Convexity, Categorical Semantics and the Foundations of Physics

Howard Barnum
Perimeter Institute for Theoretical Physics

Ross Duncan
Oxford University

Alexander Wilce
Susquehanna University

Abstract

We consider symmetric monoidal categories of convex operational models, and adduce necessary and sufficient conditions for these to be compact-closed or dagger-compact. Compact closure amounts to the condition that all processes be implementable by means of a "remote evaluation" protocol (generalizing standard conclusive quantum teleportation protocols), which amounts to a form of classical conditioning. This requires the existence, for each system, of a bipartite state involving a further system, whose corresponding conditioning map is an isomorphism, and an an effect whose corresponding map is the inverse of this isomorphism. *Degenerate* compact closure, in which systems act as their own duals in the compact structure, means that one may take this extension to be the system itself, so the isomorphism implies that systems are weakly self-dual as ordered vector spaces. Degenerate dagger compact categories emerge from a further restriction, namely, that the bipartite "isomorphism" state and effect be symmetric.

It is natural to model physical theories as categories, with objects representing physical systems and morphisms, processes. It is reasonable to expect such a category to be symmetric monoidal, to allow for a treatment of composite systems. In the categorical approach to the foundations of quantum theory, associated with Abramsky and Coecke [1], Baez [2], Selinger [8, 9] and others, the focus is on compact closed, or, more narrowly, dagger-compact, categories. As a systematization of existing quantum protocols, this approach has been very successful. However, if our aim is more broadly foundational, then these strong structural constraints require independent physical motivation, or at any rate, characterization.

It is useful to contrast the categorical approach with the older "convex operational" tradition (deriving from the work of Mackey [7], Davies and Lewis [4], Ludwig [6] and others in the 1950s and 60s). Here, state spaces of individual probabilistic systems are represented by essentially arbitrary compact convex sets, and physical processes by affine (convex-linear) mappings between these. This setting is conservative of classical probabilistic ideas, but entirely liberal as to the structures it embraces, so long as this content is respected. We show that almost *any* symmetric monoidal category can be represented as a category of concrete convex-operational models, with a "non-signaling" tensor product. This suggests that compact closure and dagger-compactness are rather special constraints. As we'll show, compact closure amounts to the condition that all processes be implementable by means of a "remote evaluation" protocol (generalizing standard conclusive quantum teleportation protocols) — which amounts to a form of classical conditioning. Degenerate dagger compact categories emerge from a further restriction, namely, that a composite of two copies of a system in the theory admit a *symmetric* bipartite "isomorphism" state that can be used to teleport.

1 Some Categorical Preliminaries

Let $\mathcal{C}$ be a symmetric monoidal category. Recall that a *dual* for an object $A \in \mathcal{C}$ consists of an object B and two morphisms, a *unit*, $\eta : I \to B \otimes A$ and *co-unit*, $\epsilon : A \otimes B \to I$, such that

$$\begin{array}{l} A \xrightarrow{\mathrm{id}A \otimes \eta} A \otimes B \otimes A \xrightarrow{\epsilon \otimes \mathrm{id}A} A \; = \mathrm{id}A \\ B \xrightarrow{\eta \otimes \mathrm{id}B} B \otimes A \otimes B \xrightarrow{\mathrm{id}B \otimes \epsilon} B \; = \mathrm{id}B. \end{array} \quad (1)$$

We shall be particularly interested in the case in which $A = A'$ (not just up to a canonical isomorphism, but on the nose). In this case, we shall say that $\mathcal{C}$ is *degenerate*.

Adjoints In any compact closed category $\mathcal{C}$, assignment $A \mapsto A'$ extends to a canonical contravariant functor $(-)' : \mathcal{C}^{\mathrm{op}} \to \mathcal{C}$ taking $\phi \in \mathcal{C}(A, B)$ to $\phi' \in \mathcal{C}(B', A')$ defined by

$$\begin{array}{ccc} B' & \xrightarrow{\eta_A \otimes \mathrm{id}B'} & A' \otimes A \otimes B' \\ \Big\downarrow{\scriptstyle \phi'} & & \Big\downarrow{\scriptstyle \mathrm{id}A' \otimes \phi \otimes \mathrm{id}B'} \\ A' & \xleftarrow[\mathrm{id}A' \otimes \epsilon_B]{} & A' \otimes B \otimes B'. \end{array} \quad (2)$$

There are natural isomorphisms $w_A : A'' \to A$, with $\sigma \circ \eta_A = (1_A \otimes w_A) \circ \eta_{A'}$ In the classic treatment of coherence for compact closed categories, one has that $\sigma \circ \eta_A = (1_A \otimes w_A) \circ \eta_{A'}$; a similar condition holds for ϵ ([5], eq. (6.4)ff.). In the case of a degenerate category, this becomes

$$\sigma \circ \eta_A = (1_A \otimes w_A) \circ \eta_A. \quad (3)$$

To say that $'$ is involutive is to say that $A'' = A$ and $\phi'' = \phi$; note that this does not imply that $w_A = \mathrm{id}_A$.

It is easy to show that if the units – or, equivalently, co-units – are symmetric, in the sense that, for every object $A \in \mathcal{C}$, $\eta_A = \sigma_{A,A} \circ \eta_A$ or, equivalently, $\epsilon_A = \epsilon_A \circ \sigma_{A,A}$, then the the functor $'$ is involutive. However the converse does not necessarily hold, unless $w_A = \mathrm{id}_A$. In general, it is a delicate matter what coherence requirements are appropriate for degenerate compact closed categories, and in particular whether the functors involved should be strict. Therefore we will establish explicitly that the involutiveness of the adjoint is equivalent to the symmetry of the unit and co-unit for the compact closed categories of convex operational models considered in this paper.

For an arbitrary degenerate compact closed category, there is no guarantee that η_A will be symmetric. (We thank Peter Selinger for supplying a nice example involving a category of planar tangles.) Thus, it is a non-trivial constraint on such a category that the canonical adjoint be an involution.

Daggers A *dagger* on $\mathcal{C}$ is an involutive functor $(-)^\dagger : \mathcal{C}^{\text{op}} \to \mathcal{C}$ acting as the identity on objects. We say that $f \in \mathcal{C}(A, B)$ is *unitary* iff $f^\dagger = f^{-1}$. A *dagger-monoidal* category is a symmetric monoidal category equipped with a dagger such that (i) all the canonical isomorphisms defining the symmetric monoidal structure are unitary, and (ii) $(f \otimes g)^\dagger = f^\dagger \otimes g^\dagger$ for all morphisms f and g in $\mathcal{C}$. Finally, a †-monoidal category $\mathcal{C}$ is *dagger-compact* if it is compact closed and $\eta_A = \sigma_{A,A'} \circ \epsilon_A^\dagger$ for every A. In the case of a degenerate compact closed category, the canonical adjoint $'$ functions as a dagger *if* it is involutive. However, as remarked above, this is a nontrivial condition.

2 (Categories of) Convex Operational Models

In the convex approach to physical theories, any compact convex set Ω can serve as a state space. A measurement outcome is represented as an affine functional $f : \Omega \to [0, 1]$ — an *effect* — with $f(\alpha)$ understood as the probability of that outcome in state $\alpha \in \Omega$. Note that the complementary event, that f not occur, is represented by the effect $u - f$, where u is the *unit effect* is the effect $u : \alpha \equiv 1$. This apparatus can be linearized in a canonical way. Let $\text{Aff}(\Omega)$ denote the space of bounded affine functionals $\Omega \to \mathbb{R}$, ordered point-wise on Ω. Embed Ω in $\text{Aff}(\Omega)^*$ by evaluation. Define $V(\Omega)$ to be the span of Ω in $\text{Aff}(\Omega)^*$, ordered by the cone $V_+(\Omega)$ consisting of all $\mu \in V$ with $\mu(f) \geq 0$ for all $f \geq 0$ in $\text{Aff}(\Omega)_+$. One can show that Ω is a base for $V_+(\Omega)$, and, indeed, that $V(\Omega)$ is a complete base-normed space. Every bounded affine functional on Ω – in particular, every effect – extends uniquely to a linear functional in $V(\Omega)^*$, so that $\text{Aff}(\Omega) = V(\Omega)^*$. It is often assumed that every effect represents a physically accessible measurement outcome; however, to allow for flexibility on this point, we make the following

Definition: A *convex operational model* is a triple $(A, A^\#, u_A)$ where A a complete base-normed space with order unit $u_A \in A^*$, and $A^\#$ is a weak-$*$ dense subspace of A^*, ordered by a generating cone $A^\#_+$ (which may be smaller than $A^*_+ \cap A^\#$). By an *effect* on A, we mean a functional $a \in A^\#_+$ with $a \leq u_A$. The *normalized state space* of A is $\Omega_A := u_A^{-1}(1) \cap A_+$. If $\alpha \in \Omega_A$ and a is an effect, we regard $a(\alpha)$ as the probability that a is observed in state α.

The basic example to bear in mind is that of a quantum-mechanical system with associated Hilbert space $\mathbf{H}$: here, A is the space of trace-class Hermitian operators on $\mathbf{H}$, ordered by the usual cone of positive operators; the order-unit is $u_A(\rho) = \text{Tr}(\rho)$. We have some flexibility in the choice of $A^\#$: we might take $A^\# = A^*$, the space of all bounded Hermitian operators on $\mathbf{H}$, but we might also take $A^\#$ to consist, say, of bounded Hermitian operators with discrete spectrum. (In finite dimensions, of course, the density assumption requires that $A^{\#} = A^*$ as a vector space. Most of the work done thus far in the categorical idiom is motivated by finite-dimensional QM.)

Mackey Triples The following construction will be important below. A *Mackey triple* is a structure (X, Σ, p) where X and Σ be any non-empty sets, understood as sets of measurement outcomes and of states, respectively, and $p : \Sigma \times X \to [0, 1]$ is a function assigning, to $x \in X$ and $\alpha \in \Sigma$, a *probability* $p(x|\alpha)$. Given a Mackey triple (X, Σ, p), let $\Omega \subseteq \mathbb{R}^X$ be the point-wise closed (hence, compact), convex hull of the functions $p(\cdot|\alpha)$, $\alpha \in \Sigma$. Construct $V(\Omega)$ as described above; identifying X with its image in $V(\Omega)^*$ and letting $A_+^{\#}$ be the cone in $V(\Omega)^*$ generated by X, we have a convex operational model.

Processes A *process* from a convex operational model A to a convex operational model B, is a positive mapping $\phi : A \to B$ such that ϕ^* is positive as a mapping $B^{\#} \to A^{\#}$, and $\phi^*(u_B) \leq u_A$ (so that if $\alpha \in \Omega_A$, $u_B(\phi(\alpha)) \leq 1$). We interpret ϕ as representing a process, for which an input state $\alpha \in \Omega_A$ is transformed to a (possibly sub-normalized) state $\phi(\alpha) \in B_+$. We interpret this last quantity as the *probability* that the process represented by ϕ takes place. It is clear that processes compose, and that every convex operational model has an identity process, so we have here a concrete category, $\mathbf{Com}$, of convex operational models and processes. We shall write $\mathbf{FdCom}$ for the full sub-category of finite-dimensional convex operational models.

Composite Systems and Monoidality If A and B are convex operational models, one can show that any joint probability weight $\omega : [0, u_A] \times [0, u_B] \to [0, 1]$ extends to a unique positive bilinear form on $A^{\#} \times B^{\#}$. We define the *maximal tensor product* of A and B to be the space $A \otimes B = \mathfrak{B}(A^{\#}, B^{\#})$, ordered by the cone of bilinear forms that are positive as functions on $A_+^{\#} \times B_+^{\#}$, and with $(A \otimes B)_+^{\#}$ the cone generated by product effects.

In finite dimensions, $A \otimes_{\text{min}} B$ and $A \otimes_{\text{max}} B$ both coincide with $A \otimes B$ *as vector spaces*, i.e., disregarding their ordered structure. However, the minimal tensor cone $(A \otimes_{\text{min}} B)_+$ is generally much smaller than the maximal tensor cone, $(A \otimes_{\text{max}} B)_+$, unless at least one of the two systems is classical. With the maximal TP *or* the minimal TP, the category $\mathbf{Com}$ is symmetric monoidal. Equipped with both, it is linearly distributive. Notice, however, that the usual tensor product of quantum systems — obtained by forming the tensor product of the underlying Hilbert spaces — coincides with neither the maximal nor the minimal tensor product. Rather, the cone of bipartite density operators of the composite quantum system lies strictly between the minimal and maximal tensor cones.

Definition: A *symmetric monoidal probabilistic theory* is a sub-category $\mathcal{C}$ of $\mathbf{Com}$, equipped with a monoidal product $\otimes$ such that, for all $A, B \in \mathcal{C}$, $A \otimes_{\text{min}} B \leq A \otimes B \leq A \otimes_{\text{max}} B$.

By way of example, the theory corresponding to finite-dimensional QM has, for objects, the real vector spaces of Hermitian operators on finite-dimensional *complex* Hilbert spaces, and, for objects, completely positive mappings between such spaces. One can also construct a symmetric monoidal category by freely combining such spaces using the maximal tensor product, and using arbitrary positive mappings as morphisms.

Representation Theorem The following observation shows that our terminology, "symmetric monoidal probabilistic theory", does not overreach. Let $\mathcal{C}$ be a SMC having the property that $\mathcal{C}(A, I) \times \mathcal{C}(B, I)$ separates points of $\mathcal{C}(I, A \otimes B)$, and let $S = \mathcal{C}(I, I)$ be the associated commutative monoid of scalars. Given any mapping $p : S \to [0, 1]$, whereby scalars can be interpreted probabilistically, we can construct a Mackey triple (X_A, Σ_A, p) where $X_A = \mathcal{C}(A, I)$, $\Sigma_A = \mathcal{C}(A, I)$, and $p(x, \alpha) = p(x \circ \alpha)$. If p is a monoid homomorphism, i.e, $p(s \otimes t) = p(s)p(t)$, the linearization construction described leads to a monoidal probabilistic theory:

Proposition 1: *Let $\mathcal{C}$ be as above. For every monoid homomorphism $p : S \to [0, 1]$, there exists a symmetric monoidal category $V_p(\mathcal{C})$ of convex operational models, and a monoidal functor $V : \mathcal{C} \to V(\mathcal{C})$ such that for all objects $A \in \mathcal{C}$, and all $a \in \mathcal{C}(A, I)$, $\alpha \in \mathcal{C}(I, A)$, we have $Va(V\alpha) = p(a \circ \alpha)$. If p is injective, V is an embedding.*

3 Remote Evaluation, Compact Closure and Dagger-compactness

Let $\mathcal{C}$ be a symmetric monoidal probabilistic theory, and let $\omega : I \to A \otimes B$ and $f : A \otimes B \to B$ be respectively a bipartite state and effect on the composite system $A \otimes B$. Then there are canonical mappings $\hat{\omega} : \mathcal{C}(A, I) \to \mathcal{C}(I, B)$ and $\hat{f} : \mathcal{C}(I, A) \to \mathcal{C}(B, I)$ given by

$$\begin{array}{ccc} I & \xrightarrow{\omega} & A \otimes B \\ & {\scriptstyle \hat{\omega}(a)} \searrow & \downarrow {\scriptstyle a \otimes \mathrm{id}_B} \\ & & B \end{array} \quad \text{and} \quad \begin{array}{ccc} B & \xrightarrow{\alpha \otimes \mathrm{id}_B} & A \otimes B \\ & {\scriptstyle \hat{f}(\alpha)} \searrow & \downarrow {\scriptstyle f} \\ & & I \end{array} . \qquad (4)$$

If $\mathcal{C}$ is *already* a category of COMs, then $\hat{\omega}(a)(b) = \omega(a, b)$, and hence, $\widehat{\omega}(a)/u_B(\widehat{\omega}(a))$ is exactly the conditional state of B given that the effect a has occurred on A. Accordingly, we call $\widehat{\omega}$ the *conditioning map* associated with ω. A consequence of the Lemma is that, for every bipartite state $\omega \in \mathcal{C}(I, A \otimes B)$, the conditional states $\omega_{B|a}$ and $\omega_{A|b}$ are legitimate states of A and B, i.e., elements of $\mathcal{C}(I, A)$ and $\mathcal{C}(I, B)$,respectively. An easy diagram-chase gives us

Lemma 2 [Remote Evaluation]: *Let $\omega : I \to B \otimes C$ and $f : A \otimes B \to I$ in $\mathcal{C}$. Then, for all $\alpha \in \mathcal{C}(I, A)$,*

$$\hat{\omega}(\hat{f}(\alpha)) \;=\; (f \otimes id_C) \circ (id_A \otimes \omega) \circ \alpha = (f \otimes id_C) \circ (\alpha \otimes \omega) \qquad (5)$$

When $\omega(1) \in A \otimes B$ is a normalized state, and $f : B \otimes C \to I$ is an effect, the Lemma says that the mapping $\hat{\omega} \circ \hat{f}$ is represented, within the category $\mathcal{C}$, by the composite morphism $(f \otimes \mathrm{id}_C) \circ (\mathrm{id}_A \otimes \omega)$. In other words, preparing $B \otimes C$ in joint state ω, and then measuring $A \otimes B$ and obtaining f, guarantees that the "un-normalized conditional state" of C is $\hat{\omega}(\hat{f}(\alpha))$, where α is the state of A. Thus, *any process that factors as $\hat{\omega} \circ \hat{f}$* can be understood as an instance of "un-normalized classical conditioning". In particular, there is no need to invoke any mysterious "collapse" of the state, nor for that matter, any other *physical* dynamics at all.

If $C = A$, and $\hat{\omega} : B^{\#} \to A$ is an isomorphism of ordered linear spaces, we can rescale $\widehat{\omega}^{-1}$ to obtain a bilinear form f on $A \otimes B$ that is positive on $A_+ \times B_+$ (i.e. an effect for the minimal tensor product), given by $f(\alpha, \beta) = c\hat{f}(\alpha)(\beta)$ for a suitable constant $c > 0$. If a measurement on $A \otimes B$ yields outcome f, with the composite system $A \otimes B \otimes C$ in state $\alpha \otimes \omega$, the un-normalized conditional state of C is $\hat{\omega}(\hat{f}(\alpha)) = c\alpha$, whence, the normalized state of C is α. In other words, we have here a conclusive, correction-free teleportation protocol. Moreover, $\hat{\omega}^{\#} : B^{\#} \simeq A$, together with its inverse, supplies a protocol for teleporting B's states through A. We shall say that A and B *teleport one another* when there exist such an isomorphism state $\omega \in A \otimes B$ and effect $f \in (A \otimes B)^{\#}$. It is clear that, in this case, ω and $c^{-1}f$ supply a unit and co-unit for a duality. This gives us the implication $(b) \Rightarrow (a)$ in:

Proposition 2: *Let $\mathcal{C}$ be a monoidal category of COMs. The following are equivalent.*

(a) $\mathcal{C}$ is compact closed.

(b) For every $A \in \mathcal{C}$ there is some $B \in \mathcal{C}$ such that A and B teleport one another.

(c) Every morphism in $\mathcal{C}$ has the form $\hat{\omega} \circ \hat{f}$ for some bipartite state ω and bipartite effect f.

Dagger Compactness In order to obtain a dagger-compact category, it is natural to consider the case in which every $A \in \mathcal{C}$ supports an *automorphism state* $\omega \in A \otimes A$ with $\widehat{\omega} : A^{\#} \simeq A$, and an *automorphism effect* f on $A \otimes A$ with $\hat{f} : A \simeq A^{\#}$, with $\hat{f}$ a multiple of $\widehat{\omega}^{-1}$. In the language used above: A teleports itself. In this case, we say that C is *weakly self-dual* (WSD). Where ω can be chosen to be symmetric, we say that $\mathcal{C}$ is *symmetrically* WSD. Note that, by Proposition 2, such a theory is a degenerate compact closed category, and we have:

Proposition 3: A monoidal category $\mathcal{C}$ of convex operational models is weakly self-dual iff it is compact closed, and can be equipped with a degenerate compact structure.

Proposition 4: *Let $\mathcal{C}$ be a WSD probabilistic theory. Then $\mathcal{C}$ is dagger-compact with respect to the canonical adjoint $'$, if and only if $\mathcal{C}$ is symmetrically WSD.*

A symmetric isomorphism state on a convex operational model A, amounts to a symmetric, non-degenerate bilinear form on A. If this form is also positive-definite, i.e, an inner product, then the cone A_+ is said to be *self-dual*. Finite dimensional *homogeneous* self-dual cones (those whose automorphism groups act transitively on their interiors) are exactly the positive cones of euclidean (or formally real) Jordan algebras, by classical results of Koecher and Vinberg. This is close to narrowing things down to standard quantum theory, as such cones are direct convex sums of the irreducible ones which are either Lorentz cones (cones with a ball as base), mixed-state spaces for real, complex, and quaternionic quantum theory, and an isolated a three-dimensional octonionic analogue of quantum theory. This connection will be pursued in another paper.

Conclusion and Prospectus:

In the present work we have established, within a relatively loose framework of symmetric monoidal categories of convex operational models, that compact closure and dagger compact closure impose very significant restrictions on the convex structure of state spaces. For example, degenerate compact closure imposes weak self-duality of the cone of unnormalized states, while dagger compact closure imposes that the bilinear form associated with the map implementing weak self-duality is *symmetric*, a stronger property that moves us in the direction of what is usually called self-duality *tout court* (for which the form is not just symmetric but in fact positive semidefinite), and hence in the direction of quantum theory. The compact closed structure is associated with information-processing capacities of the theory, notably, as expected, the capacity to teleport systems, and also with conceptual properties of the theory, notably the fact that any process can be interpreted in terms of conditioning.

Several directions for further study suggest themselves. It would be interesting to identify necessary and sufficient conditions for the COM representations of Proposition 1 to yield *finite dimensional* models – and, equally, one would like to know how far our other results can be extended, with suitable adaptation, to infinite-dimensional systems. Our definition of category of weakly self-dual state spaces assumes the existence of a state that induces, by conditioning, an isomorphism from the state cone to the effect cone, and an effect inducing its inverse; it would be interesting to investigate conditions under which this follows just from weak self-duality of the objects. As already mentioned, the consequences of homogeneity of the state-spaces should also be explored. Perhaps the most urgent task, though, is to identify operational and category-theoretic conditions equivalent to the strong self-duality of a probabilistic theory.

References

[1] S. Abramsky and B. Coecke, A categorical semantics of quantum protocols, in *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science: LICS 2004*, pp. 415–425, IEEE Computer Society (2004); also quant-ph/0402130v5 (2004).

[2] J. Baez, Quantum quandaries: a category-theoretic perspective, available from `quant-ph/0404040`, 2004.

[3] H. Barnum, J. Barrett, M. Leifer and A. Wilce, Teleportation in general probabilistic theories, in Proceedings of the Clifford Lectures, Tulane University, March 12-15, 2008, to appear in *Proceedings of Symposia in Applied Mathematics* (AMS);

[4] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, Comm. Math. Phys. **17** 239-260 (1970).

[5] G. M. Kelly and M. L. Laplaza, Coherence for compact closed categories, J. Pure Appl. Algebra (1980) 193-213.

[6] G. Ludwig, Foundations of Quantum Mechanics I, Springer, 1983.

[7] G. Mackey, Mathematical Foundations of Quantum Mechanics, Addison-Wesley, 1963

[8] P. Selinger, Dagger compact closed categories and completely positive maps (extended abstract), in Proceedings of QPL 05. Electronic Notes in Theoretical Computer Science **170**, 139–163 (2007).

[9] P. Selinger, Towards a semantics for higher-order quantum computation, in *Proceedings of the 2nd International Workshop on Quantum Programming Languages, Turku, Finland*, TUCS General Publication No. 33, pp, 127–143 (2004).