Part V

Observables and States

In this part we will go from Heisenberg's foundational ideas to the mathematics of quantum mechanics. We start with discussing Heisenberg's idea for treating observables as self-adjoint operators made precise by von Neumann. This naturally leads to identifying effects of quantum theory with projection operators on some Hilbert space. Then we will follow Birkhoff-von Neumann's quantum logic path and discuss Gleason's theorem.

Heisenberg's idea

By the end of the nineteenth century, it was clear that elementary processes obeyed some 'discontinuous' laws. There existed no mathematical model of quantum theory that would provide a unified structure for treating discrete quantities and continuous quantities. Heisenberg's solution to this problem was to use linear operators as a starting point for the description of physical quantities or observables. What von Neumann took away from Heisenberg's idea was that the mathematical objects needed for the description of observables is found in Hilbert spaces.

The pre-quantum modelling of observables was as follows, one started with an indexing set, and a state space. The observables were maps from the state space to the indexing set. If the indexing set is $\mathcal{I}$ and the state space is $\mathcal{X}$, an observable is a map,

$$A : \mathcal{X} \to \mathcal{I}.$$

By the end of twentieth century it had become apparent that there existed both continuous physical quantities and discrete physical quantities. So, the indexing set could be a discrete space such as $\mathbb{Z}$ or a continuous space such as $\mathbb{R}$. The existence of continuous physical quantities, i.e., a physical quantity whose values can be any real number, tells us that the cardinality of $\mathcal{X}$ should be atleast uncountably infinite. However now, if there exists a discrete quantity, then some of the values should have infinite multiplicity in the state space i.e., infinitely many points in $\mathcal{X}$ are mapped to the same thing in $\mathcal{I}$. This model forbids the coexistence of continuous quantities with discrete quantities.

Heisenberg's idea was to start with 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. 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. This insight initiated the mathematical foundations of quantum theory and the theory of Hilbert spaces and operator algebras. Our goal now is to relate the notion of effects and ensembles to the notion of observables and states in quantum theory.

1 | Observables and States

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 should act as the space of effects.

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

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

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

a projection valued function. 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 [9]. Their idea seems to be 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'. Though we find this to be a beautiful generalization for the future of quantum theory we don't 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 $E(\sqcup_i \epsilon_i) = \sum_i E(\epsilon_i)$, 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.

1.1 | Gleason's Theorem

The classification of such non-commutative probability measures on Hilbert spaces is given by the Gleason's theorem and could be treated as the strong 'Born rule'.

Theorem 1.1. (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 or density matrix. Conversely, every density matrix determines a state as defined in the above formula.

Sketch of Proof

Gleason's proof of the theorem is quite complicated. He starts by defining what he calls frame functions of weight W on separable Hilbert space $\mathcal{H}$. A frame function f is a real valued functions on unit sphere of $\mathcal{H}$ such that for any orthonormal basis, $\{|\varkappa_i\rangle\}_{i\in\mathbb{N}}$, $\sum_i f(|\varkappa_i\rangle) = W$. A frame function is regular if there exists a self-adjoint operator ρ such that for every unit vector $|\varkappa\rangle \in \mathcal{H}$,

$$f(|\varkappa\rangle) = \langle \varkappa | \rho \varkappa \rangle$$

Gleason proves that every frame function on two dimensional Hilbert spaces is regular. For Hilbert spaces of dimension greater than three the result holds for every two dimensional subspaces. Then he proves the continuity of frame functions. Every non-negative frame function on a Hilbert space of dimension greater than three is regular. Much of the hard work lies in this part. I will cheat and skip this hard part. A brave reader can go read Gleason's original paper [5] or H Granström's master's thesis [6] on Gleason's theorem.

Suppose $\omega : \mathcal{P}(\mathcal{H}) \to [0,1]$ be a function as described above. Let E_φ be the projection onto the subspace spanned by the unit vector φ. $f(\varphi) = \omega(E_\varphi)$ defines a non-negative frame function. By regularity there exists a self-adjoint operator ρ such that,

$$f(\varphi) = \langle \varphi | \rho \varphi \rangle.$$

Since this holds for all unit vectors, ρ is positive semi-definite. Denote by $E_\mathcal{H}$ the projection onto the whole Hlibert space i.e., the identity operator. Given an orthonormal basis $\{|\varphi_i\rangle\}_{i\in\mathbb{N}}$ of $\mathcal{H}$ we have,

$$\omega(E_\mathcal{H}) = \sum_i \omega(E_{|\varphi_i\rangle}) = \sum_i \langle \varphi_i | \rho \varphi_i \rangle = Tr(\rho).$$

For any subspace $\mathcal{K} \subset \mathcal{H}$ denote by $E_\mathcal{K}$ the corresponding projection operator. Take an orthonormal basis $\{\varkappa_i\}_{i\in I}$ for $\mathcal{K}$ and extend it to $\mathcal{H}$. Then we can write,

$$\omega(E_\mathcal{K}) = \sum_{i\in I} \omega(\varkappa_i) = \sum_{i\in I} \langle E_\mathcal{K} \varkappa_i | \rho \varkappa_i \rangle = Tr(\rho E_\mathcal{K}).$$

So we have for all projection operators $E \in \mathcal{P}(\mathcal{H})$, we have,

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

□

The proof of Gleason's theorem is unimportant. The proof requires patience to read through and high amount of problem solving skill, intelligence and insight to come up with. But we don't 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 [7].

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.

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

The extreme points of this convex set are called pure states, pure states are of the form, $\rho = \rho^2$ and corresponds to some 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 $|\varphi\rangle\langle\varphi|$.

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,

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

All the Hilbert spaces will be assumed to be separable, complex.

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,

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

The interpretation of f 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 f must be linear for each component. The universal solution $\mathcal{H}$ to this problem is the algebraic tensor product. It's the vector space $\mathcal{H}$ together with an n-linear map t such that, for any n-linear map $f : \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \to \mathcal{I}$, there exists a unique linear map $\tilde{f} : \mathcal{H} \to \mathcal{I}$,

$$\begin{tikzcd} \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \arrow[r, "t"] \arrow[dr, "f"'] & \mathcal{H} \arrow[d, "\exists!\ \tilde{f}"] \\ & \mathcal{I} \end{tikzcd}$$

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's 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.

2 | Algebraic Approach

As we saw, the observables in standard quantum mechanics correspond to self-adjoint operators on some Hilbert space. However the structure needed for the description of observables can be abstracted out.

C^*-algebra and the self-adjoint operators on it as the starting point for the description of observables. Given a C*-algebra and a self-adjoint operator $A \in \mathcal{A}$, by Spectral theorem of self-adjoint operators on a C*-algebras, we have,

$$A = \int \lambda \, d\, E_A(\lambda).$$

The spectral measure E_A carries the same physical interpretation as discussed before. The mathematical representatives of the physical states for the quantum case are the maps, $\omega : \mathcal{P}(\mathcal{A}) \to [0,1]$, such that,

$$\omega(0) = 0,\ \ \omega(E^{\perp}) = 1 - \omega(E),\ \ \ \omega(\vee_i E_i) = \textstyle\sum_i \omega(E_i),$$

for mutually orthogonal E_i. Here $\mathcal{P}(\mathcal{A})$ are projection operators on $\mathcal{A}$. For an observable with the associated self-adjoint operator A, the map

$$\mu^A = \omega \circ E_A :\ \ \mathcal{B}(\mathbb{R}) \to [0,1],$$

determines a classical probability measure.

A linear functional ω over the algebra $\mathcal{A}$ is called positive if,

$$\omega(A^*A) \geq 0$$

for all $A \in \mathcal{A}$. A positive linear functional over $\mathcal{A}$ with $\|\omega\| = 1$ is called a state. The state is called faithful if $\omega(A^*A) = 0$ implies $A = 0$. This is a generalisation of states in standard quantum mechanics. The set of all continuous linear functionals over $\mathcal{A}$ is denoted by $\mathcal{A}^{\#}$. Together with the sup norm, $\mathcal{A}^{\#}$ is a Banach space.

For a physically meaningful topology on operator algebras, we expect two observables to be close to each other if the expectation values are close. Trace class operators on a Hilbert space $\mathcal{H}$ consists of operators ρ such that $\mathrm{Tr}(\rho) < \infty$. The expectations give us seminorms on the space of operators.

$$p_\rho(A) = \mathrm{Tr}(\rho A).$$

The topology determined by them is called ultraweak topology. Operator algebras closed under ultraweak topology are called von Neumann algebras. These will be the objects of interest in algebraic quantum theory.

Further Reading

The first chapter of von Neumann's book [1] provides detailed motivation for the use of Hilbert spaces in quantum theory. Redei's book on quantum logic, [4], provides a very good introduction to quantum logic with focus on algberaic approach.

References

[1] J von Neumann, Mathematical Foundations of Quantum Theory. Princeton University Press, 1955

[2] J M Jauch, Foundations of Quantum Mechanics. Addison-Wesley Publishing Company, 1968

[3] V Moretti, Fundamental Mathematical Structures of Quantum Theory. Springer, 2018

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

[5] A M Gleason, Measures on the closed subspaces of a Hilbert space. Journal of Mathematics and Mechanics, vol. 6, no. 6, pp. 885–893, 1957

[6] H Granström, Gleason's theorem. Master's thesis, Stockholm University, `http://www.kiko.fysik.su.se/en/thesis/helena-master.pdf`, 2006

[7] P Busch, Quantum states and generalized observables: A simple proof of Gleason's theorem. Physical Review Letters 91(12):120403. `http://arxiv.org/abs/quant-ph/9909073`, 2003

[8] V S Varadarajan, Geometry of Quantum Theory. Van Nostrand, New York, 1970

[9] A Döring, C J Isham, What is a thing?. `https://arxiv.org/abs/0803.0417`, 2008