Phil Physics: Week 5

Review

Let's return to the map of "different places where quantum weirdness might be located." We've already talked about superposition (a relation on quantum states), and a little bit about entangled states. Last week with Alex you talked about what many people think is the absolute most central problem of QM: the measurement problem. Let's begin with a short discussion of how you assess the relation between the measurement problem and the previous two issues.

Cards on the table: I myself am somewhat puzzled that people put so much weight on the measurement problem. Here's a question: can you imagine a quantum world in which there were no observers whatsoever — where, by definition, there can be no measurement problem? How would you understand superpositions in this world? How would you understand entangled states in this world? Do you think that *those* questions are easy to answer, but that there's something mysterious about what happens during measurements? My feeling is that once you've done all the work to understand how quantum theory could be true in a world without observers, then it won't be an additional huge challenge to understand how quantum theory could be true in a world with observers. In other words, I would suggest that there is a problem that comes up *before* the measurement problem: how to understand superpositions and entangled states. (For further reflection: what would it mean to understand superpositions and entangled states?)

Hidden variables

The subject this week (and, to some degree, next week) are the famous "no hidden variables theorems." Some people think that these theorems are deeply significant, i.e. they tell us something important about the quantum world. Other people think that these theorems are massively overrated, and indeed misleading. To be honest, I've never seen a good *argument* for either one of those views. I think it's a matter of taste — and I happen to be among those who find these theorems to be illuminating.

Historical overview

1932 First NHV theorem proven by John von Neumann (Von Neumann, 1932)

1935 Von Neumann's proof criticized by philosopher-physicist Grete Hermann (1935). Unfortunately, Hermann was largely ignored.

1952 David Bohm constructs a hidden variable model, seemingly in contradiction with von Neumann's result. (Some would say that Bohm rediscovered Louis De Broglie's theory, and there are strong similarities between Bohm's views and the views of Schrödinger.)

1957 Mackey conjectures and Gleason proves: when $\dim H \geq 3$, every probability measure on subspaces (projections) is represented by a quantum state. (Gleason, 1957)

1964–66 John Bell criticizes von Neumann's proof; introduces idea of contextual hidden variables; proves no local hidden variables theorem (i.e. derivation of Bell inequality); uses Gleason's theorem to derive a result equivalent to the Kochen-Specker theorem

1967 Kochen and Specker prove a NHV theorem without linearity assumption (Kochen and Specker, 1967)

1972–2018 Bell tests, see `https://en.wikipedia.org/wiki/Bell_test_experiments`

1988 "The von Neumann proof, if you actually come to grips with it, falls apart in your hands! There is nothing to it. It's not just flawed, it's silly! . . . When you translate [his assumptions] into terms of physical dispositions, they're nonsense. You may quote me on that: The proof of von Neumann is not merely false but foolish!" (Bell, 1988)

1993 "Many generations of graduate students who might have been tempted to try to construct hidden-variables theories were beaten into submission by the claim that von Neumann, 1932, had proved that it could not be done. …A third of a century passed before John Bell, 1966, rediscovered the fact that von Neumann's nohidden-variables proof was based on an assumption that can only be described as silly—so silly, in fact, that one is led to wonder whether the proof was ever studied by either the students or those who appealed to it to rescue them from speculative adventures." (Mermin, 1993)

2006 Conway-Kochen free will theorem (Conway and Kochen, 2006)

Framework assumptions

To clarify one thing: you cannot prove any mathematical theorem without adopting a mathematical framework. You've got to make some assumptions about what mathematical things will be used to represent things like: propositions, properties, truth-values, states, probabilities, etc.

The NHV theorems we will look at fall under one of the following two frameworks.

1. Logical: show that we cannot simultaneously assign truth values to all propositions that can (at some time or other) be asserted about a quantum system.

2. Algebraic: show that we cannot simultaneously assign dispersion-free expectation values to all quantities that (at some time or other) a system can have.

For many of you in this class — e.g. philosophy concentrators — the logical result will be easier to understand. However, the logical framework is farther away from the actual formalism that physicist use. So here is how I will proceed. I will first explain the result in the algebraic framework. During this part, you philosophy types shouldn't worry if you have trouble following step by step. Then I'll explain how the algebraic result translates into a logical result, and that's when everybody needs to pay close attention again.

Quantities and states

What we know so far:

1. Quantum **states** are represented by vectors in some space H. These states can be superposed.

2. A **quantity** is represented by an assignment of real numbers to orthonormal basis of vectors. For example, spin-z is represented by an assignment of $+1$ to one vector, and -1 to an orthogonal vector. (Note: people often call these **observables** instead of quantities.) The numbers here are the possible values that the quantity can take.

3. With the above two conventions, the Born rule tells us how to compute, for any given quantum state, and for any given quantity, the probability that quantity will have a certain value in that quantum state.

We're now going to explain a second way of thinking about quantities, which is captured by the following correspondence:

quantity (aka observable)	assignment of real numbers to an orthonormal set of vectors	self-adjoint linear operator
property (aka proposition)	subspace of Hilbert space	projection operator
state	assignment of expectation values to quantities	density operator

Under this correspondence, the orthonormal basis of a quantity are the **eigenvectors** of the operator, and the assigned real numbers are the **eigenvalues** of the operator.

So now for some precise definitions.

Definition. Let H and K be vector spaces. A **linear operator** $A : H \to K$ is a function such that $A(x+y) = Ax+Ay$ and $A(\lambda x) = \lambda Ax$, for all $x, y \in H$ and for all $\lambda \in \mathbb{C}$.

Here we're using complex numbers $\mathbb{C}$ for our scalars.

Exercise. Given linear operators $A : H \to K$ and $B : H \to K$, define $A+B$ to be the function that assigns $Ax + Bx$ to x. Show that $A + B$ is a linear operator. Similarly, for $\lambda \in \mathbb{C}$, define $(\lambda A)x = \lambda(Ax)$, and show that λA is a linear operator. It's straightforward to verify that $L(H, K)$ is itself a vector space.

For the next result, we need to make use of the following fact: if x and y are unit vectors, then $\langle y, x\rangle = 1$ iff $x = y$. For the "only if" part, note first that $x = \langle y, x\rangle y + z$, where z is a vector such that $\langle y, z\rangle = 0$, and $|\langle y, x\rangle|^2 + \|z\|^2 = 1$. (Here z is the projection of x onto the subspace orthogonal to y.) Hence, if $\langle y, x\rangle = 1$, then $z = 0$ and $x = y$.

1 Theorem. *Let H be a finite-dimensional inner product space. For each $y \in H$, then the equation $\varphi_y(x) = \langle y, x\rangle$ defines a linear functional φ_y on H. Moreover, each linear functional on H arises, in this way, from a unique element $y \in H$.*

Proof. The first claim follows immediately from the fact that the inner product is linear in the second argument. For the second claim, let $\varphi : H \to \mathbb{C}$ be a linear functional. If $\varphi(x) = 0$ for all $x \in H$, then the result follows with $y = 0$. So suppose that $\varphi(x) \neq 0$ for some $x \in H$. Let K be the kernel of ϕ, i.e. the subspace of vectors $x \in H$ such that $\phi(x) = 0$. Since ϕ isn't constantly zero, $K \neq H$, and $K^\perp$ is nonempty. Let u be a unit vector in $K^\perp$, and note that

$$\varphi\left(\varphi(u)x - \varphi(x)u\right) \;=\; 0,$$

for each $x \in H$. Hence, $\phi(u)x - \phi(x)u \in K$, and since $u \in K^\perp$, we have

$$0 \;=\; \langle u, \varphi(u)x - \varphi(x)u\rangle \;=\; \varphi(u)\langle u, x\rangle - \varphi(x).$$

If we then set $y = \overline{\varphi(u)}$, we have

$$\varphi_y(x) \;=\; \langle \overline{\varphi(u)}u, x\rangle \;=\; \varphi(x), \qquad (x \in H).$$

Therefore, φ has the form φ_y for some $y \in H$. To show the uniqueness of y, suppose that $\varphi_y = \varphi_z$. Then $\langle y, x\rangle = \langle x, x\rangle = 1$, and from the discussion preceding this theorem, $x = y$. □

2 Proposition. *Given a linear operator $A : H \to K$, there is a unique linear operator $A^* : K \to H$ such that*

$$\langle y, Ax\rangle \;=\; \langle A^* y, x\rangle,$$

for all $x \in H$ and $y \in K$.

The operator A^* is called the **adjoint** of A.

Proof. Define $\varphi : H \to \mathbb{C}$ by $\varphi(x) = \langle y, Ax \rangle$. By the previous theorem, $\varphi = \varphi_{A^*y}$ for some vector $A^*y \in H$; that is, $\langle y, Ax \rangle = \langle A^*y, x \rangle$, for all $x \in H$. Moreover, given $y, z \in K$, we have

$$\langle A^*(y+z), x \rangle \;=\; \langle y+z, Ax \rangle \;=\; \langle y, Ax \rangle + \langle z, Ax \rangle \;=\; \langle A^*y + A^*z, x \rangle,$$

for all $x \in H$. By the previous theorem again, $A^*(y+z) = A^*y + A^*z$. A similar argument shows that $A^*(\lambda y) = \lambda A^* y$, and hence A^* is a linear operator.

To see that A^* is unique, suppose that $\langle A^*y, x \rangle = \langle By, x \rangle$, for all $x, y \in H$. Then $\langle A^*y - By, x \rangle = 0$, for all $x \in H$, from which $A^*y = By$, for all $y \in H$. □

3 Proposition. *For any $A \in B(H)$, we have $A^{**} = A$.*

Proof. Since $\langle x, Ay \rangle = \langle A^*x, y \rangle$, it follows that $\langle Ay, x \rangle = \langle y, A^*x \rangle$. □

Exercise. Show that $(A+B)^* = A^* + B^*$, and $(AB)^* = B^*A^*$.

Definition. We say that $A : H \to H$ is **self-adjoint** if $A = A^*$.

This last definition isn't so illuminating; but the important thing is that it's tantamount to saying that A has an orthonormal basis of eigenvectors, and that its eigenvalues are real numbers.

Definition. Let x be a nonzero vector in H. We say that x is an **eigenvector** for A just in case $Ax = \lambda x$ for some $\lambda \in \mathbb{C}$.

In other words, A doesn't move x outside of its ray.

Definition. We say that $\lambda \in \mathbb{C}$ is an **eigenvalue** for A just in case $Ax = \lambda x$ for some nonzero vector $x \in H$.

Definition. We say that an operator E on H is a **projection operator** just in case it is self-adjoint and $E^2 = E$. Equivalently, E is self-adjoint and has eigenvalues in the set $\{0, 1\}$. In particular, the zero operator 0 and the identity operator I are projection operators.

Seen as a "quantity", a projection operator E is kind of boring: it only has two possible values, 0 and 1. But that allows it to play the special role as representing a **property**, i.e. something that the system either has or lacks. That is, if the value of E is 1, then the system is thought to have the property E, and if the value of E is 0, then the system is thought to lack the property E. (In some literature, the projection operators are called **yes-no questions**.)

Exercise. Let $y \in H$ be a unit length vector, and define an operator $E : H \to H$ by

$$Ex = \langle y, x \rangle y, \qquad (x \in H).$$

Show that E is a projection operator, and that $z \in H$ is an eigenvector for E iff $z = cy$ for some $c \in \mathbb{C}$.

Definition. We say that projection operators E and F are **orthogonal** when $EF = 0$.

Exercise. Show that if E and F are orthogonal projection operators, then $E + F$ is also a projection operator.

Definition. Let K be a subset of the vector space H. We say that K is a **subspace** just in case for all $x, y \in K$ and all $\lambda \in \mathbb{C}$, both $x + y \in K$ and $\lambda x \in K$. We write $\dim K$ for the dimension of K, i.e. the maximal number of mutually orthogonal vectors in K.

Exercise. Suppose that K and L are subspaces of H. Show that $K \cap L$ is also a subspace.

Exercise. Let E be a projection operator on H, and let K be the set of vectors $x \in H$ such that $Ex = x$. Show that K is a subspace.

Exercise. Suppose that E and F are projection operators on H and that $EF = FE$. Show that EF is a projection operator. Let $[E]$ be the subspace onto which E projects. Show that $[EF] = [E] \cap [F]$.

Spectral Theorem. *If A is a self-adjoint linear operator, then there are orthogonal projection operators $E_1, \ldots, E_n$ and real numbers $\lambda_1, \ldots, \lambda_n$ such that*

$$A = \lambda_1 E_1 + \cdots + \lambda_n E_n.$$

For understanding the spectral representation of a self-adjoint operator, it can be helpful to think of the sum operation on orthogonal projection operators as a logical disjunction. In fact, for any projections E and F, we can define $E \vee F$ to be the projection onto the smallest subspace of H that contains both $[E]$ and $[F]$.

Exercise. Show that when E and F are orthogonal, then $E \vee F = E + F$.

But beware not to carry over all of your intuitions from classical logic. For example, if E and F are projections onto mutually orthogonal unit vectors x and y, then $E \vee F$ is the projection onto the subspace spanned by $\{x, y\}$. In that case, the state $\frac{1}{\sqrt{2}}(x+y)$ assigns 1 to $E \vee F$ even though it doesn't assign 1 to either E or to F. In other words, quantum probabilities have the strange feature that a disjunction can be certainly true even if neither disjunct is certainly true.

We now let $B(H)$ denote the set of all linear operators on H. (The letter "B" here comes from the fact that when H is infinite dimensional, we restrict to *bounded* linear operators.) The set $B(H)$ has the following operations:

addition $A, B \mapsto A + B$

scalar multiplication $\lambda, A \mapsto \lambda A$

multiplication $A, B \mapsto AB$

adjunction $A \mapsto A^*$

There are also two special elements $0 \in B(H)$ and $I \in B(H)$. The former is defined by $0x = 0$ for all $x \in H$, and the latter by $Ix = x$ for all $x \in H$. With these operations, $B(H)$ is not simply a vector space over $\mathbb{C}$, it is also an "algebra with adjunction".

Definition. For operators $A, B \in B(H)$, we define $[A, B] = AB - BA$. We say that A and B are **compatible** just in case $[A, B] = 0$.

We will now define the notion of an abstract "state" on $B(H)$. In short, a state assigns a real number to a self-adjoint operator, which can be interpreted as the expectation value of the corresponding quantity in that state.

Definition. A **linear functional** on $B(H)$ is a function $\rho : B(H) \to \mathbb{C}$ such that $\rho(A + B) = \rho(A) + \rho(B)$ and $\rho(\lambda A) = \lambda \rho(A)$.

Definition. A **state** in the abstract sense on $B(H)$ is a function $\omega : B(H) \to \mathbb{C}$ such that:

1. ω is linear, and
2. $\min[\text{sp}(A)] \leq \omega(A) \leq \max[\text{sp}(A)]$, for every self-adjoint operator $A \in B(H)$.

In particular, if ω is a state, then $\omega(I) = 1$, since 1 is the only eigenvalue of I. Similarly, if E is a projection operator, then $\omega(E) \in [0, 1]$. It's easy to see why a state ω should have this latter feature, i.e. that $\omega(A)$ lies between the smallest and largest eigenvalues of A. After all, ω is supposed to assign an expectation value to A, which should be a weighted average of the possible values that A could take.

But why assume that states are linear functionals? It is true that, in classical physics, expectation values are linear functions on random variables. So why should it be any different in quantum physics? However, it's the linearity assumption that John Bell considers to be "silly" and "meaningless."

Example. Let $x \in H$ be a unit length vector, and define a function $\omega_x : B(H) \to \mathbb{C}$ by $\omega_x(A) = \langle x, Ax \rangle$. Obviously ω_x is a linear functional. Furthermore, if $A = \lambda_1 E_1 + \cdots + \lambda_n E_n$ is a self-adjoint operator in its spectral representation, then

$$\omega_x(A) = \sum_i \lambda_i \langle x, E_i x \rangle,$$

which is a weighted average of the eigenvalues of A. Therefore, ω_x is a state.

Definition. If σ and ρ are linear functionals, and $a, b \in \mathbb{C}$, then we define $a\sigma + b\rho$ to be the function defined by

$$(a\sigma + b\rho)(A) \;=\; a\sigma(A) + b\rho(A).$$

It's easy to see that $a\sigma + b\rho$ is a linear functional.

Exercise. Suppose that σ and ρ are states on $B(H)$, and that $\lambda \in (0, 1)$. Show that $\lambda\sigma + (1 - \lambda)\rho$ is also a state on $B(H)$. We call this latter state a **mixture** of σ and ρ. If a state is *not* a mixture of other states, then we say that it is **pure**.

The previous exercise shows that if we have n unit vectors $x_1, \dots, x_n \in H$, then we can form mixed states such as

$$\lambda_1 \omega_{x_1} + \cdots + \lambda_n \omega_{x_n},$$

where $\lambda_i \in [0, 1]$ such that $\lambda_1 + \cdots + \lambda_n = 1$.

Example. Since we're working here in the finite-dimensional case, i.e. where the dimension of H is a finite number n, there is always a special state, the so-called **maximally mixed state**. The idea behind this state is simple: take any quantity, say A, that has possible values $a_1, \ldots, a_n$. Suppose that we don't know anything whatsoever about the system. Then what value should we "expect" for A? The maximally mixed state says: take the average of $a_1, \ldots, a_n$, i.e. set

$$\tau(A) \;=\; \frac{a_1 + \cdots + a_n}{n}.$$

In particular, for any projection operator E, $\tau(E) = \frac{d(E)}{n}$, where $d(E)$ is the dimension of the subspace onto which E projects. This maximally mixed state has another name: it is the **trace**, or to be more precise, the trace divided by the dimension of the vector space H. We use Tr() for the trace, so that $\mathrm{Tr}(A) = n\tau(A)$.

For those of you who have done linear algebra, you'll remember that the trace of a matrix is the sum of its diagonal elements. Here is a more formal (coordinate-free) definition.

Definition. Suppose that $\dim H = n$, and let $\{x_1, \ldots, x_n\}$ be an orthonormal basis for H. The **trace** on $B(H)$ is defined by

$$\mathrm{Tr}(A) \;=\; \sum_{i=1}^{n} \langle x_i, Ax_i \rangle,$$

for all $A \in B(H)$.

It is easy to show that trace is a linear functional. Furthermore,

$$\mathrm{Tr}(A^*) \;=\; \sum_i \langle x_i, A^* x_i \rangle \;=\; \sum_i \overline{\langle x_i, Ax_i \rangle} \;=\; \overline{\mathrm{Tr}(A)},$$

for any operator $A \in B(H)$. Finally, if $\mathrm{Tr}(A^*A) = 0$, then

$$0 \;=\; \mathrm{Tr}(A^*A) \;=\; \sum_i \|Ax_i\|^2,$$

which means that $Ax_i = 0$ for all x_i in the basis $\{x_1, \ldots, x_n\}$, hence $A = 0$. (In this case, we say that the trace is **faithful**.)

We're going to show now that the definition of the trace is in fact independent of the chosen orthonormal basis $\{x_1, \ldots, x_n\}$. But first we need to gather some more facts about Hilbert spaces. Recall that $\{x_1, \ldots, x_n\}$ is said to be an **orthonormal basis** for H just in case: (1) it's a basis, i.e. every vector $y \in H$ can be written as $y = c_1 x_1 + \cdots + c_n x_n$ for a unique sequence c_i of complex numbers, and (2) $\langle x_i, x_j \rangle = 0$ when $i \neq j$, and (3) $\langle x_i, x_i \rangle = 1$. It immediately follows then that for any vector $y \in H$, if $\langle y, x_i \rangle = 0$ for $i = 1, \ldots, n$, then $y = 0$.

Definition. For any subset K of H, we write $K^\perp$ for the set of vectors $y \in H$ such that $\langle y, x \rangle = 0$ for all $x \in K$.

Exercise. Show that $K^\perp$ is a subspace of H.

4 Proposition. *Suppose that $\{x_1, \ldots, x_n\}$ is an orthonormal basis for H. Then for any vector $y \in H$,*

$$y \;=\; \sum_{i=1}^{n} \langle x_i, y \rangle x_i.$$

Proof. Since $\{x_1, \ldots, x_n\}^\perp = \{0\}$, it will suffice to show that $y - \sum_{i=1}^{n} \langle x_i, y \rangle x_i$ is orthogonal to each x_j. For a fixed j, we have

$$\left\langle x_j \, , \, y - \sum_{i=1}^{n} \langle x_i, y \rangle x_i \right\rangle \;=\; \langle x_j, y \rangle - \langle x_j, y \rangle \;=\; 0,$$

as we needed to show. □

5 Proposition. *If $\{x_1, \ldots, x_n\}$ is an orthonormal basis, then for any vectors $y, z \in H$,*

$$\langle y, z \rangle \;=\; \sum_{i=1}^{n} \langle y, x_i \rangle \, \langle x_i, z \rangle.$$

Proof. Just write out $z = \sum_{i=1}^{n} \langle x_i, z \rangle x_i$, and take its inner product with y. □

Now we can show that the definition of the trace is independent of basis. Let $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ be orthonormal bases for H. Then

$$\begin{aligned}
\sum_i \langle x_i, Bx_i \rangle &= \sum_i \sum_j \langle x_i, y_j \rangle \langle y_j, Bx_i \rangle \\
&= \sum_i \sum_j \langle x_i, y_j \rangle \langle B^* y_j, x_i \rangle \\
&= \sum_j \sum_i \langle B^* y_j, x_i \rangle \langle x_i, y_j \rangle \\
&= \sum_j \langle B^* y_j, y_j \rangle \\
&= \sum_j \langle y_j, By_j \rangle,
\end{aligned}$$

as we needed to show.

Example. For an operator represented by a matrix, the trace can be taken by summing the entries on the diagonal. For example,

$$\mathrm{Tr} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 0.$$

This operation represents summing with the basis consisting of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Definition. Let $U : H \to K$ be a linear operator. We say that U is **unitary** just in case $\langle x, y \rangle_H = \langle Ux, Uy \rangle_K$ for all $x, y \in H$.

Exercise. Show that $U : H \to K$ is unitary iff U^*U is the identity on H, and U^*U is the identity on K.

Exercise. Suppose that $\{x_1, \ldots, x_n\}$ is an orthonormal basis for H, and that $U : H \to K$ is a unitary operator. Show that $\{Ux_1, \ldots, Ux_n\}$ is an orthonormal basis for K.

6 Proposition. *For any unitary operator* $U : H \to H$, *we have* $\mathrm{Tr}(U^*AU) = \mathrm{Tr}(A)$ *for all* $A \in B(H)$.

Proof. We have

$$\begin{aligned}\mathrm{Tr}(U^*AU) &= \sum_{i=1}^{n}\langle x_i, U^*AUx_i\rangle \\ &= \sum_{i=1}^{n}\langle Ux_i, AUx_i\rangle \\ &= \mathrm{Tr}(A),\end{aligned}$$

where the last equation follows from the fact that $\{Ux_i, \ldots, Ux_n\}$ is an orthonormal basis. □

7 Proposition. *Any operator $A \in B(H)$ is a linear combination of unitary operators.*

I'll omit the proof for now, but for the case of spin operators, it turns out that $\{S_x, S_y, S_z, I\}$ is, in fact, a linear basis for $B(H)$. In other words, for each $A \in B(H)$, there is a unique quadruple $a_i \in \mathbb{C}$ such that

$$A = a_0I + a_1S_x + a_2S_y + a_3S_z.$$

8 Proposition. *For any operators $A, B \in B(H)$, $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.*

Definition. A linear operator D on H is said to be a **density operator** just in case D is self-adjoint with eigenvalues $\lambda_i \in [0, 1]$ such that $\sum_i \lambda_i = 1$.

Example (Trace formula)**.** Suppose that D is a density operator, and define a function $\omega_D : B(H) \to \mathbb{C}$ by setting

$$\omega_D(A) = \mathrm{Tr}(DA),$$

for all $A \in B(H)$. Then ω_D is a state. Indeed, if $D = \lambda_1 E_{x_1} + \cdots + \lambda_n E_{x_n}$, then $\omega_D = \lambda_1 \omega_{x_1} + \cdots + \lambda_n \omega_{x_n}$.

9 Proposition. *For an abstract state ω on $B(H)$, the following are equivalent:*

1. *$\omega(A^2) = \omega(A)^2$ for every self-adjoint operator $A \in B(H)$.*
2. *$\omega(A)$ is an eigenvalue of A, for every self-adjoint operator $A \in B(H)$.*
3. *$\omega(E) \in \{0, 1\}$ for every projection operator $E \in B(H)$.*

Proof. Sketch of proof (1 $\Rightarrow$ 3) Let E be a projection operator. Since $EE = E$, it follows that $\omega(E)^2 = \omega(E)$, and hence $\omega(E) \in \{0, 1\}$.

(3 $\Rightarrow$ 2) Let $A = \sum_i \lambda_i E_i$, where $\sum_i E_i = I$. Since ω is linear, $\omega(E_i) = 1$ for one E_i, and $\omega(E_j) = 0$ for the others. Hence, $\omega(A) = \lambda_i$. □

Definition. When the conditions above hold, we say that ω is a **dispersion-free state** on $B(H)$.

Note: first, this definition of dispersion-free matches classical statistics, where the **variance** of a random variable X is $E[X^2] - E[X]^2$, where E is the expectation value function. In other words, a dispersion-free state is a state where all random variables have zero variance. Second, if we take the projection operators in $B(H)$ to represent propositions, then a dispersion-free state is precisely an assignment of truth values to these propositions.

We will shortly prove the von Neumann NHV theorem, which shows that when $\dim H \geq 2$, then there are no dispersion-free states on $B(H)$. The proof has two parts:

1. When $\dim H \geq 2$, then there are no dispersion-free quantum state, i.e. states of the form ω_D for some density operator $D \in B(H)$.

2. Every abstract state on $B(H)$ has the form ω_D for some density operator D on H.

   Note: it's this second part where John Bell thinks that von Neumann made a mistake. Von Neumann assumed that a "state" or "hidden variable" would have to be a linear function on $B(H)$. Bell thinks that this assumption is "silly." Bell thinks that hidden variables would only have to be linear on compatible observables (i.e. observables that can be simultaneously measured, represented by self-adjoint operators that commute with each other). Let's call that feature **sub-linearity**. Ironically, Gleason's theorem shows that when the dimension of H is 3 or more, then sub-linearity implies linearity.

Proving the first part of von Neumann's theorem is easy. Suppose that $\dim H \geq 2$, and let $x \in H$ be a unit vector. Since $\dim H \geq 2$, there is another unit vector y that is skew to x, i.e. it is neither in the ray generated by x, nor is it orthogonal to x. In particular, $0 < |\langle x, y\rangle|^2 < 1$. Now let E be the projection onto y. Then

$$\omega_x(E) \;=\; \langle x, Ex\rangle \;=\; \langle x, \langle y, x\rangle y\rangle \;=\; |\langle x, y\rangle|^2.$$

It follows that $\omega_x(E) \notin \{0, 1\}$, and hence ω_x is not dispersion free. Since x was an arbitrary unit vector in H, it follows that no state of the form ω_D is dispersion-free on $B(H)$. [[There is a little gap here in infering from there are no dispersion-free vector states to there are no dispersion-free density operator states. But the fact is that density operator states always have more dispersion than vector states.]]

Now on to the second part of von Neumann's theorem, which is a "representation" result. In particular, we show that if $\omega : B(H) \to \mathbb{C}$ is a state in the abstract sense, then there is a density operator D on H such that

$$\omega(A) \;=\; \mathrm{Tr}(DA),$$

for all $A \in B(H)$.

10 Proposition. *Every operator A in $B(H)$ is a sum of two self-adjoint operators.*

Proof. Let $A_r = \frac{1}{2}(A + A^*)$ and let $A_i = \frac{i}{2}(A^* - A)$. Then

$$2(A_r + iA_i) \;=\; A + A^* + A - A^* \;=\; 2A.$$

□

11 Theorem. *For any abstract state $\omega : B(H) \to \mathbb{C}$, there is a density operator $D \in B(H)$ such that $\omega(A) = \mathrm{Tr}(DA)$ for all $A \in B(H)$.*

Proof. Assume that $\omega : B(H) \to \mathbb{C}$ is an abstract state. Define a positive-definite inner product on $B(H)$ by

$$\langle B, A \rangle_2 \;=\; \mathrm{Tr}(B^* A),$$

for all $B, A \in B(H)$. Since ω is linear, Theorem 1 entails that there is an operator $D \in B(H)$ such that

$$\omega(A) \;=\; \langle D, A \rangle_2 \;=\; \mathrm{Tr}(D^* A),$$

for all $A \in B(H)$. We need to show then that D is a density operator, i.e. a self-adjoint operator with eigenvalues $0 \leq \lambda_1, \ldots, \lambda_m \leq 1$ such that $\sum_{i=1}^m \lambda_i = 1$.

Since every operator $A \in B(H)$ has the form $A = A_r + iA_i$, with A_r and A_i self-adjoint, and since states are linear and assign real-values to self-adjoint operators, it follows that $\omega(A) = \overline{\omega(A^*)}$ for all $A \in B(H)$. Hence,

$$\mathrm{Tr}(DA) \;=\; \overline{\mathrm{Tr}(DA^*)} \;=\; \mathrm{Tr}(D^*A),$$

for all $A \in B(H)$. Thus, $\langle D, A\rangle_2 = \langle D^*, A\rangle_2$ for all $A \in B(H)$, from which it follows that $D = D^*$.

Now let λ be an eigenvalue of D, and let E be the projection onto the corresponding eigenvector. Since $\omega(E) \in [0, 1]$, it follows that

$$\lambda \;=\; \lambda\mathrm{Tr}(E) \;=\; \mathrm{Tr}(DE) \;\in\; [0, 1].$$

Furthermore, if $E_1, \ldots, E_m$ are the spectral projections of D, then $\sum_{i=1}^{m} E_i = I$ and

$$\sum_{i=1}^{m} \lambda_i \;=\; \sum_{i=1}^{m} \mathrm{Tr}(DE_i) \;=\; \mathrm{Tr}(DI) \;=\; 1.$$

Therefore, D is a density operator. □

This completes the proof of the von Neumann NHV theorem. Every abstract state on $B(H)$ is represented by some density operator D, using the trace formula. But no such states are dispersion free. Therefore, there are no dispersion-free states on $B(H)$.

Logical version of von Neumann theorem

We now prove a second version of the von Neumann NHV theorem, this time from a logical point of view. (This proof of the theorem will also serve as a bridge to the Kochen-Specker theorem.)

Definition. For projection operators E, F, we define $E \vee F$ to be the projection unto the smallest subspace that contains both $[E]$ and $[F]$. We let $E \wedge F$ be the projection onto $[E] \cap [F]$. We let $\neg E = I - E$. Finally, we write $E \leq F$ just in case $EF = E$.

Exercise. Show that if $E \leq F$ and $F \leq G$ then $E \leq G$.

Let $L(H)$ be the set of all projection operators on the Hilbert space H. When equipped with the relation $\leq$, and the operations $\wedge, \vee$, the set $L(H)$ is a **lattice**. With the operation $\neg$, it becomes an **orthocomplemented lattice**. These lattice operations look like logical operations (conjunction, disjunction, and negation), and in some ways they behave like them. However, these operations don't satisfy all the rules as classical logic, such as distribution. For example, let E be the projection onto $|z+\rangle$, let F_1 be the projection onto $|x+\rangle$ and let F_2 be the projection onto $|x-\rangle$. Then $E \wedge F_1 = 0$ and $E \wedge F_2 = 0$, but

$$E \wedge (F_1 \vee F_2) \;=\; E \wedge I \;=\; E$$

and therefore,

$$E \wedge (F_1 \vee F_2) \;\neq\; (E \wedge F_1) \vee (E \wedge F_2).$$

These features of the lattice $L(H)$ have led some people to say that the weird thing about QM is that it violates the rules of classical logic.[1]

Definition. Let $p : L(H) \to \{0, 1\}$ be a function. We say that p is a **truth-valuation** on $L(H)$ just in case:

1. $p(I) = 1$,
2. $p(E \wedge F)$ is the minimum of $p(E)$ and $p(F)$, i.e. $p(E \wedge F) = 1$ iff $p(E) = 1$ and $p(F) = 1$, and
3. $p(E \vee F)$ is the maximum of $p(E)$ and $p(F)$.

12 Theorem. *If* $\dim H \geq 2$*, then there is no truth-valuation on* $L(H)$*.*

Proof. Suppose for reductio ad absurdum that p is a truth-valuation on $L(H)$. If $\dim H \geq 2$, then there must be a two-dimensional projection X on H such that $p(X) = 1$. Let x_1, x_2 be orthogonal unit vectors in the range of X, and let E_i be the projection onto x_i. Let F_1 be the projection onto $x_1 + x_2$, and let F_2 be the projection onto $x_1 - x_2$. Then $E_1 \vee E_2 = X$, and hence $p(E_1) = 1$ or $p(E_2) = 1$. Without loss of generality, assume that $p(E_1) = 1$. Since $E_1 \wedge F_i = 0$, it follows that $p(F_i) = 0$ for $i = 1, 2$. But then $1 = p(F_1 \vee F_2) = 0$, a contradiction. Therefore, p cannot exist. □

[1] Birkhoff and von Neumann (1936) were the first to investigate the logical features of QM. Reichenbach (1944) claims that QM demands a move to a three-valued logic. Putnam (1969) says that QM demands that we reject classical logic. For a clear discussion of the issues, see (Gibbins, 1987).

What does this result show? Some people would say that the result shows that:

> (†) Quantum systems are, of necessity, **indeterminate** — i.e. in any situation, there will be some propositions about the system that are neither true nor false.

For example, suppose that D is a density operator on H, and that we define $p : L(H) \to [0, 1]$ by $p(E) = \text{Tr}(DE)$. Then p assigns every proposition some probability, but it will assign some propositions a value strictly between 0 and 1. So, some people would say that in such a situation (represented by the density operator D), those propositions don't have a definite truth-value.

To assess if (†) is a reasonable interpretation of the result, let's note some tacit assumptions that could be questioned:

1. Every projection operator in $L(H)$ represents a proposition about the system. (If this were false, then it might still be the case that every proposition gets assigned a definite truth value. Suppose, for example, that there really weren't any such thing as spin-x, and that sentences about spin-x could be seen as employing a fiction to speak obliquely about the real thing, spin-z. This is the kind of strategy that's employed by Bohmian mechanics, where particle position is taken to be the only fundamentally real quantity.)

2. A truth assignment to all propositions would have to satisfy the conditions above.

   First of all, of course it's possible to assign all elements of $L(H)$ either "true" or "false," if we don't respect the supposed logical relations between elements.

   Suppose that E is the projection onto $|z+\rangle$ and that F is the projection onto $|x+\rangle$. According to a standard way of interpreting the formalism, we have:

   $$\begin{aligned} E &\equiv \text{spin-}z \text{ has value } +1 \\ F &\equiv \text{spin-}x \text{ has value } +1 \end{aligned}$$

   The convention we adopted was that $E \wedge F$ is the projection onto the subspace $[E] \cap [F]$, which is simply the zero vector. In other words, $E \wedge F = 0$, so that our convention presupposes that it's not possible for both S_x and S_z to have value +1 (or any other definite value) at the same time. Of course, this convention agrees with a certain way of looking at QM, where **incompatible quantities** cannot simultaneously have values.

3. A truth assignment doesn't depend on some additional index, e.g. a context. (Some say that every proposition has a truth-value relative to a measurement context.)

Was there something "silly" in the assumptions we made about truth-valuations? Well, without futher discussion, the symbols "$\wedge$" and "$\vee$" only superficially resemble conjunction and disjunction. It isn't obvious that we should think of $E \vee F$ as "either E or F." For example, if E is the projection onto $|z+\rangle$ and F is the projection onto $|z-\rangle$, then $E \vee F = I$. If we interpreted this "$\vee$" as a classical disjunction, then we would have to see that the system's state is either definitely $|z+\rangle$ or $|z-\rangle$.

Kochen-Specker theorem

Let's take a step back from the standard formalism for QM, which has built in relations between quantities such as S_x and S_z. Instead, let's think of $L(H)$ not as a set of proposition, but as a collection of sets of propositions. For example, let $L(S_z)$ be the set of all spectral projections of S_z, i.e. $L(S_z)$ is the four element Boolean lattice with 0, I, and the projections onto $|z+\rangle$ and $|z-\rangle$. Thus, $L(S_z)$ is a maximal **Boolean sublattice** of $L(H)$, and we can think of the latter as an aggregate of Boolean lattices.

The question now is whether we can assign states to the individual Boolean sublattices of $L(H)$.

When $\dim H = 2$, the answer is obviously yes. Since the lattices $L(A), L(B), \dots$ are only connected by sharing 0 and I, their states can be chosen independently of each other. To be more precise, for each maximal Boolean sublattice L of $L(H)$, we can choose a truth-valuation p_L on L. Then we can think of the aggregate $\{p_L : L \subseteq L(H)\}$ as a hidden variable. This aggregate state assigns 0 or 1 to every proposition, and it respects the logical operations on compatible projection operators (i.e. those that commute with each other).

The Kochen-Specker theorem shows that when $\dim H > 2$, these aggregate states do not exist. The key to proving the theorem is to exploit the fact that some of these maximal Boolean sublattices share elements in common. In that case, the state of one sublattice cannot be chosen independently of the state of another sublattice.

A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$	$\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$	$\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}$	$\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}$	$\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$	$\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\\-1\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\\-1\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\\1\\-1\end{pmatrix}$
$\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$	$\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\\1\\-1\end{pmatrix}$	$\begin{pmatrix}-1\\1\\1\\1\end{pmatrix}$	$\begin{pmatrix}-1\\1\\1\\1\end{pmatrix}$
$\begin{pmatrix}1\\1\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\1\\0\end{pmatrix}$	$\begin{pmatrix}1\\1\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\0\\1\end{pmatrix}$	$\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}$	$\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\1\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\0\\1\end{pmatrix}$
$\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\\1\\1\end{pmatrix}$	$\begin{pmatrix}0\\1\\0\\-1\end{pmatrix}$	$\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}$	$\begin{pmatrix}0\\1\\-1\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\\1\\1\end{pmatrix}$	$\begin{pmatrix}0\\1\\0\\-1\end{pmatrix}$	$\begin{pmatrix}0\\1\\-1\\0\end{pmatrix}$

Figure 1: Nine quantities for Kochen-Specker contradiction

Here I'll sketch the proof for Kochen-Specker in four dimensions. Following Cabello et al. (1996), we construct nine maximal Boolean sublattices of $L(H)$, each of which has four atoms. (See Figure 1.) With nine distinct Boolean sublattices, there could be thirty-six distinct atoms. However, in this collection, each atom occurs exactly two times, so there are only eighteen distinct atoms. By assumption, only one entry in each column is assigned 1. Thus, nine entries in total get assigned 1. However, each proposition occurs in the table twice, and it must be assigned the same value at both occurrences. It follows that the number of entries assigned 1 is even, a contradiction.

Contextual hidden variables

Before the 1950s, many people thought that von Neumann's theorem showed that it is not possible to supplement QM with hidden variables. The result was often interpreted as showing that QM cannot be replaced with a **deterministic theory**. Then along came Bohm's theory which is deterministic, and which reproduces the predictions of QM. Thus, it might seem that Bohm's theory shows that hidden variables are possible after all.

Indeed, it's easy to get confused here because of what John Bell said about von Neumann's theorem — i.e. that it has a silly assumption. One might be tempted to think that Bohm's theory bypasses von Neumann's theorem because its hidden variables don't satisfy von Neumann's assumption. That, however, is a misreading of the situation, for the Bell-Kochen-Specker theorem proves that as long as the state space has dimension three, then von Neumann's conclusion follows even without his silly assumption.

To be clear, then, whatever Bohmian mechanics does, it cannot be "realist" in the sense of assigning definite values to all quantities, if we assume that all self-adjoint operators on Hilbert space represent quantities. So, there are two things that Bohmian mechanics might be doing: it might be denying that all operators represent quantities, or it might be thinking of states in a different way.

We'll come back a little bit later to the first option, but let me just say here that Bohmians definitely *do* embrace this first horn of the dilemma. Most notably, see Daumer et al. (1996).

What about the second option? Is there a more liberal notion of "hidden variable" according to which Bohmian mechanics is a hidden variable theory. One proposal that has been put out there is the idea of **contextual hidden variables**. To see what that's supposed to mean, look back at the table we used for the Kochen-Specker theorem. Let E be the "proposition" that occurs in the top left hand corner, and also in the top row of the second column. The assumption of the KS theorem is that a state ω assigns E a value (0 or 1), in other words, that "$\omega(E)$" is unambiguous. Let's write this out as an official definition:

non-contextuality The value $\omega(E)$ is independent of context.

Here, we're thinking of a "context" as corresponding to the choice of a quantity to be measured.

Here's how a contextualist could sidestep the Kochen-Specker theorem: assign all propositions in the top row 1, and assign all others 0. Of course, to do so, he ends up speaking ambiguously about the projection operator that occurs in both row 1, column 5 and in row 2, column 1. He says: if you're asking me for the value of E relative to a measurement of A_1, then my answer is 0. But if you're asking me for the value of E relative to a measurement of A_5, then my answer is 1.

So, the contextualist doesn't take projections to represent propositions in the traditional sense. Instead, projections represent the sort of thing that philosophers have called **centering features** or **propositional functions**, i.e. functions from contexts to propositions (see Egan, 2006).[2] A typical propositional function is something like (S) "I am over six feet tall," which is like a function from contexts to propositions. For example, S plus the context in which Napolean is speaking returns a false proposition, whereas S plus the context in which Goliath is speaking returns a true proposition.

Now, there is one "cheap" way in which we can turn context-relativism into realism. Suppose for simplicity that there are two context-relative correct descriptions: suppose that E_1 is true relative to context C_1, and that E_2 is true relative to context C_2. Then we could say that the true story of reality is:

$$(E_1 \text{ relative to } C_1) \text{ and } (E_2 \text{ relative to } C_2) \text{ and } \ldots$$

But to me at least, this seems like a cheap kind of realism. I have a hunch that this kind of conjunctive description should not really count as yet another description. (But the conjunctive description does, in some ways, remind me of the philosopher Kit Fine's view about relativity theory.)

To be clear: I've never heard anybody say that we can get a "god's eye view" by taking a logical sum of context-relative descriptions. I have, in contrast, heard people suggest that QM can be replaced by a realist theory, so long as its hidden variables are contextual (or nonlocal). But as far as I can tell, contextualism is completely at odds with the spirit of "realism", at least the kind of realism that is expressed in the introduction to Maudlin (2019).

[2] Here's a semi-precise definition: a propositional function is an expression having the form of a proposition but containing undefined symbols for the substantive elements and becoming a proposition when appropriate values are assigned to the symbols.

The interpretation that most obviously complies with Maudlin's vision (i.e. the goal of physics is to describe matter in motion) is Bohmian mechanics. In the Bohm picture, there is a world out there whose structure is completely independent of who is looking at it. In contrast, a contextual hidden variable assignment is a kind of relativism, i.e. things only have reality relative to a choice of measured observable.

It seems to me that the contextualist approach would fit much better with the likes of Niels Bohr than with the likes of David Bohm or Tim Maudlin. Bohr frequently talks about the need to relativize the description to a choice of measurement apparatus. That sounds quite a lot like the idea of contextual hidden variables — i.e. there is a "realist" way of describing within a reference frame, but no realist description that is frame-independent.

More or less properties

There's another way of trying to sidestep the NHV theorems: identifying properties of quantum systems besides those represented by projections. To begin with an example, consider the sentence:

(P) Ebbe is in an eigenstate of S_z.

(Here "Ebbe" is fictional name I'm using for an electron.) To be in an eigenstate of S_z is to be either to have the property E_1 (the projection onto $|z+\rangle$) or to have the property E_2 (the projection onto $|z-\rangle$). What then is the correct way to represent the sentence P? On the one hand, P might be the quantum disjunction $E_1 \vee E_2$; on the other hand, P might be the set-theoretic union of $|z+\rangle$ and $|z-\rangle$. Let's first look at the problems with the first idea, and then we'll look at how we might develop the second idea.

For further reading

- For more on Grete Hermann, see (Crull and Bacciagaluppi, 2016). For an argument that von Neumann wasn't so confused, see (Bub, 2011). And the debate goes on: (Dieks, 2017; Mermin and Schack, 2018).

Appendix: Spectral representations

Suppose that H is a finite-dimensional Hilbert space and that A is a self-adjoint operator on H. We let $\mathrm{sp}(A)$ denote the set of eigenvalues of A, i.e. the **spectrum** of A.

13 Proposition. *If A is a self-adjoint operator on a finite-dimensional Hilbert space, then the spectrum of A is a finite subset of real numbers.*

One can prove this fact in a couple of different ways. On the one hand, one can use classical linear algebra. On the other hand, one can use the theory of commutative C^*-algebras. Let $C^*(A)$ be the smallest subalgebra of $B(H)$ that contains the operator A. In the case we are interested in, where H is finite-dimensional, $C^*(A)$ will consist of polynomials (over $\mathbb{C}$) in A and I. If $A - \lambda I$ is not invertible in $C^*(A)$, then it's contained in a maximal ideal. A finite-dimensional algebra has only a finite number of distinct maximal ideals. Since the proof is fairly complicated, we will omit further details.

The following result is known as the continuous spectral representation, and it holds quite generally. In our particular case, where A is a self-adjoint operator with finite spectrum, the result is rather trivial.

14 Proposition. *The C^*-algebra $C^*(A)$ generated by A is isomorphic to $C(\mathrm{sp}(A))$, the set of all continuous complex valued functions on $\mathrm{sp}(A)$.*

In the case we're interested in, $\mathrm{sp}(A)$ is a finite Hausdorff space, and "continuous" is redundant.

15 Proposition. *Let A be a self-adjooint operator on a finte-dimensional Hilbert space H. Then there is a canonical bijection between the following sets:*

1. *Eigenvalues of A*

2. *Minimal projections in the Boolean lattice $L(A)$*

3. *Pure states on the algebra generated by A [The pure states of a commutative algebra are precisely the multiplicative states.]*

4. *Truth-valuations on the Boolean lattice $L(A)$*

16 Proposition. *When $\mathrm{sp}(A)$ is finite, the following five algebras are canonically isomorphic to each other.*

1. *$C^*(A)$*

2. *$C(\mathrm{sp}(A))$*

3. *$C(\sigma A)$, where σA is the set of pure states of $C^*(A)$.*

4. *$L_\infty(\sigma A)$, the algebra of essentially bounded Borel functions from σA to $\mathbb{C}$. [Since σA is finite, "essentially bounded Borel" is redundant.]*

5. *$\ell_\infty(\sigma A)$, the algebra of bounded sequences of complex numbers indexed by σA.*

Sketch of proof. For (1) $\Leftrightarrow$ (3), we use the famous **Gelfand duality theorem**: if A is a commutative C^*-algebra, then $A \cong C(X)$, where X is the compact Hausdorff space of states on A. That result is not easy to prove in the general case. For the case where A is finite-dimensional, the result is almost trivial.

The equivalence of $(2), (4)$, and (5) is a simple consequence of the fact that $\sigma(A)$ is finite. $\square$

The fourth and fifth representations make it obvious that $C^*(A)$ has many projection operators, which correspond to step functions in $L_\infty(\mathrm{sp}(A))$. In fact, for each subset Δ of $\mathrm{sp}(A)$, there is a operator $E(\Delta) \in C^*(A)$ that projects onto the span of the eigenspaces for A with eigenvalues in Δ. [Interpretively, we would say that $E(\Delta)$ represents the proposition that the value of A lies in Δ.] In particular, let $E_i = E(\{\lambda_i\})$, and we have

$$A = \sum_{i=1}^{n} \lambda_i E_i.$$

Thus, we have the following result.

17 Proposition. *Every operator in $B(H)$ is a sum of projection operators.*

Definition. We say that $A = \sum_{i=1}^{n} \lambda_i E_i$ is the **reduced spectral decomposition** of A just in case the E_i are nonzero and mutually orthogonal projections, and $\lambda_i \neq \lambda_j$ when $i \neq j$.

Definition. Given self-adjoint operators A, B on H, we write $A \sim B$ just in case A and B have the same reduced spectral decomposition. Speaking loosely, we say that A and B have the same spectral projections.

18 Proposition. *When H is finite-dimensional, there is a bijective correspondence between:*

1. *Equivalence classes of self-adjoint operators with the same spectral projections.*

2. *Commutative subalgebras of $B(H)$.*

3. *Boolean sublattices of $L(H)$.*

19 Proposition. *Let $A, B \in B(H)$ be self-adjoint. Then $[A, B] = 0$ iff A and B have a common eigenbasis.*

20 Theorem (finite Stone-Weierstrass)**.** *Let X be a finite set of real numbers. Then every function $f : X \to \mathbb{C}$ is a polynomial in x and 1.*

Sketch of proof. Suppose that $X = \{a, b, c\}$. Then the characteristic function of b is

$$\frac{(a-x)(c-x)}{(a-b)(c-b)}.$$

All such characteristic functions are polynomials in x and 1, and every function $f : X \to \mathbb{C}$ is a linear combination of characteristic functions. □

Example. If $X = \{-1, 1\}$ then $\frac{1}{2}(1-x)$ is the characteristic function of -1 and $\frac{1}{2}(1+x)$ is the characteristic function of 1.

Similarly, if $X = \{-1, 0, 1\}$ then $-\frac{1}{2}x(1-x)$ is the characteristic function of -1, and $\frac{1}{2}x(1+x)$ is the characteristic function of 1, and $(1+x)(1-x)$ is the characteristic function of 0.

The Stone-Weierstrass theorem entails that if E is a spectral projection of A, then there is a polynomial f in x and 1 such that $f(A) = E$. Here $f(A)$ is the operator polynomial that results from replacing x with A throughout f. Furthermore, if $E_1, \ldots, E_n$ are the spectral projections of A, then for any complex numbers $c_1, \ldots, c_n$, there is a polynomial g such that

$$g(A) \;=\; g(\lambda_1)E_1 + \cdots + g(\lambda_n)E_n \;=\; c_1E_1 + \cdots + c_nE_n.$$

References

Bell, J. (1988, May). Interview. *Omni*.

Birkhoff, G. and J. von Neumann (1936). The logic of quantum mechanics. *Annals of mathematics*, 823–843.

Bub, J. (2011). Is von Neumann's 'no hidden variables' proof silly? In H. Halvorson (Ed.), *Deep Beauty: Mathematical Innovation and the Search for Underlying Intelligibility in the Quantum World*, pp. 393–408. Cambridge University Press.

Cabello, A., J. Estebaranz, and G. García-Alcaine (1996). Bell-Kochen-Specker theorem: A proof with 18 vectors. *Physics Letters A 212*(4), 183–187.

Conway, J. and S. Kochen (2006). The free will theorem. *Foundations of Physics 36*(10), 1441–1473.

Crull, E. and G. Bacciagaluppi (2016). *Grete Hermann: Between Physics and Philosophy.* Springer.

Daumer, M., D. Dürr, S. Goldstein, and N. Zanghì (1996). Naive realism about operators. *Erkenntnis 45*(2-3), 379–397.

Dieks, D. (2017). Von Neumann's impossibility proof: Mathematics in the service of rhetorics. *Studies in History and Philosophy of Modern Physics 60*, 136–148.

Egan, A. (2006). Secondary qualities and self-location. *Philosophy and Phenomenological Research 72*(1), 97–119.

Gibbins, P. (1987). *Particles and paradoxes: The limits of quantum logic.* Cambridge University Press.

Gleason, A. M. (1957). Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics 6*(6), 885–893.

Hermann, G. (1935). Die naturphilosophischen Grundlagen der Quantenmechanik (Auszug). *Abhandlungen der Fries'schen Schule 6*, 75–152.

Kochen, S. and E. Specker (1967). The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics 17*(1), 59–87.

Maudlin, T. (2019). *Philosophy of physics: Quantum theory*. Princeton University Press.

Mermin, N. D. (1993). Hidden variables and the two theorems of John Bell. *Reviews of Modern Physics 65*(3), 803.

Mermin, N. D. and R. Schack (2018). Homer nodded: von Neumann's surprising oversight. *Foundations of Physics 48*(9), 1007–1020.

Putnam, H. (1969). Is logic empirical? In *Boston studies in the philosophy of science*, pp. 216–241. Springer.

Reichenbach, H. (1944). *Philosophic Foundations of Quantum Mechanics*. University of California Press.

Von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik*. Springer.