A Royal Road to Quantum Theory (or Thereabouts)

Extended Abstract

Alexander Wilce
Department of Mathematics, Susquehanna University

Abstract

A representation of finite-dimensional probabilistic models in terms of formally real Jordan algebras is obtained, in a strikingly easy way, from simple assumptions. This provides a framework in which real, complex and quaternionic quantum mechanics can be treated on an equal footing, and allows some (but not too much) room for other alternatives. This is based on earlier work (arXiv:1206:2897), but the development here is further simplified, and also extended in several ways. I also discuss the possibilities for organizing probabilistic models, subject to the assumptions discussed here, into symmetric monoidal categories, showing that such a category will automatically have a dagger-compact structure. (Recent joint work with Howard Barnum and Matthew Graydon (arXiv:1507.06278) exhibits several categories of this kind.) An interesting open question is whether, conversely, a dagger-compact category of finite-dimensional probabilistic models must be a category of euclidean Jordan algebras

1 Introduction and Overview

A spate of recent papers, notably [7, 9, 15], have derived standard formulation of finite-dimensional QM is from various packages of axioms governing the information-carrying and information-processing capacity of finite-dimensional probabilistic systems. In this paper, I derive somewhat less, but do so in what I think is a very attractive and simple way. Specifically, I characterize *formally real Jordan algebras* as probabilistic models, in terms of a few having a straightforward probabilstic interpretation.

This has two advantages. First, it allows some — but not too much — latitude to go beyond standard complex QM. Secondly, this approach is simpler mathematically than any of those cited above. Once the various definitions are in place, the proofs of the main theorems are all quite easy, at least if one is allowed to invoke one classical mathematical result. An ordered vector space $\mathbf{E}$, with positive cone $\mathbf{E}_+$, is *self-dual* iff there exists an inner product on $\mathbf{E}$ such that $a \in \mathbf{E}_+$ iff $\langle a, b\rangle \geq 0$ for all $b \in \mathbf{E}_+$. Call $\mathbf{E}$ *homogeneous* iff the group of order automorphisms — positive linear automorphisms with positive inverses — on $\mathbf{E}$ acts transitively on the *interior* of $\mathbf{E}_+$. For an accessible proof of the following, see [10].

Theorem [Koecher 1958; Vinberg 1961]: *Let $\mathbf{E}$ be a finite-dimensional ordered vector space with a distinguished order-unit u. If $\mathbf{E}$ is homogeneous and self-dual, then there exists a unique Jordan product $\bullet$ such that $\mathbf{E}$ is a formally real Jordan algebra, u is the Jordan unit, and $\mathbf{E}_+$ is the cone of squares.*

Neither self-duality nor homogeneity is an entirely obvious property of a physical state space. If we can find a conceptually compelling way to motivate these assumptions, we will have cleared a route to (the vicinity of) quantum theory.

Much of what follows is drawn from the earlier papers [16, 17, 18], but the approach sketched here is organized somewhat differently, is (even) simpler, and goes soemwhat further.

2 Probabilistic Models

A *test space* is a collection $\mathcal{M}$ of non-empty sets, regarded as the outcome-sets of various experiments, measurements, etc. We refer to a set $E \in \mathcal{M}$ as a test. Letting $X = \bigcup \mathcal{M}$ be the *outcome-space* of $\mathcal{M}$, a *probability weight* on $\mathcal{M}$ is a mapping $\alpha : X \to [0, 1]$ that sums to unity on every test. We say that α is *non-singular* iff $\alpha(x) > 0$ for all $x \in X(A)$.[1]

Definition: A *probabilistic model* is a pair $A = (\mathcal{M}(A), \Omega(A))$ where $\mathcal{M}(A)$ is a test space and $\Omega(A)$ is a specified convex set of probability weights on $\mathcal{M}(A)$, called the *states* of the model.

It is harmless to assume that, for every outcome $x \in X(A)$, there exists at least one state $\alpha \in \Omega(A)$ wih $\alpha(x) > 0$. The span of $\Omega(A)$ in $\mathbb{R}^{X(A)}$, ordered by the cone $\mathbf{V}_+(A)$ of non-negative multiples of states, is denoted $\mathbf{V}(A)$. It will be useful below to note that the interior of the cone $\mathbf{V}(A)_+$ consists of positive multiples of non-singular states.

There is a unique positive functional $u_A \in \mathbf{V}(A)^*$ given by $u(\alpha) = 1$ for all $\alpha \in \Omega$. An *effect* on $\mathbf{V}(A)$ is positive linear functional $a \in \mathbf{V}^*$ with $0 \leq a \leq u$. For example, each $x \in X$ defines an effect $\widehat{x} \in \mathbf{V}(A)^*$ by evaluation, i.e., $\widehat{x}(\alpha) = \alpha(x)$ for all $\alpha \in \mathbf{V}(A)$. Effects can be taken to represent outcomes of *mathematically* possible measurements, but I make no assumption about which effects, other than those of the form $\widehat{x}$, are physically accessible.

Classical, Quantum and Jordan Models If E is a finite set, the corresponding *classical model* is $A(E) = (\{E\}, \Delta(E))$ where $\Delta(E)$ is the simplex of probability weights on E. If $\mathcal{H}$ is a finite-dimensional complex Hilbert space, let $\mathcal{M}(\mathcal{H})$ denote the set of orthonormal bases of $\mathcal{H}$: then $X = \bigcup \mathcal{M}(\mathcal{H})$ is the unit sphere of $\mathcal{H}$, and any density operator W on $\mathcal{H}$ defines a probability weight α_W, given by $\alpha_W(x) = \langle Wx, x\rangle$ for all $x \in X$. Letting $\Omega(\mathcal{H})$ denote the set of states of this form, we obtain the *quantum model*, $A(\mathcal{H}) = (\mathcal{M}(\mathcal{H}), \Omega(\mathcal{H}))$, associated with $\mathcal{H}$. The space $\mathbf{V}(A(\mathcal{H}))$ is isomorphic to the space $\mathcal{L}_h(\mathcal{H})$ of hermitian operators on $\mathcal{H}$, ordered as usual.

More generally, every formally real Jordan algebra $\mathbf{E}$ gives rise to a probabilistic model. Recall here that a Jordan algebra is a real commutative algebra $(\mathbf{E}, \bullet)$ with unit element u and $\bullet$ satisfying the Jordan identity $a^2 \bullet (a \bullet b) = a \bullet (a^2 \bullet b)$. $\mathbf{E}$ is *formally real* if $\sum_i a_i^2 = 0$ implies $a_i = 0$ for all i. A minimal or *primitive* idempotent of $\mathbf{E}$ is an element $p \in \mathbf{E}$ with $p^2 = p$ and, for $q = q^2 < p$, $q = 0$. A *Jordan frame* is a maximal pairwise orthogonal set of primitive idempotents. Let $X(\mathbf{E})$ be the set of primitive idempotents, let $\mathcal{M}(\mathbf{E})$ be the set of Jordan frames, and let $\Omega(\mathbf{E})$ be the set of probability weights of the form $\alpha(p) = \langle a, p\rangle$ where $a \in \mathbf{E}_+$ with $\langle a, u\rangle = 1$. This data defines the *Jordan model* $A(\mathbf{E})$ associated with $\mathbf{E}$. In the case where $\mathbf{E} = \mathcal{L}_h(\mathcal{H})$ for a finite-dimensional Hilbert space $\mathcal{H}$, this *almost* gives us back the quantum model $A(\mathcal{H})$: the difference is that we replace unit vectors by their associated projection operators, thus conflating outcomes that differ only by a phase.

[1] Material in this section is standard. See [5] for a more detailed account, and for further references. A good general reference for ordered vector spaces is [2]

Sharp models A probabilistic model A is *sharp* iff, for every outcome $x \in X(A)$, there exists a unique state $\alpha \in \Omega(A)$ with $\alpha(x) = 1$. Quantum models are evidently sharp; more generally, any Jordan model is sharp.

If A is sharp, then there is a sense in which each test $E \in \boldsymbol{\mathcal{M}}(A)$ is *maximally informative:* if we know for certain which outcome will occur, then we know the system's state exaclty, as there is only one state in which this outcome has probability 1. Conversely, sharpness can be understood as the requirement that all tests be maximally informative in this sense.

Processes We may want to regard two systems, represented by models A and B, as the input to and output from some *process*, whether dynamical or purely information-theoretic, that has some probability to destroy the system or otherwise "fail". Such a process can be represented mathematically by a positive linear mapping $T : \mathbf{V}(A) \to \mathbf{V}(B)$ taking each normalized state α of A to a possibly *sub-normalized* state $T(\alpha)$ of B, i.e., $T(\alpha) = p\beta$ where $\beta \in \Omega(B)$ and $p \in [0, 1]$ is the probability for the process to fail, given input state α. Note that p is independent of α, since T is linear. I do not suppose that every positive linear mapping of this sort represents a physically possible process.

Even if a process T has a nonzero probability of failure, it may be possible to reverse its effect with nonzero probability:

Definition: A process $T : A \to B$ is *probabilistically reversible*, or *p-reversible*, for short, iff there exists a process S such that, for all $\alpha \in \Omega(A)$, $(S \circ T)(\alpha) = p\alpha$, where $p \in (0, 1]$.

This means that there is a probability $1 - p$ of the process $S \circ T$ failing, but a probability p that it will leave the system in its initial state. (Since $S \circ T$ is linear, p must be constant.) If T preserves normalization, so that $T(\Omega(A)) \subseteq \Omega(B)$, S will also preserve normalization, and will undo the result of T with probability 1. In this case, we just say that T is *reversible.*

Every process $T : \mathbf{V}(A) \to \mathbf{V}(B)$ has a dual mapping $T^* : \mathbf{V}^*(B) \to \mathbf{V}^*(A)$, also positive, given by $T^*(b)(\alpha) = b(T(\alpha))$ for all $b \in \mathbf{V}^*(B)$ and $\alpha \in \mathbf{V}(A)$. That T takes normalized states to subnormalized states is equivalent to the requirement that $T^*(u_B) \leq u_A$, that is, that T^* maps effects to effects.

Bipartite States If A and B are two models, a (non-signaling) *bipartite state* on A and B is a mapping $\omega : X(A) \times X(B) \to \mathbb{R}$ such that (i) $\sum_{(x,y) \in E \times F} \omega(x, y) = 1$ for all $E \in \boldsymbol{\mathcal{M}}(A), F \in \boldsymbol{\mathcal{M}}(B)$; (ii) the *marginals*

$$\omega_1(x) = \sum_{y \in E} \omega(x, y) \text{ and } \omega_2(y) = \sum_{x \in E} \omega(x, y)$$

are independent of $E \in \boldsymbol{\mathcal{M}}(A)$ and $F \in \boldsymbol{\mathcal{M}}(B)$, respectively; (iii) the *conditional states*

$$\omega_{1|y}(x) := \omega(x, y)/\omega_2(y) \text{ and } \omega_{2|x} := \omega(x, y)/\omega_1(x)$$

belong to $\Omega(A)$ and $\Omega(B)$, respectively. We then have the *law of total probability*:

$$\omega_1(x) = \sum_{y \in F} \omega_2(y)\omega_{1|y}(x) \text{ and } \omega_2(y) = \sum_{x \in E} \omega_1(x)\omega_{2|x}(y).$$

(Note this implies that ω_1 and ω_2 belong to $\Omega(A)$ and $\Omega(B)$.)

It is helpful at this point to introduce another ordered vector space associated with a model A. Let $\mathbf{E}(A)_+ \subseteq \mathbf{V}(A)^*$ be the set of all linear combinations $\sum_i t_i \widehat{x}_i$ with $x_i \in X(A)$ and $t_i \geq 0$ for all i. This is a convex generating cone for $\mathbf{V}(A)^*$, generally smaller than the dual cone $\mathbf{V}(A)^*$. Write $\mathbf{E}(A)$ for the space $\mathbf{V}(A)^*$, as ordered by this smaller cone. The main utility of $\mathbf{E}(A)$ is the following observation:

Lemma 0: *If ω is a bipartite state on A and B, there exists a unique positive linear mapping $\widehat{\omega} : \mathbf{E}(A) \to \mathbf{V}(B)$ such that $\widehat{\omega}(\widehat{x})(y) = \omega(x, y)$ for all $x \in X(A)$ and $y \in X(B)$.*

The proof is straightforward (consider the map $X(B) \to \mathbf{V}(B)$ defined by $y \mapsto \omega(\,\cdot\, y)$; then dualize). Since $\widehat{\omega}(\widehat{x}) = \omega_1(x)\omega_{2|x}$, I call $\widehat{\omega}$ the *conditioning map* associated with $\widehat{\omega}$.

3 Self-duality and homogeneity for quantum models

If there exists an inner product on $\mathbf{E}(A)$ such that $\mathbf{E}(A)$ is self-dual with $\mathbf{E}(A)_+ \simeq \mathbf{V}(A)_+$, then I will say that the model A is *self-dual*. If $\mathbf{V}(A)_+$ (hence, also $\mathbf{E}(A)_+$) is homogeneous, it will follow from the KV theorem that $\mathbf{E}(A)$ carries a formally real Jordan structure.

Why should a model have either of these properties? It is intructive to look at the standard quantum model associated with a finite-dimensional Hilbert space. As discussed above, $\mathbf{V}(A(\boldsymbol{\mathcal{H}})) \simeq \mathcal{L}_h(\boldsymbol{\mathcal{H}})$, the space of self-adjoint operators on $\boldsymbol{\mathcal{H}}$. If x is a unit vector in $\boldsymbol{\mathcal{H}}$, let p_x denote the corresponding rank-one orthogonal projection operator. Consider the trace inner product $\langle a, b\rangle = \mathrm{Tr}(ab)$ on $\mathcal{L}_h(\boldsymbol{\mathcal{H}})$: By the spectral theorem, $\mathrm{Tr}(ab) \geq 0$ for all $b \in \mathcal{L}_h(\boldsymbol{\mathcal{H}})_+$ iff $\mathrm{Tr}(ap_x) = \langle ax, x\rangle \geq 0$ for all unit vectors x. So $\mathrm{Tr}(ab) \geq 0$ for all $b \in \mathbf{E}_+$ iff $a \in \mathcal{L}_h(\boldsymbol{\mathcal{H}})_+$, i.e., the trace inner product is self-dualizing. But this now leaves us with the question, *what does the trace inner product* represent*, probabilistically?*

The trace inner product as a bipartite state Let $\overline{\boldsymbol{\mathcal{H}}}$ be the conjugate Hilbert space to $\boldsymbol{\mathcal{H}}$. Suppose $\boldsymbol{\mathcal{H}}$ has dimension n. Any unit vector Ψ in $\boldsymbol{\mathcal{H}} \otimes \overline{\boldsymbol{\mathcal{H}}}$ gives rise to a joint probability assignment to effects a on $\boldsymbol{\mathcal{H}}$ and $\overline{b}$ on $\overline{\boldsymbol{\mathcal{H}}}$, namely $\langle (a \otimes \overline{b})\Psi, \Psi\rangle$. Consider the maximally entangled *EPR state* for $\boldsymbol{\mathcal{H}} \otimes \overline{\boldsymbol{\mathcal{H}}}$ defined by the unit vector

$$\Psi = \frac{1}{\sqrt{n}} \sum_{x \in E} x \otimes \overline{x} \in \boldsymbol{\mathcal{H}} \otimes \overline{\boldsymbol{\mathcal{H}}},$$

where E is any orthonormal basis for $\boldsymbol{\mathcal{H}}$. A straightforward computation shows that, for all $a, b \in \mathcal{L}_h(\boldsymbol{\mathcal{H}})$, $\langle (a \otimes \overline{b})\Psi, \Psi\rangle = \frac{1}{n}\mathrm{Tr}(ab)$. In other words, *the normalized trace inner product* just is *the joint probability function determined by the pure state vector* Ψ! As a consequence, the state represented by Ψ has a very strong correlational property: if x, y are two orthogonal unit vectors with corresponding rank-one projections p_x and p_y, we have $p_x p_y = 0$, so $\langle (p_x \otimes \overline{p_y})\Psi, \Psi\rangle = 0$. On the other hand, $\langle (p_x \otimes \overline{p_x})\Psi, \Psi\rangle = \frac{1}{n}\mathrm{Tr}(p_x) = \frac{1}{n}$. Hence, Ψ *perfectly, and uniformly, correlates* every basic measurement (orthonormal basis) of $\boldsymbol{\mathcal{H}}$ with its counterpart in $\overline{\boldsymbol{\mathcal{H}}}$. Note that Ψ is uniquely defined by this property.

Filters and homogeneity To say that the cone $\mathcal{L}_h(\boldsymbol{\mathcal{H}})_+$ is homogeneous means that any non-singular density operator can be obtained from any other by an order-automorphism. But in fact, something better is true: this order-automorphism can be chosen to represent a probabilistically reversible *physical* process, i.e., an invertible CP mapping. To see how this works, suppose W is a positive operator on $\boldsymbol{\mathcal{H}}$. Consider the pure CP mapping $\phi_W : \mathcal{L}_h(\boldsymbol{\mathcal{H}}) \to \mathcal{L}_h(\boldsymbol{\mathcal{H}})$ given by

$\phi_W(a) = W^{1/2}aW^{1/2}$. Then $\phi_W(\mathbf{1}) = W$. If W is nonsingular, so is $W^{1/2}$, so ϕ_W is invertible, with inverse $\phi_W^{-1} = \phi_{W^{-1}}$, again a pure CP mapping. Now given another nonsingular density operator M, we can get from W to M by applying $\phi_M \circ \phi_{W^{-1}}$.

All well and good, but this leaves us with a question: *What does the mapping* ϕ_W *represent, physically?* Suppose W is a density operator, with spectral expansion $W = \sum_{x \in E} t_x p_x$. Here, E is an orthonormal basis for $\boldsymbol{\mathcal{H}}$ diagonalizing W, and t_x is the eigenvalue of W corresponding to $x \in E$. Then, for each vector $x \in E$, $\phi_W(p_x) = t_x p_x$, where p_x is the projection operator associated with x. This means that ϕ_W acts as a *filter* on the test E: the *response* of each outcome $x \in E$ is attenuated by a factor $0 \leq t_x \leq 1$. That is, if M is a density operator corresponding to another state of the system, then the probability of x's occuring in the (sub-normalized) state $\phi_W(M)$ is $t_x \mathrm{Tr}(M p_x)$. If we think of the orthonormal basis E as representing a set of alternative *channels* plus *detectors*, as in the figure below, we can add a classical filter attenuating the response of one of the detectors — say, x — by a fraction t_x. The point is that What the computation above tells us is that we can achieve the same result by applying a suitable CP map to the system's state in advance of the measurement. Moeover, this can be done independently for each outcome of E.

This is illustrated below for 3-dimensional quantum system: $E = \{x, y, z\}$ is an orthonormal basis, representing three possible outcomes of a Stern-Gerlach-like experiment; the filter Φ acts on the system's state in such a way that the probability of outcome x is attenuated by a factor of $t_x = 1/2$, while outcomes y and z are unaffected.

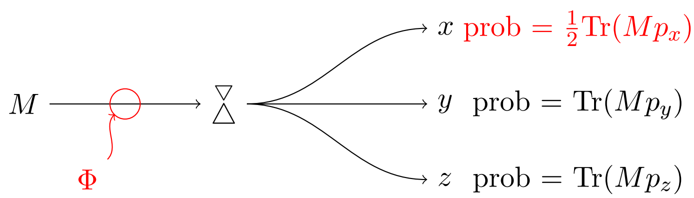

Figure 1: Φ attenuates x's sensitivity by $1/2$.

If we apply the ϕ_W to the maximally mixed state $\frac{1}{n}\mathbf{1}$, we obtain $\frac{1}{n}W$. Thus, we can *prepare* W, up to normalization, by applying the process Φ_W to the maximally mixed state.

Filters are Symmetric Here is a final observation, linking these last two: The filter Φ_W is *symmetric* with respect to the uniformly correlating state Ψ, in the sense that

$$\langle (\Phi_W(a) \otimes \overline{b})\Psi, \Psi \rangle = \langle (a \otimes \overline{\Phi}_W(b))\Psi, \Psi \rangle$$

for all effects $a, b \in \mathcal{L}_h(\boldsymbol{\mathcal{H}})_+$. As we'll soon see, this is basically *all that's needed* to recover the Jordan structure of finite-dimensional quantum theory: the existence of a conjugate system, with a uniformly correlating "EPR"-like joint state, plus the possibility of preparing non-singular states by means of p-reversible filters that are symmetric with respect to this state.

4 Conjugates and Filters

In order to abstract the two features of QM discussed above, we first need to restrict our focus very slightly. Call a test space $(X, \boldsymbol{\mathcal{M}})$ *uniform* iff all tests $E \in \boldsymbol{\mathcal{M}}$ have the same size, which we call the *rank* of A. The test spaces associated with Jordan (and so, in particular, with quantum) models all have this feature.

Definition: Let A be a uniform probabilistic model of rank n. A *conjugate* for A is a model $\overline{A}$, plus a chosen isomorphism[2] $A \simeq \overline{A}$ and a nonsignaling bipartite state η_A on A and $\overline{A}$ such that $\eta_A(x, \overline{x}) = 1/n$ for all $x \in X(A)$.

If $E \in \boldsymbol{\mathcal{M}}(A)$, we have $|E| = n$ and $\sum_{x,y \in E \times E} \eta_A(x, \overline{y}) = 1 = \sum_{x \in E} \eta(x, \overline{x})$. Hence, $\eta_A(x, \overline{y}) = 0$ for $x, y \in E$ with $x \neq y$. Thus, η_A establishes a perfect, uniform correlation between *any* test $E \in \boldsymbol{\mathcal{M}}(A)$ and its counterpart, $\overline{E} := \{\overline{x} | x \in E\}$, in $\boldsymbol{\mathcal{M}}(\overline{A})$. Since η is a bipartite state on A and $\overline{A}$, it follows that the uniformly or maximally mixed state $\rho(x) := \frac{1}{n}$ belongs to $\Omega(A)$. If $A = A(\boldsymbol{\mathcal{H}})$ is the quantum-mechanical model associated with an n-dimensional Hilbert space $\boldsymbol{\mathcal{H}}$, then we can take $\overline{A} = A(\overline{\boldsymbol{\mathcal{H}}})$ and define $\eta_A(x, \overline{y}) = |\langle \Psi, x \otimes \overline{y} \rangle|^2$, where Ψ is the EPR state, as discussed in Section 3.

So much for conjugates. We generalize the filters associated with pure CP mappings as follows:

Definition: A *filter* associated with a test $E \in \boldsymbol{\mathcal{M}}(A)$ is a process $\Phi : \mathbf{V}(A) \to \mathbf{V}(A)$ such that for every outcome $x \in E$, there is some coefficient $t_x \in [0, 1]$ with $\Phi(\alpha)(x) = t_x \alpha(x)$ for every state $\alpha \in \Omega(A)$.

Equivalently, Φ is a filter iff the dual process $\Phi^* : \mathbf{V}^*(A) \to \mathbf{V}^*(A)$ satisfies $\Phi^*(\widehat{x}) = t_x \widehat{x}$ for each $x \in E$. Just as in the quantum-mechanical case, a filter independently attenuates the "sensitivity" of the outcomes $x \in E$. We'll shortly see that the existence of a conjugate, plus the *preparability* of arbitrary nonsingular states by *symmetric*, p-reversible filters, is enough to make A a Jordan model. Most of the work is done by the easy Lemma 1, below.

Definition: Let $\Delta = \{\delta_x | x \in X(A)\}$ be any family of states, indexed by outcomes $x \in X(A)$ with $\delta_x(x) = 1$. I will say that a state $\alpha \in \Omega(A)$ is *spectral* with respect to Δ if there is a test $E \in \boldsymbol{\mathcal{M}}(A)$ such that $\alpha = \sum_{x \in E} \alpha(x) \delta_x$. The model A is spectral with respect to Δ iff all its states are.

If A is sharp and also spectral with respect to a set Δ of states, then Δ must be the unique set of states δ_x with $\delta_x(x) = 1$. Thus, for a sharp model, we can use the adjective "spectral" without qualification.

Lemma 1: *Let A have a conjugate $(\overline{A}, \eta_A)$. Suppose that every non-singular state of A is spectral with respect to the states $\delta_x := \eta_{1|\overline{x}}$, $x \in X$. Then A is sharp, and $\langle a, b \rangle := \eta_A(a, \overline{b})$ is a self-dualizing inner product on $\mathbf{E}(A)$, with respect to which $\mathbf{V}(A) \simeq \mathbf{E}(A)$. That is, A is self-dual.*

Proof: That $\langle\, , \,\rangle$ is symmetric and bilinear follows from η's being symmetric and non-signaling. Since $\widehat{\eta}$ is a positive mapping, the image $\widehat{\eta}(\mathbf{E}(A)_+)$ is contained in $\mathbf{V}(A)_+$. By the spectrality assumption, the image $\widehat{\eta}(\mathbf{E}(A)_+)$ contains the interior of $\mathbf{V}(A)_+$. It follows (recalling that we are dealing with finite-dimensional spaces) that $\widehat{\eta}(\mathbf{E}(A)_+) = \mathbf{V}(A)_+$, and, hence, that $\widehat{\eta}$ is an order-isomorphism. The spectrality assumption now implies that every a belonging to the interior of $\mathbf{E}(A)_+$ has a decomposition of the form $\sum_{x \in E} t_x x$ fo some coefficients $t_x > 0$ and some test $E \in \boldsymbol{\mathcal{M}}(A)$. It now follows that the same decomposition is available (albeit with arbitrary coefficients) for any $a \in \mathbf{E}(A)$: for a sufficiently large value of N, $a + Nu$ belongs to

[2]By an isomorphism from a model A to a model B, I mean a bijection $\phi : X(A) \to X(B)$ inducing a bijection $\boldsymbol{\mathcal{M}}(A) \to \boldsymbol{\mathcal{M}}(B)$, and such that $\alpha \in \Omega(A)$ if $\alpha \circ \phi \in \Omega(B)$.

the interior of $\mathbf{E}(A)_+$, and hence, has the desired decomposition, say $a + Nu = \sum_{x\in E} t_x x$ for some $E \in \boldsymbol{\mathcal{M}}(A)$. Since $u = \sum_{x\in E} x$, we have $a = (a + Nu) - Nu = \sum_{x\in E}(t_x - N)x$.

If $a = \sum_{x\in E} t_x x$ for some $E \in \boldsymbol{\mathcal{M}}(A)$ and coeffcients $t_x \in \mathbb{R}$, then

$$\langle a, a\rangle = \sum_{x,y\in E\times E} t_x t_y \eta_A(x, \overline{y}) = \frac{1}{n}\sum_{x\in E} {t_x}^2 \geq 0.$$

This is zero only where all coefficients t_x are zero, i.e., only for $a = 0$. So $\langle\ ,\ \rangle$ is an inner product. That this is self-dualizing, and makes $\mathbf{E}(A) \simeq \mathbf{V}(A)$, is now straightforward (using the spectral decomposition yet again, plus the fact that $\widehat{\eta}$ is an order-isomorphism). □

If A is sharp and has a conjugate $\overline{A}$, then the state $\eta_{1|\overline{x}}$ is the unique state δ_x with $\delta_x(x) = 1$, so the spectrality assumption in Lemma 1 is fulfilled if we simply say that A is spectral. Hence, *a sharp, spectral model with a conjugate is self-dual.* For the simplest systems, this is already enough to secure the desired representation in terms of a formally real Jordan algebra. Call A a *bit* iff it has rank 2 (all tests have two outcomes), and if every state $\alpha \in \Omega(A)$ can be expressed as a mixture of two sharply distinguishable states; that is, $\alpha = t\delta_x + (1 - t)\delta_y$ for some $t \in [0, 1]$ and states δ_x and δ_y with $\delta_x(x) = 1$ and $\delta_y(y) = 1$ for some test $\{x, y\}$. If a bit is sharp, then it is already spectral; so it follows from Lemma 1 that if a sharp bit A has a conjugate, it is self-dual. It follows easily that $\Omega(A)$ must then be a ball of some finite dimension d. If d is $2, 3$ or 5, we have a real, complex or quaternionic bit. For $d = 4$ or $d \geq 6$, we have a non-quantum spin factor.

Suppose now that A has arbitrary rank, and satisfies the hypotheses of Lemma 1. If $\mathbf{V}(A)$ and, hence, $\mathbf{E}(A)$ are homogeneous, then, by the Koecher-Vinberg Theorem, $\mathbf{E}(A)$ carries a canonical Jordan structure. In fact, we can say something a little stronger [18]:

Theorem 1: *Let A be spectral with respect to a conjugate system $\overline{A}$. If $\mathbf{V}(A)$ is homogeneous, then there exists a canonical Jordan product on $\mathbf{E}(A)$ with respect to which u is the Jordan unit. Moreover with respect to this product $X(A)$ is exactly the set of primitive idempotents, and $\boldsymbol{\mathcal{M}}(A)$ is exactly the set of Jordan frames.*

The homogeneity of $\mathbf{V}(A)$ can be understood as a *preparability assumption:* it says that every nonsingular state can be obtained, *up to normalization*, from the maximally mixed state by a p-reversible process. In fact, if the hypotheses of Lemma 1 hold, the homogeneity of $\mathbf{V}(A) \simeq \mathbf{E}(A)$ follows from the mere existence of p-reversible filters with arbitrary non-zero coefficients. For if a is in the interior of $\mathbf{E}(A)_+$, then $a = \sum_{x\in E} t_x x$ for some $E \in \boldsymbol{\mathcal{M}}(A)$, with $t_x > 0$ for all $x \in E$. If Φ is a p-reversible filter with $\Phi(x) = t_x x$ for all $x \in E$, then $\Phi(u) = a$.

Two paths to spectrality Some form of spectral decomposition for states is occasionally taken as an axiom [11, 6]. However, spectrality can be derived seem a good deal more transparent.[3]

Definition: A nonsignaling bipartite state ω on probabilistic models A and B is *correlating* iff there exists a pair of tests $E \in \boldsymbol{\mathcal{M}}(A)$, $F \in \boldsymbol{\mathcal{M}}(B)$ and a partial bijection $f : E \to F$ such that for $x \in E, y \in F$, $\omega(x, y) > 0$ iff $x \in E_o$ and $y = f(x)$.

This is equivalent to saying that $\omega(x, f(x)) = \omega_1(x) = \omega_2(f(x))$, or $\omega_{2|x}(f(x)) = 1$, for $\omega_1(x) \neq 0$. (Notice, too, that since ω must sum to 1 over $E \times F$, f must be non-empty.) Using

[3]A different path to spectrality is charted in a recent paper [8] by G. Chiribella and C. M. Scandolo.

these observations and the law of total probability, we have

Lemma 2: *Suppose A is sharp. Any state arising as the marginal of a correlating bipartite state between A and some model B, is spectral.*

Let us say that A satisfies the *correlation principle* iff every state of A arises as the marginal of — *dilates to* — a correlating bipartite state. It follows that, if A is sharp, it is also spectral. The correlation principle has an affinity with the purificaton postulate of [7], which requires that every state dilate to a *pure* state on a composite system. It is also related to the idea that, for every state α, there should exist a *non-disturbing, recordable measurement*: a test $E \in \boldsymbol{\mathcal{M}}(A)$ that can be made, without affecting α, in such a way that the outcome an be recorded in the state of some ancillary system B. Prior to the test E, A is in state α and B is in some "ready" state. After the test, the combined system is in some joint, non-signaling state ω. If the test is non-disturbing of α, we must have $\omega_1 = \alpha$. If the outcome of E was x, we suppose B to be in a "record state" β_x: if this record is accurate, we must have $\beta_x = \omega_{2|x}$, the conditional state of B given x. If these record states are to be readable, there must exist a test F on B and an injection $f : E \to F$ such that $\beta_x(f(x)) = 1$ for every $x \in E$ with $\alpha(x) > 0$. Thus, ω correlates E with F.

Here is another, superficially quite different, way of arriving at spectrality. Call a transformation Φ *symmetric* with respect to η_A iff, for all $x, y \in X(A)$, $\eta_A(\Phi^* x, \overline{y}) = \eta_A(x, \overline{\Phi}^* y)$. Now let $\alpha = \Phi(\rho)$ where Φ is a symmetric filter on a test $E \in \boldsymbol{\mathcal{M}}(A)$, say $\Phi(x) = t_x x$ for all $x \in E$. A direct computation then shows that $\alpha = \sum_{x \in E} t_x \delta_x$. Thus, if every nonsingular state is preparable by a symmetric filter, the spectrality assumption of Lemma 1 holds, and A is self-dual. If the preparing filter can always be taken to be p-reversible, as well as symmetric, then $\mathbf{V}(A)$ is homogeneous, and we have a Jordan model. On the other hand, as noted above, in the presence of spectrality, it's enough to *have* arbitrary p-reversible filters, as these allow one to prepare the spectral decompositions of arbitrary non-singular states. Thus, conditions (a) and (b), below, both imply that A is a Jordan model. Conversely, one can show that any Jordan model satisfies both (a) and (b), closing the loop [18]:

Theorem 2: *The following are equivalent:*

- *(a) A has a conjugate, and every non-singular state can be prepared by a p-reversible symmetric filter;*
- *(b) A is sharp, has a conjugate and arbitary p-reversible filters, and sastisfies the correlation principle;*
- *(c) A is a Jordan model.*

Remark: With some work, one can show that the assumptions of Lemma 1 imply a spectral uniqueness theorem, and hence, a functional calculus, for $\mathbf{E}(A)$. Call an effect $e \in \mathbf{E}(A)$ *sharp* iff there exists a state α with $\alpha(e) = 1$. It is easy to show that e must then have the form $e = e(D) := \sum_{x \in D} \widehat{x}$ where $D \subseteq E$ for some $E \in \boldsymbol{\mathcal{M}}(A)$. Call sharp effects $e_1, ..., e_n$ *jointly orthogonal* iff $e_i = e(D_i)$ where $D_1, ..., D_n$ are pairwise disjoint subsets of a single test $E \in \boldsymbol{\mathcal{M}}(A)$. Then every $a \in \mathbf{E}(A)$ has a unique representation $a = \sum_{i=0}^{n} t_i e_i$ where $t_o > t_1 > .... > t_n$ and $e_1, ..., e_n$ are jointly orthogonal sharp effects. Thus, for any function $f : \{t_o, ..., t_n\} \to \mathbb{R}$, one can define $f(a) := \sum_i f(t_i) e_i$. In particular, $a^2 = \sum_i t_i^2 e_i$. There is now only one candidate for a Jordan product on $\mathbf{E}(A)$, namely, $a \bullet b = (a+b)^2 - a^2 - b^2$. If $\mathbf{V}(A)$ is also homogeneous, the KV theorem implies that this *is* a Jordan product. In particular, it is bilinear. An interesting problem is whether one can prove this *without* invoking the Koecher-Vinberg Theorem.

5 Jordan Composites and Jordan Theories

A probabilistic theory is best understood as a category of probabilistic models and processes. In order to handle composite systems, one would like this to be a symmetric monoidal category. However, one wants to place some minimal restriction on how the monoidal product interacts with the probabilistic structure:

Definition: A (non-signaling) *composite* of probabilistic models A and B is a model AB, equipped with a mapping $\pi : X(A) \times X(B) \to \mathbf{V}(AB)^*$, whereby outomes $x \in X(A)$ and $y \in X(B)$ can be combined into a single effect $\pi(x, y) =: xy \in \mathbf{V}(AB)^*_+$. Moreover, we require that (i) $\sum_{(x,y)\in E\times F} \pi(x, y) = u_{AB}$ for all $E \in \boldsymbol{\mathcal{M}}(A)$, $F \in \boldsymbol{\mathcal{M}}(B)$, and (ii) $\forall \omega \in \Omega(AB)$, $\omega \circ \pi$ is a (nonsignaling) bipartite state on A and B.

By a *monoidal probabilistic theory*, I mean a symmetric monoidal category $\mathcal{C}$ in which objects are probabilistic models, morphisms are processes, and the monoidal product is a composite in the above sense. Moreover, I require the monoidal unit to be the (obvious) trivial model 1 with $\mathbf{V}(1) = \mathbb{R}$. By a *conjugage* for $A \in \mathcal{C}$, I mean a conjugate in the sense defined earlier, but with the added restriction that $\overline{A} \in \mathcal{C}$ and $\eta_A \in \Omega(A\overline{A})$.

Theorem 3: *Let $\mathcal{C}$ be a locally tomographic monoidal probabilistic theory in which every model A is spectral, and has a conjugate. Then $\mathcal{C}$ has a canonical dagger compact structure.*

The proof is straightforward. Lemma 1 gives us a self-dualizing inner product on each of the spaces $\mathbf{E}(A)$, and also sets up canonical isomorphisms $\mathbf{E}(A) \simeq \mathbf{V}(A)^* \simeq \mathbf{V}(A)$, whence, we have a canonical inner product on the latter. Thus, given a process $T : \mathbf{V}(A) \to \mathbf{V}(B)$, we have an adjoint $T^\dagger : \mathbf{V}(B) \to \mathbf{V}(A)$. Regarding $\eta_A \in \mathbf{V}(A\overline{A})$ as a mapping $\eta_A : \mathbb{R} = \mathbf{V}(1) \to \mathbf{V}(A\overline{A})$, we define $\epsilon_A := \sigma_{A,\overline{A}} \circ \eta_A^\dagger$, where $\sigma_{A,\overline{A}}$ is the swap morphism taking $A\overline{A}$ to $\overline{A}A$. Local tomography lets us expand $\epsilon_A^\dagger$ as $\sum_i x_i \otimes x_i$ where $\{x_i\}$ is an orthonormal basis for $\mathbf{E}(A)$. Using this, the verification of the "snake" identities is routine.

This raises two questions. The first is whether the local tomography assumption can be dropped. In a forthcoming paper [4] (see also [3]), Howard Barnum, Matthew Graydon and I have shown that one can construct a dagger-compact category embracing real, complex and quaternionic quantum systems *at the same time*, at the cost of modifying the composition rule for complex quantum systems to include an extra classical bit, i.e., a two-valued superselection rule. (This has the function of allowing time-reversal to be a physical operation in complex QM, as it is in the real and quaternionic cases). Composites in this category are generally not locally tomographic. On the other hand, morphisms are not processes, in the sense defined above, bu rather, certain positive mappings on the enveloping complex matrix algebras associated with these Jordan algebras.

The second question, which at present I cannot answer either, is what sort of converse, if any, holds for Theorem 3. That is, given a dagger-compact category of finite-dimensional (uniform) probabilistic models, must these models satisfy the hypotheses of Lemma 1? Must they in fact be Jordan models?

Acknowledgements Work on this paper, which is partly based on talks given in workshops and seminars in Amsterdam and Oxford in 2014 and 2015, has been supported by a grant (FQXi-RFP3-1348) from the FQXi foundation.

References

[1] S. Abramsky and B. Coecke, Abstract Physical Traces, Theory and Applications of Categories **14** (2005) 111–124 (`arXiv: arXiv:0910.3144`, 2009)

[2] C. D. Aliprantis and R. Tourky, *Cones and Duality*, American Mathematical Society, Graduate Studies in Mathematics **84**, 2007

[3] H. Barnum, M. Graydon and A. Wilce, Some nearly quantum theories, EPTCS **195** (2015), 59-705; `arXiv:1507.06278`

[4] H. Barnum, M. Graydon and A. Wilce, Composites and categories of euclidean Jordan algebras, to appear.

[5] H. Barnum and A. Wilce, Post-classical probability theory, to appear in G. Chiribella and R. Spekkens (eds.), *Quantum Theory: Informational Foundations and Foils*, Springer, 2016; `arXiv: 1205.3833`

[6] H. Barnum, M. Müller and C. Ududec, Higher-order interference and single-system postulates characterizing quantum theory, *New J. Phys.* **16** (2014) ; `arXiv:1403.4147`.

[7] G. Chiribella, M. D'Ariano and P. Perinotti, Informational derivation of quantum theory, Phys. Rev. A **84** (2011); `arXiv:1011:6451`

[8] G. Chribella, C. M. Scandolo, Operational axioms for state diagonalization, `arXiv: 1506:00380`

[9] B. Dakic and Č. Brukner, Quantum theory and beyond: is entanglement special? `arXiv:0911.0695`

[10] J. Faraut and A. Koranyi, *Analysis on Symmetric Cones*. Oxford University Press (1994).

[11] J. Gunson, On the algebraic structure of quantum mechanics, Comm. Math. Phys. **6** (1967), 262-285.

[12] H. Hanche-Olsen, On the structure and tensor products of JC-algebras, Can. J. Math. **35** (1983) 1059-1074.

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

[14] P. Jordan, J. von Neumann and E. Wigner, On an algebraic generalization of the quantum mechanical formalism, Annals of Math. **35** (1934)

[15] Ll. Masanes and M. Müller, A derivation of quantum theory from physical requirements, New J. Phys. **13** (2011); `arXiv:1004.1483`

[16] A. Wilce, 4.5 axioms for finite-dimensional quantum probability, in Y. ben-Menahem and M. Hemmo, *Probability in Physics*, Springer, 2012; `arXiv:0912.5530`

[17] A. Wilce, Symmetry, self-duality and the Jordan structure of finite-dimensional quantum mechanics, `arxiv:1110.6607` (2011)

[18] A. Wilce, Conjugates, Filters and Quantum Mechanics, `arxiv.org/pdf/1206.2897` (2012)