Why Hilbert Spaces?

Bharath Ron

Abstract

The current mathematical framework of quantum theory, initiated by von Neumann, uses the language of Hilbert spaces and operator algebras. When compared to classical theories this language is very abstract and unintuitive. Much of the difficulty physicists face when they work with quantum theory is this unintuitive abstract language. There have been many attempts at understanding why quantum theory requires such an abstract mathematical language, starting from von Neumann and Birkhoff's quantum logic [1]-[2]. The current state of the art is that of operational quantum physics, first clearly formulated by Ludwig, [4]-[5].

However 'operational quantum physics' is only known to a very small minority of people who work in quantum foundations. The goal of this essay is to explain the operational quantum theory to a slightly larger audience starting from structure of physical theories.

Keywords: Quantum foundations, Operational quantum theory, Why the quantum

1 | Structure of Physical Theories

2 | Operational Quantum Theory

The presentation here will be closest to Günther Ludwig school [4]-[6], with some stuff borrowed from quantum logic literature. A physical theory is in some sense to be interpreted from outside in terms of pre-theories not belonging to the theory in question itself. Usually when one tries to formulate quantum theory one starts with a pre-theory such as classical mechanics then 'quantizes' the theory. This makes the theory messy and the underlying physical ideas hidden and unclear. To minimize going to these pre-theories operational quantum theory adopts a purely instrumentalist view of physics. The construction and behavior of instruments will not be of interest to us. Any changes occurring in the instruments during 'measurements' will be accepted as objective events. In this point of view, the fundamental notions of quantum mechanics have to be defined operationally in terms of macroscopic instruments and prescriptions for their application. Quantum mechanics is then interpreted entirely in terms of such instruments and events which are the changes occuring to instruments. These instruments and events are our links to 'objective reality'.

2.1 | Observables & States

From this instrumentalist or operational point of view, the notion of 'state' can be defined in terms of the preparation procedure. A preparation procedure is characterized by the kind of system it prepares. The other important thing is the existence of a measuring instrument that is capable of undergoing changes upon their interaction. The observable change in the instrument is called a possible event or an observable effect or simply effect.

To simplify the procedure consider instruments that record 'hits'. These instruments perform simple 'yes-no' measurements. Any measurement can be interpreted as a combination of 'yes-no' measurements. These 'yes-no' instruments can be used to build any general instrument. Suppose we have such an instrument, label its registration procedure by R. If the experiment is conducted a lot of times, we get a relative frequency of occurrence of 'yes'. Here 'yes' is an observable change in the instrument. It is hence an observable effect. To every preparation procedure ρ and registration procedure R_i there exists a probability $\mu(\rho, R_i)$ of occurrence of 'yes' associated with the pair.

$$(\rho, R_i) \longrightarrow \mu(\rho|R_i).$$

The numbers $\mu(\rho|R_i)$ are called operational statistics. Two completely different preparation procedures may give the same probabilities for all experiments R. Such preparation procedures must be considered equivalent. Such preparation procedures are called operationally equivalent preparations. An equivalence class of preparations procedures yielding the same result is called an ensemble. Ensembles are precursors to the notion of states.

The basic mathematical structure of ensembles and effects can be understood using purely mathematical reasons, without introducing any new physical law. Denote the class of ensembles by S and the class of effects by E. The maps of interest to us are the following,

$$S \times E \xrightarrow{\mu} [0, 1].$$

There may be two experiments that give the same probabilities for every ensemble. Such apparatuses must be considered equivalent. They are called operationally equivalent observable effects. An effect is the equivalence class of apparatuses yielding the same result. In general, a registration procedure R for an experiment will have outcomes $\{R_i\}$. For an outcome, R_i of the registration procedure R, denote the corresponding equivalence class of measurement procedures by E_{R_i}. Each outcome R_i of the registration procedure corresponds to a functional E_{R_i} called the effect of R_i that acts on the ensemble of the system to yield the corresponding probability.

$$E_{R_i} : \rho \mapsto E_{R_i}(\rho) = \mu(\rho|R_i).$$

Maps of interest to us will be those that assign to each of its outcomes R_i its associated effect E_{R_i}. Since each ensemble fixes a probability distribution, we have,

$$\mu_\rho : R_i \mapsto \mu_\rho(R_i) = \mu(\rho|R_i).$$

Accounting to the fact that preparation procedures can be combined to produce a mixed ensemble, the set of ensembles is taken to be a convex set. Since a mixture of ensembles corresponds to a convex combination of probabilities, each functional E_{R_i} preserves the convex structure. Since two preparations giving the same result on every effect represent the same ensemble and two measurement procedures that cannot distinguish ensemble represent the same effect, ensembles and effects are mutually separating.

A generalized probabilistic theory is an association of a convex state space and effect vectors to a given system, such that the states and effects are uniquely determined by the probabilities they produce. This is known as the principle of tomography. The aim is to obtain a GPT from an operational theory. We are interested in embedding the ensembles inside the vector space of linear functionals on the effects and embed effects inside the vector space of linear functionals on the ensembles. More generally, one takes an operational theory and 'quotients' with operational equivalences to obtain a GPT.

Denote by $\mathcal{S}$ the set of maps, $f : E \longrightarrow \mathbb{R}$ such that $f(X) = \sum_i \alpha_i \mu(\rho_i|X)$ and denote by $\mathcal{E}$ the set of maps, $g : S \longrightarrow \mathbb{R}$ such that $g(\rho) = \sum_i \beta_i \mu(\rho|R_i)$ where ρ_i and R_i are ensembles and effects respectively and $\alpha_i, \beta_i \in \mathbb{R}$. Clearly $\mathcal{S}$ and $\mathcal{E}$ are real vector spaces. We can embed ensembles inside $\mathcal{S}$ with the map,

$$\rho \longmapsto \mu_\rho,$$

and similarly embed effects inside $\mathcal{E}$ with the map,

$$R_i \longmapsto E_{R_i}.$$

The bilinear map $\langle \cdot|\cdot \rangle : \mathcal{S} \times \mathcal{E} \to \mathbb{R}$ which coincides with μ is then uniquely determined. $\langle \mathcal{S}\,|\,\mathcal{E} \rangle$ becomes a dual pair. The completions of $\mathcal{S}$ and $\mathcal{E}$ will provide us the necessary mathematical structure for ensembles and effects. We will denote $\langle \cdot|\cdot \rangle$ by μ.

A registration procedure E_R is an effect valued function that assigns to each possible outcome R_i its effect E_{R_i},

$$E_R : \; R_i \longmapsto E_{R_i}.$$

It is important to find a mathematical structure that describes the registration procedure E_R beyond this basic vector space structure. The purpose of this section is to study the mathematical representatives of effects and ensembles in the quantum formalism.

To get the mathematical representatives of physical observables one has to study the logical relations of a set of propositions that are considered meaningful and empirically verifiable according to the theory that describes the physical system. The logic of a physical system will mean the algebraic structure that represents the equivalence classes corresponding to instruments. To simplify the procedure one initially reduces the instruments to simple 'yes-no' instruments.

For the development of any mathematical theory, the first step is the idealization of the registrations. The concept of observable which is one of the main physical objects of quantum theory can be obtained from a certain idealization of the registration procedure. Consider a registration procedure E_A whose outcomes $\{A_i\}$ are measured using the same equipment. The events of such a registration procedure should form a Boolean algebra. The aim is to arrive at the notion of observable from these special kinds of registration procedures. What we seek are maps from Boolean algebra to the effects. A mapping A of a Boolean algebra Σ into an ordered interval $[0, \epsilon]$ of a vector space, such that, $A(\mathbb{I}) = \epsilon$ where $\mathbb{I}$ is unit of Σ and

$$A(\sigma_1 \vee \sigma_2) = A(\sigma_1) + A(\sigma_2) \quad \text{for all} \;\; \sigma_1 \wedge \sigma_2 = 0,$$

is called an additive measure on Σ. A set $F \subset \mathcal{E}$ is called a set of coexistent effects if there exists a Boolean algebra Σ_A with an additive measure $A : \Sigma_A \to \mathcal{E}$ such that $F \subset A\Sigma_A$.

An observable is a special kind of registration procedure where the outcomes form a complete Boolean algebra. An observable is a pair (Σ_A, A), where Σ_A is a Boolean algebra and A is an additive measure,

$$A : \; \Sigma_A \to \mathcal{E}.$$

We will denote the observable by the map A. The complete Boolean algebra structure of Σ_A is the idealization of the registration procedure. Observables are effect-valued functions where outcomes have a Boolean lattice structure.

Suppose we have two observables A and B and there exists a homomorphism h of the Boolean ring Σ_A into the Boolean ring Σ_B then intuitively the observable B says about more possible events than A since the measurements of the observable A is contained in the observable B. Two observables are equivalent if the homomorphism h is an isomorphism. Two observables A and B are said to coexist if there exists an observable AB and two homomorphisms h and i such that $h : \Sigma_A \to \Sigma_{AB}$ and $i : \Sigma_B \to \Sigma_{AB}$. Denote by Ξ the effects that coexist with every other effect. Two observables A and B are mutually complementary if every coexistent effect is in Ξ. If two effects $E_{A_i} \in A$ and $E_{B_j} \in B$ are coexistent then at least one of them is in Ξ. The existence of such observables is a feature of quantum mechanics that wasn't the case in classical mechanics.

Suppose $A : \Sigma_A \to \mathcal{E}$ is an observable then a state μ_ρ gives us a map,

$$\mu_\rho^A : \Sigma_A \to [0, 1],$$

such that $\mu_\rho^A(0) = 0$, $\mu_\rho^A(E^\perp) = 1 - \mu_\rho^A(E)$ and whenever E_i are mutually orthogonal,

$$\mu_\rho^A(\vee_i E_i) = \sum_i \mu_\rho^A(E_i).$$

For all practical purposes, we will assume the measurement scale is separable. This assumption gives us all the nice properties needed to do mathematics.

It is important to note that preparation and registration procedures producing the same ensembles and effects are not always equal, in fact, the notion of equality won't even make sense. The transition from preparation and registration procedures to ensembles and effects is a transition from the real world to the abstract mathematical world. It should be noted that it doesn't make sense to 'prepare' closed systems, one has to assume such systems start off in some state a priori.

2.2 | Heisenberg's Idea & von Neumann's Implementation

The operational requirements force the space of effects and ensembles to be vector spaces. This is however insufficient to bring any predictivity to the theory. We need more structure than mere vector space structure to understand a physical theory. How the effects are interrelated is very important, and this is an experimental input.

By the end of the nineteenth century, it was clear that elementary processes obeyed some 'discontinuous' laws. That's to say, there exist observables whose collection of effects form discrete sets, and also observables whose collection of effects form a continuous set. If the collection of effects of an observable is labeled by a discrete set $\{R_i\}_{i\in\mathbb{N}}$, each ensemble ρ corresponding to some preparation procedure gives rise to a function,

$$\mathbb{N} \xrightarrow{\rho} [0, 1]$$

Since preparation procedures are independent of which observable is measured, each ensemble should also give us a map,

$$\mathcal{B}(\mathbb{R}) \xrightarrow{\rho} [0, 1]$$

corresponding to continuous observable. Corresponds to the fact that the total probability of occurance of the effects of an observable in the collection $\{R_i\}$ should be 1, the above described functions should be summable of integrable.

Before quantum theory, observables were modeled as functions on a manifold. Each state corresponds to a point in the manifold. So each state can be viewed as an evaluation map. This is however problematic if we take discrete observables into account. In such a case, certain values of the observable will have infinite multiplicity, and hence will not be summable. Hence, the pre-quantum modeling of observables cannot unify discrete and continuous physical variables. Heisenberg's radical idea was to start rethinking how we should model observables themselves. Heisenberg used linear operators as models for observables, and the values of the observables corresponded to the eigenvalues of these operators. So, the effects of the observable correspond to the projection operators to the eigenspaces. von Neumann took away from this the following idea, instead of considering the relation between discrete space and continuous space, von Neumann compared the relation between the functions on the discrete space and continuous space. The space of square integrable functions on $\mathbb{R}$ is isomorphic to the space of square summable sequences, which are functions on $\mathbb{Z}$. This isomorphism allows us to develop a unified mathematical model for observables and states. The square summable functions correspond to ensembles, and the squares correspond to probabilities. The necessary structure for the abstract mathematical framework of quantum theory is found in Hilbert spaces and operator algebras. For a careful discussion of this in terms of preparation procedure and measurements, see, §6.7.1 of [5].

2.2.1 | Lattice of Subspaces in Hilbert Space

Let $(\mathcal{H}, \langle\cdot|\cdot\rangle)$ be a complex Hilbert space. $\mathcal{P}(\mathcal{H})$ denote the set of all closed subspaces. Denote $\mathcal{H}_i \leq \mathcal{H}_j$ if and only if $\mathcal{H}_i \subseteq \mathcal{H}_j$. The relation $\leq$ is a partial ordering in $\mathcal{P}(\mathcal{H})$. Join $\vee$ of a family $\{\mathcal{H}_i\}_{i\in I}$ is the linear span of the family denoted $\vee_i \mathcal{H}_i$. Meet $\wedge$ of a family $\{\mathcal{H}_i\}_{i\in I}$ is the intersection of the family, denoted $\wedge_i \mathcal{H}_i$. The orthocomplement of $\mathcal{H}_i$ in $\mathcal{P}(\mathcal{H})$ denoted by $\mathcal{H}_i^\perp$ is the closed subspace of vectors $\varphi \in \mathcal{H}$ such that $\langle\varphi|\mathcal{H}_i\rangle = 0$. Since there is a bijection between closed subspaces of a Hilbert space and projection operators acting on the Hilbert space, the set of all projection operators on the Hilbert space inherits a lattice structure from the lattice of closed subspaces. Abusing notation, we will denote the projection operators on $\mathcal{H}$ by $\mathcal{P}(\mathcal{H})$. The orthocomplement of the projection E is the projection onto the orthogonal complement of the subspace corresponding to the projection operator E and is denoted by $E^\perp$. The lattice structure of $\mathcal{P}(\mathcal{H})$ coming from the above relations gives us the necessary structure to get the mathematical representatives of physical observables. The non-Boolean lattice $\mathcal{P}(\mathcal{H})$ of projections acts as the space of effects.

$$\mathcal{E} \equiv \mathcal{P}(\mathcal{H})$$

In the quantum logic literature the central objects that model events are the so called orthomodular lattices. Though these can be studied abstractly, we will not do so in this document as we do not find that to be particularly helpful.

For a family of projection operators to represent an observable, we should make sure that the family forms a Boolean algebra. A quantum mechanical observable is an additive measure of the form,

$$E_A : \Sigma_A \to \mathcal{P}(\mathcal{H}),$$

a projection valued function. Usually in physical experiments, the statements that can be made are of the type 'the value of the observable lies in some set ϵ_i of real numbers'. To accommodate the fact that the measurement scale is composed of real numbers, we identify Σ_A with the Borel sets of $\mathbb{R}$. It should be noted that the observables need not be real, the physics community has historically decided to use real numbers to label the outcomes of

experiments. Any other labeling should work equally well. Döring and Isham have done an interesting generalization of this scheme [**?**]. Their idea is to replace the Boolean structure in Σ_A with a more general propositional language system and question if values of the system should be more general than 'real'. They use the abstract categorical notion of a topos which can formalize the notion of 'where'. The idea is to formulate where the value of an observable lies in this abstract language. Though we find this to be a beautiful generalization for the future of quantum theory we do not think this is the part needing fixing for solving the foundational problems in quantum theory. We believe we can get a lot of work done with real measurement scales themselves.

The quantum observables are analogous to classical random variables, namely, that of a projection valued measure,

$$E_A : \mathcal{B}(\mathbb{R}) \to \mathcal{P}(\mathcal{H}).$$

This generalizes the classical case, for which mathematical representatives were the measure space $(\Omega, \Sigma(\Omega), \mu)$, where the σ-algebra, $\Sigma(\Omega)$ is a class of subsets of the set Ω which correspond to events and μ is a probability measure. A classical random variable is defined as a map $X : \Omega \to \mathbb{R}$. The map doing the work in assigning necessary probabilities is its inverse, considered as a set map,

$$X^{-1} : \mathcal{B}(\mathbb{R}) \to \Sigma(\Omega).$$

A spectral measure is a projection operator-valued function E defined on the sets of $\mathbb{R}$ such that, $E(\mathbb{R}) = I$ and $E(\coprod_i \epsilon_i) = \sum_i E(\epsilon_i)$, where ϵ_is are disjoint Borel sets of $\mathbb{R}$. The spectral theorem says that every self-adjoint operator A corresponds to a spectral measure E_A such that,

$$A = \int \lambda \, dE_A(\lambda),$$

and conversely, every spectral measure corresponds to a self-adjoint operator. In the finite-dimensional case this reduces to $A = \sum_i \lambda_i E_i$ where E_is are projections onto eigenspaces of λ_is. Observables in quantum theories are represented by self-adjoint operators on some complex Hilbert space and the orthogonal projections of the self-adjoint operator correspond to the events. The values of the observable are the spectrum of the operator. The characteristic feature of quantum theory is that the space of effects is a non-commutative entity.

The mathematical representatives of the physical states for the quantum case are the maps, $\omega : \mathcal{P}(\mathcal{H}) \to [0, 1]$, such that $\omega(0) = 0$, $\omega(E^\perp) = 1 - \omega(E)$ and $\omega(\vee_i E_i) = \sum_i \omega(E_i)$ for mutually orthogonal E_i. For an observable with the associated self-adjoint operator A, the map

$$\mu^A = \omega \circ E_A : \Sigma_A \to [0, 1],$$

determines a classical probability measure. The existence and classification of such non-commutative probability measures on Hilbert spaces is given by the Gleason's theorem.

Theorem 2.1. (Gleason) *If the complex separable Hilbert spaces $\mathcal{H}$ of dimension greater than 2, then every ω is of the form*

$$\omega(E) = Tr(\rho E).$$

where ρ is a positive semidefinite self-adjoint operator of unit trace, also called density matrices. Conversely, every density matrix determines a state as defined in the above formula.

The proof of Gleason's theorem is unimportant. The proof requires patience to read through and high amount of problem solving skill, and insight to come up with. But we do not need

those to understand what it is saying. We probably will never need the methods used in the proof of Gleason's theorem for understanding quantum mechanics.

More general quantum experiments correspond to positive operator-valued measures. The effects E are given by positive operators, $O \leq E \leq I$ as probabilities are positive quantities. Since these should sum to 1 for an experiment, it will be a resolution of identity $\sum_i E_{A_i} = I$, where E_{A_i}s are effects. The resolution of identity $E_A : A_i \to E_{A_i}$ is called positive operator-valued measure (POVM). General quantum mechanical experiments are represented by pairs (ρ, E_A). For Gleason's theorem in this setting see [**?**].

We call ω a Gleason measure. Every state corresponds to a positive semidefinite self-adjoint operator of unit trace. We denote the set of all states on the Hilbert space $\mathcal{H}$ by $\mathcal{S}(\mathcal{H})$. The mathematical representatives of ensembles are states.

$$\mathcal{S} \equiv \mathcal{S}(\mathcal{H})$$

This is a convex set. The extreme points of this convex set are called pure states, pure states are of the form, $\rho = \rho^2$ and corresponds to some unit vector $|\varphi\rangle$ in the Hilbert space $\mathcal{H}$ and ρ is the projection onto the subspace generated by $|\varphi\rangle$. Such a state is denoted by $|\varphi\rangle\langle\varphi|$. Note that the existence requires Krein-Milman theorem.

For an observable with an associated self-adjoint operator A the probability that the observable takes a value lying in the interval ϵ is given by,

$$\mu_\rho^A(\epsilon) = Tr(\rho E_A(\epsilon)).$$

The expectation value of the observable will be,

$$\langle A \rangle = \int \lambda \, d\mu_\rho^A(\lambda) = Tr(\rho A).$$

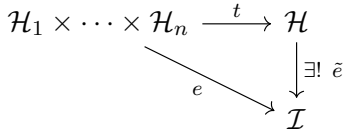
In quantum theory, all the Hilbert spaces are assumed to be separable, complex.

Tensor Product

Given a finite number of Hilbert spaces $\mathcal{H}_i$, for n quantum systems, the problem is to describe the Hilbert space appropriate to the 'product' system. Let $\mathcal{I}$ denote a possible solution to this problem: that is the states of $\mathcal{I}$ are supposed to be states for the product system. Then, at the very least, some of the preparation procedures for the product system should be obtainable by arranging in some manner the preparation procedures on the individual systems. We should be able to construct a certain function,

$$e : \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \to \mathcal{I}.$$

The interpretation of e is that, it introduces a component from each individual system into the product system. Accounting to the superpositions, the product system should inherit the structure from the components. The map e must be linear for each component. The universal solution $\mathcal{H}$ to this problem is the algebraic tensor product. It is the vector space $\mathcal{H}$ together with an n-linear map t such that, for any n-linear map $e : \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \to \mathcal{I}$, there exists a unique linear map $\tilde{e} : \mathcal{H} \to \mathcal{I}$,



This vector space inherits a canonical inner product from the component Hilbert spaces. The completion of $\mathcal{H} = \otimes_{i \in I} \mathcal{H}_i$ under the canonical inner product will serve as the Hilbert spaces for product systems.

It is important to note that when we are given a closed system, the notion of preparation of state doesn't make sense. So in such cases, we are stuck with states given by nature.

References

[1] J von Neumann, Mathematical Foundations of Quantum Mechanics, New edition, Princeton University Press, 2018

[2] M Redei, Quantum Logic in Algebraic Approach. Springer, 1998

[3] G Ludwig, G Thurler, A New Foundation of Physical Theories. Springer-Verlag, 2006

[4] G Ludwig, Foundations of Quantum Mechanics I. Springer-Verlag, 1983

[5] G Ludwig, Foundations of Quantum Mechanics II. Springer-Verlag, 1983

[6] K Kraus, States, Effects, and Operations. Lectures Notes in Physics, Springer-Verlag, 1983