Quantum Theory in Hilbert Space: a Review

Jeremy Butterfield (jb56@cam.ac.uk) and Bryan Roberts (b.w.roberts@lse.ac.uk):
Philosophical Aspects of QFT, 13 Oct and 20 Oct 2020

To prepare for philosophical aspects of quantum field theory, we begin with a review of mathematical quantum theory, with some interspersed Philosophical Remarks. Apart from these Remarks, the main themes here will be to build up to some intuition for the functional analysis ideas that underpin the algebraic approach, which we will often see applied to quantum field theory. We will also see some high points of quantum theory *per se*, such as Stone's theorem, Gleason's theorem and superselection.

The main books 'in the background' of this review are:

T. Jordan. *Linear Operators for Quantum Mechanics*: Chapters 1 to 5. Wiley 1969; Dover 2006.

E. Prugovecki. *Quantum Mechanics in Hilbert Space*: Parts III, IV. Academic 1981; Dover 2006.

We especially recommend for this review, and for foundations of quantum theory, as a whole:

N. Landsman. Foundations of Quantum Theory. Springer 2017: especially Chapters 5, 6,7,9,10.
Open access: downloadable at: https://link.springer.com/book/10.1007/978-3- 319-51777-3

We also recommend for the early history of mathematical quantum theory, Landsman's recent survey: 'Quantum theory and functional analysis', arxiv: 1911.06630.

1 Vector spaces and linear functionals

1: Vector spaces (over $\mathbb{R}$, or over $\mathbb{C}$). Inner products. For a vector space over $\mathbb{C}$, we write $(z\psi, w\phi) = z^* w (\psi, \phi)$. The Cauchy-Schwarz inequality: $|(\psi, \phi)| \leq ||\psi||.||\phi||$, with equality iff the vectors are linearly dependent.

2: *Hilbert space*:—

The convergence of vectors: $\psi_n \to \psi := ||\psi_n - \psi|| \to 0$ as $n \to \infty$. So infinite linear combinations of vectors are defined on analogy with $\Sigma_{n=1}^\infty \omega_n$ with $\omega_n \in \mathbb{C}$. The vector space is *complete* iff every Cauchy sequence converges. A *Hilbert space* is a complete inner product space. It is *separable* iff it has a countable (finite or denumerable) basis.

By the way: One similarly says that a *metric space* (X, d) (i.e. X is a set; $d : X^2 \to \mathbb{R}^+ := \{r \in \mathbb{R} \mid r \geq 0\}$ with $d(x, x) = 0, d(x, y) = d(y.x)$ and triangle inequality) is *complete* iff every Cauchy sequence converges. In fact, any metric space has a 'canonical completion'. We define an equivalence relation between Cauchy sequences of X. Roughly speaking: $\{x_n\} \sim \{x'_n\}$ iff $\{x_n\}$ and $\{x'_n\}$ are 'trying to converge to the same point that is trying to be in X'. The set of equivalence classes inherits the metric from (X, d) (I.e. in a representative-independent way); and (X, d) can be isometrically embedded in the set of equivalence classes.

Similarly: given an *incomplete* inner product space (sometimes called a 'pee-Hilbert space'), we can build its 'canonical completion' : which is a Hilbert space.

A subset of a vector space that is itself a vector space (so: closed under linear combination) is a *linear manifold*. A linear manifold that is closed, i.e. that contains the limit vector of every Cauchy sequence of vectors, is a *subspace*. For a finite-dimensional Hilbert space, every linear manifold is a subspace. A subspace of a separable Hilbert space is itself a separable Hilbert space.

Example: $l^2 := \{(x_1, x_2, ...) \mid x_n \in \mathbb{C}, \Sigma |x_n|^2 < \infty\}$ has an orthonormal basis $(1, 0, 0, ...), (0, 1, 0, 0, ...), ...$ $=: \{\phi_n\}$. So each vector is $\Sigma x_n \phi_n$. The partial sums are $(x_1, x_2, ..., x_N, 0, 0, 0, ...)$, and these converge

to $(x_1, x_2, ..., x_N, x_{N+1}, ...)$.

In general: each vector ψ has a unique expression in terms of an orthonormal basis $\{\phi_n\}$: $\psi = \Sigma\,(\phi_n, \psi)\phi_n$. One uses Schwarz, applied to partial sums, to show this .

Any two Hilbert spaces (over $\mathbb{R}$, or over $\mathbb{C}$) of equal dimension are isomorphic: ("just map one ortho-basis onto another"). So any infinite-dimensional separable Hilbert space can be identified with l^2.

Ortho-complements and projectors: if $\mathcal{M}$ is a subspace of a Hilbert space $\mathcal{H}$, then $\mathcal{M}^\perp := \{\psi \in \mathcal{H} \mid (\psi, \phi) = 0 \; \forall \phi \in \mathcal{H}\}$ is a linear manifold, indeed a subspace of $\mathcal{H}$. Every $\psi \in \mathcal{H}$ has a unique expression as a sum of two components in $\mathcal{M}$ and in $\mathcal{M}^\perp$ respectively: $\psi = \psi_{\mathcal{M}} + \psi_{\mathcal{M}^\perp}$.

3: *Spaces of functions*:—
Treating *spaces of functions* needs care, for two main reasons. We *want to* say: $\int \psi^*(x)\phi(x)\,dx$ is an inner product.
(1): But an inner product requires: $(\psi, \psi) \geq 0$ with equality only if $\psi = 0$. And there are many non-zero functions on, say $[0, 1]$, $\psi : [0, 1] \to \mathbb{C}$, with $\int_0^1 |\psi|^2 dx = 0$.
(2): Secondly, in Riemann (i.e. elementary) integration theory, there are Cauchy sequences that do not converge. Define $f_n : [0, 1] \to \{0, 1\} \subset \mathbb{R}$ by $f_n(x) := 1$ iff x is of the form $\frac{m}{2^{-n}}$, with m an integer between 0 and 2^n; and otherwise $f_n(x) := 0$. Then any two functions $f_n, f_{n'}$ differ at only finitely many points; and for every n, $\int f_n = 0$. But the limit of the sequence $\{f_n\}$ is the function f that takes the value 1 on every integer-multiple of a reciprocal of a power of 2, and is otherwise 0. f is not Riemann-integrable.

Both problems are solved by adopting Lebesque integration. We will not give details of this and the associated measure theory. But we note that the function f just defined (value 1 on every integer-multiple of a reciprocal of a power of 2, and otherwise 0) *is* Lebesque-integrable and $\int_{[0,1]} f\,dx = 0$: an intuitive result in that f takes value 1 on a 'merely' denumerable set of arguments.

For us, the benefits of adopting Lebesque integration can be summed up, in terms of our two problems: as follows ...

As to (1): We define an equivalence relation between functions on, say $[0, 1]$, $\psi : [0, 1] \to \mathbb{C}$: $f \sim f'$ iff f and f' are equal *almost everywhere* (a.e.), meaning 'equal everywhere except on a set of (Lebesque) measure 0'. Then the equivalence classes $[f]$ themselves form a vector space, in a natural way. For example, the equivalence class of the pointwise sum $f + g$ of two representative elements, $f \in [f]$ and $g \in [g]$, is independent of the representatives chosen. Besides, the equivalence classes $[f]$ of those functions f, whose square integral $\int |f|^2 dx$ is finite, form an inner product space in a natural way. That is: the inner product we intuitively want to have, viz. $\int f^*(x)g(x)\,dx$ is well-defined on the equivalence classes, since the integral is independent of the representatives f, g that are chosen. Thus returning to the original problem (1): the equivalence class of the zero-function, $[0]$ is the unique vector with norm zero.

As to (2): This inner product space whose elements are equivalence classes (under: almost everywhere equality) of Lebesque-integrable functions f with finite square integral on, say $[0, 1]$, i.e. $\int_0^1 |f|^2\,dx < \infty$, is *complete*. That is: it is a Hilbert space. Similarly for square-integrable functions on the whole real line. These spaces are called L^2 spaces. Thus we write, understanding the equivalence relation to be: almost everywhere equality:—

$$L^2([0,1]) := \{[\psi] \mid \psi : [0,1] \to \mathbb{C}, \int_0^1 |\psi|^2\,dx < \infty\} \;\; ; \;\; L^2(\mathbb{R}) := \{[\psi] \mid \psi : \mathbb{R} \to \mathbb{C}, \int_{\mathbb{R}} |\psi|^2\,dx < \infty\}$$

$$(1)$$

4: *Philosophical Remarks*:—

(A): We of course recognise $L^2(\mathbb{R})$ as the (rigorous version of) the quantum state-space of a spinless non-relativistic particle confined to the real line: the state-space with which we all first learnt wave mechanics. Since the classical configuration space of such a particle is $\mathbb{R}$, we see here the basic idea that the *state-space of a quantum system consists of assignments of complex amplitudes to the possible configurations of the corresponding classical system*. This will later lead to *quantization theory*. A general slogan, and notation, would be: "we replace a classical configuration space $\mathcal{Q}$ by its L^2 space: $L^2(\mathcal{Q})$".

(B): We already see here, in embryo, the *measurement problem*: "how can we extract—how does Nature deliver to us— a single classical fact, e.g. a particle being in a position $X \in \mathbb{R}$, from a function $\psi : \mathbb{R} \to \mathbb{C}$?" ... about which this document (this course?!) will be silent...

(C): Even if one sets aside the physical and philosophical measurement problem, the question arises: *'Why Hilbert space?'* That is: what motivations can be given for assuming the state space of a physical system is a Hilbert space?
Various research traditions make this precise, and offer an answer. Here is a glimpse of three:

(1): *Quantum logic.* This was initiated by Birkhoff and von Neumann in their 'Logic of Quantum Mechanics' (1936), and flourished especially in the 1960s (the Geneva school of J. Jauch) and later. Experimental 'Yes-No' propositions about a physical system are partially ordered by a sort of logical strength, endowed with logical operations of conjunction (and), disjunction (or) and negation (not), subject to certain (judiciously chosen!) conditions, to make them a *lattice*, (usually: an *orthomodular lattice*). This lattice is then shown in a representation theorem to be represented by the lattice of subspaces/projectors of a Hilbert space, partially ordered by inclusion.

(2): *'Algebras of quantities': C*-algebras.* We have not yet discussed *quantities*. (In Section 2 below, we will review the usual treatment of them as self-adjoint operators on Hilbert space.) So we have not yet seen the quantum-classical contrast as a matter of quantum theory allowing non-commutation of quantities. But a good case can be made for thinking of a system as given primarily by an algebra of quantities: an algebra that is commutative for classical systems, but non-commutative for quantum systems. States are then introduced as mathematical superstructure on top of the algebra of quantities: namely, as linear expectation functionals—details below. The most developed version of this approach uses *C*-algebras*. More details later: here, we just note that in such an algebra, you can multiply any two elements even if they do not commute—an allowance about which you might well raise a philosophical eyebrow ... Then states on C*-algebras are shown to be representable in the traditional i.e. familiar ways—phase space for classical systems, Hilbert space for quantum systems. The buzzword is: *the GNS construction*, applying to commutative, respectively non-commutative, algebras.

(3): *Information-theoretic and operational approaches.* Inspired by studies of quantum non locality and Bell's theorem, and the rise of quantum information theory, with its protocols for eg. teleportation: various approaches take as primitive a set of probability distributions, for various quantities (normally with a finite discrete set of outcomes), on various *individual and joint* systems (normally finite-dimensional). Thus conditions like *no signalling/parameter independence* are prominent. Again, the Hilbert space formalism (normally finite-dimensional) is recovered with appropriate representation theorems. Example: D'Ariano, Chiribella and Perinotti, *Quantum Theory from First Principles*, CUP.

(D): The question 'Why should quantities be represented by self-adjoint operators?' also has a very different aspect or meaning, that is not touched on in the literature under (C). After all, there is nothing to prevent one associated experimental outcomes with complex, e.g. imaginary, numbers: 'I can paint '5i metres' on the dial of an apparatus measuring position!' In Section 2, we will report a helpful classification (due to Roberts) of the possibilities for a quantum physical quantity to be represented by an operator that is *not* self-adjoint.

(E): The pilot-wave theory is a noble tradition for solving the measurement problem. But it has been developed entirely using intuitive wave mechanics, not L^2 spaces.

5: *Linear functionals*:—
Given a vector space V over the field $\mathbb{R}$ or $\mathbb{C}$, the *dual space* V^* consists of the linear functionals $F : V \to \mathbb{R}$ (or $\mathbb{C}$). We recall that for a finite dimensional V, $\dim(V) = \dim(V^*)$; but there is no natural (canonical, i.e. basis-independent) isomorphism between them. But between V and its second dual V^{**}, there is a natural isomorphism.

With the extra structure of an inner product space, there is a natural isomorphism between V and its dual V^*. (This underlies how in relativity theory, the metric raises and lowers indices.) Thus each ψ in an inner product space defines a linear functional F_ψ by:

$$F_\psi(\phi) := (\psi, \phi). \tag{2}$$

And if V is finite-dimensional, with $\{\phi_i\}$ an orthonormal basis: we assign to each $F \in V^*$, the vector $\psi_F := \Sigma_i \, F(\phi_i)^* \phi_i \in V$. Applying the definition in eq. 2 to this ψ_F yields F again. That is: for any vector $\phi = \Sigma_i \, (\phi_i, \phi)\phi_i$, we have:

$$F(\phi) = \Sigma_i \, (\phi_i, \phi)F(\phi_i) = (\psi_F, \phi). \tag{3}$$

To get a corresponding basis-independent correspondence for an infinite-dimensional inner product space, we must require the linear functionals to be *continuous*, defined in the obvious way. Namely: that F is continuous iff: $\psi_n \to \psi$ implies that $F(\psi_n) \to F(\psi)$. Then we have the *Riesz representation theorem*:—
For every continuous linear functional F on a separable Hilbert space, there is a unique $\psi_F \in \mathcal{H}$, such that $F(\phi) = (\psi_F, \phi)$.

Of course, in Dirac notation the correspondence between linear functionals and vectors induced by the inner product is built in to the notation. The linear functional F_ψ is denoted by $\langle\psi|$, and the two sides of eq. 2 are written as $\langle\psi|\phi\rangle$.

2   Linear operators on a Hilbert space

1: Linear operators and matrices: the elements recalled ... On a suitable space of functions, a linear operator might be defined by

$$(A\psi)(x) := \frac{d\psi(x)}{dx} \quad ; \quad (A\psi)(x) := \int a(x, y)\psi(y)dy \tag{4}$$

We shall generally assume that all Hilbert spaces are separable.

2: *Bounded operators*:—
A linear operator A is *continuous* iff: $\psi_n \to \psi$ implies that $A(\psi_n) \to A(\psi)$. A linear operator A is *bounded* iff there is a positive number b such that for all ψ, $||A\psi|| < b||\psi||$. The infimum of such numbers is the *norm* of A, written $||A||$.

Theorem: A linear operator A is continuous iff it is bounded. (If $\mathcal{H}$ is finite-dimensional, then every linear operator is continuous, and so bounded.)

The norms of linear operators obey:

$$||A + B|| \le ||A|| + ||B|| \; ; \; ||aA|| = |a|||A|| \; ; \; ||A|| = 0 \text{ iff } A = 0 \; ; \; ||AB|| \le ||A||.||B|| \, . \tag{5}$$

The first three assertions follow straightforwardly from the same properties of the norm for vectors in $\mathcal{H}$. So we have a normed vector space of linear operators.

A bounded linear operator on a separable Hilbert spaces can be represented by a matrix.

3: *Inverses:*—

A linear operator A has an *inverse* if there is linear operator B with $AB = 1 = BA$. Theorem: A has an inverse iff: $\forall \psi, \exists ! \phi$ with $\psi = A\phi$.

For a finite-dimensional $\mathcal{H}$, $\dim(\mathcal{H}) = n$, with $\{\phi_i\}$ any basis:— Each of the following is necessary and sufficient for A to have an inverse:

(i): there is no non-zero vector ϕ such that $A\phi = 0$;
(ii): the set $\{A\phi_1, ..., A\phi_n\}$ is linearly independent;
(iii): there is a linear operator B such that $BA = 1$;
(iv): the matrix corresponding to A has a non-zero determinant.

But for an infinite-dimensional $\mathcal{H}$, (i)-(iii) are not sufficient—even together. For consider the "right-shift" on l^2: $A : l^2 \to l^2$, with $A(x_1, x_2, x_3, ...) := (0, x_1, x_2, x_3, ...)$. Then (i) and (ii) hold. Also: define B as "delete the first component and left-shift": $B(x_1, x_2, x_3, ...) := (x_2, x_3, ...)$; then (iii) holds. But A has no inverse. For if $\psi = (x_1, x_2, x_3, ...)$ with $x_1 \neq 0$ then there is no ϕ such that $\psi = A\phi$.

4: *Unitaries:*—

A linear operator U is *unitary* iff: both (a) U has an inverse and (b) $||U\psi|| = ||\psi||$ for all ψ. (Incidentally: the example of the "right-shift" on l^2, just above, shows that for a infinite-dimensional $\mathcal{H}$, condition (a) is needed.)

Every unitary operator is bounded, with $||U|| = 1$.

Theorem: If U is unitary, then $(U\psi, U\phi) = (\psi, \phi)$ for all ψ, ϕ. Corollary: It follows that the unitary image of an ortho-basis is an ortho-basis.

A "partial converse to the Corollary": If U is bounded, and the U-image of some ortho-basis is an ortho-basis, then U is unitary.

5: *Adjoints and Hermitian operators:*—

Let A be bounded, and so continuous. Then for each $\psi \in \mathcal{H}$, the linear functional $F^{[\psi]}$ defined by

$$F^{[\psi]}(\phi) := (\psi, A\phi) \tag{6}$$

is continuous. And so, by the Riesz theorem (Section 1: 5), there is a unique vector, call it $A^\dagger(\psi)$ such that $F^{[\psi]}(\phi) := (A^\dagger(\psi), \phi)$.

$A^\dagger$ is trivially linear. Using the Schwarz inequality (applied to $||A^\dagger(\psi)||^2 = (\psi, AA^\dagger\psi)$), one has:

Theorem: If A is bounded, then $A^\dagger$ is bounded, and $||A^\dagger|| = ||A||$.

One checks that: $A^{\dagger\dagger} = A$; $(AB)^\dagger = B^\dagger A^\dagger$; $(aA)^\dagger = a^* A^\dagger$; $(A + B)^\dagger = A^\dagger + B^\dagger$.

A *bounded* linear operator A is *self-adjoint* or *Hermitian* iff $A^\dagger = A$. NB: for an infinite-dimensional $\mathcal{H}$, it is impossible to define an unbounded Hermitian operator on all vectors: see later. This means: $(\phi, A\psi) = (A\phi, \psi)$. That is: $(\phi, A\psi) = (\psi, A\phi)^*$. So for all ψ, we have: $(\psi, A\psi) \in \mathbb{R}$.

Example: On $L^2([0, 1])$, we define $(A\psi)(x) := x\psi(x)$. This A is bounded with $||A\psi||^2 \leq ||\psi||^2$; and so $||A|| = 1$. And A is Hermitian, since $\int_0^1 \phi^*(x).x\psi(x)\,dx = \int_0^1 [x\phi(x)]^*.\psi(x)\,dx$.
But a "corresponding definition" on $L^2(\mathbb{R})$ is of course not bounded. But multiplying by a suitable

"damping factor" gives a bounded and Hermitian operator on $L^2(\mathbb{R})$, e.g. we define $(V\psi)(x) := (\exp|x|)\psi(x)$.

Theorem: If A is bounded, and has a bounded inverse A^{-1}, then $(A^\dagger)^{-1}$ exists and $(A^\dagger)^{-1} = (A^{-1})^\dagger$.

Corollary: If A is bounded and Hermitian, and has a bounded inverse A^{-1}, then A^{-1} is Hermitian.

Hermitian operators are analogues of real numbers. As in: if A is bounded, then its real and imaginary parts defined by

$$\mathrm{Re}A := \frac{1}{2}(A + A^\dagger) \ ; \ \mathrm{Im}A := \frac{-i}{2}(A - A^\dagger) \tag{7}$$

are bounded and Hermitian; and $A = \mathrm{Re}A + \mathrm{Im}A$.

In the same way, unitary operators are analogues of complex numbers of absolute value one. We have:–
Theorem: A linear operator U is unitary iff $U^\dagger U = 1 = UU^\dagger$.

For bounded operators A, B, one readily checks using the adjoints that for any ortho-basis $\{\phi_n\}$, the representing matrix (c_{jk}) of the product $C := AB$ is the product of the representing matrices, that is: $c_{jk} = \Sigma_i a_{ji} b_{ik}$.

6: *Projection operators*:—
Recall from Section 1:2 that if $\mathcal{M}$ is a subspace of $\mathcal{H}$, then $\mathcal{M}^\perp := \{\psi \in \mathcal{H} \mid (\psi, \phi) = 0, \ \forall \phi \in \mathcal{H}\}$ is also a subspace; and every $\psi \in \mathcal{H}$ has a unique expression as a sum of the two components: $\psi = \psi_{\mathcal{M}} + \psi_{\mathcal{M}^\perp}$. So we define the projection/projector $E_{\mathcal{M}} : \mathcal{H} \to \mathcal{H}$, by $E_{\mathcal{M}}(\psi) := \psi_{\mathcal{M}}$.

Theorem: A bounded linear operator E is a projector iff $E^2 = E = E^\dagger$. (To prove the leftward implication, one defines the set $\mathcal{M}$ to be the range of E, shows it to be a subspace, and shows that for any vector ψ, $(1 - E)(\psi) \in \mathcal{M}^\perp$ etc.)

7: *Unbounded operators*:—
Example: The quantity *position* in the Schrödinger representation (of the canonical commutation relations; cf. later) on $L^2(\mathbb{R})$. We define $(Q\psi)(x) := x\psi(x)$. Then $||Q\psi||^2 \equiv \int_{\mathbb{R}} |x\psi(x)|^2 \, dx$ can be arbitrarily larger than $||\psi||^2 \equiv \int_{\mathbb{R}} |\psi(x)|^2 \, dx$. So Q is unbounded. In fact we have the following ...

Theorem: If a linear operator A is defined for all vectors, and if $(\phi, A\psi) = (A\phi, \psi)$ for all ϕ, ψ, then: A is bounded.

Quantum theory needs unbounded operators A with the algebraic Hermitian property, i.e. $(\phi, A\psi) = (A\phi, \psi)$ for all ϕ, ψ in the domain of A. So it needs operators A with domains of definition less than all of $\mathcal{H}$. Hence the jargon of: *dense domain*, and *extension*, of an operator.

If A has a dense domain, we can define $A^\dagger$. Namely:
    $\mathrm{dom}(A^\dagger) := \{\psi \in \mathcal{H} \mid \text{there is a vector } \tilde{\psi} \text{ such that } \forall \phi \in \mathrm{dom}(A) : (\phi, \tilde{\psi}) = (A\phi, \psi) \,.\}$
Then we define $A^\dagger$ by $A^\dagger : \psi \in \mathrm{dom}(A^\dagger) \mapsto \tilde{\psi}$. This defines $A^\dagger(\psi)$ uniquely (because $\mathrm{dom}(A)$ is dense); and $A^\dagger$ is linear, and $\mathrm{dom}(A^\dagger)$ is a linear manifold.

We say that a linear operator A is *symmetric* iff: A has a dense domain, and $(\phi, A\psi) = (A\phi, \psi)$ for all ϕ, ψ in the domain of A. Then by the discussion just above, we conclude that: for all ψ in the domain of A, $A^\dagger(\psi)$ is defined, and $A^\dagger(\psi) = A(\psi)$. That is: $A^\dagger$ is an *extension* of A.

If in fact $A^\dagger = A$, then we say A is *self-adjoint* or *Hermitian*.

Example: We define Q on $L^2(\mathbb{R})$ by specifying that $\mathrm{dom}(Q)$: =. $\{\psi \mid \int_{\mathbb{R}} |x\psi(x)|^2 \, dx < \infty\}$. This domain is dense. (For we can approximate an arbitrary $\psi \in L^2(\mathbb{R})$ by the "truncated" functions ψ_n

($n \in \mathbb{Z}$) that are defined to be equal to ψ on the interval $[-n, n]$, and to take the value 0 outside that interval. Clearly $\psi_n \to \psi$; and $x\psi_n(x)$ is square-integrable.) Then on this domain, we define: $(Q\psi)(x) := x\psi(x)$. Then Q is clearly symmetric, since $\int_{\mathbb{R}} \phi^*(x).x\psi(x)\,dx = \int_{\mathbb{R}} [x\phi(x)]^*.\psi(x)\,dx$. So $Q^\dagger$ is defined and extends Q. But is the domain of $Q^\dagger$ in fact larger than the domain of Q? In fact it is not larger: (cf. Jordan, p.31). So $Q^\dagger = Q$, and so Q is self-adjoint, aka: Hermitian.

A symmetric operator that cannot be extended to a larger domain is called *maximal symmetric*. Theorem: Every self-adjoint operator is maximal symmetric; (but not conversely)

An unbounded operator cannot be continuous. The "next best thing to continuity" is being *closed*, as follows.

We say an operator A is *closed* iff: if (i) a sequence of vectors ψ_n in dom(A) converges to a vector ψ and (ii) the sequence of vectors $A(\psi_n)$ converges to a vector ϕ, then $\psi \in$ dom(A) and $A\psi = \phi$.

Theorem: If dom(A) is dense, then $A^\dagger$ is closed. So every self-adjoint operator is closed.

It is natural to ask: *'What are the conditions for a symmetric operator to be self-adjoint?'* As just noted, being maximally symmetric is not sufficient. What about being symmetric and closed? This also turns out to be not sufficient. But in paragraph 4 of the next Section, we will get an answer: an answer that relates to Philosophical Remark (D) in paragraph 4 of Section 1: about the idea that a quantity can be represented by an operator that is *not* self-adjoint.

3   Diagonalizing operators

We assume a complex separable Hilbert space.

1: *Eigenvalues and eigenvectors*:—
The definition of eigenvalue and of eigenvector is exactly the same for an infinite-dimensional Hilbert space, as for the finite-dimensional case. The following elementary but important results are derived exactly as for the finite-dimensional case:

1: If T is a linear operator with an inverse T^{-1}, then A and TAT^{-1} have the same eigenvalues.

2: The eigenvalue of a Hermitian, respectively unitary, operator is real, respectively of absolute value 1.

3: Two eigenvectors of a Hermitian, or of a unitary, operator, corresponding to different eigenvalues are orthogonal.

Let A be Hermitian or unitary. Let $a_1, a_2, ..., a_k, ...$ be its eigenvalues, with eigenspaces $\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_k, ...$ Then the orthogonal sum $Eig(A) := \oplus_k \mathcal{M}_k$ is the subspace of $\mathcal{H}$ spanned by eigenvectors of A. Of course, for a finite-dimensional complex Hilbert space, $Eig(A) = \mathcal{H}$. This is the *spectral decomposition* or *eigen-decomposition* of the operator A.

We say that a subspace $\mathcal{M} < \mathcal{H}$ *reduces* the linear operator A iff both $\mathcal{M}$ and $\mathcal{M}^\perp$ are invariant under A. This turns out to be equivalent to A commuting with the projector onto $\mathcal{M}$:

Theorem: Let $E_{\mathcal{M}}$ be the projector onto $\mathcal{M}$. Then $\mathcal{M}$ reduces A iff $E_{\mathcal{M}}A = AE_{\mathcal{M}}$ iff $(1-E_{\mathcal{M}})A = A(1 - E_{\mathcal{M}})$.

Theorem: Let A be Hermitian or unitary: then $Eig(A)$ reduces A. So a Hermitian or unitary operator splits in to two separate parts: one part acting on $Eig(A)$, and represented there, with respect to an eigenbasis, as a diagonal matrix; the other part acting on $Eig(A)^\perp$.

2: *Eigenvalue decomposition*:—
We will generalise the spectral decomposition of a Hermitian or unitary operator A to the infinite-dimensional case, i.e. address the question of how A acts on $Eig(A)^\perp$. We begin by rewriting in a suggestive way the spectral decomposition of a Hermitian operator A for the finite-dimensional case.

So let A's real eigenvalues be, in ascending order: $a_1 < a_2 < ... < a_k < ... < a_m$, with corresponding eigenspaces $\mathcal{M}_k$. Let I_k be the projector on to $\mathcal{M}_k$. Then the spectral decomposition of A is just

$$A = \Sigma_{k=1}^m a_k I_k \equiv \oplus_{k=1}^m a_k I_k \tag{8}$$

Now we define for each real number x, $E_x := \oplus_{a_k < x} I_k$. So: $E_x = 0$ for $x < a_1$, and $E_x = 1$ for $x > a_m$. And if $x < y$, then $E_x E_y = E_x = E_y E_x$, i.e. $E_x \leq E_y$.

For each $x \in \mathbb{R}$, we also define $dE_x := E_x - E_{x-\varepsilon}$ with ε chosen so small that there is no a_j such that $x - \varepsilon < a_j, x$.

So dE_x is not zero only when x is an eigenvalue a_k; and in that case $dE_x = I_k$.

So for $\Sigma_{k=1}^m I_k = 1$, we can write: $\int_{\mathbb{R}} dE_x = 1$. And for $A = \Sigma_{k=1}^m a_k I_k$, we can write: $A = \int_{\mathbb{R}} x \, dE_x$.

Besides: $(\phi, E_x \psi)$ is a complex function of $x \in \mathbb{R}$ that is continuous from the right, but which jumps in value by $(\phi, I_k \psi)$ at $x = a_k$. So we have (a sordinary Riemann integrals):

$$(\phi, \psi) = \int_{\mathbb{R}} d(\phi, E_x \psi) \; ; \; (\phi, A\psi) = \int_{\mathbb{R}} x \, d(\phi, E_x \psi) \,. \tag{9}$$

Similarly, for a unitary operator U in the finite-dimensional case. Its eigenvalues are $u_k \equiv e^{i\theta_k}$, where in ascending order: $0 < \theta_1 < \theta_2 < ... < \theta_k < ... < \theta_m \leq 2\pi$. Then we define for each real number x, $E_x := \oplus_{\theta_k < x} I_k$. So we can write:

$$U = \int_0^{2\pi} e^{ix} dE_x \; ; \; (\phi, U\psi) = \int_0^{2\pi} e^{ix} \, d(\phi, E_x \psi) \,. \tag{10}$$

3: *Spectral decomposition:*—
A family of projectors $\{E_x\}_{x \in \mathbb{R}}$ is called a *spectral family* iff:
    (i) if $x \leq y$ then $E_x \leq E_y$, i.e. $E_x E_y = E_x = E_y E_x$;
    (ii) continuity from the right: for all ψ and for all x: if $\varepsilon > 0$, then $E_{x+\varepsilon} \psi \to E_x \psi$, as $\varepsilon \to 0$;
    (iii) for all ψ: $E_x \psi \to 0$ as $x \to -\infty$, and $E_x \psi \to \psi$ as $x \to +\infty$. The main theorem is then...

The Spectral Theorem: For each self-adjoint operator, there is a unique spectral family $\{E_x\}_{x \in \mathbb{R}}$ such that for all $\psi, \phi \in \mathcal{H}$:

$$(\phi, A\psi) = \int_{\mathbb{R}} x \, d(\phi, E_x \psi) \; ; \quad \text{so we write} \quad A = \int_{\mathbb{R}} x \, dE_x \,. \tag{11}$$

Similarly for unitary operators U, with $E_x = 0$ for $x \leq 0$, and $E_x = 1$ for $x > 2\pi$:

$$(\phi, U\psi) = \int_0^{2\pi} e^{ix} \, d(\phi, E_x \psi) \; ; \quad \text{so we write} \quad U = \int_0^{2\pi} e^{ix} \, dE_x \,. \tag{12}$$

Example: "Position" on $L^2([0,1])$. We define E_x on $L^2([0,1])$ as "chopping the function off above x". That is: $(E_x \psi)(y) := \psi(y)$ for $y \leq x$, and $(E_x \psi)(y) := 0$ for $y > x$. Then

$$||E_{x+\varepsilon} \psi - E_x \psi)||^2 = \int_x^{x+\varepsilon} |\psi(y)|^2 \, dy \to 0, \text{ as } \varepsilon \to 0 \,. \tag{13}$$

and $\{E_x\}_{x \in \mathbb{R}}$ is a spectral family. Now define A on $L^2([0,1])$ by $(A\psi)(x) := x\psi(x)$. Then A is bounded and self-adjoint. For all $\psi, \phi \in \mathcal{H}$, we have

$$\int_{\mathbb{R}} x \, d(\phi, E_x \psi) = \int_{\mathbb{R}} x \, d \int_0^1 \phi(y)^* (E_x \psi)(y) \, dy = \int_{\mathbb{R}} x \, d \int_0^x \phi(y)^* \psi(y) \, dy = \int_0^1 \phi(x)^* x \psi(x) \, dx = (\psi, A\phi) \,. \tag{14}$$

8

So $\{E_x\}_{x\in\mathbb{R}}$ gives the spectral decomposition of A.

This $\{E_x\}_{x\in\mathbb{R}}$ is continuous from the left, as well as from the right. We have

$$(\psi, E_x\phi) - (\psi, E_{x-\varepsilon}\phi) \equiv \int_{x-\varepsilon}^{x} \psi^*(y)\phi(y)dy \;\rightarrow 0, \;\; \text{as } \varepsilon \to 0 \,. \tag{15}$$

We say that $\{E_x\}_{x\in\mathbb{R}}$ *jumps in value at* x if for some vector ψ, $(E_x - E_{x-\varepsilon})\psi$ does *not* converge to 0, as $\varepsilon \to 0$. Otherwise, we say that $\{E_x\}_{x\in\mathbb{R}}$ *is continuous* at x.

So in the above example, $\{E_x\}_{x\in\mathbb{R}}$ is continuous at all x, since A has no eigenvalues/eigenvectors. Similarly of course for position on $L^2(\mathbb{R})$.

Theorem: Let A be a self-adjoint operator with spectral decomposition $A = \int_{\mathbb{R}} x \, dE_x$. Then $\{E_x\}_{x\in\mathbb{R}}$ jumps in value at a iff a is an eigenvalue of A. And with I_a the projector onto the eigenspace fo a, we have: $E_x I_a = 0$ for $x < a$; and $E_x I_a = I_a$ for $x \geq a$; and for any ψ, $E_a\psi - E_{a-\varepsilon}\psi \to I_a\psi$, as $\varepsilon \to 0$.

Accordingly, we define:—-
(1): the *spectrum* of $A := \mathrm{sp}(A) := \{x \in R \,|\, E_x \text{ increases}\} \equiv \{x \in R \,|\, x \text{ is not in an interval } (a,b) \text{ on which } E_x \text{ is c}$
(2): the *point spectrum* of $A := \{x \in R \,|\, E_x \text{ jumps}\} \equiv \{x \in R \,|\, x \text{ is an eigenvalue of } A\}$;
(3): the *continuous spectrum* of $A := \{x \in R \,|\, E_x \text{ increases continuously}\}$.

Theorem:: A self-adjoint operator is bounded iff its spectrum is bounded.

We say that self-adjoint operator A is *positive* iff for all ψ, $(\psi, A\psi) \geq 0$.

Theorem:: A self-adjoint operator is positive iff its spectrum is non-negative.

4: *Philosophical remark*:—
We return to the Philosophical Remark (D) in paragraph 4 of Section 1: the idea that a quantum physical quantity can be represented by an operator that is *not* self-adjoint. Cf. B. Roberts (2018): 'Observables, disassembled', *Studies in History and Philosophy of Modern Physics* **63**, 150– 162. (Preprint: http://philsci-archive.pitt.edu/14449/). On p. 153, Roberts reports that being self-adjoint is a "two out of three" property. That is:
A closed, densely-defined linear operator A is self-adjoint if it satisfies *any two* of the following three properties:

(1): A is *normal*. That is: $AA^\dagger = A^\dagger A$;

(2): A is symmetric: $(\phi, A\psi) = (A\phi, \psi)$ for all ϕ, ψ in the (dense) domain of A.

(3): A has real Spectrum, where we define 'Spectrum' in a more general way than we did 'spectrum' above: namely as the set $\mathrm{Spec}(A) := \{z \in \mathbb{C} \,|\, (A - z.\mathbb{1}) \text{ has no inverse}\}$. So the condition is: $\mathrm{Spec}(A) \subset \mathbb{R}$.

This yields four ways that a closed, densely-defined linear operator A can *fail to be* self-adjoint: having just one of the above three properties, or having none of them. Roberts then explores each of these four ways, finding for each of them: conceptual issues and circumstances in which it is a reasonable notion of physical quantity.

5: *Functions of an Operator*:—
Let A be a self-adjoint operator, with spectral decomposition $A = \int_{\mathbb{R}} x \, dE_x$. Let f be a complex-valued function on the real line: $f : \mathbb{R} \to \mathbb{C}$. We define the operator $f(A)$, by

$$(\phi, f(A)\psi) := \int_{\mathbb{R}} f(x) \, d(\phi, E_x\psi) \,; \tag{16}$$

which for f continuous can be taken as an ordinary Riemann integral. Then we have:

(1): for $f(x) = x$, $f(A) = A$;

(2): for $f(x) = 1$, $f(A) = 1$, since $\int_\mathbb{R} d(\phi, E_x \psi) = (\phi, \psi)$;

(3): $(f + g)(A) = f(A) + g(A)$ and $(cf)(A) = c(f(A))$;

(4): we define $(fg)(x) := f(x)g(x)$, so that $(\phi, (fg)(A)\psi) := \int_\mathbb{R} (fg)(x) d(\phi, E_x \psi) = \int_\mathbb{R} f(x)g(x) d(\phi, E_x \psi)$, and then we compute that

$$\int_\mathbb{R} f(x)g(x) d(\phi, E_x \psi) = \int_\mathbb{R} f(x) d \int_{-\infty}^{x} g(y) d(\phi, E_y \psi) = \int_\mathbb{R} f(x) d_x \int_{-\infty}^{+\infty} g(y) d_y(E_x \phi, E_y \psi) = \quad (17)$$

$$\int_{-\infty}^{+\infty} f(x) d(\phi, E_x g(A)\psi) = (\phi, f(A)g(A)\psi) . \quad (18)$$

So we conclude that $(fg)(A) = f(A)g(A)$, and thus that functions of the operator A commute with each other.

So (5): polynomial functions are defined in the natural way. If $f(x) = c_0 + c_1 x + c_2 x^2 + ... + c_n x^n$, then $f(A) = c_0 + c_1 A + c_2 A^2 + ... + c_n A^n$.

(6) If we define $(f^*)(x) := (f(x))^*$, then we compute that

$$(\phi, [f(A)]^\dagger \psi) = (\psi, f(A)\phi)^* = \int_{-\infty}^{+\infty} f(x)^* d(\psi, E_x \phi)^* = \int_{-\infty}^{+\infty} (f^*)(x) d(\phi, E_x \psi) ; \quad (19)$$

so that $[f(A)]^\dagger = (f^*)(A)$. So if f is a real-valued function, $f; \mathbb{R} \to \mathbb{R}$, $f(A)$ is also self-adjoint. And if $f^* f = 1$, then $f(A)$ is a unitary operator since $[f(A)]^\dagger f(A) = 1 = f(A)[f(A)]^\dagger$.

(7): $f(A)$ is positive if $f(x) \geq 0$ on the spectrum of A. For just consider: $(\phi, f(A)\phi) = \int_{-\infty}^{+\infty} f(x)d||E_x \phi||^2$. Similarly, we deduce:

(8): $f(A)$ is bounded if $|f(x)|$ is bounded on the spectrum of A.

6: *Stone's Theorem* :—

Given a self-adjoint operator $H = \int_{-\infty}^{+\infty} x dE_x$, we define for all $t \in \mathbb{R}$: $(\phi, U_t \psi) := \int_{-\infty}^{+\infty} e^{itx} d(\phi, Ex\psi)$. Then U_t is an operators, viz. $U_t = e^{itH}$ and U_t is unitary since $(e^{itx})^* . e^{itx} = 1$ (cf. the end of (6) above). Evidently, $U_0 = 1$; and since $e^{itx} e^{it'x} = e^{i(t+t')x}$, we have $U_t U_{t'} = U_{t+t'}$. The converse of this is...

Stone's Theorem: Suppose that for all $t \in \mathbb{R}$, U_t is a unitary operator, such that:

(i): for all vectors ϕ, ψ: $(\phi, U_t \psi)$ is a continuous function of t

(ii): $U_0 = 1$ and $U_t U_{t'} = U_{t+t'}$: so the family $\{U_t\}_{t \in \mathbb{R}}$ is a *unitary representation* of the group $(\mathbb{R}, +)$.

Then: there is a unique a self-adjoint operator H such that $U_t = e^{itH}$ for all $t \in \mathbb{R}$, and

(1): the domain of H is $\{\psi \in \mathcal{H} \mid \frac{1}{it}(U_t - 1)\psi$ converges as $t \to 0\}$; and then the limit vector is $H\psi$;

(2): if a bounded operator commutes with all of the U_t, then it commutes with H.

Using (1), we infer: If $U_t \psi \in \text{dom}(H)$, then: $\frac{1}{i\Delta t}(U_{\Delta t} - 1)U_t \psi \to HU_t \psi$, as $\Delta t \to 0$. That is:

$$\frac{1}{i\Delta t}(U_{t+\Delta t} - U_t)\psi \to HU_t \psi, \text{ as } \Delta t \to 0 \quad (20)$$

which we write as the "Schrödinger equation":

$$-i\frac{d}{dt}(U_t \psi) = HU_t \psi . \quad (21)$$

7: *Functions of commuting operators:*—

Theorem: Let A be self-adjoint with spectral decomposition $A = \int x dE_x$; and let B be bounded, self-adjoint and $AB = BA$. Then $BE_x = E_x B$.

Proof: If A has pure point spectrum, $A = \Sigma_k a_k I_k$, the proof is elementary. If A is unbounded, one needs a more careful statement to deal with issues about domains. Indeed: if two self-adjoint operators A_1, A_2 are both unbounded, then we take $E_x^1 E_y^2 = E_y^2 E_x^1$ for all $x, y \in \mathbb{R}$ to be the *definition* of $[A_1, A_2] = 0$.

Given $f : \mathbb{R}^2 \to \mathbb{C}, (x, y) \mapsto f(x, y)$, we define $f(A_1, A_2)$ for two commuting self-adjoint operators A_1, A_2 by

$$(\phi, f(A_1, A_2)\psi) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) \, d_x d_y (\phi, E_x^1 E_y^2 \psi) \, , \quad \text{for all} \quad \phi, \psi \in \mathcal{H} \,. \tag{22}$$

Sums, scalar multiples, and products of such functions are defined in the obvious way. One shows that:

$[f(A_1, A_2)]^\dagger = (f^*)f A_1, A_2)$; so that real functions f define self-adjoint operators $f(A_1, A_2)$ and functions f for which $f^* f = 1$ define unitary operators $f(A_1, A_2)$;

$f(A_1, A_2)$ is positive if $f(x, y) > 0$ on the cartesian product of the spectra of A_1 and A_2 ;

$f(A_1, A_2)$ is bounded if $f(x, y)$ is bounded on the cartesian product of the spectra of A_1 and A_2.

Compare items (5) to (8) at the end of Paragraph 5 above.

Example: the Schrödinger representation of position in $\mathbb{R}^3$. Consider $L^2(\mathbb{R}^3) \ni \psi(\mathbf{x}) \equiv \psi(x_1, x_2, x_3)$; with inner product $(\phi, \psi) = \int_{\mathbb{R}^3} \phi^*(\mathbf{x})\psi(\mathbf{x}) \, d\mathbf{x}$. For $r = 1, 2, 3$, we define the self-adjoint operator Q_r by $Q_r \psi(\mathbf{x}) = x_r \psi(\mathbf{x})$. Then in the spectral decomposition

$$Q_r = \int_{\mathbb{R}} x \, dE_x^r \tag{23}$$

the projectors E_x^r are defined by : $(E_x^r \psi)(\mathbf{y}) = \psi(\mathbf{y})$ for $\mathbf{y}$ with $y_r \leq x$, and $(E_x^r \psi)(\mathbf{y}) = 0$ for $\mathbf{y}$ with $y_r > x$. Then one has

$$(f(Q_1, Q_2, Q_3)\psi)(\mathbf{x}) = f(x_1, x_2, x_3)\psi(\mathbf{x}) \,. \tag{24}$$

8: *Complete sets of commuting operators* :—

Let $A_1, A_2, ..., A_N$ be mutually commuting self-adjoint operators with pure point spectra; each with their spectral decomposition, $r = 1, 2, ..., N$

$$A_r = \Sigma_k a_k^{(r)} I_k^{(r)} \,. \tag{25}$$

Then for all r, s and $j.k$, we have: $I_k^{(r)} I_j^{(s)} = I_j^{(s)} I_k^{(r)}$. And for any $j, k, ...l$, the product $I_j^{(1)} I_k^{(2)} ... I_l^{(N)}$ is a projector. Namely, the projector onto the subspace of simultaneous eigenvectors with corresponding eigenvalues, i.e. the space of vectors ψ with $A_1 \psi = a_j^{(1)} \psi$, $A_2 \psi = a_k^{(2)} \psi, ..., A_N \psi = a_l^{(N)} \psi$. Some of these projectors may be zero: corresponding to combinations of eigenvalues that are not, in philosophical jargon!, *co-possible* or *compossible*.

These subspaces are orthogonal, i.e.

$$I_j^{(1)} I_k^{(2)} ... I_l^{(N)} I_{j'}^{(1)} I_{k'}^{(2)} ... I_{l'}^{(N)} = \delta_{jj'} \delta_{kk'} ... \delta_{ll'} \, I_j^{(1)} I_k^{(2)} ... I_l^{(N)} \tag{26}$$

and complete, i.e.

$$\Sigma_j \Sigma_k ... \Sigma_l \, I_j^{(1)} I_k^{(2)} ... I_l^{(N)} = 1 \,. \tag{27}$$

11

If none of these projects onto a subspace of dimension larger than one, we say that $\{A_1, A_2, ..., A_N\}$ is a *complete set* of commuting operators. Then choosing eigenvectors of length one, and labelling them with their eigenvalues, i.e. choosing an *orthonormal eigenbasis*, we can write in Dirac notation:

$$I_j^{(1)} I_k^{(2)} ... I_l^{(N)} = |a_j^{(1)} a_k^{(2)} ... a_l^{(N)}\rangle \langle a_j^{(1)} a_k^{(2)} ... a_l^{(N)}| \tag{28}$$

Theorem: Let $A_1, A_2, ..., A_N$ be mutually commuting self-adjoint operators with pure point spectra. This is a complete set iff: every bounded operator B that commutes with all the $A_1, A_2, ..., A_N$ is a function of them: $B = f(A_1, A_2, ..., A_N)$.

The orthonormal eigenbasis $\{|a_j^{(1)} a_k^{(2)} ... a_l^{(N)}\rangle\}$ gives a *spectral representation* of $A_1, A_2, ..., A_N$ and of the functions $f(A_1, A_2, ..., A_N)$ as diagonal matrices:

$$\langle a_j^{(1)} a_k^{(2)} ... a_l^{(N)} | f(A_1, A_2, ..., A_N)\psi \rangle = f(a_j^{(1)} a_k^{(2)} ... a_l^{(N)}) \langle a_j^{(1)} a_k^{(2)} ... a_l^{(N)} | \psi \rangle. \tag{29}$$

9: *Philosophical Remarks*:—
The selection of a self-adjoint operator with pure point spectrum, and so of a complete family of orthogonal eigenspaces, is at the heart of not just
    (i) the mathematics; but also:
    (ii) the physics; and
    (iii) the interpretation/philosophy
of quantum theory. Besides, the main issues are already clear in the case of finite dimensional Hilbert spaces. (Indeed, they are clear for real Hilbert spaces: for which, of course, 'self-adjoint' is replaced by 'symmetric' in the sense of elementary matrix theory, i.e. the matrix elements a_{ij} obey: $a_{ij} = a_{ji}$. Of course, this is not the sense we defined above!)

As to (i), the mathematics: we here note just:
Recall Philosophical Remark (C) (1), about quantum logic, in Paragraph 4 of Section 1. That is: the lattice $\mathcal{L}$ of projectors (equivalently: subspaces) of Hilbert space. Selecting a complete family of orthogonal eigenspaces, $\{E_i\}$ say, amounts to picking a sublattice of L: one that is *Boolean*. More precisely: a complete family of orthogonal eigenspaces (equivalently: projectors) is the set of atoms (smallest, logically strongest, least-in-the-partial-order—but non-zero—elements) of a Boolean sublattice of $\mathcal{L}$. The other elements of the sublattice are given by all the possible sums (orthogonal sums, $\oplus$) of these atoms. Roughly speaking: 'Boolean' means that the distributive laws, of intersection $\cap$ over $+$, and vice versa, hold. That is with E, F, G being three subspaces, $E, F, G < \mathcal{H}$, the laws are:

$$E \cap (F + G) = (E \cap F) + (E \cap G) ; \text{ and } E + (F \cap G) = (E + F) \cap (E + G) \tag{30}$$

These equations are easily proved for E, F, G mutually orthogonal (as are the atoms); cf. elementary projector algebra. But they also hold for the various possible sums of atoms. And they are very easily *disproved* for E, F, G 'skew'. Just take three mutually skew rays in the Euclidean plane $\mathcal{H} = \mathbb{R}^2$. Then the first equation of eq. 30 would read: $E = 0$, and the equation of eq. 30 would read: $E = \mathcal{H}$. Both of which are false.

We say that *roughly speaking* 'Boolean' means that the distributive laws hold: because we are here neglecting conditions about the behaviour of the complement, i.e. the unary operation on subspaces, $E \mapsto E^\perp$, or equivalently for projectors, $E \mapsto 1 - E$.

Three final remarks about this lattice-theoretic perspective:
    (1): Since a function f is in general many-one, i.e. two arguments can map to the same value, a self-adjoint operator B being a function of another A, i.e. $B = f(A)$, means that the sublattice of $\mathcal{L}$ that is defined, as sketched above, by B (with B's eigenspaces as its atoms) will be a sublattice of

the sublattice defined by A. In short: *coarse-graining*.

Note also that B can be a function of two operators A, C that do *not* commute: i.e. . $B = f(A) = g(C)$ with $[A, C] \neq 0$. This point will be crucial in the discussion of FUNC and of 'no go' theorems against 'hidden variable' supplementations of orthodox quantum theory, that we discuss under interpretation/philosophy below. For it means that imposing conditions on the assignment of values to the operator B (and perhaps some other operator with which B commutes) can have consequences for the values of operators A and C that do *not* commute.

(2): Later, when we discuss states i.e. probability distributions for the values of quantities (and so for the values 1 and 0 for projectors), this Booleanness of the sub-lattice will secure there being a classical (Kolmogorov) probability space for the distribution to be defined on.

(3): When we consider the algebraic approach to quantum theory, Booleanness of a lattice will correspond to the abelianness of the algebra of quantities. The rough picture, here stated only for bounded operators, is:

(3A): Although the elements of an algebra 'go beyond' projectors, by including also self-adjoint operators (indeed, arbitrary real linear combinations of projectors: which can then be spectrally resolved), and then also skew-adjoint operators, the projectors are the building blocks of the algebra; and so the abelianness of the algebra is caught by the mutual commutation of the projectors.

(3B): The commutation of projectors $[E, F] = 0$ is equivalent to a neat lattice-theoretic expression of their ranges (1-eigenspaces). If we now write E, F for the ranges, it is (using $\oplus$ to signal that the summands are orthogonal):

$$E = (E \cap F) \oplus (E \cap F^{\perp}) \; ; \; \text{ and } \; F = (E \cap F) \oplus (E^{\perp} \cap F) \, . \tag{31}$$

As to (ii), the physics: we note:

The selection of such an operator can be interpreted as choosing to measure the quantity it represents, i.e. choosing an experimental context. And the non-commutation of two operators represents it being impossible to measure them both simultaneously with arbitrary accuracy. *Recall Bohr on mutually exclusive experimental contexts!* Cf. e.g. his essay in P.A. Schilpp ed. *Albert Einstein: Philosopher-Scientist*

Later, after we introduce states, we will make simultaneous measurability more rigorous. As we will see: this can be done *without* committing us to a version of the Projection Postulate or a similar "collapse of the wave-packet". For the moment, we just note that for a finite dimensional (indeed real or complex) Hilbert space, it is natural to define:

(the quantities represented by) two self-adjoint operators A and B are *co-measurable* along the following lines:

'a measurement of one quantity does not disturb a pre-existing value of the other quantity'. And it is natural to make 'pre-existing value' more precise in terms of an outcome/result ("pointer-reading") from an immediately preceding measurement process. That is: we imagine a measurement of A yields some outcome/eigenvalue a; then an immediately succeeding measurement B yields some outcome/eigenvalue b; and then an immediately succeeding *second* measurement A is done and yields some outcome/eigenvalue a'. In this scenario with its three successive acts of measurement, it is natural to define:

'measuring B does not disturb the measurement of A'
as follows:

'Whatever are the outcomes/eigenvalues a, b of the first two measurements (and whatever probabilities our theory may ascribe to them), the third measurement (i.e. the final = second measurement A) is bound/certain to give the same outcome/eigenvalue as the first one did: $a' = a$'.

So let us say that A and B are *co-measurable* iff: measuring B does not disturb the measurement

of A in this operational sense, *and vice versa*; that is: also, measuring A does not disturb the measurement of B (i.e. in an obvious notation, we must have: $b' = b$).

Then it is easy to connect this definition of co-measurability to *commutation of operators* for finite dimensional Hilbert space, *if* we also say that measurement processes are described by the elementary Projection Postulate, i.e. the postulate that a measurement of the system, in any state (vector) ψ, for the quantity represented by A

      (i) projects the system's state instantaneously into the eigenspace of the eigenvalue obtained as the measurement's outcome; and then of course

      (ii) re-normalizes the state by dividing it by its own length.

Thus, the Projection Postulate says there is an instantaneous state transition, which is, in a notation adapted from the above discussion:

$$\psi \mapsto \frac{E_a^A \psi}{||E_a^A \psi||} \tag{32}$$

The connection, for finite dimensional Hilbert space, of the above definition of co-measurability, understood with this Projection Postulate, to commutation of operators is the readily proved equivalence:

      Theorem: Two self-adjoint operators, with pure point spectra $A = \Sigma\, a_j E_j^A$ and $B = \Sigma\, b_k E_k^B$ are co-measurable in this sense iff they commute, i.e. $[A, B] = 0$. (Of course, this latter condition is equivalent to all pairs of spectral projectors sommuting: i.e. $[E_j^A, E_k^B] = 0$, for all j, k.)

As to (iii), the interpretation/philosophy: we note:

The orthodox view in the quantum textbooks (a kind of 'precipitate' of Bohr, Dirac, Heisenberg, von Neumann; 'Copenhagen') is of course that a quantum system in state ψ only has values for those physical quantities of which ψ is an eigenstate/eigenvector (the value being the corresponding eigenvalue). (These quantities are sometimes called ψ's *eigenquantities*.)

In particular, the lack of common eigenstates for non-commuting quantities like position and momentum (suppressing here their having continuous spectra...) means that no system has a value for both position and momentum. Thus we are faced with the measurement problem, i.e. the appalling possibility that this lack of values, though it seems acceptable in the atomic realm which is after all unvisualizable etc etc, could propagate to the macro-realm—and so conflict with the supreme success of classical physics' ascription to systems of values for both position and momentum. (Cf. Schrödinger's amazing 'cat' paper of 1935!) . . .

So it is natural to propose that we supplement the orthodox quantum state: we are to ascribe values beyond the orthodox eigenvalues. The natural hope is that there are states that ascribe to every self-adjoint operator an element of its spectrum, subject to natural conditions. What natural conditions? The obvious one (sometimes called 'FUNC') is that if A is ascribed a value a, then $f(A)$ is ascribed the value $f(a)$. (After all, we often envisage measuring $f(A)$ by measuring A and applying f to the outcome.) But even if we consider only operators with pure point spectrum on a finite dimensional Hilbert space, there are problems.

That is: there are *'no-go' theorems* that such an assignment, for all the self-adjoint operators, satisfying FUNC for them all, is *impossible*. Indeed, it is provably impossible even for some judiciously— the aspiring solver of the measurement problem might say: 'unfortunately!'—chosen *finite* sets of projectors on all Hilbert spaces of (complex) dimension 3 or more (i.e. $\mathbb{C}^3, \mathbb{C}^4, \dots$).

These theorems are mostly associated with the names of Gleason (a theorem of 1957), and Kochen and Specker (a joint paper of 1967). But beware: what is usually called 'Gleason's theorem' is a *positive* result. It is a representation theorem for probability distributions on the set of all subspaces (projectors) of a Hilbert space. We will state it later, when we discuss states rigorously. The no-go

theorem is a *corollary* to this positive result.

The history has various ironies: one might say, sadnesses. For:

First: One can prove this corollary directly. And J S Bell did so in his paper, 'The problem of hidden variables in quantum mechanics', published in *Reviews of Modern Physics*, in 1966. This paper was written in 1963: *before* Bell proved the Bell inequalities—which he did, and published, in 1965. (The publication delay was due to the typescript being lost in the back of a filing cabinet at the offices of *Reviews of Modern Physics*!) And the 1966 paper was written as a *defence* of the programme of supplementing the orthodox quantum state, not as an argument against it. For Bell proved the corollary, i.e. no-go theorem, as part of an analysis urging that the assumptions of it were *eminently deniable.*

Second: Indeed, he pointed out in his 1966 paper that the assumptions are denied by the pilot-wave theory: which he considered tenable, indeed *eminently reasonable*, despite having a manifest non-locality in its guidance equation for a bipartite system. That is, in the traditional and best-developed non-relativistic version of the theory: the deterministic spatial trajectory of one part of such a system (one point-particle, according to the pilot-wave theory) is sensitive to where in space the other point-particle is: instantaneously sensitive, in the manner of action-at-a-distance, though without any fall-off with distance as one has in Newtonian gravitation.

Third: In fact, Bell's 1966 paper ends by making this point: that the pilot-wave theory is manifestly non-local. He also there stresses that, of course, Bohm in 1952 was well aware of this non-locality; and he ends by raising the question whether any supplementation of quantum theory must be in some way non-local. (A footnote added in proof then mentions his previously published 1965 paper as having answered this question, for one notion of non-locality. Not so much a case of backwards causation *a la* science fiction, but merely a disordered filing cabinet …)

Fourth: Gleason's *positive* theorem—the representation theorem for probability distributions on the set of all subspaces (projectors) of a Hilbert space—generalizes a theorem of von Neumann, in his 1932 book. Historically, this latter theorem was very influential in persuading physicists between ca 1932 and ca 1966 that this sort of supplementation of quantum theory could not work. It is standardly called von Neumann's 'no hidden variables theorem'.

Here again, Bell's 1966 paper is gold. For it articulates the Achilles heel of von Neumann's theorem. Namely: von Neumann assumes that, whatever the advocate of hidden variables envisages as a state, they should accept that states ascribe expectations to self-adjoint operators, subject to the following linearity condition. Namely: a state ascribes as the expectation $\mathrm{Exp}(A+B)$ of the sum of self-adjoint operators A, B, the sum of their individual expectations. So $\mathrm{Exp}(A + B) = \mathrm{Exp}(A) + \mathrm{Exp}(B)$.

Bell points out that any advocate of hidden variables, i.e. of postulated dispersion-free states (i.e. states ascribing values to all quantities) can, indeed should!, deny this condition. In a judo-like manoeuvre ('use your opponent's momentum to defeat them!'), Bell invokes the Bohrian idea that if A and B do not commute, and so cannot be measured together, one cannot measure $A + B$ by measuring both A and B and adding the outcomes. So although $A + B$ is self-adjoint, and can in general be measured, doing so will require some other experimental arrangement, different from both that for A and that for B. Agreed: *quantum states*—whether vector states as discussed so far, or density matrices, to be discussed later—*do* obey this linearity condition, even if A and B do not commute. But, says Bell, that is a peculiarity of the quantum formalism, and by no means a compulsory feature of states as ascriptions of expectation values. (Cf 1966, Section III, p. 449, column 1. Incidentally, Einstein pointed out the same Achilles heel to Bargmann in conversation in the 1940s …)

Beware: the entire algebraic approach to quantum theory will blithely endorse von Neumann's assumption. As we say in England: 'swallow it, hook line and sinker' . . . We will return to this irony …

Fifth: Finally, there is a further irony in relation to the first one above, about Bell's 1966

paper proving the relevant corollary of Gleason's positive theorem. This final irony is that the main drift of the Kochen-Specker paper of the following year (1967) is also to prove this corollary. This paper is rightly lauded. Its merits include: ...

(a) It connects the corollary to the quantum logic, lattice-theory, approach sketched above. The non-Boolean lattice of projectors with its delicately interlaced Boolean sub-lattices, is treated in a kindred manner to differential geometry's treatment of a manifold with its delicately interlaced charts. In particular, the operations, like taking the sum of two projectors, are *partial*. They are restricted to the summands being both in some Boolean sub-lattice. Thus the buzzword: *partial Boolean algebra*. ...

(b) It exhibits a specific finite set of projectors to which the values 1 and 0 cannot be assigned without violating the (apparently natural) assumptions like FUNC, above. It even relates these projectors to a specific quantum system (orthohelium). ...

(c) It is mathematically elegant and has engendered an enormous literature, pursuing e.g. the physical idea of 'contextuality', and mathematical fields like topos theory ...

But the Kochen-Specker paper falls in the quantum logic tradition. It does not discuss, as Bell does, that the apparently natural assumptions are in fact deniable . . .

10: *Complete (sets of commuting) operators, with continuous spectra* :—

For operators $A_1, A_2, ..., A_N$ with *continuous* spectra, the *definition* of a complete set is given by the condition in the Theorem at the end of Paragraph 8, above: viz. that every bounded operator B that commutes with all the $A_1, A_2, ..., A_N$ is a function of them: $B = f(A_1, A_2, ..., A_N)$. In fact, the Schrödinger representation of position is, by itself, a complete set. That is: if on $L^2(\mathbb{R})$, we define the self-adjoint operator Q by $Q\psi(x) = x\psi(x)$, then:

Theorem: every bounded operator B that commutes with Q is a function of Q.

We briefly connect with the Dirac notation which brings out the analogy with a complete set of commuting operators, with pure point spectra. Thus we write:

$$\langle x \,|\, \psi \rangle = \psi(x) \quad \text{and} \quad \langle x \,|\, Q\psi \rangle = x \langle x \,|\, \psi \rangle \quad \text{and} \quad \langle x \,|\, f(Q)\psi \rangle = f(x) \langle x \,|\, \psi \rangle \,. \tag{33}$$

Of course, Q has no eigenvectors. (For if $x\psi(x) = a\psi(x)$, then $\psi(x) = 0$ for $x \neq a$, and so: $||\psi||^2 = 0$.) But we use delta functions, so that writing

$$a\delta(x{-}a) = x\delta(x{-}a) \tag{34}$$

"justifies" our writing

$$Q| a \rangle = a| a \rangle \quad \text{and} \quad \langle a \,|\, \psi \rangle = \psi(a) = \int_{\mathbb{R}} \delta(x{-}a)\psi(x) \, dx \,. \tag{35}$$

And similarly

$$\psi(x) = \int_{\mathbb{R}} \psi(a)\delta(x{-}a) \, da \quad \text{``justifies''} \quad | \psi \rangle = \int_{\mathbb{R}} \langle a \,|\, \psi \rangle | a \rangle \, da \,; \tag{36}$$

so that any vector ψ can be "thought of" as a linear combination of delta-functions Thus delta-functions are like an orthonormal basis of eigenfunctions.

We similarly use delta-functions to express operators, especially their spectral resolutions. Thus for each $a \in \mathbb{R}$, let us define the 'dyad' $| a \rangle\langle a |$ by its action

$$(| a \rangle\langle a \,|\, \psi)(x) := \psi(a)\delta(x{-}a) \quad \text{i.e.} \quad (| a \rangle\langle a |) \, \psi := \langle a \,|\, \psi \rangle | a \rangle \,. \tag{37}$$

Then with E_x in the spectral family for Q, we can write for all ψ: $(E_x\psi)(y) \equiv \int_{-\infty}^{x} \psi(a)\delta(y{-}a) \, da = \int_{-\infty}^{x} (| a \rangle\langle a \,|\, \psi)(y) \, da$. So we can write

$$E_x = \int_{-\infty}^{x} | a \rangle\langle a \,|\, da \quad \text{; and similarly} \quad Q \equiv \int_{\mathbb{R}} x \, dE_x = \int_{\mathbb{R}} x \,|\, x \rangle\langle x \,|\, dx \,. \tag{38}$$