Prolegomena to the CBH Theorem

Hans Halvorson

November 23, 2005

(Version 0.1, Revised 7:08pm, November 19, 2005.)

Contents

1 A Common Framework for Classical and Quantum Theories

The mathematical formalisms for classical and quantum physics look nothing alike. This is surely one reason why physics students find it difficult to learn quantum mechanics. It is surely also one of the reasons why QM fascinates philosophers — because its formalism seems to resist comparison with the formalism of classical physics.

Our goal: We want to overcome this situation of apparent Kuhnian incommensurability. We want to find a common playing field in which we can discuss both sorts of theory. We want to find a formalism that encompasses *both* classical physics (particle mechanics, classical field theories) and quantum mechanics.

To reach our goal, we begin by describing again the prototypical formalisms for classical and quantum mechanics. Then we will abstract away until we find the common elements.

$\boxed{\textbf{System Q}}$ A single particle alone in the world is described by a vector space (Hilbert space) $\mathcal{H}$, where its state at a given time t is given by a vector $x(t) \in \mathcal{H}$. The physical quantities (misleadingly called "observables") correspond to *linear operators* on $\mathcal{H}$. Actually, we only consider those linear operators whose expectation values:

$$\langle x, Ax \rangle$$

are real numbers, for each $x \in \mathcal{H}$. Such linear operators are called *self-adjoint*.

$\boxed{\textbf{System C}}$ A single particle alone in the world is described by a "phase space" $\mathbb{R}^6$, where its state at a given time t is given by a pair $(\mathbf{q}(t), \mathbf{p}(t)) \in \mathbb{R}^6$, its position and momentum. Here position and momentum are the basic physical quantities, and other physical quantities are functions of position and momentum. What sort of mathematical object represents position, momentum, and the other physical quantities? Position and momentum are coordinates, but coordinates are just functions from phase space into real numbers, and so functions of position and momentum are also just functions from phase space into real numbers.

The state spaces of system Q and system C look, at first glance, to be similar — they are both vector spaces. But that analogy is purely superficial, because the physical quantities do not relate in the same way to the structure of the state space.

We will look elsewhere for our analogy — to the physical quantities. The physical quantities of System Q have the following structure:

- We can add two linear operators. If A and B are linear operators on $\mathcal{H}$, then we can set $(A+B)v = Av + Bv$ (using the addition operation on $\mathcal{H}$) to define another linear operator. We can also multiply a linear operator by a complex number: $(cA)v = c(Av)$. Plus there is a zero linear operator, namely the operator that maps each vector to the zero vector. The overall result is that the set of linear operators on a vector space is itself a vector space!
- The composition of two linear operators

$$x \longmapsto Bv \longmapsto A(Bv),$$

has the characteristics of a multiplication operation. i.e. we define an operation $\bullet$ by $(A \bullet B)v = A(Bv)$. There is also an operator I (that maps each vector to itself) that serves as a two-sided multiplicative identity. (But note that this multiplication operation is not commutative, i.e. in some cases $A \bullet B \neq B \bullet A$.) Furthermore, multiplication "plays nicely" with the vector space structure on the set of linear operators; e.g. $A \bullet (B + C) = (A \bullet B) + (A \bullet C)$ and $c(A \bullet B) = (cA) \bullet B = A \bullet (cB)$.

- There is a unary operation on linear operators called "taking the adjoint." We denote the adjoint of A by A^*, and it can be checked that this operation is conjugate linear and reverses multiplication. The physical quantities are represented by self-adjoint operators, i.e. those operators satisfying $A = A^*$.

The physical quantities of System C have the following structure:

- We can add two functions. If f and g are functions on $\mathbb{R}^n$ then we can set $(f + g)(x) = f(x) + g(x)$, and for $a \in \mathbb{R}$, $(af)(x) = a(f(x))$. The function that is constantly zero also serves as a zero element for this vector space over $\mathbb{R}$.

  Now there is a basic disanalogy here that we need to work around. While the linear operators on $\mathcal{H}$ are a complex vector space, the functions on $\mathbb{R}^n$ are a real vector space. It turns out that this disanalogy is superficial, and we can use a formal trick to avoid it. Look at a *larger* class of functions: the complex-valued functions on $\mathbb{R}^n$. Now we have a complex vector space. But we still say that the physical quantities are those complex-valued functions that just happen always to have values in the real numbers (i.e. the imaginary part is zero). So, this is just like the quantum case: The physical quantities are embedded in slightly larger space, the "complexification" of the original space.

- Functions can be multiplied "pointwise". That is, we define a function $f \bullet g$ by $(f \bullet g)(x) = f(x)g(x)$. (This multiplication operation is commutative, i.e. $f \bullet g = g \bullet f$, because multiplication of complex numbers is commutative.) There is a multiplicative identity, namely the function that is constantly 1. The multiplication operation also plays nicely with the vector space structure on the set of functions.

- For each complex-valued function f, there is another function f^* defined by $f^*(x) = \overline{f(x)}$. That is, the value of f^* at x is the complex conjugate of the value of f at x. This unary operation $f \mapsto f^*$ is conjugate linear and reverses multiplication. The physical quantities are represented by self-adjoint functions, i.e. those functions satisfying $f = f^*$.

1.1 $*$-algebras

So now let's summarize the common structure of the space $\mathfrak{A}$ of physical quantities:

- $\mathfrak{A}$ is a vector space over the complex numbers.
- $\mathfrak{A}$ has a multiplication operation that plays nicely with the vector space structure. $\mathfrak{A}$ also has a multiplicative identity (we henceforth designate this identity by I). In this case, the mathematicians tell us to say that $\mathfrak{A}$ is an *algebra over* $\mathbb{C}$.
- There is a unary operation $*$ on $\mathfrak{A}$ that is conjugate linear:

$$(cA + B)^* = \overline{c}A^* + B^*,$$

  and antimultiplicative:

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

  An algebra over $\mathbb{C}$ with such an operation is called a $*$-*algebra*.

Note. For historical reasons, the elements of a $*$-algebra are called "operators." This does not mean that they actually operate on anything. e.g. in the case of System C, the elements of $\mathfrak{A}$ are functions, not operators! This potentially confusing terminology persists only because there is a deep theorem — called the Gelfand-Naimark-Segal representation theorem — that shows that a $*$-algebra can always be embedded into the set of linear operators on some vector space. (The branch of mathematics that studies $*$-algebras is called *operator algebras*.)

Note. The CBH paper talks about C^*-algebras rather than $*$-algebras. A C^*-algebra is a $*$-algebra plus something more: a norm (a special function that gives a notion of distance between operators) that plays nicely with the algebraic structure. But we need to specify a norm only in the case where the algebra $\mathfrak{A}$ is infinite dimensional (as a vector space). In the finite dimensional case, the elements of $\mathfrak{A}$ can be represented as matrices of complex numbers, and there is only one possible notion of distance between the operators, namely the notion we get by looking at the complex number distance between matrix entries. (More precisely, there is a unique norm — up to equivalence — on $\mathfrak{A}$ relative to which addition and scalar multiplication are continuous.)

Just in case you aren't yet motivated to study $*$-algebras, we discuss one more case.

Claim. *Every classical probability space corresponds to a $*$-algebra (and so the theory of $*$-algebras generalizes classical probability theory).*

Definition. A classical probability space is a triple $\langle X, \Sigma, \mu \rangle$ where X is a set, Σ is a Boolean σ-algebra of subsets of X, and μ is a probability measure on Σ. The "events" are the elements of Σ, and the probability of an event $S \in \Sigma$ is given by $\mu(S)$.

Proof. Let (X, Σ, μ) be a classical probability space. The construction of the corresponding $*$-algebra is straightforward. Let $\mathfrak{A}$ consist of all complex-valued functions on X that are measurable relative to Σ. That is, for each open subset U of complex numbers, we require that $f^{-1}(U) \in \Sigma$. By using the pointwise definitions of addition, scalar multiplication, multiplication, and $*$, it is easy to verify that $\mathfrak{A}$ is a $*$-algebra. The "events" from the classical probability space now correspond to operators in $\mathfrak{A}$. In particular, $S \in \Sigma$ corresponds to the characteristic function e_S of S:

$$e_S(x) = \begin{cases} 1 & x \in S, \\ 0 & x \notin S. \end{cases}$$

Note that $e_S \bullet e_S = e_S$ and $(e_S)^* = e_S$. That is, e_S behaves just like a projection operator. $\square$

1.2 States

We have found a common framework in the structure of the physical quantities. But now we need to find a place for the <u>states</u>. Let's recall the role of the states relative to the quantities in the two different sorts of theories.

$\boxed{\textbf{System Q}}$ Let $\mathfrak{A}$ denote the algebra of linear operators on a vector space $\mathcal{H}$. A state is a unit-length vector v, which assigns some complex number $\langle v, Av \rangle$ to each $A \in \mathfrak{A}$. If A represents a physical quantity, i.e. $A = A^*$, then

$\langle v, Av \rangle$ is a real number. So, a state assigns a real number (the "expectation value") to each physical quantity. Moreover, this assignment of real numbers is linear:

$$\langle v, (A+B)v \rangle = \langle v, Av \rangle + \langle v, Bv \rangle,$$

by the very definition of the $+$ operation on $\mathfrak{A}$. So, a state is a linear mapping from $\mathfrak{A}$ into $\mathbb{C}$, and it assigns real numbers to the self-adjoint elements of $\mathfrak{A}$. Furthermore,

$$\langle v, A^*Av \rangle = \langle Av, Av \rangle = \|Av\|^2 \geq 0.$$

So, this linear mapping always assigns operators of the form A^*A a positive real number. (Naturally enough, such linear mappings are called "positive.") Finally, $\langle v, Iv \rangle = \|v\|^2 = 1$. So, a state is a positive linear functional on $\mathfrak{A}$ that assigns 1 to I.

$\boxed{\textbf{System C}}$ Let $\mathfrak{A}$ denote the algebra of complex-valued functions on phase space X. A state is a point in X. A point $x \in X$ is not a linear mapping on $\mathfrak{A}$, but it is naturally related to one — define a complex-valued function $\hat{x}$ on $\mathfrak{A}$ by

$$\hat{x}(f) = f(x), \qquad\qquad (f \in \mathfrak{A}),$$

and check that it is linear (i.e. it is a linear functional on $\mathfrak{A}$). Moreover, we have $\hat{x}(f^*f) = \overline{f(x)}f(x) = |f(x)|^2 \geq 0$. That is, the linear functional $\hat{x}$ is positive. And of course $\hat{x}(I) = 1$, where I is the function on X that takes the value 1 everywhere.

$\boxed{\textbf{Classical Probability Theory}}$ Let $\langle X, \Sigma \rangle$ be a classical probability space, and let $\mathfrak{A}$ be the $*$-algebra of complex-valued measurable functions. What now is a state? A state is a probability measure on $\langle X, \Sigma \rangle$. But each probability measure ρ on $\langle X, \Sigma \rangle$ gives rise to a mapping from $\mathfrak{A}$ into $\mathbb{C}$, namely

$$f \longmapsto \int_X f \, d\rho.$$

That is, we integrate f by the measure ρ. (The integral of a complex-valued function is just the complex number obtained by integrating its real and imaginary parts separately and putting them back together again.) This assignment of complex numbers is linear, is positive, and assigns 1 to the identity function (since ρ is a probability measure).

Abstracting away from these cases, we settle on the following definition.

Definition. If $\mathfrak{A}$ is a $*$-algebra, then a *state* (or expectation functional) on $\mathfrak{A}$ is a linear functional ω such that:

- ω is positive, i.e. $\omega(A^*A) \geq 0$ for all $A \in \mathfrak{A}$;
- $\omega(I) = 1$.

1.3 Composition of systems

In both the classical and the quantum settings, we must be able to represent the composite of two systems. Here we show that the two notions of composition correspond to a common notion in the framework of $*$-algebras.

$\boxed{\textbf{System Q}}$ Given two state spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, we construct the composite system as the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$. Now the physical quantities of the composite system are linear operators on $\mathcal{H}_1 \otimes \mathcal{H}_2$. But it turns out that the $*$-algebra of linear operators on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is a tensor product of the $*$-algebras of linear operators on the individual vector spaces. To be more precise, let $\mathfrak{A}$ denote the $*$-algebra of linear operators on $\mathcal{H}_1$, and let $\mathfrak{B}$ denote the $*$-algebra of linear operators on $\mathcal{H}_2$. Then the mapping $A \mapsto A \otimes I$ gives an embedding of $\mathfrak{A}$ into $L(\mathcal{H}_1 \otimes \mathcal{H}_2)$, and similarly the mapping $B \mapsto I \otimes B$ gives an embedding of $\mathfrak{B}$ into $L(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Furthermore, the set of elements of the form $A \otimes B$ spans $L(\mathcal{H}_1 \otimes \mathcal{H}_2)$. That is, every element of $L(\mathcal{H}_1 \otimes \mathcal{H}_2)$ can be written in the form

$$(A_1 \otimes B_1) + \cdots + (A_n \otimes B_n),$$

with $A_i \in \mathfrak{A}$ and $B_i \in \mathfrak{B}$.

$\boxed{\textbf{System C}}$ Let X and Y be two classical phase spaces. The phase space of the composite system is the Cartesian product of the two sets $X \times Y$. That is, we just take the set of ordered pairs (x, y), where x is a state of the first system and y is a state of the second system. Thus, the physical quantities of the composite system are functions on $X \times Y$.

Let $\mathfrak{A}$ denote the $*$-algebra of functions on X, and let $\mathfrak{B}$ denote the $*$-algebra of functions on Y. Then a wonderful thing happens: the $*$-algebra of functions on $X \times Y$ is naturally isomorphic to the tensor product $\mathfrak{A} \otimes \mathfrak{B}$ of $\mathfrak{A}$ and $\mathfrak{B}$. This is easy to check in simple cases. For example, if X and Y

are finite sets, then obviously each complex-valued function on $X \times Y$ is a sum of functions of the form

$$(x, y) \longmapsto f_i(x) g_i(y).$$

In the general case, some delicacies are involved — we have to require that X and Y are compact Hausdorff spaces, and we have to restrict to those functions on X, Y, and $X \times Y$ that are continuous and bounded. (But this is no real restriction, because each phase space you can imagine can be embedded into a compact Hausdorff space. The embedding operation is called a "compactification" of the space.)

Classical Probability Theory The previously discussed case also covers classical probability theory. If we have two classical probability spaces $\langle X_1, \Sigma_1 \rangle$ and $\langle X_2, \Sigma_2 \rangle$, then the product probability space is $\langle X_1 \times X_2, \Sigma \rangle$, where Σ is the smallest Boolean σ-algebra that contains all sets of the form $S_1 \times S_2$ with $S_i \in \Sigma_i$. Then the $*$-algebra constructed from $\langle X_1 \times X_2, \Sigma \rangle$ will just be the tensor product of the $*$-algebras constructed from $\langle X_1, \Sigma_1 \rangle$ and $\langle X_2, \Sigma_2 \rangle$.

Definition. Let φ be a mapping from a $*$-algebra $\mathfrak{A}$ to another $*$-algebra $\mathfrak{B}$. Then φ is said to be a *homomorphism* if φ is linear, multiplicative, and $\varphi(A^*) = (\varphi(A))^*$ for all $A \in \mathfrak{A}$. If, additionally, φ is one-to-one and onto, then φ is said to be a *isomorphism*.

2 Distinguishing between Classical and Quantum

When does a $*$-algebra describe a classical system, and when does a $*$-algebra describe a quantum system?

The first question has a nice crisp answer.

Definition. A $*$-algebra $\mathfrak{A}$ is said to be *abelian* just in case $AB = BA$ for all $A, B \in \mathfrak{A}$.

Theorem 1 (Gelfand representation). *A finite-dimensional $*$-algebra $\mathfrak{A}$ is abelian iff $\mathfrak{A}$ is isomorphic to the $*$-algebra $C(X)$ of complex-valued functions on a finite set X.*

Sketch of proof. The key here is identifying the set X. Let X be the set of pure states on $\mathfrak{A}$, where a pure state is one that cannot be written as a mixture of any two distinct states. The set X must be finite, because $\mathfrak{A}$ is a finite-dimensional vector space, and the dual vector space (i.e. the linear functionals on $\mathfrak{A}$) has the same dimension as $\mathfrak{A}$. Furthermore, each $A \in \mathfrak{A}$ corresponds to a complex-valued function on X, namely the function that takes the value $\omega(A)$ at the point $\omega \in X$. The remainder of the proof is straightforward. □

For the fully general case, we have additional complications, and so are forced to talk about C^*-algebras.

Theorem 2 (Gelfand representation). *A C^*-algebra $\mathfrak{A}$ is abelian iff $\mathfrak{A}$ is isomorphic to the C^*-algebra $C(X)$ of continuous complex-valued functions on a compact Hausdorff space X.*

Sketch of proof. As above. But putting a topology on X is tricky. See a standard textbook for details. □

So, the above result tells us that the phase space picture — i.e. functions on a set X — is equivalent to having an abelian $*$-algebra. So, a classical system is one where the physical quantities are commutative.[1]

The second question is harder to answer. But we are going to stipulate that (quantum = not-classical). In other words, (quantum = nonabelian algebra of observables). However, there is no analogue here of the Gelfand representation. It is *not* true that each nonabelian $*$-algebra corresponds to the $*$-algebra of linear operators on some vector space.

We now have to do some work in order to define the distance between two states on a $*$-algebra. We will only consider the case of a finite-dimensional $*$-algebra. For the general case, we would need to talk about C^*-algebras.

Definition. For each $A \in \mathfrak{A}$ we write $A \geq 0$ just in case $A = B^*B$ for some $B \in \mathfrak{A}$. We write $A \geq B$ just in case $A - B \geq 0$.

Definition. Let ω be a linear functional on $\mathfrak{A}$. Then we set

$$\|\omega\| = \sup\{\, |\omega(A)| : -I \leq A \leq I \,\}.$$

[1] This oversimplifies issues. Having a commutative algebra of observables is at best a necessary condition on being "classical." For example, it is plausible to say that Bohmian mechanics has a commutative algebra of observables (viz. just the functions of the position observable), but is not a classical theory.

When $\mathfrak{A}$ is finite dimensional, $\|\omega\| < \infty$ (as can be seen by checking the values of ω on a basis for $\mathfrak{A}$). The norm $\|\cdot\|$ gives us a notion of the distance between two states: if ω and ρ are states, then $\omega - \rho$ is a linear functional, and we define $\|\omega - \rho\|$ to be the "distance" between ω and ρ. (One can verify that $d(\omega, \rho) = \|\omega - \rho\|$ is a metric on the state space of $\mathfrak{A}$.)

We now define the "transition probability" between two states of a $*$-algebra.

Recall that the transition probability between two vectors v, w is given $|\langle v, w\rangle|^2$. We need to translate this into a purely algebraic notion. We have the following result.

Proposition 3. *Let $\mathfrak{A}$ be the $*$-algebra of linear operators on a finite dimensional vector space V. For unit vectors $x, y \in V$, let ω_x, ω_y denote the corresponding states of $\mathfrak{A}$. Then we have*

$$|\langle x, y\rangle|^2 = 1 - \frac{1}{4}\|\omega_x - \omega_y\|^2.$$

Proof. Why spoil your fun? □

Definition. If ω and ρ are states of $*$-algebra, then we define their transition probability by

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

Note. The transition probability takes values in $[0, 1]$, is symmetric, and equals 1 iff $\omega = \rho$.

Definition. We say that ω and ρ are *orthogonal* just in case $\mathbf{p}(\omega/\rho) = 0$, which is equivalent to $\|\omega - \rho\| = 2$.

Theorem 4. *Let $\mathfrak{A}$ be a $*$-algebra. Then the following are equivalent.*

1. *$\mathfrak{A}$ is nonabelian.*
2. *$\mathfrak{A}$ has non-uniquely decomposable mixtures.*
3. *There are states ω, ρ of $\mathfrak{A}$ such that $0 < \mathbf{p}(\omega/\rho) < 1$.*

Sketch of proof. The proof of (1) $\Rightarrow$ (2) is in the CBH paper. The proof of (2) $\Rightarrow$ (1) follows from the Gelfand representation in conjunction with the fact that the states on $C(X)$ are in one-to-one correspondence with measures on X. Also, the set of measures on X is a "simplex": each measure on X gives a unique weighting to the points in X, which are the pure states of $C(X)$. The proof that (1) $\Leftrightarrow$ (3) is also discussed in the CBH paper. □