(Yet Another) Reconstruction of Quantum Theory

Bharath Ron

The physical systems of interest to us are those whose observed phenomena can be described by the notion of an 'observable'. The starting idea here is that the world can be described by sentences of the form 'the observable R has the value R_i'. The next step is to articulate this idea in the mathematical language. This means that we need to construct an appropriately predictive mathematical theory where sentences like the one above can be represented. There have been mainly two successful approaches for this.

In classical mechanics, the approach is to think of observables as continuous maps from a manifold Ω, called the phase space, to the set of real numbers

$$A : \Omega \longrightarrow \mathbb{R}.$$

As we will discuss later, this model of an observable fails when the application domain is expanded to simultaneously include both discrete and continuous case. Although quantum theory has the same idea, it is formulated in a different way that makes it compatible with the extended application domain. To understand how the quantum approach is different from the classical approach we will develop the notion of observable from an abstract operational perspective.

1 | Generalised Probabilistic Theories

The discussion here will be closest to the Ludwig school [1]-[4], with some stuff borrowed from the quantum logic literature [10]-[9] and algebraic quantum theory. This can be thought of as a modern formulation of Heisenberg's original idea. Hans Primas [5] characterises approaches to the formulation of quantum theory into three main categories, the quantum logic approach, algebraic approach, and the convex state-space approach.

The convex state-space approach was developed by Ludwig and his peers. This approach starts with measuring instruments and preparation instruments and is the most general approach among the three. A physical theory is in some sense 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 and then 'quantizes' the theory. This makes the theory messy and the underlying physical ideas hidden and unclear. In Ludwig's approach the pre-theory is the theory of preparation and measuring instruments. The assumed structure in this instrumentalist view of physics can be motivated with simple thought experiments making it foundationally sound, and hence we will adopt this approach to the formulation of quantum theory. 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 outcomes. In this point of view, the fundamental notions of quantum theory 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 outcomes which are the changes occurring to instruments. These instruments and outcomes are our links to 'objective reality'.

1.1 | Effects & Ensembles

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 measuring instrument that is capable of undergoing changes when an experiment is performed for evaluating a collection of possible outcomes. The possible results of such an experiment are called outcomes of the experiment. The observable change in the instrument is called an 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 a measuring instrument, label its measuring instrument by R_i. 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 measuring instrument R_i there exists a probability $\mu(\rho|R_i)$ of the 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 operational statistics, that is, they may give the same probabilities for every experiments. Such preparation procedures must be considered equivalent operationally. An equivalence class of preparation procedures yielding the same operational statistics for experiments is called an ensemble. Similarly, there may be two measuring instruments that have the same chances of undergoing a change for similarly prepared systems. Such measuring instruments must be operationally considered to be equivalent. An equivalence class of change for measuring instruments is called an effect. An effect is the equivalence class of all instruments that undergo a change for the same possible outcome. By considering equivalence classes we have obtained the structure of sets. Without introducing any new physical law we have obtained the basic mathematical structure for modelling preparation procedures and measuring instruments in terms of the ensembles and effects they describe. Since we are considering equivalence classes, the construction and behavior of the instruments are irrelevant.

Denote the set of ensembles by $\mathcal{S}$ and the set of effects by $\mathcal{E}$. The maps of interest to us are the following,

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

For a possible outcome R_i, we will denote the corresponding effect by E_{R_i}. Each effect acts on the ensembles of the system, and each ensemble acts on the effects of the system to yield the corresponding operational statistic

$$E_{R_i} : \rho \mapsto \mu(\rho|R_i), \quad \mu_\rho : R_i \mapsto \mu(\rho|R_i).$$

$\mathcal{E}$ and $\mathcal{S}$ together with the pairing $\mu : \mathcal{E} \times \mathcal{S} \to [0, 1]$ is called an operational theory.

Since any two preparations giving the same result on every effect represent the same ensemble and two measurement procedures that cannot distinguish ensembles represent the same effect, ensembles and effects are mutually separating with respect to the pairing $\mu(\cdot|\cdot)$. Hence studying the mathematical structures of one of these sets also already tells us a lot about the mathematical structure of the other. It is hence sufficient to focus our attention on understanding the structure of either the space of effects or the space of ensembles. At this stage note that ensembles are not the most primitive concepts in this framework. What measuring instruments an experimenter is allowed to apply to a preparation procedure already constrains the collection of allowed preparation procedures. So the space of effects which model the measuring instruments is a more primitive concept than the space of ensembles. Although this might seem like a trivial reason, the two paths diverge significantly.

On a historic note, this is where we see theoretical physics branch off into two different groups. The first group does in the direction of understanding the structure of the space of ensembles, following Schrödinger, Dirac, Feynman, and others; since state spaces in classical theories are also closely related to topology and geometry this group has been able to make a lot of progress in developing the standard approach to quantum field theories behind the celebrated standard model by utilising the geometric intuition behind classical theories and by understanding and generalising gauge theoretic ideas of Maxwell and others. The other group branched out in the direction of understanding the structure of the space of effects, following Heisenberg, Jordon, Born, von Neumann, Haag, Kastler, Araki, Borchers and others. This group has been able to make much less progress comparitively, possibly because the space of effects is very strange and unfamiliar compared to the more familiar and intuitive classical theories. This lack of pre-theories such as classical theory has left this group stranded in the middle of nowhere, and the ones who did pursue this new path are forced to build up starting from almost nothing.

Most physicists avoid the second path possibly due to unfamiliarity. We will however follow the second path due to its stronger foundations and embrace the risk of losing the centuries old intuition, for classical theories.[1]

1.2 | Algebraic Quantum Theory

In order to be able to work with effects and ensembles we now embed them inside sets with mathematical structure that respects the expected operational relations. We now describe how the collection of effects and the collection of ensembles inherit mathematical structures.

1.2.1 | Locally Convex Embeddings

In our case the operational requirement is that we should be able to make sense of taking mixtures of effects and ensembles. Suppose we can prepare systems in different ensembles by varying settings of a preparation procedures, then by adjusting the settings we should be

[1] The need for quantum theory already indicates that atleast some of the ideas behind classical theories are not compatible with observed phenomena. Keeping that in mind we feel it is justified to not blindly trust the intuition we have developed for classical theories.

able to prepare systems in mixtures of ensembles. We expect the space of ensembles to have the structures to describe mixtures

As discussed in [6], the notion of mixing corresponds mathematically to the notion of convex combination, and hence we expect the set $\mathcal{S}$ to have the structures for making sense of convexity and we expect each functional E_{R_i} preserves the convex structure since we expect the operational statistics to preserve this convexity. To make sense of taking convex linear combinations we need to be able make sense of linearity combinations. Hence we must embed the set of effects and the set of ensembles in a vector space with respect to a field that is atleast as large as the real numbers. It is convenient to choose the field to be the field of complex numbers since the field of complex numbers is algebraically closed and allows us sufficient use of analytic tools.[2] It is important to note that preparation and measuring instruments producing the same ensembles and effects are not equal, in fact, the notion of equality will not even make sense. The transition from preparation and measuring instruments to ensembles and effects is a transition from the real world to the abstract mathematical world. It should also be noted that it does not make sense to 'prepare' closed systems, one has to assume such systems start off in some state a priori.

A generalized probabilistic theory (GPT) is the embedding of $\mathcal{E}$ and $\mathcal{S}$ inside vector spaces, such that the ensembles and effects are uniquely determined by the operational statistics they produce. This uniqueness of operational statistics is called the principle of tomography in quantum foundations and quantum information literature. To construct a generalised probabilistic theory for $\mathcal{E}$ and $\mathcal{S}$ we start by thinking of effects and ensembles as linear functionals on a suitable space. Denote by $\mathcal{A}_*$ the set of maps

$$\omega(R) = \textstyle\sum_{i\in I} \alpha_i \mu(\rho_i|R), \quad \forall R \in \mathcal{E},$$

where ρ_i are ensembles and α_i are scalars with I a finite set. Since any linear combination of such maps also belongs to $\mathcal{A}_*$, it follows that $\mathcal{A}_*$ is a complex vector space. Denote by $\mathcal{A}$ the set of maps,

$$F(\rho) = \textstyle\sum_{j\in J} \beta_j \mu(\rho|R_j), \quad \forall \rho \in \mathcal{S},$$

where R_j are effects and β_j are scalars with J a finite set. It follows that $\mathcal{A}$ is also a complex vector spaces. We embed ensembles inside $\mathcal{A}_*$ with the map,

$$\mu_{(\cdot)} : \mathcal{S} \rightarrowtail \mathcal{A}_*,$$

which sends ρ to the functional $\mu_\rho : R_i \mapsto \mu(R_i|\rho)$. We embed effects in $\mathcal{A}$ with the map,

$$E_{(\cdot)} : \mathcal{E} \rightarrowtail \mathcal{A},$$

which sends an effect R_i to the functional $E_{R_i} : \rho \mapsto \mu(R_i|\rho)$. Abusing notation we will denote the images of $\mathcal{E}$ and $\mathcal{S}$ by the same and they are convex subsets of $\mathcal{A}$ and $\mathcal{A}_*$ respectively. The elements of the convex subsets $\mathcal{E}$ of $\mathcal{A}$ and $\mathcal{S}$ of $\mathcal{A}_*$ are called effects and states respectively.

The vector space structure of $\mathcal{A}_*$ and $\mathcal{A}$ allows us to develop algebraic tools for studying effects and ensembles. So we embedded an operational theory by viewing the effects and ensembles as linear functionals on each other and took 'quotients' with operational equivalences. Since we expect the ensembles and effects to be determined by the probabilities they produce we must also expect the vector spaces $\mathcal{A}$ and $\mathcal{A}_*$ to inherit a relation between each other from the pairing $\mu(\cdot|\cdot)$ of $\mathcal{E}$ and $\mathcal{S}$. So we must have a pairing

$$\langle\cdot|\cdot\rangle : \mathcal{A}_* \times \mathcal{A} \to \mathbb{C}$$

[2] Quantum mechanics can be done in more complicated fields such as the field of quarternions. We will however avoid discussions on such choice as it will digress too much from the main goal of this thesis.

that coincides with μ for effects and ensembles, that is, $\langle E_{R_i}|\mu_\rho\rangle = \mu(R_i|\rho)$. We require the scaling of both effects and ensembles by a unit length complex number to give the same pairing as the unscaled pairing. This requirement along with real linearity in each argument forces the pairing $\langle\cdot|\cdot\rangle$ to be sesquilinear.

We assume the existence of a unique element 1 which corresponds to the trivial instrument which is always true for any preparation procedure the uniqueness is assumed because if there are two instruments which are always 'yes' they are operationally indistinguishable and are hence operationally equivalent. Operationally, the element $1 \in \mathcal{A}$ can be thought of as the 'existence' element for the system, since it corresponds to the effect which is always true, which we can assume its existence as soon as we assume the system itself exists. Similarly we will assume the existence of the unique 0 element which outputs 'no' for every preparation procedure, and can be thought of as a faulty instrument which never undergoes any changes.

1.2.2 | Algebraic Structures

The operational requirements force the space of effects and ensembles to be vector spaces. The vector space structure is however insufficient for doing analysis with effects and ensembles which would allow the theory to be predictive. We now discuss how to operationally introduce additional mathematical structures on the space of effects and ensembles.

The Banach Space Structure

Suppose we have two measuring instruments for an outcome where one is more accurate than the other. In such a case, the less accurate measuring instrument will readily undergo a change compared to the more accurate instrument. If the instrument is too accurate, it becomes difficult to observe any changes occuring in the measuring instrument. In this sense, the more accurate instrument should be 'closer' to the instrument which never undergoes any changes. So, we must expect $\mathcal{A}$ to have some structure that quantifies how readily an instrument undergoes a change. The instrument which always undergoes a change should be the easiest to notice changes and the instrument which never undergoes any changes should be the hardest. Suppose we have two measuring instruments which undergo changes for outcomes R and S, and let E_R and E_S be the corresponding effects. For combined instruments, we expect the accuracy change for the combined instruments to depend on the accuracy of the component instruments. Hence the changes in accuracy must be expected to vary proportionally to the changes in accuracy of its components. This gives rise to continuity requirements on the space of effects.

The operational interpretation expectations for accuracy of measuring instruments gives rise to a norm like function on the space of effects. We can describe this function more conveniently as a supremum. Since accuracy represents the maximal likelihood of undergoing a change it must correspond to the supremum of undergoing a change over all preparation procedures. This gives us a definition in terms of supremum over preparation procedures, and can be defined on all of $\mathcal{A}$,

$$\big\|E\big\|_{\mathcal{A}} \coloneqq \sup_{\omega\in\mathcal{S}} \big|\langle E|\omega\rangle\big|, \;\; \forall E \in \mathcal{A}$$

where the supremum is taken over all preparation procedures and corresponds exactly to the convex embedded subset $\mathcal{S}$ of $\mathcal{A}_*$. It immediately follows that $\|\cdot\|_{\mathcal{A}}$ satisfies the triangle

inequality, and scaling condition,

$$\|E+F\|_{\mathcal{A}} \leq \|E\|_{\mathcal{A}} + \|F\|_{\mathcal{A}}.$$
$$\big\|\lambda E\big\|_{\mathcal{A}} = |\lambda|\big\|E\big\|_{\mathcal{A}},$$

for all $E, F \in \mathcal{A}$ and $\lambda \in \mathbb{C}$. By definition of the norm we have $\|1\|_{\mathcal{A}} = 1$. By the assumption that there will always exist some preparation procedure with non-trivial operational statistics, we deduce that $\|\cdot\|_{\mathcal{A}}$ is non-trivial, that is, $\|E\|_{\mathcal{A}} = 0$ if and only if $E \equiv 0$. Hence $\|\cdot\|_{\mathcal{A}} : \mathcal{A} \to \mathbb{R}$ defines a norm on $\mathcal{A}$.

For the operational interpretation of the triangle inequality consider two instruments. If the instruments can undergo changes under some similar situations, it means that we have some redundancies, that is, they can measure some common possible outcomes. If the accuracy of either of the instruments is changed, the accuracy of the combined instrument should also change, and this change should be proportional to the change in accuracy of the instrument. If changes in accuracy are modelled using addition like operation, we must have, $\|E_R + E_S\|_{\mathcal{A}} \leq \|E_R\|_{\mathcal{A}} + \|E_S\|_{\mathcal{A}}$. The notion of accuracy of instruments naturally gives rise to pre-triangle inequality like properties.

Similarly, $\mathcal{A}_*$ also inherits a norm from $\mathcal{A}$. We can have two preparation procedures, where the preparations settings are slightly varied. We can say two ensembles are close to each other if the operational statistics are close to each other for all effects. Hence how similar two preparation procedures are can also be described as a supremum, and this can be described on all of $\mathcal{A}_*$.

$$\big\|\sigma\big\|_{\mathcal{A}_*} \coloneqq \sup_{E\in\mathcal{E}} \big|\langle E|\sigma\rangle\big|, \quad \sigma \in \mathcal{A}_*$$

where the supremum is taken over effects and corresponds to the convex embedded subset $\mathcal{E}$ of $\mathcal{A}$. We can hence require both $\mathcal{A}$ and $\mathcal{A}_*$ to be a complex, normed spaces. Under the idealization that there exist instruments of every level of accuracy, $\mathcal{E}$ must be closed with respect to the norm $\|\cdot\|_{\mathcal{A}}$. Since $\mathcal{E}$ generates $\mathcal{A}$ we also assume $\mathcal{A}$ to be closed with respect to taking limits with respect to the norm $\|\cdot\|_{\mathcal{A}}$. Hence we may assume that the spaces $\mathcal{A}_*$ and $\mathcal{A}$ are Banach spaces. This allows us to talk about taking limits and allows us to do mathematical analysis with ensembles and effects. In particular we now have access to Banach space tools such as the Hahn-Banach theorem, the Krein-Milman theorem, the Banach-Alaoglu, and others.[3]

By the boundedness of operational statistics we note that every effect gives rise to a $\|\cdot\|_{\mathcal{A}_*}$-bounded linear functional on $\mathcal{A}_*$. Similarly, every effect gives rise to a bounded linear functional on $(\mathcal{A}_*, \|\cdot\|_{\mathcal{A}_*})$. The spaces $\mathcal{A}$ and $\mathcal{A}_*$ are generated by $\mathcal{E}$ and $\mathcal{S}$ respectively. Since we assume $\mathcal{S}$ and $\mathcal{E}$ separate each other with respect to the pairing $\mu(\cdot|\cdot)$, and since

$$\langle\cdot|\cdot\rangle : \mathcal{A} \times \mathcal{A}_* \to \mathbb{C},$$

extends $\mu(\cdot|\cdot)$ while respecting the complex-linearity properties, it follows that $\langle\mathcal{A}|\mathcal{A}_*\rangle$ is a dual pair. Since $\mathcal{A}$ and $\mathcal{A}_*$ separate each other, any bounded linear functional on $\mathcal{A}_*$ can be obtained as a limit of linear combinations of functionals corresponding to effects.

[3]Operationally, the Hahn-Banach theorem can be described as follows: For every non-trivial measuring instrument there must exist an ensemble with non-trivial operational statistics, because otherwise such a measuring instrument would be equivalent to zero. Let us choose (by axiom of choice) such a preparation procedure. We can now use the measuring instrument as a 'filter' to create a new preparation procedure which only selects the system if it makes the instrument undergo a change. This would give for each effect E a preparation procedure μ_E such that $\langle E|\mu_E\rangle = 1$.

By completeness assumptions on $\mathcal{A}$ we may assume that each bounded linear functional corresponds to an element of $\mathcal{A}$. Hence we must have

$$\mathcal{A} \cong (\mathcal{A}_*)^*$$

as Banach spaces, where $(\mathcal{A}_*)^*$ is the space of $\|\cdot\|_{\mathcal{A}_*}$-bounded linear functionals on $\mathcal{A}_*$.

The Banach Algebra Structure

Any measuring instrument can be applied to systems once it has been prepared following a preparation procedure. We can however also view the combined setup consisting of the measurement instrument applied to the preparation procedure as a preparation procedure itself. We can then perform measurements on this composite preparation procedure.

For any effects E and F in $\mathcal{A}$ we denote the above described concatenation of measurements by EF. This gives rise to a product structure on $\mathcal{A}$. We expect varying accuracy of either of the measuring instruments corresponding to E or F to vary the accuracy of EF accordingly. Hence we expect the algebraic structure on $\mathcal{A}$ to respect the topological structure on $\mathcal{A}$ coming from the norm $\|\cdot\|_{\mathcal{A}}$. We now study the topological properties of this product structure, following [14].

We may perform measurement before or after performing some other measurement. This composite measurements can itself be thought of as a measuring instrument. Hence we can view application of a measuring instrument before or after both as operators on $\mathcal{A}$. $\mathcal{A}$ comes equipped with a natural left-action of effects on it given by

$$\mathcal{L}_E F = EF.$$

This composite element corresponds to a well-defined measuring instrument, and the accuracy of the combined instrument is given by $\|EF\|_{\mathcal{A}} = \sup_{\omega\in\mathcal{S}} |\langle EF|\omega\rangle|$. Since we are allowed to apply E after applying any other measuring instrument the accuracy of left-composition by E is,

$$\big\|\mathcal{L}_E\big\|_{\mathcal{B}(\mathcal{A})} = \sup_{\substack{F\in\mathcal{E}\\ \omega\in\mathcal{S}}} \big|\langle EF|\omega\rangle\big| = \sup_{F\in\mathcal{E}} \Big\{ \sup_{\omega\in\mathcal{S}} \big|\langle EF|\omega\rangle\big| \Big\} = \sup_{F\in\mathcal{E}} \big\|EF\big\|_{\mathcal{A}}.$$

Similarly we have a natural right-action given by

$$\mathcal{R}_F E = EF.$$

The norm of $\mathcal{R}_F$ is given by,

$$\big\|\mathcal{R}_F\big\|_{\mathcal{B}(\mathcal{A})} = \sup_{E\in\mathcal{E}} \big\|EF\big\|_{\mathcal{A}}.$$

So, both $\mathcal{L}_E$ and $\mathcal{R}_F$ belong to the space of bounded operators $\mathcal{B}(\mathcal{A})$ on $\mathcal{A}$.

Since we expect the accuracy of the composite to change proportionally to the change in accuracy of the composing instrument, we expect both $\mathcal{L}_E$ and $\mathcal{R}_F$ to be continuous. This must relate the norm $\|\cdot\|_{\mathcal{B}(\mathcal{A})}$ and the norm $\|\cdot\|_{\mathcal{A}}$. To relate the two norms we start with the following relation; $\mathcal{R}_F E = EF = \mathcal{L}_E F$. By continuity of $\mathcal{R}_F$ we have, $\|\mathcal{R}_F E\|_{\mathcal{A}} = \|\mathcal{L}_E F\|_{\mathcal{A}} \leq \|\mathcal{L}_E\|_{\mathcal{B}(\mathcal{A})} \|F\|_{\mathcal{A}}$. Hence $\{\mathcal{R}_F E\}_{F\in\mathcal{A}}$ is a bounded set for every E.

By the Banach-Steinhaus theorem or the uniform boundedness principle, pointwise boundedness of the set $\{\mathcal{R}_F E\}_{F\in\mathcal{A}}$ implies that $\{\mathcal{R}_F\}_{F\in\mathcal{E}}$ is uniformly bounded, with the uniform bound $|||\mathcal{R}|||$. Hence for every $E \in \mathcal{A}$,

$$\big\|\mathcal{R}_F\big\|_{\mathcal{B}(\mathcal{A})} \leq |||\mathcal{R}||| \, \big\|F\big\|_{\mathcal{A}},$$

Hence $\mathcal{R}$ is a bounded linear map,

$$\begin{aligned}\mathcal{R} : \mathcal{A} &\to \mathcal{B}(\mathcal{A}) \\ F &\to \mathcal{R}_F.\end{aligned}$$

Hence $\mathcal{R}$ continuously maps the Banach space $\mathcal{A}$ into the Banach algebra $\mathcal{B}(\mathcal{A})$ of bounded linear maps on $\mathcal{A}$. $\mathcal{R}$ is a continuous algebra homomorphism. Assuming $\|1\|_{\mathcal{A}} = 1$, we have,

$$\left\|E\right\|_{\mathcal{A}} = \left\|\mathcal{R}_E(1)\right\|_{\mathcal{A}} \leqq \left\|\mathcal{R}_E\right\|_{\mathcal{B}(\mathcal{A})} \leqq |||\mathcal{R}||| \left\|E\right\|_{\mathcal{A}}.$$

We also have $\|\mathcal{R}_1(1)\| = \|1\| = 1$, hence it follows that $|||\mathcal{R}||| = 1$.

Hence $\mathcal{R}(\mathcal{A})$ is a norm-closed subalgebra of $\mathcal{B}(\mathcal{A})$. $\mathcal{R}$ is an algebraic isomorphism from $\mathcal{A}$ to $\mathcal{R}(\mathcal{A})$. Since for any two bounded linear operators S and T on $\mathcal{E}$ we have, $\|TS\| \leq \|T\|\,\|S\|$, $\mathcal{R}(\mathcal{A})$ inherits this property and we have,

$$\|EF\|_{\mathcal{A}} \leq \|E\|_{\mathcal{A}}\|F\|_{\mathcal{A}}$$

Since $\|\mathcal{R}_1\|_{\mathcal{B}(\mathcal{A})} = 1$ we have, $\|\mathcal{R}_E(1)\|_{\mathcal{A}} = \|E\|_{\mathcal{A}} \leq \|\mathcal{R}_E\|_{\mathcal{B}(\mathcal{A})} \leq \|\mathcal{R}_1\|_{\mathcal{B}(\mathcal{A})}\,\|E\|_{\mathcal{A}} = \|E\|_{\mathcal{A}}$. Hence it follows that

$$\mathcal{R} : \mathcal{A} \to \mathcal{R}(\mathcal{A})$$

is an isometric isomorphism and $\mathcal{R}(\mathcal{A})$ is a Banach algebra.

At this stage, it is important to note what it means to measure for the same outcomes. For an outcome R, the preparation procedure concatenated by the measuring instrument corresponding to the effect E_R is itself a preparation procedure. To such preparation procedure, an application of the same measuring instrument again will not change the ensemble the procedure produces.

Since the concatenation of measurements corresponds to product structure of $\mathcal{R}(\mathcal{A})$, this operational requirement about repeated measurements corresponds to requiring that whenever E in $\mathcal{R}(\mathcal{A})$ corresponds to an effect it must correspond to a projection

$$E^2 = E,$$

Hence effects correspond to projection operators in $\mathcal{R}(\mathcal{A})$.

We obtained the Banach algebra structure by assumptions about the continuity properties of compositions of effects, since we expect the accuracies of measuring instruments of compositions to change proportionally to that of its components. Hence the projections E corresponding to effects must be continuous projections in $\mathcal{R}(\mathcal{A})$.

The *-Structure

Suppose we have a measuring instrument which can measure an outcome R, the change occuring in the instrument corresponds to the effect for R. If the measuring instrument does not undergo any change, then we can infer that the outcome R had not occurred. Since the outcome R not occuring is also an outcome by itself, denoted by $\neg R$, and the same measuring instrument also be used for measuring $\neg R$, the instrument not undergoing a change corresponds to the effect for $\neg R$. Hence we must have an operation on the space of effects which assigns to every effect E_R corresponding to the outcome R, the effect $E_{\neg R}$ corresponding to the outcome $\neg R$. We call $E_{\neg R}$ the complement of E_R and we denote it by

$$E_{\neg R} = E_R^{\perp}.$$

Hence we obtain a pairing of the space of effects with itself. This notion of complement of effects is closely related to the notion of topological complementation of closed subspaces on Banach spaces.

The Banach space structure on $\mathcal{R}(\mathcal{A})$ allows us to make sense of closed subspaces of $\mathcal{R}(\mathcal{A})$, with respect to the norm $\|\cdot\|_{\mathcal{B}(\mathcal{A})}$ inherited from $\mathcal{B}(\mathcal{A})$. Let $\mathcal{M}$ be a closed subspace of $\mathcal{R}(\mathcal{A})$, we say that $\mathcal{M}$ is topologically complemented if there exists a closed subspace $\mathcal{M}^\perp$ of $\mathcal{R}(\mathcal{A})$ such that

$$\mathcal{M} \oplus \mathcal{M}^\perp \equiv \mathcal{R}(\mathcal{A}).$$

In such a case every vector A in $\mathcal{R}(\mathcal{A})$ can be uniquely decomposed as a direct sum of vectors $A_\mathcal{M} + A_{\mathcal{M}^\perp}$ where $A_\mathcal{M}$ is in $\mathcal{M}$ and $A_{\mathcal{M}^\perp}$ is in $\mathcal{M}^\perp$. Hence we can consider the projection map on $\mathcal{R}(\mathcal{A})$,

$$E_\mathcal{M} : \mathcal{R}(\mathcal{A}) \to \mathcal{R}(\mathcal{A})$$

which sends each A in $\mathcal{R}(\mathcal{A})$ to the vector $A_\mathcal{M}$ in $\mathcal{M}$. Since we do not have to decompose $A_\mathcal{M}$ any further we must have $E_\mathcal{M} = E_\mathcal{M}^2$.

The closed graph theorem suggests that the continuity properties of operators on Banach spaces and the topology of the product spaces are closely related, and says that an operator is continuous if and only if its graph is closed. Since both $\mathcal{R}(\mathcal{A})$ and $\mathcal{M}$ are closed, it follows that $E_\mathcal{M}$ has a closed graph, and hence the projection,

$$E_\mathcal{M} : \mathcal{R}(\mathcal{A}) \to \mathcal{R}(\mathcal{A}),$$

onto the closed complemented subspace $\mathcal{M}$ of $\mathcal{R}(\mathcal{A})$ is a continuous operator on $\mathcal{R}(\mathcal{A})$. Conversely, every continuous projection $E_\mathcal{M}$, corresponds to a unique closed complemented subspace $\mathcal{M}$. Since the space of effects is contained in $\mathcal{R}(\mathcal{A})$ and since each effect correspond to a continuous projection operator, it follows that every effect corresponds to a unique closed complemented subspace of $\mathcal{R}(\mathcal{A})$, hence the notation for the complement effect $E^\perp$ is closely related to the notion of topological complementation on $\mathcal{R}(\mathcal{A})$.[4]

Since $\mathcal{R} : \mathcal{A} \to \mathcal{R}(\mathcal{A})$ is an isometric isomorphism, the space of states acts as continuous linear functionals on $\mathcal{R}(\mathcal{A})$, given by map,

$$\omega(R_A) = \big\langle A, \omega \big\rangle,$$

with $\omega(R_A) \leq \|R_A\|_{\mathcal{B}(\mathcal{A})}$, hence the elements of $\mathcal{A}_*$ respect the topological properties of $\mathcal{R}(\mathcal{A})$, and hence the notion of topological complementations.

We now consider operations on the space of effects that are physically meaningful. Suppose we perform a rearrangement or readjustments of all the measuring instruments. Then we can consider all the measuring instruments which have been turned to non-functioning instruments, these instruments never undergo any changes, and the re-arrangement or readjustment of instruments acts as a switch-off for those instruments. Assuming such manipulations keep the system within itself, it must keep the space of effects invariant. However the only way to find out what changes we have made to the system is by performing experiments on the system. Since the space of states is separating for the space of effects, any manipulation we do to the space of effects are 'seen' by the space of states. Hence we consider the equivalence class of such manipulations on the space of effects, which the affect the pairing

[4] A theorem of Lindenstrauss & Tzafriri from 1971 says that if for a Banach space $\mathcal{H}$, every closed subspace is topologically complemented, then the norm on $\mathcal{H}$ is induced by an inner product $\langle\cdot,\cdot\rangle_\mathcal{H}$. Such Banach spaces are Hilbert spaces. The theorem provides a reason for why Hilbert spaces play an important role in quantum theory.

$\langle\cdot|\cdot\rangle$ of $\mathcal{A}$ with $\mathcal{A}_*$ in the same way. Each such equivalence class of manipulation of the system corresponds to a bounded linear map,

$$T : \mathcal{A} \to \mathcal{A}$$

In particular, we obtain an element $T(E)$ of $\mathcal{A}$ for each effect E of the system.

The support of a linear functional ω on $\mathcal{A}$ is defined to be the subspace $\mathrm{supp}(\omega) \subset \mathcal{A}$ of all A such that $\omega(A)$ is non-zero. If ω belongs to $\mathcal{A}_*$, it must be continuous, and the support corresponds to a closed subspace of $\mathcal{A}$. The support can also be described in terms of an effect, as the 'largest' effect E_ω such that $\omega(E_\omega)$ is zero.[5] The closed subspace $E_\omega^\perp \mathcal{A}$ is a closed complemented subspace and corresponds to the support of ω. Hence we obtain a map $\omega \mapsto E_\omega$ which takes ensembles to effects.

Since closed complemented subspaces are identified with effects, the map T also gives rise to a map of closed complemented subspaces of $\mathcal{R}(\mathcal{A})$. Using the pairing $\langle\cdot,\cdot\rangle$ we obtain a new pairing of the space of effects with effects,

$$(A, E_\omega) = \big\langle A, \omega \big\rangle$$

This gives rise to a 'conjugate map' of effects, such that for every ω in $\mathcal{A}_*$ we have, $\omega(T^\dagger(A)) = (A, T(E_\omega))$. Using this we obtain a map T^* on the space $\mathcal{A}_*$, which is defined using the map T as follows, $\lambda\omega + \mu\sigma \mapsto \lambda E_\omega + \mu E_\sigma \mapsto \lambda T(E_\omega) + \mu T(E_\sigma)$. This map on $\mathcal{A}_*$ gives us a map,

$$T^\dagger : \mathcal{R}(\mathcal{A}) \to \mathcal{R}(\mathcal{A}),$$

It is determined by the pairing $\langle\cdot,\cdot\rangle$ of $\mathcal{A}$ and $\mathcal{A}_*$ by,

$$\big\langle T^\dagger A, \omega \big\rangle = \big\langle A, T^*\omega \big\rangle$$

By the sesquilinearity of the pairing $\langle\cdot,\cdot\rangle$ it follows for all physically relevant operators T, S on the space of effects $\mathcal{A}$ and for all scalars λ, and μ that,

$$(\lambda T + \mu S)^\dagger = \overline{\lambda} T^\dagger + \overline{\mu} S^\dagger,$$
$$(TS)^\dagger = S^\dagger T^\dagger.$$

By linearity, the operator $(T^\dagger)^\dagger$ is determined by the pairing $\omega((T^\dagger)^\dagger(F)) = (F, T^\dagger(E_\omega))$ for all effects F belonging to the support of ω, and hence also for $E_\omega^\perp$. By the uniqueness of the pairing of closed complemented subspaces with its complement, it follows that

$$(T^\dagger)^\dagger = T.$$

For all physically relevant transformations T, the mapping $\dagger$ acts as an involution. The space of effects $\mathcal{A}$ is a subset of this collection, and we can hence consider this space of physically relevant operators on the space of effects as the object of study. Abusing notation we denote this algebra of operators by $\mathcal{A}$. From the discussion so far, we expect $\mathcal{A}$ to be a Banach algebra equipped with an involution

$$\dagger : \mathcal{A} \to \mathcal{A}.$$

[5] Note that we assume such largest E_ω to exist, because, we can consider the collection of all E such that $\omega(E)$ is zero. This collection is a partially ordered set, and by Zorn's lemma, we can assume the existence of the largest element E_ω within the space of effects by its completeness properties.

Such Banach algebras are called involutive Banach algebras, or Banach*-algebra. We now consider the operation $T^\dagger T$. For every ω, the action of $T^\dagger T$ is given by, $\langle T^\dagger TA, \omega\rangle = \langle TA, T^*\omega\rangle$. Hence the norm of $T^\dagger T$ is given by,

$$\left\|T^\dagger T\right\| = \sup_{\substack{E\in\mathcal{E}\\ \omega\in\mathcal{S}}} \left\{\left\langle TE, T^*\omega\right\rangle\right\} = \left\|T\right\|^2$$

A Banach *-algebra $\mathcal{A}$ such that for every T, $\|T^\dagger T\| = \|T\|^2$, is called a C^*-algebra. The identity $\|T^\dagger T\| = \|T\|^2$ is called the C^*-identity.

Since all the physically relevant maps were described as equivalence classes with respect to the pairing with the space of states, it follows that $\mathcal{A}_*$ remains separating for $\mathcal{A}$ if we think of $\mathcal{A}$ as physically relevant operations of the system. Hence we must have

$$(\mathcal{A}_*)^* \cong \mathcal{A}$$

A C^*-algebra $\mathcal{A}$ which is a dual Banach space of $\mathcal{A}_*$ is called a von Neumann algebra, and in such a case the Banach space $\mathcal{A}_*$ is called the predual of $\mathcal{A}$. Von Neumann algebras act as the appropriate mathematical objects for the study of physical systems. The von Neumann algebra $\mathcal{A}$ modeling the space of effects of a system is called the algebra of observables.

1.3 | What are Observables?

In general, an experiment can measure a collection of possible outcomes, the notion of an observable is an idealisation of experiments. The experiments on a quantum system can be described by a collection of outcomes it can measure. For each such possible outcome we can assign its corresponding effect. If $R \equiv \{R_i\}$ is a collection of outcomes, then each outcome R_i of the experiment R corresponds to an effect E_{R_i} called the effect of R_i. Hence each experiment $R \equiv \{R_i\}$ can be modelled as an effect valued map,

$$E_R : \{R_i\} \mapsto E_{R_i}.$$

The notion of observables corresponds to the experiments $R = \{R_i\}$ which can be performed with a single measuring instrument, in which case the outcomes $\{R_i\}$ can be simultaneously measured. We now extrapolate the algebraic structures for collection of effects for such possible outcomes by studying the expected properties of such measuring instruments, and model observables in the context of algebraic quantum theory.

1.3.1 | Boolean Algebras

Consider a collection of possible outcomes $R \equiv \{R_i\}$ that can be measured simultaneously using the same instrument. In this case a single experiment is performed. Let $\{E_{R_i}\}$ be the collection of effects corresponding to the possible outcomes $\{R_i\}$. We are interested in understanding the algebraic structure that should be expected of such effects. To understand the algebraic structure of $\{E_{R_i}\}$, we can ask what are the queries that can be answered after an experimenter performs a measurement with the same instrument. We should obviously expect whether the outcome was R_i for each $R_i \in R$. The experimenter can also ask if the outcome was not R_i. So, if $\neg R_i$ denotes the outcome corresponding to the experimental outcome not being R_i, then we should expect $E_{\neg R_i} \in \{E_{R_i}\}$. We will denote such elements by the notation

$$E_{\neg R_i} \equiv E_{R_i}^\perp.$$

Similarly, the experimenter can ask if the outcome was R_i or R_j, and if the outcome was R_i and R_j for $R_i, R_j \in R$. This tells us about the algebraic structure of effects when they can be measured by a single instrument. The effects of a measuring instrument measured by a single instrument should be expected to have classical logical operations such as meet, join and not, corresponding to whether the measurement detects outcome R_i *or* R_j, R_i *and* R_j, the outcome is *not* R_i. Denote these operations by $R_i \vee R_j$, $R_i \wedge R_j$, and $\neg R_i$ respectively.

Clearly we should have for any $R_i \in R$,

$$R_i \wedge R_i = R_i,\ R_i \vee R_i = R_i.$$

The composite operations involving $\vee$, $\wedge$, and $\neg$ can be figured out by simple thought, and $R \equiv \{R_i\}$ together with these operations is a Boolean algebra, denoted by Σ_R.

We list below the properties of Boolean algebras without attempting to be minimal. We expect the operations $\wedge$ and $\vee$ to be commutative and associative, $\wedge$ and $\vee$ satisfy the absorption property, that is,

$$R_i \vee (R_i \wedge R_j) = R_i,\ R_i \wedge (R_i \vee R_j) = R_i.$$

We also expect distributivity between $\wedge$ and $\vee$,

$$R_i \vee (R_j \wedge R_k) = (R_i \vee R_j) \wedge (R_i \vee R_k)$$
$$R_i \wedge (R_j \vee R_k) = (R_i \wedge R_j) \vee (R_i \wedge R_k).$$

Let I denote the outcome which is always true when the measurement is performed. I and 0 are called the top and bottom elements respectively. We say $\{R_i\}$ is exhaustive if $\vee_i R_i = I$. Then we expect

$$R_i \wedge I = R_i,\ \ R_i \vee \neg R_i = I.$$

Similarly, let 0 denote the outcome which never undergoes a change. We say that R_i and R_j are disjoint if $R_i \wedge R_j = 0$, and we have

$$R_i \vee 0 = R_i,\ \ R_i \wedge \neg R_i = 0.$$

If R is closed under $\wedge$, $\vee$ and $\neg$ and contains the elements I and 0 it is called a Boolean algebra, denoted by Σ_R. These axioms forces the complements to be unique.

Associativity of Boolean operations ensures that we can make sense of $\wedge_k R_{i_k}$ and $\vee_k R_{i_k}$ for any finite collection $\{R_{i_k}\}$. Σ_R is said to be a complete Boolean algebra if we can make sense of $\vee_k R_{i_k}$ and $\wedge_k R_{i_k}$ for arbitrary collections. We say Σ_R is σ-complete if we can make sense of $\vee_k R_{i_k}$ and $\wedge_k R_{i_k}$ for countable collections.[6]

The effects of a measuring instrument corresponding to a single measuring instrument should also inherit Boolean algebra structure, so the maps of interest to us are those which preserve the Boolean algebra structure. If Σ_R and Σ_S are Boolean algebras, a mapping

$$E_R : \Sigma_R \to \Sigma_S$$

[6]Every Boolean algebra can also be thought of as a lattice and vice-versa with the order give by, $R_i \leq R_j \Leftrightarrow R_i \vee R_j = R_j$. The notion of completeness can be described in terms of the lattice order and in terms of existence of least upper bounds, and this lattice theoretic notion of completeness is closer to the completeness of $\mathbb{R}$ in terms of order coming from the notion of positivity on real numbers.

The space $\mathcal{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$ is a σ-complete Boolean algebra. If we assume every measuring instrument has its values labelled by real numbers, $\mathcal{B}(\mathbb{R})$ acts as an appropriate Boolean algebra for modeling observables.

is called a Boolean map if it preserves the Boolean operations. Observables must correspond to such Boolean maps.

From the discussion above it is clear that the space of effects of a system must be modeled using a von Neumann algebra $\mathcal{A}$. Since observables correspond to collections of simultaneously measurable outcomes, they correspond to maps from Boolean algebras to the space of effects that respects the Boolean algebra structure.[7] An observable in algebraic quantum theory is a pair (Σ_R, E_R), where Σ_R is a Boolean algebra and E_R is a map

$$E_R : \Sigma_R \to \mathcal{A}.$$

which is a Boolean map to its image $E_R(\Sigma_R) \subseteq \mathcal{A}$. In particular $E_R(\Sigma_R)$ will have to be a commutative subalgebra of $\mathcal{A}$. We will denote the observable by its corresponding Boolean map E_R.

Since $E_R(\Sigma_R)$ as a subalgebra must be commutative, commutative subalgebras of the algebra of observables are of special interest to us, the study of commutative C^*-algebras gives us the notion of spectrum, and the study of commutative von Neumann algebras leads to a characterisation of observables. Although von Neumann algebras can be abstractly studied, for this note, we will focus on the von Neumann algebra of bounded operators on Hilbert spaces.

1.4 | Heisenberg's Approach

Before we go on to quantum mechanics on Hilbert spaces, we describe how it was traditionally developed starting from Heisenberg's idea, which was formalised by von Neumann via the notion of Hilbert spaces. By the end of the nineteenth century, it was clear that certain elementary processes obeyed 'discontinuous' laws. That is to say, there exist observables whose collection of effects forms discrete sets, and also observables whose collection of effects forms a continuous set.

If the collection of effects of an observable Q can be labeled by a discrete set, the observable E_Q corresponds to a map, which assigns to each collection of values of the observable $Q_F \equiv \{Q_i\}_{i\in F\subseteq\mathbb{N}}$ the effect $\sum_{i\in F} E_{Q_i}$ for the measuring the value to be in the set $\{Q_i\}_{i\in F}$. Hence the observable is a map,

$$E_Q : \mathcal{B}(\mathbb{N}) \to \mathcal{A}.$$

Where $\mathcal{B}(\mathbb{N})$ consists of the subsets of $\mathbb{N}$. Each ensemble μ_ρ is a map $\mu_\rho : \mathcal{A} \to \mathbb{R}$. The composite map is

$$\mu_\rho \circ E_Q \equiv \mu_\rho^Q : \mathcal{B}(\mathbb{N}) \longrightarrow \mathbb{R}.$$

[7] Döring and Isham, in [16], revisit this approach to modeling observables. They model observables at an even higher level of generality by replacing the Boolean algebra structure on Σ_R with Heyting algebras. If Boolean algebra operations corresponds to the classical logical operations, the Heyting algebra operations correspond to intuitionistic logic, which excludes the law of excluded middles, or the idempotency of $\neg$-operation. The Heyting algebras can be modeled categorically in terms of toposes, similar to how Boolean algebras can be modelled by sets. A topos, as a category is equivalent to the category of sheaves on a space, and hence each object of a topos corresponds to some local property of a space.

This model takes into account all categories where the notion of 'locality' makes sense, this allows us to talk about 'where' the values of the observable lies from a very abstract and general point of view. This approach to the formulation of quantum theory, starting with the remodeling the structure expected of simultaneously emasurable outcomes is known as topos quantum theory. We will however not discuss topos quantum theory in here and refer an interested reader to [16]-[17].

Note that in this case, we cannot obtain the †-operation on the algebra of observables from operational assumptions as we did before. We needed idempotency of $\neg$ for uniqueness of topological complementation. In such a case the existence of †-operation should be assumed, or should be derived with different operational assumptions.

Which assigns to the possible event Q_F its probabilty of occurance $\mu_\rho(Q_F)$ as described above. If the observable is measured, the sum total probability of occurrence of atleast one of the values should be 1. This means that

$$\sum_{i\in\mathbb{N}} \mu_\rho^Q(Q_i) = 1$$

where $\mu_\rho^Q(Q_i)$ is the probability the observable has a value Q_i when the state is μ_ρ. In particular, the sum should make sense. The expected value of the observable R for the state μ_ρ is given by, $\langle Q\rangle_\rho = \sum_{i\in\mathbb{N}} Q_i \mu_\rho^Q(Q_i)$.

On the other hand if the collection of effects of an observable can be labeled by a continuous set, that is, the measurements are labelled by countable unions of open intervals of $\mathbb{R}$, such an observable corresponds to a map, which assigns to intervals R_i of $\mathbb{R}$ the effect E_{R_i} where R_i is the interval in which the value of the observable R lies. Each ensemble μ_ρ corresponds to a function, $\mu_\rho : \mathcal{A} \to \mathbb{R}$. The gives us a composite map,

$$\mu_\rho^R : \mathcal{B}(\mathbb{R}) \longrightarrow \mathbb{R}.$$

Which assigns to the interval R_i a probabilty $\mu_\rho^R(R_i)\lambda(R_i)$ where λ is the standard Lebesgue measure on $\mathbb{R}$. Again as before, if the observable is measured, the sum total probability of occurrence of one of the values should be 1. This means that

$$\int_{\mathbb{R}} \mu_\rho^R(x) d\lambda(x) = 1$$

In particular the integral should make sense. Each state corresponds to the measure $\mu_\rho^R d\lambda$. The expected value of the observable R for the state μ_ρ is given by, $\langle R\rangle_\rho = \int_{\mathbb{R}} x\mu_\rho^R(x) d\lambda(x)$.

1.4.1 | Observables in Classical Theories

Classical theories model observables using continuous functions on a manifold. It is sufficient to assume integrability of the functions for this model. If the manifold is Ω, observables in classical physics correspond to integrable functions of the form $R : \Omega \to \mathbb{R}$. The state in classical physics corresponding to the ensemble ρ is modeled by a positive Radon measure μ_ρ on Ω. The probability measure the state μ_ρ associates with the observable R is

$$\mu_\rho^R : \mathcal{B}(\mathbb{R}) \to \mathbb{R}$$

This is the composite map,

$$\mathcal{B}(\mathbb{R}) \xrightarrow{R^{-1}} \mathcal{B}(\Omega) \xrightarrow{\mu_\rho} \mathbb{R}.$$

Which assigns to the effect E_{R_i} the probabilty $\mu_\rho^R(R_i) = \mu_\rho(R^{-1}(R_i))$. Note that $R^{-1}(R_i)$ is a measurable set in Ω and corresponds to the set of all points in Ω such that the evaluation of the function R lies in the interval R_i. Since the total probability should be 1 we have

$$\int_{R^{-1}(\mathbb{R})} d\mu_\rho(x) = 1.$$

The expected value of the observable R for an ensemble ρ is given by

$$\langle R\rangle_\rho = \int_{R^{-1}(\mathbb{R})} R(x) d\mu_\rho(R^{-1}(x))$$

The characteristic functions which assign to each Borel set in Ω the value 1 correspond to effects in this model. The collection of all characteristic functions on Ω is itself an observable and since every integrable function can be approximated as combinations of simple function, every observable in the classical model coexists with each other.

The space of all measures $\mathcal{M}(\Omega)$ forms a vector space. The set of measures with the above property of which ensures the integral is 1 ensures that it is a convex subset of $\mathcal{M}(\Omega)$. The extreme points of this convex set are the delta distributions which correspond to the points of the manifold Ω. If we denote the convex set of regular probability measures on Ω by $\mathcal{S}(\Omega)$ the state space of the classical theory is

$$\mathcal{S} \equiv \mathcal{S}(\Omega).$$

Since preparation procedures are independent of which measuring instrument is applied the states must assign a probability distribution to each observable. The collection of all integrable functions $L^1(\Omega)$ acts as the space of effects.

This is a good model as long as all the observables take continuum of values. The problem starts when there exists a discrete observable simultaneously alongside a continuous observable. Suppose there exists a discrete observable, then each state $\mu_\rho \in \mathcal{S}(\Omega)$ should also correspond to a probability distribution,

$$\mu_\rho^Q : \mathcal{B}(\mathbb{N}) \to \mathbb{R}$$

such that,

$$\sum_{i\in\mathbb{N}} \mu_\rho^Q(Q_i) = 1$$

where $\mu_\rho^Q(Q_i) = \mu_\rho(Q^{-1}(Q_i))$. Such functions correspond to summable sequences denoted by $l^1(\Omega)$. The coexistence is not possible because the space of integrable functions $L^1(\Omega)$ is not isomorphic to the space of summable sequences $l^1(\Omega)$.[8]

The conclusion we can draw is that if the observables are modeled as maps from a manifold, the state space of the physical system cannot give a probability distribution for each observable. There does not exist a common state space that can accommodate both discrete and continuous observables. So, the problem with classical theory is that it fails to consider some of the observed phenomena. So the physical ideas behind classical theory are too constraining. We now have to figure out what this hidden extra idea is and construct a new mathematically model for observables that does not take this extra idea into account.

1.4.2 | Hilbert Spaces

Heisenberg's radical solution to the problem of mathematical modeling of observables was to think of observables as operators on a vector space. The spectrum of the operator is to be thought of as values of the observable. Compared to classical theories, this is an extremely abstract and radical change. The classical model of an observable was geometric and hence intuitive. Von Neumann, Hilbert, and others were able to figure out the underlying idea and reformulated it clearly in terms of 'Hilbert spaces'. The key to von Neumann's articulation of Heisenberg's model lies in the isomorphism between the space of square summable functions and the space of square-integrable functions as Hilbert spaces, due to Riesz & Fischer.

[8] One way to prove this is via the so called Schur's property, which says that if a sequence is weakly convergent then it is also convergent in the norm. $l^1(\Omega)$ has the Schur's property and $L^1(\Omega)$ will not, proving that they cannot be isomorphic.

As noted before what we want is a convex state space $\mathcal{S}$ that can provide a probability function for both discrete and continuous variables. Each element $\mu_\rho \in \mathcal{S}$ should give rise to functions, $\mu_\rho^Q : \mathcal{B}(\mathbb{N}) \to \mathbb{R}$, and $\mu_\rho^R : \mathcal{B}(\mathbb{R}) \longrightarrow \mathbb{R}$, such that,

$$\sum_{i\in\mathbb{N}} \mu_\rho^Q(Q_i) = 1,$$

and

$$\int_{\mathbb{R}} \mu_\rho^R(x) d\lambda(x) = 1.$$

At this stage, von Neumann defined the notion of a Hilbert space which provides the appropriate mathematical language to get all the ingredients together. Abstractly a Hilbert space is a pair $(\mathcal{H}, \langle\cdot|\cdot\rangle_{\mathcal{H}})$ where $\mathcal{H}$ is a vector space and $\langle\cdot|\cdot\rangle_{\mathcal{H}}$ an inner product on it, and is topologically a complete normed space. Every Hilbert space has an orthonormal basis and an element of the Hilbert space can be uniquely specified by its coordinates with respect to a complete orthonormal system. A Hilbert space is said to be separable if it has a countable basis, and any two separable Hilbert spaces are isomorphic.

A sequence of complex numbers or a function from natural numbers to complex numbers is said to be square summable if

$$\sum_{i\in\mathbb{N}} \big|f(i)\big|^2 < \infty.$$

With pointwise addition and scalar multiplication, the set of all square summable sequences is a complex vector space. It can be endowed with an inner product,

$$\big\langle f|g\big\rangle_{l^2} = \sum_{i\in\mathbb{N}} \overline{f(i)} g(i).$$

Together with this inner product the space of square summable sequences of complex numbers is a Hilbert space denoted by $l^2(\mathbb{N})$. A measurable function from the real line to complex numbers is called square integrable if,

$$\int_{\mathbb{R}} \big|f(x)\big|^2 d\lambda(x) < \infty,$$

where λ is the standard Lebesgue measure on $\mathbb{R}$. The collection of all square-integrable functions is a complex vector space. This vector space can be endowed with an inner product,

$$\big\langle f|g\big\rangle_{L^2} = \int_{\mathbb{R}} \overline{f(x)} g(x) d\lambda(x).$$

Two square integrable functions are equivalent if they are same almost everywhere with respect to λ. The collection of equivalence classes of square integrable functions inherits a vector space structure from the space of square integrable functions. Together with the inner product $\langle\cdot|\cdot\rangle_{L^2}$ the space of equivalence classes of square-integrable functions is a Hilbert space, denoted by $L^2(\mathbb{R})$.

Theorem 1.1. (Riesz-Fischer) *$L^2(\mathbb{R}) \cong l^2(\mathbb{N})$ as Hilbert spaces.*

This isomorphism provides us with a state space that allows for coexistence of discrete and continuous observables. See §1.4 [10] for a proof of Riesz-Fischer theorem. This isomorphism acts as the starting point for von Neumann's reformulation of Heisenberg's model. Von

Neumann's approach was to compare the the space of functions on discrete and continuous spaces instead of comparing discrete space and continuous space themselves. The isomorphism as Hilbert spaces of the space of square integrable functions and the space of square summable sequences allows us to develop a unified mathematical model where coexistence of both discrete and continuous observables is possible. If observables are treated as self-adjoint operators on a separable Hilbert space $(\mathcal{H}, \langle\cdot|\cdot\rangle_{\mathcal{H}})$, then the collection of all unit length vectors can be used as the extreme points of a state space common to both discrete and continuous observables. The values of the observable correspond to the spectrum of the operator. The interpretation is that the self-adjoint operator via spectral theorem, which we will discuss later on, gives rise to a collection of projection operators which correspond to the collection of effects for the observable, and the length of the projection for a state corresponds to the probability of the state undergoing a change for that effect.

2 | Quantum Theory in Hilbert Spaces

The discussion of the previous section suggests that if we assume the coexistence of continuous and discrete observables, the algebra structure on the space of effects can be mathematically modeled using the structure of a von Neumann algebra. We now describe the special case of quantum theory in the Hilbert space formalism, where the space of effects is modelled using bounded operators on a Hilbert space.

2.1 | The Lattice of Closed Subspaces

We need to study the algebra of operators to understand the structure of the space of effects $\mathcal{E}$. Let $(\mathcal{H}, \langle\cdot|\cdot\rangle)$ be the complex separable 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 the Hilbert space $\mathcal{H}$ 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 $E^{\perp}$. It follows that

$$E^{\perp} = I - E.$$

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.[9] The non-Boolean lattice $\mathcal{P}(\mathcal{H})$ of projections should act as the space of effects

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

In terms of measuring instruments, and possible outcomes we have the following associations,

$$E_{\neg R_i} = E_{R_i}^{\perp}, \;\; E_{R_i \vee R_j} = E_{R_i} + E_{R_j}, \;\; E_{R_j \wedge R_i} = E_{R_j} E_{R_i}.$$

[9] In the quantum logic literature the central objects that model possible events are the so called orthomodular lattices. Orthomodular lattices are studied abstractly in [12].

Here $\neg R$, $R \wedge S$ and $R \vee S$ correspond to whether the value of the measurement is '*not* R', 'R *and* S' and 'R *or* S' respectively. When instruments for R and S are different $\wedge$ corresponds to 'R *after* S', and the order of measurement is important. For observables this order is unimportant since all of its possible outcomes are simultaneously measured. Since all measuring instruments cannot be simultaneously applied to a prepared system, we cannot expect commutative structure with respect to $\wedge$.

For a family of projection operators to represent an observable, the family should form a Boolean algebra. A quantum mechanical observable is an Boolean map of the form,

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

Such maps are called a projection valued function. In physical experiments, the statements usually are of the form 'the value of the observable lies in some set ϵ_i of real numbers'. To accommodate the fact that the measurement scale is consisting of real numbers, we identify Σ_A with the Boolean algebra of Borel sets of $\mathbb{R}$.

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 an integrable map $X : \Omega \to \mathbb{R}$. The map doing the work in assigning necessary probabilities is its pre-image,

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

Which has some similarity with the observables in the quantum case.

A spectral measure is a projection operator-valued function E defined on the sets of $\mathbb{R}$ such that, $E(\mathbb{R}) = I$ and $E(\coprod_{i\in\mathbb{N}} \epsilon_i) = \sum_{i\in\mathbb{N}} E(\epsilon_i)$, 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,

$$A = \int_{\mathbb{R}} \lambda \, dE_A(\lambda),$$

and conversely, every spectral measure corresponds to a self-adjoint operator given by the above formula. In the finite-dimensional case this reduces to $A = \sum_i \lambda_i E_i$ where E_i is the projections onto eigenspaces of the eigenvalues λ_i. The special feature of quantum theory is that, unlike classical mechanics, the space of effects is a non-commutative algebra.

2.1.1 | Gleason's Theorem

States are linear functionals on the space of effects. In case of quantum theory the space of effects is identified with projection operators on $\mathcal{H}$ and hence the mathematical representatives of the physical states in the quantum case are the maps,

$$\omega : \mathcal{P}(\mathcal{H}) \to \mathbb{R}$$

such that $\omega(0) = 0$, $\omega(E^\perp) = 1 - \omega(E)$ and $\omega(\sum_i E_i) = \sum_i \omega(E_i)$, for mutually orthogonal E_i. They naturally extend linearly to the space of bounded operators $\mathcal{B}(\mathcal{H})$ on $\mathcal{H}$.

A positive linear functional on the space of bounded operators $\mathcal{B}(\mathcal{H})$ on the Hilbert space $\mathcal{H}$ is called a non-commutative probability measure or a normal state or a Gleason state if for every projection operators $E \in \mathcal{P}(\mathcal{H})$,

$$\omega(E^\perp) = 1 - \omega(E)$$

and for a countable collection of mutually orthogonal projection operators $\{E_i\}_{i\in\mathbb{N}}$,

$$\omega\big(\textstyle\sum_{i\in\mathbb{N}} E_i\big) = \displaystyle\sum_{i\in\mathbb{N}} \omega(E_i).$$

This is called the 'operational independence', It says that the operational statistics is independent of how we split up $\sum_{i\in\mathbb{N}} E_i$. It is closely related to the operational assumption that the preparation procedure are independent of which measuring instrument it is applied to, and measuring instruments undergoing changes for same outcome must be equivalent, and do not depend on how the measuring instrument can be obtained as a combination of different measuring instruments.

When restricted to commutative subalgebras, non-commutative probability measures correspond to probability measures. For each observables there exists a spectral measure

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

By taking the composition, we obtain,

$$\mathcal{B}(\mathbb{R}) \xrightarrow{E_A} \mathcal{P}(\mathcal{H}) \xrightarrow{\omega} [0,1],$$

This composition satisfies the conditions of a probability measure, and hence, each state ω assigns to an observable E_A a probability measure, $\mu^A_\omega \equiv \omega \circ E_A : \mathcal{B}(\mathbb{R}) \to [0,1]$. $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu^A_\omega)$ is a classical probability space, where μ^A_ω is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Gleason's theorem characterises Gleason states on Hilbert spaces of dimension larger than or equal to 3.

Theorem 2.1. (Gleason) *If ω is a Gleason state on $\mathcal{H}$ for $\dim(\mathcal{H}) \geq 3$ then there exists an operator ρ on $\mathcal{H}$ with $\rho \geq 0$ and $Tr(\rho) = 1$ such that*

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

Sketch of Proof

The idea of Gleason's proof is to reduce the general case to the case of two dimensions, where one can use relatively simpler tools. The starting point is the definition of the so called frame functions. A frame function f of weight W on the separable Hilbert space $\mathcal{H}$ is a real valued function on the 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 the restriction to every two dimensional subspace. Then the continuity of frame functions is proved. Every non-negative frame function on a Hilbert space of dimension greater than three is regular. Much of the difficulty lies in this part. See Gleason [18],[19] for details.

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

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

Since this holds for all unit vectors, ρ is positive semi-definite. Denote by $E_{\mathcal{H}}$ the projection onto the whole Hilbert space which is 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 closed subspace $\mathcal{K} \subset \mathcal{H}$ denote by $E_{\mathcal{K}}$ the corresponding projection operator. Take an orthonormal basis $\{|\varkappa_i\rangle\}_{i\in I}$ for $\mathcal{K}$ and extend it to one of $\mathcal{H}$. Then we can write

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

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

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

□

Positive semidefinite operators of unit trace are called density matrices. Note that the converse of the theorem clearly holds, that is, every density matrix determines a state as defined in the above formula. More general quantum experiments correspond to positive operator-valued measures. The effects E are given by positive operators, $O \leq E \leq I$. Since these should sum to 1 for an experiment, it will be a resolution of identity $\sum_i E_{A_i} = I$, where the 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 like theorem in this setting see [20].

We call ω satisfying the equivalent conditions of Gleason's theorem a Gleason measure. They correspond to the states of quantum systems. 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 closed convex set. The topology on $\mathcal{S}(\mathcal{H})$ comes from the sup-norm on operators coming from the inner product on $\mathcal{H}$.

The Krein-Milman theorem says that a compact convex subset of a Hausdorff locally convex topological vector space is equal to the closed convex hull of its extreme points. The extreme points of the set of all states are called pure states, pure states satisfy $\rho = \rho^2$ and correspond 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,

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

The expectation value of the observable will be,

$$\langle A\rangle = \int_{\mathbb{R}} \lambda \, d\mu_\rho^A(\lambda) = \mathrm{Tr}(\rho A).$$

All Hilbert spaces will be assumed to be separable and complex.

Tensor Product

Given a collection quantum systems $\mathcal{H}_i$ labelled by a finte set F with Hilbert spaces $\mathcal{H}_i$, we are interested in describing the Hilbert space appropriate to the 'product' quantum 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,

$$\prod_{i\in F} \mathcal{H}_i \longrightarrow \mathcal{I}.$$

The interpretation of this map is that it introduces a component from each individual system into the product system. Accounting for the superpositions, the product system should inherit the structure from the components. The map above must be linear for each component. The universal solution $\mathcal{H}$ to this problem is the algebraic tensor product.

Hence, it is the vector space $\mathcal{H}$ together with an $|F|$-linear map such that, for any n-linear map $\prod_{i\in F} \mathcal{H}_i \to \mathcal{I}$, there exists a unique linear map : $\mathcal{H} \to \mathcal{I}$,

$$\begin{array}{ccc} \prod_{i\in F} \mathcal{H}_i & \longrightarrow & \mathcal{H} \\ & \searrow & \downarrow \exists! \\ & & \mathcal{I} \end{array}$$

This vector space inherits a canonical inner product given by the product of inner products of the component Hilbert spaces. The completion of

$$\mathcal{H} = \bigotimes_{i\in F} \mathcal{H}_i$$

under the canonical inner product will serve as the Hilbert spaces for product quantum systems. Such a topology is justified because if two states are near to each other in a component, then we expect them to be near to each other in the product quantum system.

References

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

[2] G Ludwig, Foundations of Quantum Mechanics II. Springer-Verlag, 1985

[3] G Ludwig, An Axiomatic Basis for Quantum Mechanics Volume 1. Springer-Verlag, 1985

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

[5] H Primas, Chemistry, Quantum Mechanics and Reductionism. Springer-Verlag, 1983

[6] S Gudder, Convexity and Mixtures. SIAM Review, Vol. 19, No. 2, April 1977

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

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

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

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

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

[12] G Kalmbach, Orthomodular Lattices. Academic Press, 1983

[13] A Peres, Quantum Theory: Concepts and Methods, Kluwer Academic Publishers, 2002

[14] R Kadison, J Ringrose, Fundamentals of the Theory of Operator Algebras. Volume I, Academic Press, 1983

[15] J S Lomont, P Mendelson, The Wigner Unitarity-Antiunitarity Theorem. Annals of Mathematics, Second Series, Vol. 78, No. 3, pp. 548-559, 1963

[16] A Döring, C Isham, What is a thing? https://arxiv.org/abs/0803.0417v1, 2008

[17] C Flori, Lectures on Topos Quantum Theory https://arxiv.org/abs/1207.1744, 2012

[18] 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

[19] R Cooke, M Keane, W Moran, An Elementary Proof of Gleason's Theorem. https://doi.org/10.1017/S0305004100063313. 1984

[20] 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