the set of possible results of an A-experiment, $\mathcal{M}_A$, a σ-algebra of subsets of A, which are the measurable subsets of $\mathrm{sp}(A)$. For any measurable set $\Delta \in \mathcal{M}_A$, we can define $p_\Delta^A : \Omega \to \mathbb{R}$ by,

$$p_\Delta^A(\omega) = \alpha_A(\omega)(\Delta).$$

This is an affine mapping yielding the probability, in state ω, that a measurement of A will yield a result in Δ. Since $0 \leq p_\Delta^A(\omega) \leq 1$ for all $\omega \in \Omega$, p_Δ^A is an effect. Functions of an observable are readily definable; for any measurable function f on $\mathrm{sp}(A)$, $f(A)$ is defined to be the observable whose outcome space is $f(\mathrm{sp}(A))$, with the probability that $f(A) \in \Delta$ equal to the probability that $A \in f^{-1}(\Delta)$. A set $\{A_i\}$ of observables is said to be a *compatible* set iff there is an observable C and measurable functions $\{f_i\}$ on $\mathrm{sp}(C)$ such that $A_i = f_i(C)$. Where A is real-valued—that is, $\mathrm{sp}(A) \subseteq \mathbb{R}$—we will write $\omega(A)$ for the expectation value of A in state ω.

One also specifies a set of affine linear mappings of the state space into itself, representing the physically possible operations on the system. Given a transformation $T : \Omega \to \Omega$, we define the conjugate transformation $T^\dagger : E(\Omega) \to E(\Omega)$, which maps effects to effects, by

$$T^\dagger a(\omega) = a(T\omega)$$

for all $\omega \in \Omega$. Obviously, the identity will be among the possible transformations, and it is assumed that transformations can be composed. It is not assumed that all transformations are invertible.

To sum up: with a physical system is associated a triplet $\langle \Omega, \mathcal{O}, \mathcal{T} \rangle$, where Ω is a convex set representing the state space of the system, $\mathcal{O}$ is a set of observables on Ω, and $\mathcal{T}$ is a semi-group of transformations of Ω.

Given two physical systems $S = \langle \Omega, \mathcal{O}, \mathcal{T} \rangle$, $S' = \langle \Omega', \mathcal{O}', \mathcal{T}' \rangle$, there is a plurality of choices of state space for the composite system. At minimum the set of experiments that can be performed on the combined system should include experiments performed on S, S' separately. Consider the set of experiments in $\mathcal{O}$ with outcome space $\{0, 1\}$ (the set of 'yes-no' experiments). For any such experiment A, there is an effect $a \in E(\Omega)$ such that,

for any $\omega \in \Omega$, $a(\omega)$ is the expectation value of A in state ω. Since we want to consider theories in which the set of permitted observables falls short of all observables definable on the state space Ω, we want to leave open the possibility that not every effect in $E(\Omega)$ yields the expectation value for some yes-no experiment. Let $\mathcal{E}(\Omega)$ be the set of effects on Ω that *do* figure in this way in some experiments in $\mathcal{O}$. We specify a state in a tensor product space $\Omega \otimes \Omega'$ by specifying, for every pair $a \in \mathcal{E}(\Omega)$, $b \in \mathcal{E}(\Omega')$, an expectation value for the product observable $a \otimes b$: $a \otimes b$ is measured by measuring a and b separately, and the outcome of the $a \otimes b$ experiment is taken to be the product of the component experiments. A state ω in $\Omega \otimes \Omega'$ is defined to be an affine bilinear mapping $\omega : \mathcal{E}(\Omega) \times \mathcal{E}(\Omega') \to [0,1]$, normalized so that $\omega(u, u') = 1$.[3] Given $\alpha \in \Omega, \beta \in \Omega'$, we define the *product state* $\alpha \otimes \beta$ by

$$(\alpha \otimes \beta)(a, b) = \alpha(a)\,\beta(b)$$

for all $a \in \mathcal{E}(\Omega)$, $b \in \mathcal{E}(\Omega')$. The *minimal tensor product* $\Omega \otimes_{sep} \Omega'$ is defined to be the convex hull of the set of product states. A state in $\Omega \otimes \Omega'$ is said to be *separable* iff it is in this minimal tensor product; *entangled*, otherwise. The *maximal tensor product* $\Omega \otimes_{max} \Omega'$ contains all normalized affine bilinear mappings $\omega : E(\Omega) \times E(\Omega') \to \mathbb{R}$. In general, a theory may specify a tensor product space that is a convex proper subset of the maximal tensor product space. The quantum tensor product, since it includes entangled states, exceeds the minimal tensor product. It falls short of the maximal tensor product, however; see Barnum *et al.* (2005) for discussion.

The possible transformations of the composite system should, at minimum, include transformations performed separately on the individual systems. For any transformations $T \in \mathcal{T}, T \in \mathcal{T}'$, define the transformation $T \otimes T' : \Omega \otimes_{max} \Omega' \to \Omega \otimes_{max} \Omega'$ by

$$(T \otimes T')\omega(a, b) = \omega(T^\dagger a, T'^\dagger b). \tag{1}$$

We will require our tensor product space to be closed under $T \otimes T'$, for all $T \in \mathcal{T}, T \in \mathcal{T}'$.

In his opening lecture at the Boston conference, Abner Shimony raised the question whether there could be entanglement without potentiality. If by 'potentiality' we mean that there are pure states in which some pure observables do not take on definite values, then it becomes a theorem in the convex set framework that there is no entanglement without potentiality.

Theorem 1 *Let A, B be physical systems, with state spaces Ω_A, Ω_B. If ψ is an entangled pure state in $\Omega_A \otimes \Omega_B$, then there exists a pure effect $a \in E(\Omega_A \otimes \Omega_B)$ such that $0 < \psi(a) < 1$.*

Proof. Let ψ_A, ψ_B be the marginals of ψ, defined by

$$\psi_A(a) = \psi(a, u_B)$$
$$\psi_B(b) = \psi(u_A, b)$$

---

[3]Here we depart slightly from Barnum *et al.* (2006), who define a tensor product state as an affine bilinear mapping defined on all pairs of effects on the component spaces.

for all effects $a \in E(\Omega_A)$, $b \in E(\Omega_B)$. Here u_A and u_B are the unit effects in $E(\Omega_A)$ and $E(\Omega_B)$, respectively. By Lemma 3 of Barnum *et al.* (2006), if either marginal is pure, then ψ is a product state. Since, by assumption, ψ is entangled, both marginals are mixed.

As ψ_A is a mixed state, let it be a nontrivial mixture of distinct states $\rho, \sigma \in \Omega_A$. If $\rho(a) = \sigma(a)$ for all pure effects a, then $\rho(a) = \sigma(a)$ for *all* effects a, and $\rho = \sigma$. Therefore, there exists a pure effect a such that $\rho(a) \neq \sigma(a)$. It follows from this that $\psi_A(a)$ is not equal to 0 or 1, as it could take on one of these extremal values only if both $\rho(a)$ and $\sigma(a)$ took on the same extremal value. Thus, $\psi(a, u_B) = \psi_A(a) \in (0, 1)$. ∎

The state space of an algebra, be it a C^*-algebra, a JB-algebra, or a Segal algebra, is a convex set. Alfsen and Shultz (1978) characterize the state spaces of JB-algebras among convex sets, and Alfsen, Hanche-Olsen, and Shultz (1980) characterize those convex sets that are the state spaces of C^*-algebras. This gives us a way of linking the convex-set approach to the algebraic approach. We will here not go into the full details of these characterizations, which are presented in Alfsen and Shultz (2003), but focus instead on a central feature of these characterizations.

First, some definitions. A *face* of a convex set Ω is a convex subset $F \subseteq \Omega$ that is such that, if any state in F is a mixture of states $\rho, \omega \in \Omega$, then ρ and ω are also in F. The face generated by ρ, ω, is the smallest face containing ρ and ω. In a simplex, the face generated by two pure states is just the one-dimensional simplex consisting of ρ, ω, and all mixtures of the two. The face generated by two pure quantum states, represented by state vectors $|\psi\rangle$, $|\phi\rangle$, consists of a set of pure states comprising all linear superpositions of $|\psi\rangle$ and $|\phi\rangle$, together with all mixtures of these. Thus, it is affinely isomorphic to the Bloch sphere. A face $F \subseteq \Omega$ is *norm exposed* iff there is an effect $a \in E(\Omega)$ such that F is precisely the set of states on which a takes on the value 0. A *Hilbert ball* is a convex set of states that is affinely isomorphic to the closed unit ball of some (finite- or infinite-dimensional) real Hilbert space. The Hilbert ball B^n is the closed unit ball in $\mathbb{R}^n$.

If Ω is the state space of a JB-algebra, or, as a special case, a C^*-algebra, then the faces generated by pairs of pure states take on a simple form.

Theorem 2 *If a convex set Ω is the state space of a JB-algebra, then for any distinct pure states $\rho, \omega \in \Omega$, the face generated by ρ, ω is a norm exposed Hilbert ball.*

For proof, see Alfsen *et al.* (1978), or Alfsen and Shultz (2003, Proposition 9.10). For discussion of the significance of this result for the Spekkens and Schr*dinger theories, see Halvorson (2004).

Theorem 3 *If a convex set Ω is the state space of a C^*-algebra, then, for any distinct pure states $\rho, \omega \in \Omega$, the face generated by ρ, ω is norm exposed, and is either B^1 or B^3.*

See Alfsen *et al.* 1980, or Alfsen and Shultz (2003, Theorem 11.59).

Theorem 3 entails that, in the state space of a C^*-algebra, the face generated by a pair of pure states ρ and ω is either the one-dimensional simplex consisting of ρ and ω and

all mixtures of these, or else is affinely isomorphic to the Bloch sphere. Thus, within the C^*-algebra framework, the state space of quantum mechanics can be completely characterized by the condition that any two pure states ρ, σ are connected by a continuous path through the set of pure states within the face generated by ρ and σ — a cousin of Hardy's (2001) continuity axiom. For a classical theory, on the other hand, the set of pure states is totally disconnected: any state that is "intermediate" between two classical states is a mixture of the two.

Consider a system S, with state space Ω_S, and let A be some subsystem of S. The state space Ω_A will be $\Omega_A \otimes \Omega_E$, where E is the "environment" of A, that is, everything in S that is not included in the subsystem A. For any pure state $\varepsilon \in \Omega_E$, the set $\Omega_A \otimes \varepsilon$ is a face of Ω that is affinely isomorphic to Ω_A. Thus, the state space of any subsystem of S will be affinely isomorphic to a face of Ω_S. The simplest subsystems will be those whose state spaces are the faces generated by a pair of distinguishable states. Theorem 3 therefore says that, in the state space of a C^*-algebra, these simplest systems will have state spaces that are equivalent either to the state space of a classical bit, or to that of a qubit. Composite systems might be hybrids of the two.

This is another reason for regarding the C^*-algebraic framework as too restrictive for a framework meant to embrace physical theories in general. This limitation is also a strength. Though it can make no claims to physical generality, the C^*-framework is a useful one for comparing classical and quantum theories, and for discussing information processing and transmission possibilities afforded by a combination of classical and quantum operations, precisely because it is a framework in which the simplest systems are constrained to be either classical bits or qubits.

The work of Clifton, Bub and Halvorson (2003) was motivated by a suggestion, which they attribute to Chris Fuchs and Gilles Brassard, that quantum mechanics can be derived from an appropriately chosen set of cryptographical principles. The C^*-algebraic framework is too restrictive a starting point for such a project. A more appropriate starting point is the convex set approach. The question arises: Is there a set of principles, with clear operational (cryptographic or otherwise) significance, that distinguishes quantum state spaces among the convex sets?

Alfsen and Shultz (1978) and Araki (1980) tell us how to single out, among convex sets, those that are the state spaces of JB algebras. Alfsen *et al.* (1980) tell us how to distinguish, among these, those that are state spaces of C^*-algebras. The further condition that the face generated by any pair of pure states is affinely isomorphic to the Bloch sphere distinguishes the quantum state spaces among these. These theorems, therefore, could form a useful starting-point for a project that seeks to characterize quantum mechanics in terms of information-theoretic principles. Araki's work may be useful in this regard because his conditions are expressed in terms more directly connected with operations than those of Alfsen *et al.*. The information-theoretical principles proposed by Clifton, Bub and Halvorson do not suffice to characterize quantum mechanics, because, even within the C^*-framework, they do not entail that there are *no* states that are dispersion-free in all observables. The

task of characterizing quantum mechanics in terms of possibilities of information processing and transfer remains an open one.

5   The Spekkens Toy Theory

Consider a ball that can be in one of four boxes. A state of this system is specified by four probabilities (p_1, p_2, p_3, p_4). Extremal states are those in which the ball is definitely in one of the boxes. The state space is therefore a simplex consisting of four pure states and mixtures of these, and can be depicted as a tetrahedron.

Now, suppose we impose a restriction on state preparations, to the effect that the probability of the ball being in any one of the boxes cannot exceed $\frac{1}{2}$. This restriction chops the vertices off our tetrahedron, leaving an octahedron whose vertices are the midpoints of each edge of our original tetrahedral state space (see Figure 1). These six states, which are the extremal points of the new state space, are those in which the probability is equally divided between two of the boxes. These are the pure states of elementary systems in the toy theory constructed by Spekkens (2004, 2007), the 'states of maximal knowledge,' as Spekkens puts it. In Spekkens' theory, arbitrary mixtures of these pure states are not permitted; Spekkens permits as a mixture only the state in which the probability is equally divided among all four boxes. This seems to be an inessential restriction. Moreover, it is one of questionable coherence, if the states are interpreted, as Spekkens would have them be, as possible states of knowledge. A theory whose states are intended as states of knowledge ought to include states of less-than-maximal knowledge. Suppose Alice's belief state is a Spekkens maximal-knowledge state about an elementary system of Spekkens' theory, and that Bob knows nothing about the system except that Alice's belief state is either ρ_1 or ρ_2, and that Bob's degree of belief in the former is λ, in the latter $1 - \lambda$. Then Bob's epistemic state about the system is the corresponding mixture of states of maximal knowledge.

We will, therefore, take as a state space the convex hull of Spekkens' pure states, as suggested by Halvorson (2004). This is the octahedron we have just described, depicted again in Figure 2. Note that, though we started with a simplex, the state space we have ended up with is *not* a simplex. In particular, the maximally mixed state in the center (call it ω_0), can be decomposed as an equally weighted mixture of $\{x+, x-\}$, or of $\{y+, y-\}$, or of $\{z+, z-\}$.

Given this state space, one can define observables that do not extend to observables on our original tetrahedral simplex, and hence do not correspond to observable quantities in the Spekkens theory. There is, for example, an affine functional on the octahedron that takes on the value 0 on the $(x+, y+, z+)$ face, and the value 1 on the $(x-, y-, z-)$ face. Were such a function extended to the embedding tetrahedron, it would have to take on a value greater than 1 on vertex 4. The set of permitted observables will be a proper set of

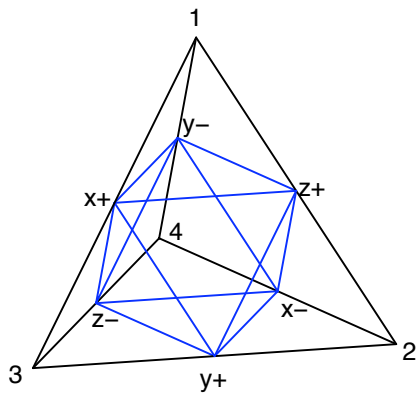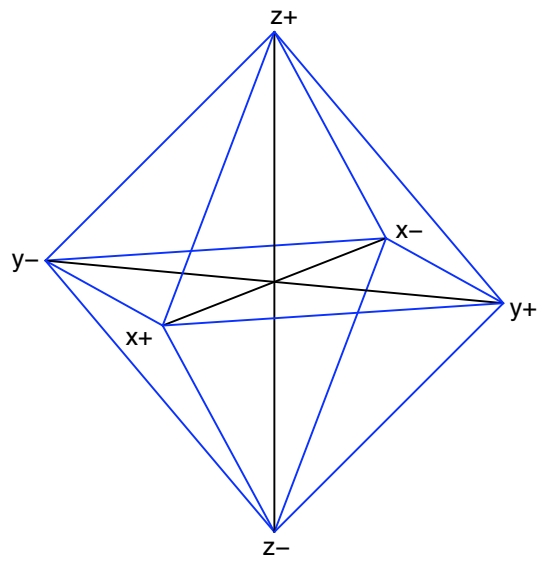Figure 1: The state space of the Spekkens toy theory

Figure 2: The octahedron re-oriented.

the observables definable on the state space. Define the affine functionals:[4]

$$a_x(x+) = 1, \quad a_x(x-) = 0, \quad a_x(y+) = a_x(z+) = \tfrac{1}{2}$$

$$a_y(y+) = 1, \quad a_y(y-) = 0, \quad a_y(x+) = a_y(z+) = \tfrac{1}{2} \qquad (2)$$

$$a_z(z+) = 1, \quad a_z(z-) = 0, \quad a_z(x+) = a_z(y+) = \tfrac{1}{2}$$

We define the observables $\{\sigma_\mu \,|\, \mu = x, y, z\}$, analogous to the quantum spin observables. The outcome space of σ_μ is $\{-1, +1\}$, and the probability in state ω of obtaining outcome $+1$ upon measuring σ_μ is $a_\mu(\omega)$. Take as our set of observables these three, together with the identity observable I, which yields with certainty the result $+1$ in any state.

Now suppose we start with the Bloch sphere, which represents the state space of a single qubit (equivalently, a spin-$\tfrac{1}{2}$ particle), and impose a restriction that the only preparable pure states are spin eigenstates in one of three mutually orthogonal directions: our state preparation device is a Stern-Gerlach apparatus whose mounting permits only these three orientations. Call these directions (x, y, z). Label the six preparable pure states $\{x+, x-, y+, y-, z+, z-\}$. The convex hull of these states is once again our octahedron. We place the same restrictions on measurement: only spin measurements in the same three directions are permitted. What we thereby obtain is a state space and set of observables that are isomorphic to those of elementary systems of the Spekkens theory. Thus, the same structure, consisting of state space plus set of permitted observables, is obtainable on the one hand by starting from a simplex, that is, an essentially classical state space, and imposing restrictions on state preparation, and on the other by starting from a quantum state space and imposing restrictions on state preparation and measurement. Moreover, as Spekkens (2001) demonstrates in detail, many of the features of quantum theories that one might be tempted to think of as characteristic of quantum mechanics already exist within this reduced state space.

Though we have obtained the same structure of state space and observables in the two cases, the resulting theories, will nevertheless be different if the permitted transformations of the state spaces are different, and, since we require tensor product spaces to be closed under transformations of the component systems, this will have consequences for the state spaces of composite systems. Unitary transformations of the quantum state space of a qubit correspond to rotations of the Bloch sphere. The transformations permitted in the Spekkens theory are permutations of the four boxes. We thus have two natural choices for the group of permitted transformations of our octahedral state space: $\mathcal{T}_S$, the group of transformations resulting from permutations of the four boxes, or $\mathcal{T}_q$, the set of rotations that take the octahedron into itself. Each of these groups has 24 distinct elements, but, as we shall see, the two groups are not isomorphic.

---

[4]The affine dimension of the Spekkens state space, like that of the Bloch sphere, is four. That is, to specify any affine function, it suffices to specify its value on a set of four linearly independent states. We could have chosen the same set of four for the definition of each of these functionals, say, $\{x+, y+, z+, \omega_0\}$.

Obviously, any automorphism of the state space will have to leave the unique totally mixed state ω_0 fixed. To specify an automorphism, therefore, it suffices to specify its action on the set of pure states $\{x+, y+, z+\}$, which any automorphism will take into a sets of pure states containing no pair of opposite states. Both groups of transformations include cyclic permutations of $x+, y+, z+$. Let T_1 be the transformation $\langle x+ \rightarrow y+, y+ \rightarrow z+, z+ \rightarrow x+ \rangle$. This corresponds, in the quantum domain, to a rotation of $2\pi/3$ about the axis pointing in the direction of $\hat{x} + \hat{y} + \hat{z}$. It is achieved by the Spekkens transformation that leaves box 4 invariant and permutes the other three by $\langle 1 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rangle$. Let T_2 be the transformation that takes $x+$ to $x-$ and $y+$ to $y-$, and leaves $z+$ invariant. This is achieved, in the group of rotations, by a rotation of π about the z-axis, and, in the group of Spekkens permutations, by the swap $\langle 1 \leftrightarrow 2, 3 \leftrightarrow 4 \rangle$. Rotations about the other axes can be achieved by combining this rotation with cyclic permutations of the axes; we thereby achieve any transformation that inverts two axes while leaving the other invariant.

Let T_3^S be the transformation that leaves $z+$ invariant while swapping $x+$ and $y+$. This belongs to the Spekkens group of transformations, as it is effected by the swap $1 \leftrightarrow 2$. As it involves parity inversion, it is not a transformation resulting from a rotation of the Bloch sphere, and so does not belong to the group of quantum transformations. We do, however, have in the quantum group a transformation that leaves $z+$ invariant and takes $x+$ to $y+$ and $y+$ to $x-$; this is a counterclockwise rotation of $\pi/2$ about the z-axis. Call this T_3^q.

The reader will be able to verify that the sets $\{T_1, T_2, T_3^S\}$, $\{T_1, T_3^q\}$ suffice to generate the Spekkens and quantum groups of transformations, respectively. We summarize the generating sets of the Spekkens and quantum transformations in Table 1. For each automorphism T of the state space Ω, we include also the conjugate automorphism $T^\dagger$.

Table 1: Generators of $\mathcal{T}_S$ and $\mathcal{T}_q$.

	T			$T^\dagger$		
	$x+$	$y+$	$z+$	σ_x	σ_y	σ_z
T_1	$y+$	$z+$	$x+$	σ_z	σ_x	σ_y
T_2	$x-$	$y-$	$z+$	$-\sigma_x$	$-\sigma_y$	σ_z
T_3^S	$y+$	$x+$	$z+$	σ_y	σ_x	σ_z
T_3^q	$y+$	$x-$	$z+$	$-\sigma_y$	σ_x	σ_z

Let us now begin to construct tensor product spaces for the two theories. A state is specified by specfying its value on the effects $a_\mu \otimes a_\nu$, $\mu, \nu \in \{0, x, y, z\}$ (where we take a_0 to be the unit effect u), or, equivalently, by specifying the expectation value it assigns to the observables $\sigma_\mu \otimes \sigma_\nu$, $\mu, \nu \in \{0, x, y, z\}$, where σ_0 is the observable yielding outcome 1 in all states. The minimal tensor product contains all product states and convex combinations of product states. In addition to these, Spekkens allows states whose marginals are maximally mixed, but with maximal correlations between the states of the component systems. One

Figure 3: The maximally correlated state χ.
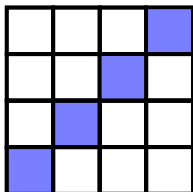
such state is the state (call it χ) in which it is certain that both balls are in boxes with the same label (that is, if ball 1 is in its box 1, then ball 2 is in its box 1 also, *etc.*), with probabilities equally divided among the four possibilities. This state is symbolized in Figure 3. In this 4×4 array, system 1's box states are along the vertical axis, and system 2's box states along the horizontal; for the joint system to occupy the box (i, j) thus represents a state in which ball 1 is in box j and ball 2 is in box i. The state χ is defined by

$$\chi(\sigma_\mu \otimes \sigma_\nu) = \delta_{\mu\nu}, \quad \mu, \nu \in \{0, x, y, z\}. \tag{3}$$

That is, the three observables $\{\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y, \sigma_z \otimes \sigma_z\}$ take on dispersion-free values, equal to 1. Note that, since, quantum-mechanically,

$$(\sigma_x \otimes \sigma_x)(\sigma_y \otimes \sigma_y) = -\sigma_z \otimes \sigma_z, \tag{4}$$

this is not a state in the standard quantum tensor product.

If we include the state χ in our tensor product space, then all states obtainable from it by permutations of the boxes will also be in the tensor product space (we can restrict our attention to permutations on one subsystem, as we will not obtain any new states by considering permutations of both subsystems). This gives us a set of 24 pure, maximally entangled states, each of which can be represented by a 4×4 array with one element of each row and each column shaded. For example, since we have the state χ, application of T_1 produces the state in which $\sigma_x \otimes \sigma_x$ and $\sigma_y \otimes \sigma_y$ have definite value -1, and $\sigma_z \otimes \sigma_z$ has definite value $+1$, and, by combining T_1 with T_2, we obtain common eigenstates of these observables for other combinations of eigenvalues that multiply to $+1$ (note that, since the only sign-changing element of our generating set for the Spekkens group is T_1, and this changes two signs, we will not be able to change the sign of the product of the three eigenvalues). These four states are depicted in Figure 4.

By application of T_3^S to χ, we get the $+ + +$ common eigenstate of $\{\sigma_y \otimes \sigma_x, \sigma_x \otimes \sigma_y, \sigma_z \otimes \sigma_z\}$, and via applications of T_1 and T_2, the other eigenstates with eigenvalues that multiply to $+1$. These states are depicted in Figure 5. Continuing the process, we obtain, for each of the rows and each of the columns of the following array of observables, four common eigenstates of the observables in that row (column), corresponding to each combination of eigenvalues multiplying to $+1$.

18

Figure 4: Common eigenstates of $\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y, \sigma_z \otimes \sigma_z$.
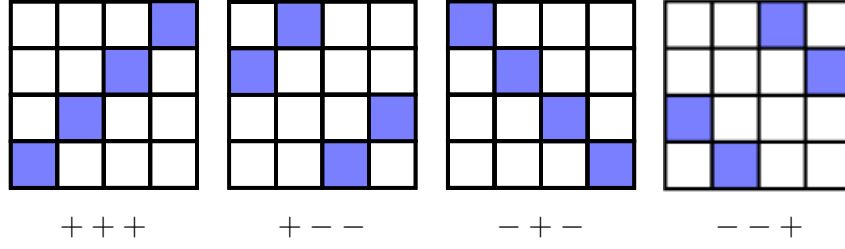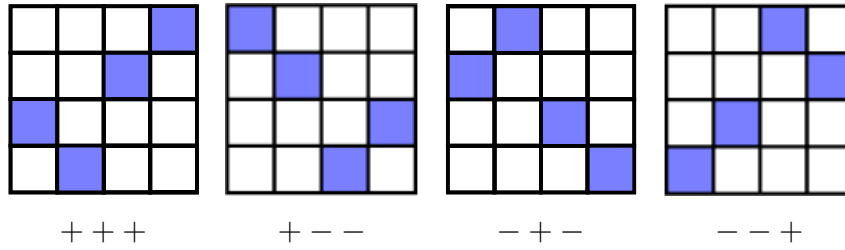

$$+ + + \qquad + - - \qquad - + - \qquad - - +$$

Figure 5: Common eigenstates of $\sigma_x \otimes \sigma_y, \sigma_y \otimes \sigma_x, \sigma_z \otimes \sigma_z$.

$$+ + + \qquad + - - \qquad - + - \qquad - - +$$

$$\sigma_x \otimes \sigma_x \quad \sigma_y \otimes \sigma_y \quad \sigma_z \otimes \sigma_z$$

$$\sigma_z \otimes \sigma_y \quad \sigma_x \otimes \sigma_z \quad \sigma_y \otimes \sigma_x \tag{5}$$

$$\sigma_y \otimes \sigma_z \quad \sigma_z \otimes \sigma_x \quad \sigma_x \otimes \sigma_y$$

These states are the 24 pure entangled states of the Spekkens theory. Call the tensor product space whose pure states consist of the product states together these 24 maximally entangled states, Ψ_S. Note that the rows and columns of (5) are, in quantum theory, also sets of observables having joint eigenstates. In the quantum theory, however, eigenvalues of observables in a row will multiply, not to $+1$, but to -1.

Suppose that we start constructing a quantum-like tensor product space for our theory by starting with a pair of systems having our octahedral state space, and endowing it with a state that *is* in the quantum tensor product. For example, we could start with the state ξ, common to both the quantum and the Spekkens tensor product, defined by

$$\xi(\sigma_x \otimes \sigma_y) = \xi(\sigma_y \otimes \sigma_x) = \xi(\sigma_z \otimes \sigma_z) = +1$$

$$\xi(\sigma_\mu \otimes \sigma_\nu) = 0, \text{ all other combinations.} \tag{6}$$

19

Application of the quantum group of transformations again yields common eigenstates for each of the rows and columns of the array (5). For example, applying T_3^q to ξ yields the state ϕ^-, with

$$\phi^-(\sigma_x \otimes \sigma_x) = -1 \quad \phi^-(\sigma_y \otimes \sigma_y) = +1 \quad \phi^-(\sigma_z \otimes \sigma_z) = +1$$

Pairwise sign changes (combining applications of T_1 and T_2) yield the the other common eigenstates of $\{\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y, \sigma_z \otimes \sigma_z\}$, forming the set of Bell states $\{\phi^+, \phi^-, \psi^+, \psi^-\}$, which are defined by the conditions,

$$\phi^+(\sigma_x \otimes \sigma_x) = +1 \quad \phi^+(\sigma_y \otimes \sigma_y) = -1 \quad \phi^+(\sigma_z \otimes \sigma_z) = +1$$

$$\phi^-(\sigma_x \otimes \sigma_x) = -1 \quad \phi^-(\sigma_y \otimes \sigma_y) = +1 \quad \phi^-(\sigma_z \otimes \sigma_z) = +1$$

$$\psi^+(\sigma_x \otimes \sigma_x) = +1 \quad \psi^+(\sigma_y \otimes \sigma_y) = +1 \quad \psi^+(\sigma_z \otimes \sigma_z) = -1$$

$$\psi^-(\sigma_x \otimes \sigma_x) = -1 \quad \psi^-(\sigma_y \otimes \sigma_y) = -1 \quad \psi^-(\sigma_z \otimes \sigma_z) = -1,$$

$$(7)$$

together with the condition that, for $\mu \neq \nu$, $\phi^+(\sigma_\mu \otimes \sigma_\nu) = \phi^-(\sigma_\mu \otimes \sigma_\nu) = \psi^+(\sigma_\mu \otimes \sigma_\nu) = \phi^-(\sigma_\mu \otimes \sigma_\nu) = 0$. Let Ψ_q be the tensor product space whose set of pure entangled states is the closure of the set of Bell states under $\mathcal{T}_q \otimes I$. Like the Spekkens tensor product, Ψ_q contains 24 pure entangled states, consisting of 4 common eigenstates for each of the rows and columns of (5). It shares with Ψ_S the eigenstates of observables in the colums of the array; it differs from Ψ_S in having eigenvalues of observables in any row of (5) that multiply to -1.

An alternative tensor product (call it $\tilde{\Psi}_q$) can be constructed by including the state χ and taking the closure under the quantum set of transformations. In this tensor product, joint eigenstates of the rows of (5) will have eigenvalues multiplying to $+1$, with eigenvalues of the columns multiplying to -1; this set of entangled states are those obtainable from the quantum states by a parity inversion on one of the component systems. Choosing one or the other of these tensor products amounts to a choice of relative orientation on the two systems— that is, given an orientation of one, a choice of which orientation of the other will be regarded as the "same" orientation. We can also construct a tensor product by including *both* sets of entangled states.

Similar remarks apply to tensor products closed under the Spekkens transformations: there is one such space, Ψ_S, containing χ and eigenstates of all rows and columns of (5) with sets of eigenvalues multiplying to $+1$, another, $\tilde{\Psi}_S$, containing the Bell states and eigenstates of all rows and columns of (5) with eigenvalues multiplying to -1, and one that is the convex hull of $\Psi_S \cup \tilde{\Psi}_S$.

Which tensor product we use will have consequences for whether or not a Kochen-Specker obstruction will be forthcoming—it will be possible to construct a Kochen-Specker obstruction in the spaces Ψ_q and $\tilde{\Psi}_q$, but not in Ψ_S or $\tilde{\Psi}_S$.

Mermin's simple Kochen-Specker obstruction (Mermin 1990), applicable to the Hilbert space of a pair of qubits, begins with the observation that, in the following array, each row and each column consists of mutually compatible quantum observables.

$$\sigma_x \otimes I \qquad I \otimes \sigma_x \qquad \sigma_x \otimes \sigma_x$$

$$I \otimes \sigma_y \qquad \sigma_y \otimes I \qquad \sigma_y \otimes \sigma_y \tag{8}$$

$$\sigma_x \otimes \sigma_y \qquad \sigma_y \otimes \sigma_x \qquad \sigma_z \otimes \sigma_z$$

Because of the algebraic relations,

$$\sigma_x \, \sigma_y = -\sigma_y \, \sigma_x = i \, \sigma_z, \tag{9}$$

we have,

$$(\sigma_x \otimes \sigma_x) \, (\sigma_y \otimes \sigma_y) = -\sigma_z \otimes \sigma_z$$
$$(\sigma_x \otimes \sigma_y) \, (\sigma_y \otimes \sigma_x) = \sigma_z \otimes \sigma_z. \tag{10}$$

The product of the elements each row of the array (8) is the identity $I \otimes I$, as is the product of the elements of each of the first two columns, whereas the product of the elements of the third column is $-I \otimes I$. This means that, if we tried to assign definite values to these nine observables, satisfying the product rule that, whenever A, B are compatible observables, $v(AB) = v(A)v(B)$, we cannot succeed; the product of all nine of these values would have to be both $+1$ and -1.

Return now to the convex set framework and the tensor products we have constructed for pairs of octahedral state spaces. We have not introduced any notion of multiplication of observables, and it is not clear how to make sense, in the general setting, of multiplication of incompatible observables. Algebraic relations among compatible observables, however, *do* make sense. Recall that a set $\{A_i\}$ of observables is a compatible set if and only if there is an observable C such that each A_i is a function of C. This will give rise to functional relations among the observables A_i.

Consider, now, the four Bell states $\{\phi^+, \phi^-, \psi^+, \psi^-\}$. For each of these states, there is an affine functional that takes on the value 1 on that state and 0 on the others: define $\{a_{\phi^+}, a_{\phi^-}, a_{\psi^+}, a_{\psi^-}\}$ by,

$$a_{\phi^+}(\omega) = \tfrac{1}{4} \left(1 + \omega(\sigma_x \otimes \sigma_x) - \omega(\sigma_y \otimes \sigma_y) + \omega(\sigma_z \otimes \sigma_z)\right)$$

$$a_{\phi^-}(\omega) = \tfrac{1}{4} \left(1 - \omega(\sigma_x \otimes \sigma_x) + \omega(\sigma_y \otimes \sigma_y) + \omega(\sigma_z \otimes \sigma_z)\right)$$

$$\tag{11}$$

$$a_{\psi^+}(\omega) = \tfrac{1}{4} \left(1 + \omega(\sigma_x \otimes \sigma_x) + \omega(\sigma_y \otimes \sigma_y) - \omega(\sigma_z \otimes \sigma_z)\right)$$

$$a_{\psi^-}(\omega) = \tfrac{1}{4} \left(1 - \omega(\sigma_x \otimes \sigma_x) - \omega(\sigma_y \otimes \sigma_y) - \omega(\sigma_z \otimes \sigma_z)\right)$$

These four affine functionals sum to the unit functional,

$$a_{\phi+} + a_{\phi-} + a_{\psi+} + a_{\psi-} = u. \tag{12}$$

If they are effects—that is, if their values on *any* state are confined to the interval $[0, 1]$—then we will be able to define a discrete observable C with outcome space $\mathrm{sp}(C) = \{1, 2, 3, 4\}$, with probabilities for the four outcomes yielded by the four effects $\{a_{\phi+}, a_{\phi-}, a_{\psi+}, a_{\psi-}\}$. If there is such an observable C, the Bell states are a distinguishable set of states. The observables $\{\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y, \sigma_z \otimes \sigma_z\}$ will be functions of C: take f_i to be the functions on $\mathrm{sp}(C)$, defined by

$$\begin{aligned}
f_1(n) &= 2(\delta_{1n} + \delta_{3n}) - 1 \\
f_2(n) &= 2(\delta_{2n} + \delta_{3n}) - 1 \\
f_3(n) &= 2(\delta_{1n} + \delta_{2n}) - 1
\end{aligned} \tag{13}$$

Then we will have

$$\begin{aligned}
\sigma_x \otimes \sigma_x &= f_1(C) \\
\sigma_y \otimes \sigma_y &= f_2(C) \\
\sigma_z \otimes \sigma_z &= f_3(C)
\end{aligned} \tag{14}$$

Moreover, since

$$f_1(x)f_2(x) = -f_3(x), \tag{15}$$

for all $x \in \mathrm{sp}(C)$, we will have the functional relation

$$f_1(C)f_2(C) = -f_3(C), \tag{16}$$

or,

$$(\sigma_x \otimes \sigma_x)(\sigma_y \otimes \sigma_y) = -\sigma_z \otimes \sigma_z. \tag{17}$$

It is easy to check that, on the tensor product Ψ_q, the functionals $\{a_{\phi+}, a_{\phi-}, a_{\psi+}, a_{\psi-}\}$ are indeed effects. Therefore, we can add to the observables on Ψ_q an observable C, that distinguishes the Bell states, and, having done so, we will obtain the algebraic relations (17). In the tensor product $\tilde{\Psi}_q$, each of these functionals takes on the value $-1/2$ on some state, and so they are not effects. We will, however, be able to define effects on $\tilde{\Psi}_q$ that lead to the "anti-quantum" functional relation, $(\sigma_x \otimes \sigma_x)(\sigma_y \otimes \sigma_y) = \sigma_z \otimes \sigma_z$.

In a similar manner, we can define affine functionals that distinguish Ψ_q's common eigenstates of $\{\sigma_x \otimes \sigma_y, \sigma_y \otimes \sigma_x, \sigma_z \otimes \sigma_z\}$ and, on Ψ_q, are effects, and use these to add an observable that induces the functional relation,

$$(\sigma_x \otimes \sigma_y)(\sigma_y \otimes \sigma_x) = \sigma_z \otimes \sigma_z. \tag{18}$$

It is worth noting in passing that, if we move to the larger tensor product space that is the convex hull of $\Psi_q \cup \tilde{\Psi}_q$, it will no longer be possible to introduce an observable that distinguishes the Bell states. This can be seen from the fact that any affine functional that

takes on the value 1 on ϕ^+ and 0 on the other Bell states will take on the value $-1/2$ on the state $\tilde{\phi}^+$, defined by

$$
\begin{aligned}
\tilde{\phi}^+(I \otimes \sigma_i) &= \tilde{\phi}^+(\sigma_j \otimes I) = 0 \\
\tilde{\phi}^+(\sigma_i \otimes \sigma_j) &= -\phi^+(\sigma_i \otimes \sigma_j).
\end{aligned} \tag{19}
$$

for $i, j \in \{x, y, z\}$. This follows from the fact that the totally mixed state $\omega_0 \otimes \omega_0$ can be written either as an equally weighted mixture of the Bell states, or as an equally weighted mixture of ϕ^+ and $\tilde{\phi}^+$.

$$
\omega_0 \otimes \omega_0 = \frac{1}{4} \left(\phi^+ + \phi^- + \psi^+ + \psi^- \right) = \frac{1}{2} \left(\phi^+ + \tilde{\phi}^+ \right). \tag{20}
$$

Similar remarks hold for the convex hull of $\Psi_S \cup \tilde{\Psi}_S$. On this space, there can be no positive affine functional that takes on the value 1 on the state ξ and 0 on all the other eigenstates of $\{\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y, \sigma_z \otimes \sigma_z\}$.

By adding to the set of observables on Ψ_q an observable that distinguishes the Bell states, and one that distinguishes the common eigenstates of $\{\sigma_x \otimes \sigma_y, \sigma_y \otimes \sigma_x, \sigma_z \otimes \sigma_z\}$, we induce the algebraic relations (17), (18), that lead to a Mermin-style Kochen-Specker obstruction. It is worth noting that we have obtained these relations, not by reference to the quantum relation (9), which involves products of incompatible observables, but by considerations of relations between *compatible* observables.

In a similar manner, we can induce algebraic relations among the same observables in the Spekkens tensor product Ψ_S. We will, in this space, have the relation (18). However, instead of the quantum relation (17), we will have

$$
(\sigma_x \otimes \sigma_x)(\sigma_y \otimes \sigma_y) = \sigma_z \otimes \sigma_z, \tag{21}
$$

and no Kochen-Specker obstruction will be forthcoming.

Thus, starting with the same state space for elementary systems, we obtain, depending on which tensor product space we bestow on pairs of systems, either a theory, whose pure states are those of the Spekkens toy theory, whose states which admit of a representation as mixture of classical states, or a theory, obtained by using one of the quantum tensor products, that admits of no non-contextual hidden-variables theory. Similarly, it is clear that whether or not our theory will contain states with correlations that violate a Bell inequality will depend on the tensor product used—though the tensor products we have so far constructed contain no Bell inequality-violating correlations, we can, without contradiction (and without disrupting the algebraic relations between observables we have obtained) extend our tensor space (be it Spekkens or quantum) to include them.

All this begins to suggest that Schrödinger was right to find the essential distinction between classical and quantum theory in the way that the latter treats compound systems. However, it is not entanglement *per se* that is distinctively quantum; it matters *which* entangled states the theory contains.

6   Acknowledgments

7   References

Alfsen, Erik M., Harald Hanche-Olsen, and Frederick W. Shultz (1980). "State Spaces of C^*-algebras." *Acta Mathematica* **144**, 267–305.

Alfsen, Erik M., and Frederick W. Shultz (1978). "State Spaces of Jordan Algebras." *Acta Mathematica* **144**, 267–305.

—— (2003). *Geometry of State Spaces of Operator Algebras.* Boston: Birkhäuser.

Araki, Huzihiro (1980). "On a Characterization of the State Space of Quantum Mechanics." *Communications in Mathematical Physics* **75**, 1–24.

Barnum, Howard, Jonathan Barrett, Matthew Leifer, and Alexander Wilce (2006). "Cloning and Broadcasting in Generic Probabilistic Models." quant-ph/0611295.

Barnum, Howard, Christopher A. Fuchs, Joseph M. Renes, and Alexander Wilce (2005). "Influence-free states on compound quantum systems." quant-ph/0507108.

Bell, John S. (1964). "On the Einstein-Podolsky-Rosen paradox." *Physics* **1**, 195–200. Reprinted in Bell (1987), 14–21.

—— (1966). "On the Problem of Hidden Variables in Quantum Mechanics." *Reviews of Modern Physics* **38**, 447–52. Reprinted in Bell (1987), 1–13.

—— (1987). *Speakable and unspeakable in quantum mechanics.* Cambridge University Press.

Beltrametti, E. G., and S. Bugajski (1997). " Effect algebras and statistical physical theories." *Journal of Mathematical Physics* **38**, 3020–3030.

Clifton, Rob, Jeffrey Bub, and Hans Halvorson (2003). "Characterizing Quantum Theory in Terms of Information-Theoretic Constraints." *Foundations of Physics* **33**, 1561–1590.

D'Ariano, Giacom Mauro (2006) "Operational Axioms for Quantum Mechanics". quant-ph/0611094.

Fuchs, Christopher A. (2002). "Quantum Mechanics as Quantum Information (and only a little more)." quant-ph/0205039.

Fuchs, Christopher A., and Rüdiger Schack (2004). "Unknown Quantum States and Operations, a Bayesian View." quant-ph/0404156.

Greenberger, D.M., M. Horne, and A. Zeilinger (1989). "Going Beyond Bell's Theorem," in M. Kafatos, ed., *Bell's Theorem, Quantum Theory, and Conceptions of the Universe* (Dordrecht: Kluwer Academic), 73–76.

Halvorson, Hans (2004). "On information-theoretic characterizations of physical theories." *Studies in History and Philosophy of Modern Physics* **35**, 277-293.

Hardy, Lucien (1992). "Quantum Mechanics, Local Realistic Theories, and Lorentz-Invariant Realistic Theories," *Physical Review Letters* **68**, 2981–2984.

—— (1993). "Nonlocality for Two Particles without Inequalities for Almost All Entangled States." *Physical Review Letters* **71**, 1665–1668.

—— (2001). "Quantum Theory from Five Reasonable Axioms." quant-ph/0101012.

—— (2002). "Why Quantum Theory?" in T. Placek and J. Butterfield, eds., *Non-Locality and Modality: Proceedings of the NATO Advanced Research Workshop on Modality, Probability, and Bell's Theorems, Cracow, Poland, August 19-23, 2001* (Springer), 61–74. Also quant-ph/0111068.

Holevo, A. S. (1982). *Probabilistic and Statistical Aspects of Quantum Theory.* North-Holland Publishing Company.

Lewis, David (1980). "A Subjectivist's Guide to Objective Chance," in Richard C. Jeffrey, ed., *Studies in Inductive Logic and Probability*, Vol. II (Los Angeles: University of California Press), 263–293. Reprinted in Lewis 1986, 83–113.

—— (1986). *Philosophical Papers*, Vol. II. Oxford: Oxford University Press.

Ludwig, G. (1983). *Foundations of Quantum Mechanics I.* Springer-Verlag.

Mermin, David N. (1990). "Simple Unified Form for the Major No-Hidden-Variables Theorems." *Physical Review Letters* **65**, 3373–3376.

—— (1993). "Hidden variables and the two theorems of John Bell." *Reviews of Modern Physics* **65**. 803–815.

Schrödinger, E. (1936). "Probability Relations Between Separated Systems." *Proceedings of the Cambridge Philosophical Society* **32**, 446–452.

Shimony, Abner (2003). "Reality, Causality, and Closing the Circle," in *Search for a Naturalistic World View, Volume I: Scientific Method and Epistemology.* Cambridge University Press.

Spekkens, Robert W. (2001). "In defense of the epistemic view of quantum states: a toy theory." quant-ph/0401052.

(2007). "Evidence for the epistemic view of quantum states: A toy theory." *Physical Review A* **75**, 032110.

8 Appendix

8.1 C^*- and JB algebras

A *Banach space* is a normed linear vector space that is complete with respect to the norm. That is, every Cauchy sequence converges to a limit.

A *Banach algebra* $\mathfrak{A}$ is a Banach space that is also an algebra with identity I, such that the operation of multiplication is separately continuous. That is, for each $B \in \mathfrak{A}$, if $A_n \to A$, then $A_n B \to AB$, and, for each $A \in \mathfrak{A}$, if $B_n \to B$, then $AB_n \to AB$.

An *involution* is a mapping $A \to A^*$ such that

i). $(aA + bB)^* = \bar{a}A^* + \bar{b}B^*$

ii). $(AB)^* = B^*A^*$

iii). $(A^*)^* = A$

A C^*-*algebra* is a complex Banach algebra with an involution that satisfies $\| A^*A \| = \| A \|^2$.

A *Jordan algebra* is a vector space with a commutative bilinear product $\circ$ satisfying

$$(a^2 \circ b) \circ a = a^2 \circ (b \circ a)$$

A *JB-algebra* is a Jordan algebra equipped with a norm $\| \ \|$ satisfying

i). $\|a \circ b\| \leq \|a\| \, \|b\|$

ii). $\|a^2\| = \|a\|^2$

iii). $\|a^2\| \leq \|a^2 + b^2\|$,

which is complete with respect to this norm.

The self-adjoint elements of any C^*-algebra form a JB-algebra, with the symmetric product,

$$a \circ b = \frac{1}{2}(ab + ba).$$

8.2 C^*-algebraic no-cloning

Theorem 4 *If $\{\rho, \omega\}$ are a cloneable pair of distinct pure states of a C*-algebra then they are orthogonal.*

Proof. We define the *transition probability* of a pair of pure states by

$$p(\rho, \omega) = 1 - \frac{1}{4}\|\rho - \omega\|^2$$

A nonselective operation cannot increase the norm distance between two states; therefore, the transition probability cannot decrease under a nonselective operation. Hence, if there is a nonselective operation that clones $\{\rho, \omega\}$,

$$p(\rho \otimes \rho, \omega \otimes \omega) \geq p(\rho, \omega).$$

However, it follows from CBH's Lemma 2 that, for pure states ρ, ω,

$$p(\rho \otimes \rho, \omega \otimes \omega) = p(\rho, \omega)^2.$$

This gives us

$$p(\rho, \omega)^2 \geq p(\rho, \omega),$$

or,

$$p(\rho, \omega)(1 - p(\rho, \omega)) \leq 0.$$

Since $p(\rho, \omega)$ lies in the interval $[0, 1]$, this is possible only if $p(\rho, \omega)$ is equal to 0 or 1.