# Notes on
# Algebraic Quantum Theory

Bharath Ron

The physical systems of interest to us are those whose observed phenomenon 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 $\Omega$, 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 first try to understand what is meant by an 'observable'.

## 1 | Generalised Probabilistic Theories

The discussion here will be closest to the Ludwig school [?]-[4], with some stuff borrowed from the quantum logic literature [8]-[9]. This can be thought of as a modern formulation of Heisenberg's original idea. Hans Primas in §4 [5] characterises approaches to formulation of quantum theory into three main categories, the quantum logic approach, algebraic approach, and the convex state-space approaches.

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 of and measuring instruments. The assumed structure in this instrumentalist view of physics can be motivated with simple thought experiments 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 events. 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 events which are the changes occurring to instruments. These instruments and events 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 events. 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 $\rho$ 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 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

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

In general, an experiment can measure a collection of possible outcomes, hence the measuring instrument used for the experiment 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 measuring instrument, then each outcome $R_i$ of the measuring instrument $R$ corresponds to an effect $E_{R_i}$ called the effect of $R_i$. Hence each measuring instrument $R \equiv \{R_i\}$ can be modelled as an effect valued map.

Similarly, each ensemble fixes the operational statistics for possible outcomes, and hence we have,

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

$\mathcal{E}$ 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 the 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 ro the more familiarity 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 its stronger foundations and embrace the risk of losing the centuries old, and possibly false intuition, for classical theories.[1]

### 1.1.1 | Convex Embeddings

In order to be able to work with effects and ensembles we now embed them inside a set with mathematical structure that respects the expected operational relations and also allows

[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.

us to do mathematics with. In our case the operational requirement is that we should be able to make sense of taking mixtures of effects and ensembles. Accounting for the fact that preparation procedures can be combined to produce a mixed ensemble, the set of ensembles should be closed under forming of mixtures.

The notion of mixing corresponds mathematically to the notion of convex combination, see [6], we expect the set $\mathcal{S}$ to have structure that enables us to take convex combinations 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 convinient 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. 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 $\rho_i$ are ensembles and $\alpha_i$ 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,

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

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

$$\rho \mapsto \mu_\rho,$$

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

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

Abusing notation we will denote the images of $\mathcal{E}$ and $\mathcal{S}$ by the same. 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. equip $\mathcal{A}$ and $\mathcal{A}_*$ with mathematical structure sufficient for their study.[2]

[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.

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

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

that coincides with $\mu$ 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 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.

### 1.1.2 | The Banach *-Structure

The operational requirements force the space of effects and ensembles to be vector spaces. This is however insufficient for analysis. To bring any predictivity to the theory we need more structure than mere vector space structure. Since we can make sense of one measurement after another, the space of effects should come equipped with an algebra structure. Similarly, the notion of accuracy of instruments gives rise to a notion of nearness, and gives rise to a topology. We intend to understand the relation between the algebraic and the topological structures.

#### Construction of a Norm

We can have two measuring instruments for an event 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, there needs to be a way of quantifying the ease of noticing the changes occuring to the instruments. 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.

Let $\|\cdot\|_{\mathcal{E}}$ be the overall likelihood of undergoing change. Suppose we have two measuring instruments which undergo changes for outcomes $R_i$ and $R_j$, and let $E_{R_i}$ and $E_{R_j}$ be the corresponding effects. Suppose the two instruments could undergo changes under some similar situations, then it means that we have more redundancy. The more redundancies, the easier it is to undergo changes, and more likely it is to notice changes. Hence we must have,

$$\|E_{R_i} + E_{R_j}\|_{\mathcal{E}} \leq \|E_{R_i}\|_{\mathcal{E}} + \|E_{R_j}\|_{\mathcal{E}}.$$

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. Hence we expect $\|\cdot\|_{\mathcal{E}}$ to be continuous under addition. Hence $\|\cdot\|_{\mathcal{E}}$ satisfies the triangle inequality.

When a measuring instrument is applied to any preparation procedure it can potentially undergo a change. Even though this gives us an operational interpretation of $\|\cdot\|_{\mathcal{E}}$ in terms of accuracy of the measuring instruments it is more convinient to define it as a supremum. Since $\|\cdot\|_{\mathcal{E}}$ represents the overall likelihood of undergoing a change it must correspond to the highest chance of undergoing over all preparation procedures. Hence for every $F \in \mathcal{E}$, $\|F\|_{\mathcal{E}}$ can be defined as the supremum,

$$\big\|E\big\|_{\mathcal{E}} \coloneqq \sup_{\mu_\rho \in \mathcal{S}} \big|\langle E|\mu_\rho\rangle\big|,$$

where the supremum is taken over all preparation procedures which is represented by the convex embedded subset $\mathcal{S}$ of $\mathcal{E}_*$. By the assumption that there will always exist some preparation procedure with non-trivial operational statistics, we deduce that $\|\cdot\|_{\mathcal{E}}$ is non-trivial, that is, $\|E\|_{\mathcal{E}} = 0$ if and only if $E \equiv 0$. By definition it also follows that

$$\|\lambda E\|_{\mathcal{E}} = |\lambda| \|E\|_{\mathcal{E}}, \quad \forall \lambda \in \mathbb{C}.$$

Since $\|\cdot\|_{\mathcal{E}}$ satisfies triangle inequality, and satisfies the above scaling condition it defines a norm on $\mathcal{E}$. Similarly, $\mathcal{E}_*$ also inherits a norm from $\mathcal{E}$ such that for any $\sigma \in \mathcal{E}_*$,

$$\|\sigma\|_{\mathcal{E}_*} := \sup_E |\langle E|\sigma\rangle|$$

where the supremum is taken over effects which is represented by the convex embedded subset $\mathcal{P}$ of $\mathcal{E}$. We can hence require both $\mathcal{E}$ and $\mathcal{E}_*$ to be a complex, normed spaces. By the boundedness of the functionals

$$E_{R_i} \mapsto \langle E_{R_i}|\mu_\rho\rangle,$$

we note that every ensemble gives rise to a continuous linear functional on $\mathcal{E}$ with respect to the topology induced by the norm $\|\cdot\|_{\mathcal{E}}$ and similarly for $\mathcal{E}_*$. Under the idealisation that there exist instruments of every level of accuracy, we can assume $\mathcal{E}$ to be closed with respect to taking limits in the above defined norms. Hence we will assume the spaces $\mathcal{E}_*$ and $\mathcal{E}$ to be Banach spaces. This allows us to talk about limits and allows us to do mathematical analysis with ensembles and effects. We assume the existence of a unique element $1 \in \mathcal{E}$ 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. By definition of the norm we have $\|1\|_{\mathcal{E}} = 1$. In this sense, the element $1 \in \mathcal{E}$ is the 'existence' element for the system. Similarly we will assume the existence of the unique 0 element which outputs 'no' for every preparation procedure.

Since $\mathcal{E}$ and $\mathcal{E}_*$ inherit topologies from each other we expect $\mathcal{E}$ and $\mathcal{E}_*$ to be related to each other topologically also. Since $\mathcal{S}$ and $\mathcal{P}$ separate each other with respect to the pairing $\langle\cdot|\cdot\rangle$, we will assume $\langle\mathcal{E}|\mathcal{E}_*\rangle$ is a dual pair.

### Topological vs Algebraic Data

Let the outcome corresponding to measuring the outcome $R_i$ after $R_j$ be $R_i \wedge R_j$, the corresponding effect is denoted by $E_{R_i}E_{R_j}$. If we vary the accuracy of either of these measuring instruments we expect the accuracy of the combined effect to vary accordingly, that is to say the map

$$(E_{R_i}, E_{R_j}) \mapsto E_{R_i}E_{R_j}$$

is continuous. This is condition imposes conditions on $\mathcal{E}$, making it a Banach algebra. We make this heuristic argument precise following [**?**].

We need to show the product structure is compatible with the norm $\|\cdot\|_{\mathcal{E}}$ on $\mathcal{E}$. The idea that $\mathcal{E}$ has a natural action of the space of effects and we can embed effects inside the bounded operators on $\mathcal{B}(\mathcal{E})$, which comes equipped with an algebra structure given by composition of operators and the sup-norm which we will denote by $\|\cdot\|$. So, we should expect the Banach space $\mathcal{E}$ to be an algebra such that $\mathcal{L}_E F = EF$ is continuous, and similarly, $\mathcal{R}_F E = EF = \mathcal{L}_E F$ is continuous, so $\mathcal{L}_E$ and $\mathcal{R}_F$ are both bounded.

By continuity of $\mathcal{R}_F$ we have,

$$\|\mathcal{R}_E F\|_{\mathcal{E}} = \|\mathcal{L}_F E\|_{\mathcal{E}} \leq \|\mathcal{L}_F\|_{\mathcal{E}} \|E\|_{\mathcal{E}}.$$

Hence $\{\mathcal{R}_E F\}_{E\in\mathcal{E}}$ is a bounded set for every $F$. By uniform boundedness principle, pointwise boundedness implies that the set $\{\mathcal{R}_E\}_{E\in\mathcal{E}}$ is uniformly bounded, with the uniform bound $|||\mathcal{R}|||$. Hence for every $E \in \mathcal{E}$,

$$\|\mathcal{R}_E\| \leq |||\mathcal{R}|||\, \|E\|_{\mathcal{E}},$$

Hence the map

$$\mathcal{R} : E \to \mathcal{R}_E$$

defines a bounded and hence continuous linear map from $\mathcal{E}$ to $\mathcal{B}(\mathcal{E})$ of bounded linear maps on $\mathcal{E}$. $\mathcal{R}$ is a continuous algebra homomorphism. Assuming $\|1\|_{\mathcal{E}} = 1$, we have,

$$\|E\|_{\mathcal{E}} = \|\mathcal{R}_E(1)\|_{\mathcal{E}} \leqq \|\mathcal{R}_E\| \leqq |||\mathcal{R}|||\, \|E\|_{\mathcal{E}}.$$

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

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

Since $\|\mathcal{R}_1\| = 1$ we have, $\|\mathcal{R}_E(1)\|_{\mathcal{E}} = \|E\|_{\mathcal{E}} \leq \|\mathcal{R}_E\| \leq \|\mathcal{R}_1\|\,\|E\|_{\mathcal{E}} = \|E\|_{\mathcal{E}}$. Hence $\mathcal{R}$ is an isometric isomorphism and $\mathcal{R}(\mathcal{E})$ is a Banach algebra. Hence abusing notation we will denote $\mathcal{R}(\mathcal{E})$ by $\mathcal{E}$ itself.

Every preparation procedure can be followed up by a measurement. The preparation procedure together with the measurement can itself be thought of as a preparation procedure, let us denote such an operation on the ensembles by

$$(E^\dagger \mu_\rho)(F) = \mu_\rho(FE).$$

The action of $E$ on $\mathcal{E}$ corresponds to performing the experiment corresponding to the effect $E$ before, and the action of $E^\dagger$ on ensembles corresponds to the composite preparation where one performs the experiment for $E$ after initial preparation. Hence $\mathcal{E}_*$ has an action of $\mathcal{E}$ given by,

$$(\lambda E + \mu F)^\dagger \mu_\rho = \lambda E^\dagger \mu_\rho + \mu F^\dagger \mu_\rho$$

So we can think of $\mathcal{E}_*$ as a $\mathcal{E}$-module.

The structure of the state space and its relation with the space of effects gives extra structure on the space of effects. For every linear map $f$ on $\mathcal{E}$ the sesquilinear pairing $\langle\cdot|\cdot\rangle$ gives us a map on $\mathcal{E}_*$ defined by

$$\left\langle E_{R_i} | f^\dagger \mu_\rho \right\rangle = \left\langle f E_{R_i} | \mu_\rho \right\rangle.$$

For any linear operators $f, g$ on $\mathcal{E}$ we must have $(f \circ g)^\dagger = g^\dagger \circ f^\dagger$ and $(\lambda f)^\dagger = \overline{\lambda} f^\dagger$. This gives us a map on the space of effects, $\dagger : \mathcal{E} \to \mathcal{E}$.[3] Since the norm is given by $\|f\| = \sup\left\{\left|\langle fE|\mu_\rho\rangle\right|\right\}$, we have,

$$\left\|f^\dagger f\right\| = \sup\left\{\left|\left\langle f^\dagger f E|\mu_\rho\right\rangle\right|\right\} = \sup\left\{\left|\left\langle fE|f\mu_\rho\right\rangle\right|\right\}.$$

[3]Since we are looking at effects as operators, we can also obtain the †-structure by starting with $\perp$ operation on the space of effects, and use the equality between kernel of an operator and $\perp$ of image of the adjoint of the operator as the starting point. We have avoided this path since it involves more steps to relate projection to a subspace corresponding to the kernel of an operator with the operator.

For effects we have, $\langle EF|E^\dagger\mu_\rho\rangle = \mu_\rho(E^2F)$. Hence we expect the norm $\|\cdot\|$ on $\mathcal{E}$ to be such that $\|f^\dagger f\| = \|f\|^2$. An algebra $\mathcal{A}$ together with a norm $\|\cdot\|$ is said to be a Banach algebra if it is complete with respect to the norm $\|\cdot\|$ and satisfies

$$\|AB\| \leq \|A\|\|B\|, \quad \forall A, B \in \mathcal{A}$$

for every $A, B \in \mathcal{A}$. For Banach algebras it follows by triangle inequality that for fixed $B$, for any $\epsilon$, we can find $A_i$ and $A_j$ with $\|A_i - A_j\| < \epsilon/\|B\|$, and hence it follows that $\|A_iB - A_jB\| \leq \|A_i - A_j\|\|B\| < \epsilon$. Hence it follows that multiplication is jointly continuous with respect to the norm topology.

If $\mathcal{A}$ also has an involution $*$, it is called a Banach $^*$-algebra. A $C^*$-algebra is a Banach $^*$-algebra $\mathcal{A}$ which satisfies the $C^*$-identity, that is,

$$\|A^*A\| = \|A\|^2, \quad \forall A \in \mathcal{A}$$

If $\mathcal{A}$ is a $C^*$-algebra we have $\|A^*A\| = \|A\|^2 \leq \|A\|\|A^*\|$ and similarly by symmetry we will also have $\|A^*\|^2 \leq \|A\|\|A^*\|$, and hence we have

$$\|A\| \leq \|A^*\| \leq \|A\|.$$

Hence the $*$-operation preserves the norm, that is $\|A\| = \|A^*\|$ for every $A \in \mathcal{A}$ and must be continuous with respect to the topology induced by the norm.

From now on we will denote the algebra that can be used to model the space of effects by $\mathcal{A}$, called the algebra of observables which will be assumed to be a $C^*$-algebra,

$$\mathcal{E} \equiv \mathcal{A}.$$

From the discussion earlier we expect every 'yes-no' effects $E$ to satisfy $E^2 = E$. Hence the collection of all 'yes-no' effects corresponds to the collection of all projection operators on $\mathcal{A}$ denoted by $\mathcal{P}(\mathcal{A})$, and we have $\mathcal{P} \equiv \mathcal{P}(\mathcal{A}) \subset \mathcal{A}$. Since we expect that every experiment can be decomposed in terms of more elementary 'yes-no' instruments, we must expect $\mathcal{A}$ to be generated by $\mathcal{P}(\mathcal{A})$.

We expect two effects to undergo changes more often for different preparation procedures if they were closer to each other. Hence for every preparation procedure and nearby effects we expect the corresponding operational statistics to also be nearby. Any linear functionals satisfying this continuity requirement is called normal. The collection of all normal linear functionals is itself a Banach space equipped with sup-norm. We will denote the space of normal linear functionals on $\mathcal{A}$ by $\mathcal{A}_*$. Hence we have,

$$\mathcal{E}_* \equiv \mathcal{A}_*.$$

Since an ensemble assigns to each effects the corresponding operational statistics, we expect ensembles to assign numbers in the interval $[0, 1]$ for every projection in $\mathcal{A}$ and by assumption must assign the value 1 for the element 1, the norm of such functionals must be 1. Hence the state space corresponds to the space $\mathcal{S}(\mathcal{A})$ of positive norm 1 normal functionals on $\mathcal{A}$. We have, $\mathcal{S} \equiv \mathcal{S}(\mathcal{A}) \subset \mathcal{A}_*$. The elements of the state space $\mathcal{S}(\mathcal{A})$ generate $\mathcal{A}_*$.

We have a natural pairing of $\mathcal{A}$ and linear functionals on $\mathcal{A}$ given by

$$\langle\cdot|\cdot\rangle : (A, \omega) \mapsto \omega(A),$$

where $A \in \mathcal{A}$ and $\omega$ is a linear functional on $\mathcal{A}$. From the discussion earlier, we require $\langle\mathcal{A}|\mathcal{A}_*\rangle$ to be a dual pair. Hence $\mathcal{A}$ is the topological dual space of the Banach space $\mathcal{A}_*$ of normal linear functionals on $\mathcal{A}$.

A $C^*$-algebra $\mathcal{A}$ which is the dual space of a Banach space $\mathcal{A}_*$ is called a von Neumann algebra or a $W^*$-algebra. $\mathcal{A}_*$ is called the predual of $\mathcal{A}$ and we will prove later the uniqueness of predual. If we denote the space of continuous linear functionals on a Banach space $V$ by $V^*$, for a von Neumann algebra $\mathcal{A}$ we have

$$(\mathcal{A}_*)^* = \mathcal{A}.$$

The weak*-topology defined by the semi-norm on $\mathcal{A}$ by taking its modulus of linear functionals in $\mathcal{A}_*$ is called the weak topology on the von Neumann algebra $\mathcal{A}$ denoted by $\sigma(\mathcal{A}, \mathcal{A}_*)$.

## 1.2 | What are Observables?

In general a measuring instrument can simultaneously measure a collection of possible outcomes. The mathematical representatives of such generalised measuring instruments are called observables. We now extrapolate the algebraic structures for collection of effects for such possible outcomes by studying the expected properties of such measuring instruments.

### 1.2.1 | Boolean Algebras & Maps

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 $R_i \in R$,

$$\begin{aligned} R_i \wedge R_i &= R_i \\ R_i \vee R_i &= R_i. \end{aligned}$$

The composite operations involving $\vee$, $\wedge$, and $\neg$ can be figured out by simple thought. We list these properties below without attempting to be minimal. We expect the operations $\wedge$ and $\vee$ to be commutative and associative, that is, for every $R_i, R_j \in R$,

$$\begin{aligned} R_i \vee R_j &= R_j \vee R_i \\ R_i \wedge R_j &= R_j \wedge R_j \\ R_i \wedge (R_i \wedge R_k) &= (R_i \wedge R_j) \wedge R_k \\ R_i \vee (R_i \vee R_k) &= (R_i \vee R_j) \vee R_k. \end{aligned}$$

$\wedge$ 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 $\vee$, that is, for every $R_i, R_j, R_k \in R$

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

Let $\mathbb{I}$ denote the outcome which is always true when the measurement is performed. Similarly let 0 denote the outcome which never undergoes a change. $\mathbb{I}$ and 0 are called the top and bottom elements respectively. We say $\{R_i\}$ is exhaustive if $\vee_i R_i = \mathbb{I}$. We say that $R_i$ and $R_j$ are disjoint if $R_i \wedge R_j = 0$. Then we expect

$$R_i \wedge \mathbb{I} = R_i$$
$$R_i \vee 0 = R_i$$
$$R_i \vee \neg R_i = \mathbb{I}$$
$$R_i \wedge \neg R_i = 0.$$

If $R$ is closed under $\wedge$, $\vee$ and $\neg$ and contains the elements $\mathbb{I}$ and 0 it is called a Boolean algebra, denoted by $\Sigma_R$. These axioms of Boolean algebra forces the complements to be unique. Associativity ensures that we can make sense of $\wedge_k R_{i_k}$ and $\vee_k R_{i_k}$ for any finite collection $\{R_{i_k}\}$. $\Sigma_R$ 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.[4]

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 $\Sigma_R$ and $\Sigma_S$ are Boolean algebras, a mapping

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

is called a Boolean map if it preserves the Boolean operations.

An observable is a collection of effects corresponding to such measuring instruments. So, each observable should correspond to a map from a Boolean algebra to the space of effects that respects the Boolean algebra structure. An observable is a pair $(\Sigma_R, E_R)$, where $\Sigma_R$ is a Boolean algebra and $E_R$ is a map

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

[4]The notion of competion can be described intuitively by viewing Boolean algebras in lattice theoretically. We denote

$$R_i \leq R_j \Leftrightarrow R_i \wedge R_j = R_j$$

An upper bound of a finite collection $\{R_{i_k}\}$ is an element $R_j$ such that $R_{i_k} \leq R_j$ for all $k$. The supremum $\sup_k\{R_{i_k}\}$ for the collection is the lowest upper bound for the collection, that is, for any $U$ with $R_{i_k} \leq U$ for all $k$ implies $\sup_k\{R_{i_k}\} \leq U$. We have $\sup_k\{R_{i_k}\} = \vee_k R_{i_k}$ for any finite collection.

A Boolean algebra is said to be complete if every collection has a supremum with respect to the above order. For complete Boolean algebras the existence of supremum allows us to use the notation

$$\sup_k\{R_{i_k}\} \equiv \vee_k R_{i_k}.$$

for arbitrary collection $\{R_{i_k}\}$. We can similarly make define lower bounds and infimum for a collection $\{R_{i_k}\}$, and the infimum can be defined as $\inf_k\{R_{i_k}\} = \wedge_k R_{i_k}$. Existence of supremums combined with Boolean algebra requirements guarantees existence of infimums and vice-versa.

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 an observable by its corresponding Boolean map $E_R$. We will use complete Boolean algebra structure as an idealization of outcomes measurable by a single measuring instrument, since the labels used for possible outcomes in physical theories is the set of real numbers which has this completeness property.

### 1.2.2 | Stone Duality

The definition of an observable is closely related to the algebraic structures on Boolean algebras. Understand the structure of Boolean algebras can be useful for understanding the mathematical modelling of observables. We now prove Stone's characterisation of Boolean algebras, and describe the notion of a Stonean topological space.

The relation between Stonean spaces and von Neumann algebras, in §**??**, gives us a characterisation of observables in quantum theories.

#### Stone Space of $\Sigma$

Let $\Omega$ be a set. Equipped with union, intersection and complements, the powerset $\mathcal{P}(\Omega)$ of $\Omega$ is a Boolean algebra, where the empty set and the whole set $\Omega$ are the bottom and top elements respectively. Stone's characterisation of Boolean algebras says that every Boolean algebra can be thought of as a Boolean algebra of subsets of a set.

We now start with an abstract completel Boolean algebra $\Sigma$. Which for the sake of heuristics and intuition maybe assumed to be the Boolean algebra $\Sigma_R$ for simultaneously measurable outcomes $R \equiv \{R_i\}$ for heuristics and intuition.

An ideal of $\Sigma$ is a set $I \subset \Sigma$ containing the bottom element of $\Sigma$, and closed under the operation $\vee$, such that

$$R_i \wedge R_j \in I, \quad \forall R_i \in I,\ R_j \in \Sigma.$$

If the Boolean algebra $\Sigma_R$ corresponds to a collection of simultaneously measurable outcomes $R \equiv \{R_i\}$, and we can think of each $R_i$ as an instrument in itself. The outcome $R_i \vee R_j$ describes the outcome where a prepared system is passed through $R_i$ and $R_j$ in parallel. The closure under $\vee$ says that all such parallel instruments also represent an outcome in $I_R$.

Similarly, $R_i \wedge R_j$ describes the outcome in which both $R_i$ and $R_j$ are true. Hence the condition $R_i \in I_R$, $R_j \in \Sigma_R$ implies $R_i \wedge R_j \in I_R$, says that $I_R$ contains all such elementary outcomes, describing all of the most accurate instruments.

#### The Space of Maximal Ideals

The notion of an ideal of a Boolean algebra says that outcomes belonging to the ideal can be expressed by a combination of most accurate instruments. Since the ideal need not be the whole Boolean algebra, it might be the case that the collection does not include all such 'most accurate instruments'.

The notion of maximal ideals helps us distinguish such instruments. An ideal of $I \subset \Sigma$ is called a proper ideal if $I \neq \Sigma$. Maximal ideal is a proper ideal which is not contained in any other proper ideal. Clearly proper ideals cannot contain the top element $\mathbb{I}$, because that would mean all the most accurate instruments would be contained in the ideal. if $R_i \in I$ then $I$ cannot contain $\neg R_i$, because otherwise

$$R_i \vee \neg R_i = \mathbb{I} \notin I$$

By Zorn's lemma we may assume every proper ideal is contained in a maximal ideal.

**Theorem 1.1.** *$I \subset \Sigma$ is a maximal ideal iff $\forall R_i \in \Sigma$, either $R_i \in I$ or $\neg R_i \in I$.*

**Lemma 1.2.** *Every proper ideal is contained in a maximal ideal.*

The idea for construction of $S(\Sigma)$ is to view the elements of $\Sigma$ as functions. Let $\mathbf{2}$ be the set $\{0, 1\}$.

Similarly, A field of a set is a collection of subsets of the set which is closed under union, intersections and complements. Stone's representation theorem or Stone duality for Boolean algebras suggests that every Boolean algebra $\Sigma$ is a field of sets of a space, called the Stone space $S(\Sigma)$ for $\Sigma$.

The completion properties of $\Sigma$ are closely related to the topological properties of $S(\Sigma)$

**Lemma 1.3.** *Every maximal ideal is the kernel of a Boolean homomorphism.*

A topological space is called totally disconnected if every open set is the union of all its subsets that are simultaneously open and closed. A Hausdorff topological space $\Omega$ that is compact and totally disconnected is called a Stonean space.

## 1.3 | Heisenberg's Approach

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 $\mu_\rho$ is a map $\mu_\rho : \mathcal{A} \to \mathbb{R}$. The composite map is

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

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

### 1.3.1 | 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 $\Omega$, 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 $\rho$ is modeled by a positive Radon measure $\mu_\rho$ on $\Omega$. The probability measure the state $\mu_\rho$ 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 $\Omega$ and corresponds to the set of all points in $\Omega$ 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 $\rho$ 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 $\Omega$ the value 1 correspond to effects in this model. The collection of all characteristic functions on $\Omega$ 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 $\Omega$. If we denote the convex set of regular probability measures on $\Omega$ 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 $l^1(\Omega)$.[5]

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.

[5]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.

### 1.3.2 | 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.

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

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

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,

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

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,

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

where $\lambda$ 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,

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

Two square integrable functions are equivalent if they are same almost everywhere with respect to $\lambda$. 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 1.4. (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 [9] 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 | Structure of $C^*$-Algebras

For this subsection we will assume $\mathcal{A}$ is a $C^*$-algebra with identity. The goal of this section is to use the tools for the study of complex numbers to study elements of $\mathcal{A}$. We begin by introducing the notion of spectrum which allows us to import and use complex analytic tools for the study of elements of $\mathcal{A}$. The starting point is the group of invertible elements of $\mathcal{A}$ denoted by $\mathcal{G}(\mathcal{A})$, this is an open set in $\mathcal{A}$.[6]

### 2.1 | Analysis with the Spectrum

The openness of $\mathcal{G}(\mathcal{A})$ allows us to introduce the tools of complex analysis to the study of elements of $\mathcal{A}$ so these tools do not mess with the already existing structures on $\mathcal{A}$. The complex analytic tools are to be introduced from the complex numbers to $\mathcal{A}$, such that the analytic functions correspond to maps on $\mathcal{A}$ with suitable analyticity.

We introduce the complex analytic tools via the notion of resolvent. For every element $A$ in $\mathcal{A}$ define a map $R_A$ from $\mathbb{C}$ to $\mathcal{A}$ by

$$R_A(\lambda) = (A - \lambda)^{-1}.$$

Let $R(A)$ be the set of all $\lambda \in \mathbb{C}$ for which the above definition makes sense inside $\mathcal{A}$, that is, $(A - \lambda)$ has an inverse in $\mathcal{A}$. $R(A)$ is called the resolvent of $A$. Since $\mathcal{G}(\mathcal{A})$ is open, it follows

[6] If $A \in B_1(\mathbb{I})$, that is, $\|\mathbb{I} - A\| < 1$ then we can make sense of the sum $B = \sum_{\mathbb{N}}(A - \mathbb{I})^i$ since $\sum_{\mathbb{N}} \|A - \mathbb{I}\|^i$ converges as a geometric series. By the completeness of Banach algebras, $B$ is an element of $\mathcal{A}$, and by computing the product $AB$ and $BA$ it follows that $A^{-1} = B$. For any invertible element $A \in \mathcal{A}$, by taking product with the neighborhood we obtain an open neighborhood of $A$ of invertible elements $AB_\epsilon(\mathbb{I}) \subset \mathcal{G}(\mathcal{A})$. Which proves that $\mathcal{G}(\mathcal{A})$ is open.

that $R(A)$ is open. The spectrum of $A$ in $\mathcal{A}$ defined to be the closed set

$$\sigma(A) = \mathbb{C} \backslash R(A) \subset \mathbb{C}$$

By rearranging we can check that for all $\lambda, \mu \in R(A)$,

$$\frac{R_A(\lambda) - R_A(\mu)}{\lambda - \mu} = R_A(\lambda) R_A(\mu).$$

Hence $R_A$ defines an analytic function with $R'_A(\lambda) = -R_A(\lambda)^2$ from the open set $R(A)$ of $\mathbb{C}$ to the $C^*$-algebra $\mathcal{A}$. We note that if $\sigma(A)$ is empty, then $R_A$ is an analytic function on all of $\mathbb{C}$ and as $\lambda$ tends to infinity we have, $\|R_A(\lambda)\| = \|(\lambda^{-1}A - 1)^{-1}\|/|\lambda|$. Hence it follows that $\|R_A(\lambda)\| \to 0$ as $\lambda$ tends to infinity.

For any bounded linear functional $\varphi$ by the continuity and linearity we have

$$\begin{aligned} (\varphi \circ R_A)'(\lambda) &= \lim_{\lambda \to \mu} \Big( \tfrac{\varphi(R_A(\lambda)) - \varphi(R_A(\mu))}{\lambda - \mu} \Big) \\ &= \varphi\Big( \lim_{\lambda \to \mu} \tfrac{(R_A(\lambda) - R_A(\mu))}{(\lambda - \mu)} \Big) = \varphi(R_A(\lambda)^2). \end{aligned}$$

The composition $\varphi \circ R_A$ defines a bounded entire function. By Liouville's theorem every bounded entire function corresponds to a constant, and since $R_A(\lambda) \to 0$, it follows that $\varphi \circ R_A \equiv 0$ which is absurd since $\varphi$ was arbitrarily chosen. Hence $\sigma(A)$ must be non-empty for every $A \in \mathcal{A}$.

Suppose every element of $\mathcal{A}$ is invertible, and has a non-constant element, say $A$. Then by assumption $A - \lambda$ is always invertible in $\mathcal{A}$. This implies that $\sigma(A)$ is empty which cannot happen as discussed above. Hence we have proved the following lemma;

**Lemma 2.1. (Gelfand-Mazur)** *If $\mathcal{G}(\mathcal{A}) = \mathcal{A}$ as Banach algebra then $\mathcal{A} \cong \mathbb{C}$.*

Although the notion of spectrum maybe defined on general algebra, the import of complex analytic tools is possible due to the topology on the Banach algebra. Spectrum depends on the Banach algebra. If $\mathcal{B} \subseteq \mathcal{A}$ then more elements may be invertible in $\mathcal{A}$ and hence we must have $R_{\mathcal{B}}(A) \subseteq R_{\mathcal{A}}(A)$ or equivalently

$$\sigma_{\mathcal{B}}(A) \supseteq \sigma_{\mathcal{A}}(A)$$

However when $\mathcal{A}$ is a $C^*$-algebras $(A - \lambda)^{-1}$ belongs to the algebra generated by the elements $A, A^\dagger$ and $\mathbb{I}$. Hence the notion of spectrum is an intrinsic property for $C^*$-algebra.

The map $p(z) \mapsto p(A)$ an algebra homomorphism from the space of polynomials on $\mathbb{C}[z]$ to the algebra $\mathbb{C}[A]$ generated by $\mathbb{I}$ and $A$. For any fixed complex number $\lambda$, by fundamental theorem of algebra, the polynomial can be decomposed as the product,

$$p(z) - \lambda = a_N \prod_{i \leq N} (z - \lambda_i)$$

where $\lambda_i$ are the roots of the polynomial. Since $p(z) \to p(A)$ is an algebra homomorphism it follows that

$$p(A) - \lambda = a_N \prod_{i \leq N} (A - \lambda_i)$$

Since the product of invertible elements is always invertible, we have that $p(A) - \lambda$ is invertible if and only if each of $A - \lambda_i$ is invertible. Similarly if $(A - \lambda)$ is invertible then $(A^\dagger - \lambda^*) = (A - \lambda)^\dagger$ is invertible. Hence we have proved the following theorem

**Theorem 2.2. (Spectral Mapping Theorem)**

$$\forall p \in \mathbb{C}[z],\ \sigma(p(A)) = p(\sigma(A)),\ \sigma(A^\dagger) = \sigma(A)^*.$$

We now relate the norm $\|A\|$ of an element $A$ to the 'size' of its spectrum. For every $\lambda$ with $\|A\| < |\lambda|$, we can make sense of the sum $-\sum_i A^i/\lambda^{i+1}$ and it is the inverse of $(A - \lambda)$. Hence we have

$$\rho(A) = \sup_{\lambda \in \sigma(A)} \big\{|\lambda|\big\} \subseteq \big[-\|A\|, \|A\|\big].$$

$\rho(A)$ is called the spectral radius of $A$.

By the spectral mapping theorem, whenever $\lambda$ is $\sigma(A)$, it follows that $\lambda^n$ is in $\sigma(A^n)$. Hence we have $\rho(A)$.

In Banach algebras we must have $\|A^k\| \leq \|A\|^k$ for all $A$. Hence the limit infimum must be bounded from above by $\|A\|$.

$$\|A\| \leq \liminf_{k\to\infty} \|A^k\|^{\frac{1}{k}}.$$

Since

If $A_{\epsilon^+} = (\rho(A)+\epsilon)^{-1}A$ we must have $\|A_{\epsilon^+}\| = \|A\|/(\rho(A)+\epsilon)$ and $\rho(A_{\epsilon^+}) \leq \rho(A)/(\rho(A)+\epsilon) < 1$. Since $\|A^2\| \leq \|A\|^2$ for all $A \in \mathcal{A}$ the sequence $A_\epsilon^i$ converges to zero. Hence $\|A_{\epsilon^+}^j\| < 1$ for $j$ large enough. Similarly for $A_{\epsilon^-} = (\rho(A) - \epsilon)A$, we have $\|A_{\epsilon^-}^j\| > 1$ for $j$ large enough. This means that for large enough $j$,

$$(\rho(A) - \epsilon)^j < \big\|A^j\big\| < (\rho(A) - \epsilon)^j.$$

Since choice of $\epsilon$ was arbitrary we have proved the following formula for spectral radius;

**Theorem 2.3. (Spectral Radius Formula)**

$$\forall A \in \mathcal{A},\ \rho(A) = \lim_n \big\|A^n\big\|^{\frac{1}{n}}.$$

An element $A$ of a Banach $*$-algebra $\mathcal{A}$ is said to be self-adjoint if $A^\dagger = A$. In a $C^*$-algebra self-adjoint element $A$ satisfies $\|A^2\| = \|A^\dagger A\| = \|A\|^2$ and by applying spectral radius formula we get

$$A = A^\dagger \Rightarrow \|A\| = \rho(A).$$

Since $\sigma(A^\dagger) = \sigma(A)^*$ the spectrum of a self-adjoint element consists of real numbers. A positive element $A$ of a $C^*$-algebra, denoted as $A \geq 0$ is a self-adjoint operator whose spectrum consists of non-negative real numbers. $\geq$ gives rise to an order on $C^*$-algebras, which we will call the spectral order.

### 2.1.1 | The Gelfand Transform

While the notion of spectrum brings in structures of complex numbers to a $C^*$-algebra, linear functionals take the structures on $C^*$-algebra to the complex numbers. We are interested in studying the case when all of the $C^*$-structure is preserves, and for the product can only be preserved if $\mathcal{A}$ is a commutative $C^*$-algebra since product structure on $\mathbb{C}$ is commutative.

Let $\mathcal{A}$ be a commutative $C^*$-algebra, a linear functionals on $\mathcal{A}$ is called a character if it is also a $^*$-algebra homomorphism. The collection of all characters on $\mathcal{A}$ is called the

Gelfand spectrum of $\mathcal{A}$ denoted by $\chi(\mathcal{A})$. If $\varphi$ is a character then the linearity and $^*$-algebra homomorphism requirements can be summarised by

$$\varphi(AB + \lambda C^\dagger) = \varphi(A)\varphi(B) + \lambda\big(\varphi(C)\big)^*.$$

This definition immediately implies that $\varphi(\mathbb{I}) = \varphi(AA^{-1}) = \varphi(A)\varphi(A)^{-1} = 1$. Hence $\varphi(A)$ is non-zero whenever $A$ is invertible. Hence the existence of inverse of $(A - \lambda)$ implies that $\varphi(A) - \lambda$ is non-zero. Hence $\varphi(A) \in \sigma(A)$. Since we have

$$\sigma(A) \subseteq \big[-\|A\|, \|A\|\big],$$

it follows that

$$|\varphi(A)| \leq \|A\|, \ \ \forall A \in \mathcal{A}.$$

Hence every character $\varphi$ is a bounded linear functionals with $\|\varphi\| = \sup_{\|A\|\leq 1} |\varphi(A)| = 1$. $\chi(\mathcal{A})$ is a closed subset of the unit ball in the space of continuous linear functionals with respect to the weak$^*$ topology. By Banach-Alaoglu theorem, $\chi(\mathcal{A})$ is a compact Hausdorff space.

**Theorem 2.4.** *$\chi(\mathcal{A})$ is a compact Hausdorff space.*

The space of continuous functions on $\chi(\mathcal{A})$, denoted by $C(\chi(\mathcal{A}))$ contains the topological data of the space $\chi(\mathcal{A})$, and equipped with the sup-norm it is a commutative $C^*$-algebra. Due to the weak$^*$-continuity of characters, the map $\varphi \mapsto \varphi(A)$ is a continuous function on $\chi(\mathcal{A})$ for each $A \in \mathcal{A}$. Since every character $\varphi$ is a $^*$-homomorphism it follows that

$$\begin{aligned} \Gamma : \mathcal{A} &\to C(\chi(\mathcal{A})) \\ A &\mapsto \Gamma(A) \end{aligned} \tag{Gelfand transform}$$

where $\Gamma(A)$ is a continuous function on $\chi(\mathcal{A})$ given by $\varphi \to \varphi(A)$, is a $^*$-homomorphism. $\Gamma$ is called the Gelfand transform. We now relate $\mathcal{A}$ and $C(\chi(\mathcal{A}))$ topologically, by relating the spectrum of an element $A$ in $\mathcal{A}$ with the range of its Gelfand transform $\Gamma(A)$;

If $\lambda$ is in the range of the continuous function $\Gamma(A)$ there must exist a character $\varphi$ such that $\Gamma(A)(\varphi) = \lambda$, and hence $\varphi(A - \lambda) = 0$. Since characters can never be zero for invertible elements $A - \lambda$ must not be invertible. Hence

$$\sigma(A) \subseteq \big\{\Gamma(A)(\varphi)\big\}_{\varphi \in \chi(\mathcal{A})}.$$

If $\lambda$ is in the spectrum of $A$, that is, if $A - \lambda$ is not invertible, then $(A - \lambda)$ is contained in the maximal ideal defined by $\mathcal{I}_\lambda = (A - \lambda)\mathcal{A}$. The maximal ideal $\mathcal{I}_\lambda$ cannot contain any invertible elements. Consider the quotient map

$$\pi_\lambda : \mathcal{A} \to \mathcal{A}/\mathcal{I}_\lambda.$$

The quotient space $\mathcal{A}/\mathcal{I}_\lambda$ is itself a Banach$^*$-algebra with the norm

$$\big\|B + \mathcal{I}_\lambda\big\| = \inf_{I \in \mathcal{I}_\lambda} \big\{\big\|B + I\big\|\big\}.$$

$\pi_\lambda$ is a $C^*$-algebra homomorphism. $\mathcal{A}/\mathcal{I}_\lambda$ cannot have any non-invertible elements, because otherwise it would be contained in a proper ideal $\mathcal{J} \subset \mathcal{A}/\mathcal{I}_\lambda$ in which case $\pi^{-1}(\mathcal{J}) \supset \mathcal{I}_\lambda$ will be a proper ideal contradicting the maximality of $\mathcal{I}_\lambda$.

Since every element of $\mathcal{A}/\mathcal{I}_\lambda$ is invertible, by the Gelfand-Mazur lemma there exists a $^*$-isomorphism, $\psi_\lambda : \mathcal{A}/\mathcal{I}_\lambda \to \mathbb{C}$. Their composition $\varphi_\lambda = \psi_\lambda \circ \pi_\lambda$ will also be a $^*$-homomorphism. Hence $\varphi_\lambda$ is a character. By construction the image of $A - \lambda$ under $\pi_\lambda$ must be zero, and hence we have $\varphi_\lambda(A - \lambda) = 0$. Hence we have proved that

$$\sigma(A) \subseteq \{\Gamma(A)(\varphi)\}_{\varphi \in \chi(\mathcal{A})} \subseteq \sigma(A).$$

By definition of the sup-norm we have, $\|\Gamma(A)\|_{\text{sup}} = \sup_{\varphi \in \chi(\mathcal{A})}\{|\Gamma(A)(\varphi)|\} = \rho(A)$.

**Theorem 2.5. (Gelfand-Naimark)** $\Gamma$ *is an isometric $^*$-isomorphism.*

**Proof**

Since $\mathcal{A}$ is a $C^*$-algebra we have for every self-adjoint element $\|A^\dagger\| = \|A\|$. To reduce the problem to the case of self-adjoint elements, for any $A \in \mathcal{A}$ we consider the element $A^\dagger A$ which is always self-adjoint. Hence we have, $\|A^{2^k}\| = \|(A^{2^k-1})^\dagger(A^{2^k-1})\| = \|A^{2^k-1}\|^2$ by repetition it follows that $\|A^{2^k}\|^{1/2^k} = \|A\|$. By the spectral radius formula we have

$$\|\Gamma(A)\|_{\text{sup}} = \rho(A) = \lim_k \|A^{2^k}\|^{\frac{1}{2^k}} = \|A\|.$$

The Gelfand transformation is an isometry.

Note that $\Gamma(\mathcal{A})$ defines a self-adjoint subalgebra of $C(\chi(\mathcal{A}))$ with identity, and also separates points of $\chi(\mathcal{A})$ since by definition two characters $\varphi$ and $\psi$ can only be different if there exists some $A$ with

$$\Gamma(A)(\varphi) = \varphi(A) \neq \psi(A) = \Gamma(A)(\psi).$$

By Stone-Weierstrass theorem, it follows that $\Gamma(\mathcal{A})$ is a dense subset of $C(\chi(\mathcal{A}))$. Hence the image of $\Gamma$ is dense in $C(\chi(\mathcal{A}))$, and $\Gamma$ is a isometric $^*$-isomorphism. □

Spectral theory relates properties of elements of a $C^*$-algebra $\mathcal{A}$ with the properties of its spectrum which is a compact subset of complex numbers. This in turn allows us to import many of the tools for the study of complex numbers to studying the elements of $\mathcal{A}$. In particular, self-adjoint elements behave similarly to real numbers.

Let $C^*[A]$ be the commutative $C^*$-algebra generated by the identity and $A$. By Hahn-Banach theorem every character on $C^*[A]$ can be extended to a linear functional on $\mathcal{A}$, with the same norm. Hence we can describe the norm of an element in terms of linear functionals on $\mathcal{A}$. For any self-adjoint element,

$$\|A\| = \sup_{\omega \in \mathcal{S}} |\omega(A)|, \quad \forall A = A^\dagger,$$

where $\mathcal{S}$ is the unit ball of the space of all linear functionals on $\mathcal{A}$.

Every continuous function on the Gelfand spectrum $\chi(C^*[A])$ gives rise to an operator in $C^*[A]$. We will denote the image of a continuous function $f$ under the Gelfand inverse by $f(A)$. The operator $A$ corresonds to the identity function on the Gelfand spectrum;

$$f(A) = A, \text{ if } f(\lambda) = \lambda,$$

for all $\lambda \in \chi(C^*[A])$. The constant function corresponds to the identity in $\mathcal{A}$. For compositions we obtain the corresponding element by iteratively applying the above described process.

$$(g \circ f)(A) = g(f(A)).$$

For a general commutative $C^*$-algebra $\mathcal{A}$ without identity, $\chi(\mathcal{A})$ will be a locally compact space, and the elements of $\mathcal{A}$ correspond to continuous functions, vanishing at infinity. By taking function that are 'mostly' constant on the Gelfand spectrum we can obtain approximate identities for every $C^*$-algebra.

## 3 | Structure of $W^*$-Algebras

Mathematical objects can be studied by studying function from the mathematical object which respect the structures we wish to study. The study of a mathematical object is particularly simplified if these function spaces are well-behaved and are themselves simple mathematical objects. The Riesz-Frechet representation theorem states that for a Hilbert spaces $(\mathcal{H}, \langle\cdot|\cdot\rangle_{\mathcal{H}})$ any continuous linear functional $\varphi$ on $\mathcal{H}$ corresponds to a vector $f_\varphi$ in $\mathcal{H}$ such that

$$\varphi(\nu) = \langle f_\varphi|\nu\rangle_{\mathcal{H}}, \quad \forall \nu \in \mathcal{H},$$

such that $\|\varphi\|_{\mathcal{H}^*} = \|f_\varphi\|_{\mathcal{H}}$. A Hilbert space, as a Banach space, is its own dual Banach space. This allows for an intrinsic study of its linear and topological structures, and hence Hilbert spaces are better understood than general Banach spaces.

The von Neumann algebras are special $C^*$-algebras similarly to how Hilbert spaces are special Banach spaces, and the structures of interest to us is the Banach $^*$-structure. $\mathcal{A}$ is called a von Neumann algebra if it is the dual Banach space of a Banach space $\mathcal{A}_*$

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

$\mathcal{A}_*$ called a predual of $\mathcal{A}$. By definition, $\mathcal{A}$ is the dual Banach space of $\mathcal{A}_*$, and hence the norm $\|\cdot\|_{\mathcal{A}}$ can be thought of as the operator norm,

$$\|A\|_{\mathcal{A}} = \sup_{\|\omega\|_{\mathcal{A}_*}=1} \left\{ \big|\omega(A)\big| \right\}.$$

This suggests that the norm $\|\cdot\|_{\mathcal{A}}$ also describes the overall topological data about $\mathcal{A}$ contained in the predual $\mathcal{A}_*$. We can however also study the data contained in individual elements of $\mathcal{A}_*$, and extract more data about the structure of $\mathcal{A}$.

### 3.1 | The Weak Topology on $\mathcal{A}$

The goal is to construct a weaker topology on $\mathcal{A}$ using individual elements of the predual $\mathcal{A}_*$, and use this weaker topology to extract more data about $\mathcal{A}$. Heuristically, this weaker topology is a stronger relation between the von Neumann algebra and its predual. By constructing a locally convex topology on $\mathcal{A}$ using the elements of $\mathcal{A}_*$, we can use tools such as the separation theorems to study the relation between the elements of $\mathcal{A}$.

Let $\mathcal{A}$ be a von Neumann algebra, and let $\mathcal{A}_*$ be its predual. By definition $\mathcal{A}$ is the collection of continuous linear functionals on $\mathcal{A}_*$. Hence for every element $\omega$ of the predual $\mathcal{A}_*$ we obtain a semi-norm on $\mathcal{A}$ given by,

$$p_\omega(A) = |\omega(A)|.$$

The locally convex topology induced by this collection of semi-norms is called the weak topology on $\mathcal{A}$, denoted by $\sigma(\mathcal{A}, \mathcal{A}_*)$. This is a topology on $\mathcal{A}$ induced by semi-norms constructed from $\mathcal{A}_*$. A net $A_\alpha$ converges to $A$ if and only if $\omega(A_\alpha)$ converges to $\omega(A)$ for all $\omega$ in $\mathcal{A}_*$.[7]

[7] The difference between the norm topology and $\sigma(\mathcal{A}, \mathcal{A}_*)$ is similar to the difference between uniform convergence and almost everywhere convergence with respect to Radom measures on the space. Heuristically the topology $\sigma(\mathcal{A}, \mathcal{A}_*)$ forgets the topological data about $\mathcal{A}$ which cannot be 'seen' by the elements of $\mathcal{A}_*$.

### 3.1.1 | Weak Continuity & Normality

We can now use the Hahn-Banach separation theorems, for closed convex sets of $\mathcal{A}$. By the Banach-Alaoglu theorem the closed unit sphere of $\mathcal{A}$ is compact with respect to $\sigma(\mathcal{A}, \mathcal{A}_*)$. By the Krein-Milman theorem, there exist extreme points in the unit sphere, hence we have proved the following lemma,

**Lemma 3.1.** *Every von Neumann algebra contains an identity element.*

Von Neumann algebras inherit the notion of spectrum of elements for $C^*$-algebras. We now describe how the topology $\sigma(\mathcal{A}, \mathcal{A}_*)$ on $\mathcal{A}$ respects the properties described in terms of spectrum. This allows us keep using the tools of spectral theory for the study of von Neumann algebras.

The notion of positivity of elements. We start by discussing how certain subsets of $\mathcal{A}$ remain stable with respect to the topology $\sigma(\mathcal{A}, \mathcal{A}_*)$ under limits.

**Theorem 3.2.** *If $A \in \mathcal{A}$, then*

## 3.2 | The Lattice of Projections

We now prove that for a von Neumann algebra $\mathcal{A}$, the Gelfand spectrum of any maximally abelian $C^*$-subalgebras of $\mathcal{A}$ is a Stone space. By Stone's duality, every Stone space gives rise to a Boolean algebra, and conversely every Boolean algebra can be thought of as a Boolean algebra of subsets of a Stonean space. Since observables are Boolean maps into $\mathcal{A}$, we obtain a characterisation of observables when $\mathcal{A}$ is a von Neumann algebra.

### 3.2.1 | Weakly-Closed Ideals

The starting point for the characterisation observables is again a discussion on ideals, and relate maximal ideals to the collection of elementary effects. We now use the $\sigma(\mathcal{A}, \mathcal{A}_*)$-topology to describe closures, as opposed to the norm topology, which was the case for the discussion of the Gelfand transform.

For any projection $E$ in $\mathcal{A}$, the collection $\mathcal{A}E$ is a left-ideal of $\mathcal{A}$. Since the left multiplication is $\sigma(\mathcal{A}, \mathcal{A}_*)$-continuous it follows that $\mathcal{L}_E = \mathcal{A}E$ is a $\sigma(\mathcal{A}, \mathcal{A}_*)$-closed left-ideal of $\mathcal{A}$. We now characterise all left-ideals in terms of such projections.

Suppose $\mathcal{L}$ is a $\sigma(\mathcal{A}, \mathcal{A}_*)$-closed right-ideal of $\mathcal{A}$. Let $\mathcal{N}_\mathcal{L}$ be the largest $C^*$-algebra contained in $\mathcal{L}$. It is given by

$$\mathcal{N}_\mathcal{L} = \mathcal{L} \cap \mathcal{L}^*.$$

Since the $^*$-operation is $\sigma(\mathcal{A}, \mathcal{A}_*)$-continuous, $\mathcal{N}_\mathcal{L}$ is $\sigma(\mathcal{A}, \mathcal{A}_*)$-closed. Hence $\mathcal{N}_\mathcal{L} \subset \mathcal{A}$ is a von Neumann subalgebra of $\mathcal{A}$. Let $E_\mathcal{L}$ be the identity element of the von Neumann subalgebra $\mathcal{N}_\mathcal{L}$. Hence we must have $E_\mathcal{L}^2 = E_\mathcal{L}$. Since $\mathcal{N}_\mathcal{L}$ is a left-ideal of $\mathcal{A}$ it follows that

$$E_\mathcal{L}\mathcal{A} \subset \mathcal{L}.$$

Let $L$ be any element of $\mathcal{L}$. Since $L^\dagger L$ is a self-adjoint element, it belongs to $\mathcal{N}_\mathcal{L}$. Since $E_\mathcal{L}$ is the identity element of $\mathcal{N}_\mathcal{L}$ we have, $E_\mathcal{L}L^\dagger L E_\mathcal{L} = E_\mathcal{L}L^\dagger L = L^\dagger L E_\mathcal{L} = L^\dagger L$. Hence $(1 - E_\mathcal{L})L^\dagger L(1 - E_\mathcal{L}) = E_\mathcal{L}L^\dagger L E_\mathcal{L} - E_\mathcal{L}L^\dagger L - L^\dagger L E_\mathcal{L} + L^\dagger L = 0$. Hence we have

$$L(1 - E_\mathcal{L}) = 0, \quad \forall L \in \mathcal{L}.$$

Hence we must have

$$\mathcal{L} = \mathcal{A}E_\mathcal{L}.$$

By the uniqueness of identity element, $E_\mathcal{L}$ is uniquely determined. Hence every $\sigma(\mathcal{A}, \mathcal{A}_*)$-closed right-ideal $\mathcal{L}$ of $\mathcal{A}$ is of the form $\mathcal{A}E_\mathcal{L}$ for a unique projection $E_\mathcal{L}$ in $\mathcal{A}$.

Let $\mathcal{P}(\mathcal{A})$ be the set of all projections in $\mathcal{A}$. Every projection $E$ gives rise to a $\sigma(\mathcal{A}, \mathcal{A}_*)$-closed left ideal,

$$\mathcal{L}_E = \mathcal{A}E$$

Let $\{E_\alpha\}_{\alpha \in I}$ be a set of projections in $\mathcal{A}$. Consider the $\sigma(\mathcal{A}, \mathcal{A}_*)$-closed left ideal generated by the $\{\mathcal{A}E_\alpha\}_{\alpha \in I}$, there must exist a projection, denoted by $\vee_\alpha E_\alpha$, such that the ideal is given by $\mathcal{A}(\vee_\alpha E_\alpha)$.

Similarly, since arbitrary intersection of closed sets is also closed the intersection of the collection $\{\mathcal{A}E_\alpha\}_{\alpha \in I}$ is a closed ideal, and hence corresponds to a projection, denoted by $\wedge_\alpha E_\alpha$. Clearly we have

$$\mathcal{A}(\wedge_\alpha E_\alpha) \subseteq \mathcal{A}E_\alpha \subseteq \mathcal{A}(\vee_\alpha E_\alpha).$$

Hence we have,

$$(\wedge_\alpha E_\alpha) \leq E_\alpha \leq (\vee_\alpha E_\alpha).$$

If $E_\alpha \leq E$, then we also have $\mathcal{A}E_\alpha \subseteq \mathcal{A}E$. Hence $\mathcal{P}(\mathcal{A})$ inherits an order from $\sigma(\mathcal{A}, \mathcal{A}_*)$-closed left-ideals of $\mathcal{A}$. Since we can make sense of $\vee_\alpha E_\alpha$ and $\wedge_\alpha E_\alpha$ for any collection $\{E_\alpha\}$ it follows that $\mathcal{P}(\mathcal{A})$ equipped with $\leq$ has the structure of a complete lattice. Hence we have proved the following theorem:

**Theorem 3.3.** *If $\mathcal{A}$ is a von Neumann algebra then $\mathcal{P}(\mathcal{A})$ is a complete lattice.*

The complete lattice $\mathcal{P}(\mathcal{A})$ is also called the von Neumann lattice, and from the point of view of physics the elements of $\mathcal{P}(\mathcal{A})$ corresponds to 'yes-no' measuring instruments.

### 3.2.2 | The Spectral Theorem

Let $A$ be an element in $\mathcal{A}$. By $\sigma(\mathcal{A}, \mathcal{A}_*)$-continuity of $\mathcal{R}_A$, the set $\mathcal{I}_A \equiv \ker(\mathcal{R}_A)$ is a $\sigma(\mathcal{A}, \mathcal{A}_*)$-closed left-ideal. Hence there exists some projection $E_{\mathcal{I}_A}$ such that $\mathcal{I}_A = \mathcal{A}E_{\mathcal{I}_A}$. By construction of $E_{\mathcal{I}_A}$ above it follows that $E_{\mathcal{I}_A}$ is the largest projection with $E_{\mathcal{I}_A}A = 0$. Let $L_A \equiv 1 - E_{\mathcal{I}_A}$. $E_{\mathcal{I}_A}$ corresponds to the identity element of the largest $C^*$-algebra in $\mathcal{I}_A$. Let $E$ be a projection such that $EA = A$ then

$$L_A \equiv (1 - E_{\mathcal{I}_A}) \leq E.$$

The projection $L_A$ is called the left support of $A$. We can similarly define right support of $A$. starting with $\mathcal{J}_A = \ker(\mathcal{I}_A)$. If $A$ is self-adjoint then $A^\dagger = A$, and we have $\mathcal{N}_{\mathcal{I}_A} = \mathcal{N}_{\mathcal{J}_A}$, and it follows that $L_A = R_A$, and is called the support of the self-adjoint element $A$.

If $A$ is self-adjoint element of the von Neumann algebra $\mathcal{A}$. Let $W[A]$ be the von Neumann algebra generated by $A$. Since $W[A]$ is a von Neumann algebra, it must contain the identity element. If $E_{W[A]}$ is the corresponding identity element, we observe that

$$E_{W[A]}A = A.$$

Hence it follows that $S_A \leq E_{W[A]}$, and that $(E_{W[A]} - S_A)A = 0$. Since left-multiplication is $\sigma(\mathcal{A}, \mathcal{A}_*)$-continuous it follows that $E_{W[A]} - S_A$ annihilates the weak closure of the collection of polynomials in $A$. Since $S_A$ is the smallest such projection, it follows that

$$E_{W[A]} = S_A, \quad \forall A = A^\dagger \in \mathcal{A}.$$

#### Spectral Measure

For each real $\lambda$, the element

**THEOREM 3.4.** *For a self-adjoint element $A$ in $\mathcal{A}$, there exists a system of projections $E_A(\lambda)$ for $\lambda \in \mathbb{R}$ called the spectral resolution of identity such that,*

$$E_A(\lambda) \leq E_A(\mu) \quad \forall \lambda \leq \mu.$$

$$\lambda_i \to \lambda \Rightarrow \omega(E_A(\lambda_i)) \to \omega(E_A(\lambda)), \quad \forall \omega \in \mathcal{A}_*.$$

**PROOF**

Clearly for $\lambda > \|A\|$, we have $E_A(\lambda) = 1$

**Theorem 3.5. (Stone)** *Let $\Sigma_R$ be a Boolean algebra. Then $\Sigma_R$ is isomorphic to a subalgebra of $\mathcal{P}(\mathcal{F}_I(\Sigma_R))$. Moreover the map $\varphi : \Sigma_R \to \mathcal{P}(\mathcal{F}_I(\Sigma_R))$ defined by*

$$\varphi : R_i \mapsto \{F \in \mathcal{F}_I(\Sigma_R) \mid R_i \in F\}$$

*is a Boolean algebra embedding.*

**Theorem 3.6. (Stone)** *Every Boolean algebra $\Sigma_R$ is isomorphic to a certain field of sets $R$.*

A compact Hausdorff space $K$ is called a Stonean space if the closure of every open set if also open. Let $f \in C(K)$, without loss of generality, we can assume $f$ is a positive function. We can divide up the range of

**Lemma 3.7.** *Let $X$ be Stonean, then $f \in C(K)$ can be uniformly approximated by finite linear combinations of projections.*

**Theorem 3.8.** *Let $K$ be a compact Hausdorff space, such that every bounded increasing directed set of real valued non-negative functions $\{f_\alpha\} \subset C(K)$ has a least upper bound in $C(K)$. Then $K$ is Stonean.*

A typical measuring instrument can however measure a collection of simultaneously measurable outcomes, and as discussed before such instruments correspond to Boolean maps into $\mathcal{A}$.

Such maps will be called $\sigma$-additive. $\sigma$-additivity allows us to develop tools of measure theory with states. If $E_R : \Sigma_R \to \mathcal{A}$ is an observable then a state $\mu_\rho$ gives us a composite map,

$$\mu_\rho \circ E_R \equiv \mu_\rho^R : \Sigma_R \to [0, 1],$$

such that $\mu_\rho^R(0) = 0$, $\mu_\rho^R(E^\perp) = 1 - \mu_\rho^R(E)$ and whenever $\{E_{R_i}\}_{i \in \mathbb{N}}$ are mutually orthogonal,

$$\mu_\rho^R(\vee_i R_i) = \sum_{i \in \mathbb{N}} \mu_\rho^R(E_{R_i}).$$

The $\sigma$-additivity allows us to think of states as probability measures for each observable.

We are interested in characterising observables, which are maps,

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

where the image $E_A(\Sigma_A)$ is Boolean subalgebra of $\mathcal{P}(\mathcal{A})$. Hence $E_A(\Sigma_A)$ must generate a commutative $C^*$-algebra. For an exhaustive collection of outcomes of a measurement, we expect the join of the corresponding effects must be the identity of $\mathcal{A}$, since atleast one of the result will have to occur.

## 3.3 | The Predual

We now construct the predual $\mathcal{A}_*$ for a von Neumann algebra $\mathcal{A}$ using the properties of the von Neumann algebra $\mathcal{A}$. We have proved that von Neumann algebras are complete with respect to the order coming from positivity of operators, that is, given any bounded increasing net $\{A_\alpha\}$ of positive elements, $A_\alpha$ converges to its supremum in $\mathcal{A}$. We now show that the collection of all linear functionals which respect this completeness property, is the predual $\mathcal{A}_*$.

**Theorem 3.9. (Sakai)** *Let $\varphi$ be a positive linear functional on a von Neumann algebra $\mathcal{A}$. If $\varphi$ is normal then $\varphi$ is $\sigma(\mathcal{A}, \mathcal{A}_*)$-continuous.*

### Proof

Let $E_\alpha$ be an increasing net of projections with the supremum $E \equiv \vee_\alpha E_\alpha$. If $\omega$ is a normal linear functional. Since the product is $\sigma(\mathcal{A}, \mathcal{A}_*)$-continuous, it follows that then the map $A \mapsto \omega(AE_\alpha)$, as a composition of is $\sigma(\mathcal{A}, \mathcal{A}_*)$-continuous.

Then for all $A$ with $\|A\| = 1$, we have by Cauchy-Schwartz inequality,

$$\left|\omega\big(A(E_\alpha - E)\big)\right| \leq \omega(A(E - E_\alpha)A^\dagger)^{\frac{1}{2}}\omega(E - E_\alpha)^{\frac{1}{2}} \leq \omega(1)^{\frac{1}{2}}\omega(E - E_\alpha)^{\frac{1}{2}}.$$

Hence $\omega(AE)$ is a uniform limit of the directed set $\omega(AE_\alpha)$

### 3.4 | Effects & Ensembles in $\mathcal{A}$

The lattice of projections $\mathcal{P}(\mathcal{A})$ comes equipped with the operations $\wedge$, $\vee$ as discussed earlier, and $\mathcal{P}(\mathcal{A})$ also has the complement operation where $\neg E = 1 - E$. Since the outcomes of instruments of physical experiments are usually labelled by real numbers, and since every interval contains a rational number, we may assume for all practical purposes that the collection of simultaneously measurable outcomes is separable. So, instead of completeness with respect to arbitrary sets, it is sufficient to develop tools for countable collections.

A spectral measure in a von Neumann algebra $\mathcal{A}$ is a $\sigma$-complete Boolean map $E_A$ from a $\sigma$-complete Boolean algebra $\Sigma_A$ into its image in $\mathcal{P}(\mathcal{A})$. If the Boolean algebra $\Sigma_A \equiv \mathcal{B}(\mathbb{R})$, then the spectral measure $E_A$ defines a self-adjoint element in $\mathcal{A}$ by

$$\omega(A) := \int_{\mathbb{R}} \lambda d\omega(E_A(\lambda)),$$

for every element $\omega \in \mathcal{A}_*$. The normality of the linear functional $\omega$ ensures that $\mu_\omega(\lambda) = \omega(E_A(\lambda))$ is a complex Radon measure on $\mathbb{R}$. If $\omega$ is a normal state, then $\mu_\omega$ will be a probability measure. Since the elements in $\mathcal{A}_*$ is separating for $\mathcal{A}$,

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

makes sense as an element in $\mathcal{A}$.

The instruments corresponding to 'yes-no' outcomes correspond to the elements of the projection lattice of a von Neumann algebra $\mathcal{A}$. Every observable corresponds to a self-adjoint element of the von Neumann algebra $\mathcal{A}$.

# 4 | Representation Theory

Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{B}(\mathcal{H})$ be the algebra of all bounded operators on $\mathcal{H}$. Since Hilbert spaces come equipped with a lot of useful structure, it is convinient to view $C^*$-algebras as subalgebras of such operator algebras.

A $*$-homomorphism between two $C^*$-algebras $\mathcal{A}$ and $\mathcal{B}$ is a mapping

$$\pi : \mathcal{A} \to \mathcal{B},$$

such that $\pi(\alpha A + \beta B) = \alpha\pi(A) + \beta\pi(B)$, $\pi(AB) = \pi(A)\pi(B)$ and $\pi(A^*) = (\pi(A))^*$. It is an algebra homomorphism which also preserves the $*$-operation. Given a $*$-homomorphism $\pi : \mathcal{A} \to \mathcal{B}$, we have, $\pi(A)\pi(A^{-1}) = \pi(AA^{-1}) = 1$. So, $\pi$ maps invertible elements to invertible elements, and $\pi(A^{-1}) = \pi(A)^{-1}$. Hence we observe that

$$\sigma(\pi(A)) \subset \sigma(A).$$

This immediately tells us that, for self-adjoint operators, $\|\pi(A)\| = \mathrm{rad}(\pi(A)) \leq \mathrm{rad}(A) = \|A\|$. Since $A^*A$ is self-adjoint, we have,

$$\|\pi(A)\|^2 = \|\pi(A^*)\pi(A)\| = \|\pi(A^*A)\| \leq \|A^*A\| = \|A\|^2.$$

A representation $(\mathcal{H}, \pi)$ of a unital $C^*$-algebra $\mathcal{A}$ is a $*$-homomorphism,

$$\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$$

which is unital, that is, $\pi(1) = 1$ for some Hilbert space $\mathcal{H}$. Two representation $(\mathcal{H}_1, \pi_1)$ and $(\mathcal{H}_2, \pi_2)$ of an algebra $\mathcal{A}$ are said to be equivalent if there exists a unitary operator $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that

$$\pi_1(A) = U\pi_2(A)U^*,$$

for all $A \in \mathcal{A}$.

If $\pi$ is an isomorphism between $\mathcal{A}$ and $\pi(\mathcal{A})$ it is called a faithful representation. Suppose we have a faithful representation of $\mathcal{A}$ then by injectivity we have, $\ker(\pi) = \{0\}$. There exists $\pi^{-1}$ from the range of $\pi$ into $\mathcal{A}$.

$$\|A\| = \|\pi^{-1}(\pi(A))\| \leq \|\pi(A)\| \leq \|A\|.$$

So whenever $\pi$ is a faithful representation, then for every $A \in \mathcal{A}$,

$$\|A\| = \|\pi(A)\|.$$

If $(\mathcal{H}, \pi)$ is a representation of $\mathcal{A}$, a subspace $\mathcal{H}_1$ of $\mathcal{H}$ is said to be invariant under $\pi$ if $\pi(A)\mathcal{H}_1 \subseteq \mathcal{H}_1$ for all $A \in \mathcal{A}$. If $\mathcal{H}_1$ is closed and $P_{\mathcal{H}_1}$ is the orthogonal projection with range $\mathcal{H}_1$ then the invariance implies,

$$P_{\mathcal{H}_1}\pi(A)P_{\mathcal{H}_1} = \pi(A)P_{\mathcal{H}_1}.$$

for all $A \in \mathcal{A}$. Hence,

$$\begin{aligned}
\pi(A)P_{\mathcal{H}_1} &= (P_{\mathcal{H}_1}\pi(A^*)P_{\mathcal{H}_1})^* \\
&= (\pi(A^*)P_{\mathcal{H}_1})^* \\
&= P_{\mathcal{H}_1}\pi(A).
\end{aligned}$$

for all $A \in \mathcal{A}$. Hence $\mathcal{H}_1$ is invariant under $\pi$ if and only if, $\pi(A)P_{\mathcal{H}_1} = P_{\mathcal{H}_1}\pi(A)$ for all $A \in \mathcal{A}$. If we define $\pi_1$ by,

$$\pi_1(A) = P_{\mathcal{H}_1}\pi(A)P_{\mathcal{H}_1},$$

then $(\mathcal{H}_1, \pi_1)$ is a representation of $\mathcal{A}$. It is called a subrepresentation of $(\mathcal{H}, \pi)$. This procedure of going to subrepresentation gives a decomposition of $\pi$. If $\mathcal{H}_1$ is invariant under $\pi$ then so is $\mathcal{H}_1^\perp$. Setting $\mathcal{H}_2 = \mathcal{H}_1^\perp$ one can define a second subrepresentation. Now the original Hilbert space $\mathcal{H}$ can be written as a direct sum, $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ and each operator $\pi(A)$ then decomposes as a direct sum $\pi(A) = \pi_1(A) \oplus \pi_2(A)$. So the representation can be written as $(\mathcal{H}, \pi) = (\mathcal{H}_1, \pi_1) \oplus (\mathcal{H}_2, \pi_2)$.

Given a family of representations $(\mathcal{H}_\alpha, \pi_\alpha)_{\alpha \in I}$ of $\mathcal{A}$ the direct sum of representations $\mathcal{H}_\alpha$ is defined as follows,

$$\mathcal{H} = \bigoplus_{\alpha \in I} \mathcal{H}_\alpha,$$

consisting of vectors of the form $\varphi = \{\varphi_\alpha\}_{\alpha \in I}$ such that $\lim_F [\sum_{\alpha \in F} \|\varphi_\alpha\|^2] < \infty$ where $F$ is a finite subset of $I$. The purpose of this definition is so that norm is definable nicely. This Hilbert space together with the representation map,

$$\pi = \bigoplus_{\alpha \in I} \pi_\alpha,$$

is called direct sum of representations $\{(\mathcal{H}_\alpha, \pi_\alpha)\}_{\alpha \in I}$, denoted by, $\sum_{\alpha \in I}^{\oplus} \{(\mathcal{H}_\alpha, \pi_\alpha)\}$. The operators $\pi(A)$ on $\mathcal{H}$ are bounded because $\|\pi_\alpha(A)\| \leq \|A\|$ for each $\alpha \in I$.

A representation is trivial if $\pi(A) = 0$ for every $A \in \mathcal{A}$. These are uninteresting representations. A representation can however have a trivial part.

$$\mathcal{D} = \{\varphi \in \mathcal{H} \mid \pi(A)\varphi = 0 \, \forall A \in \mathcal{A}\}.$$

It follows that $\pi_\mathcal{D} = P_\mathcal{D} \pi P_\mathcal{D} = 0$ where $P_\mathcal{D}$ is the projection onto the subspace $\mathcal{D}$. A representation $(\mathcal{H}, \pi)$ is non degenerate if $\mathcal{D} = \{0\}$.

A vector $|\Omega\rangle$ in a Hilbert space $\mathcal{H}$ is called cyclic for $\mathcal{A}$ if $\{A|\Omega\rangle\}_{A \in \mathcal{A}}$ is dense in $\mathcal{H}$. A cyclic representation of $\mathcal{A}$ is a triple $(\mathcal{H}, \pi, |\Omega\rangle)$ where $(\mathcal{H}, \pi)$ is a representation of $\mathcal{A}$ and $|\Omega\rangle$ is a cyclic for $\pi(\mathcal{A})$.

Let $(\mathcal{H}, \pi)$ be a nondegenerate representation of $\mathcal{A}$. Take a maximal family of nonzero vectors $|\{\Omega_\alpha\rangle\}_{\alpha \in I}$ in $\mathcal{H}$ such that,

$$\langle \pi(A)\Omega_\alpha | \pi(B)\Omega_\beta \rangle = 0,$$

for all $A, B \in \mathcal{A}$ and $\alpha \neq \beta$. Define, $\mathcal{H}_\alpha = \overline{\{\pi(A)|\Omega_\alpha\rangle\}}_{A \in \mathcal{A}}$. This is an invariant subspace of $\mathcal{H}$. Define $\pi_\alpha = P_{\mathcal{H}_\alpha} \pi P_{\mathcal{H}_\alpha}$ where $P_{\mathcal{H}_\alpha}$ is projection onto $\mathcal{H}_\alpha$. Then by construction each $\mathcal{H}_\alpha$ are mutually orthogonal and hence the representation $(\mathcal{H}, \pi)$ is of the form,

$$\mathcal{H} = \bigoplus_{\alpha \in I} \{(\mathcal{H}_\alpha, \pi_\alpha)\}$$

So every nondegenerate representation can be written as a direct sum of a family of cyclic subrepresentations. If no invariant subspaces the representation $(\mathcal{H}, \pi)$ of $\mathcal{A}$ is called irreducible. If $\mathcal{A}$ be a self-adjoint algebra of operators on Hilbert space $\mathcal{H}$,

## 4.1 | Gelfand-Naimark-Segal Construction

Denote the dual space of $\mathcal{A}$ that is, the set of all continuous linear functionals over $\mathcal{A}$ by $\mathcal{A}^*$. The norm of a functional $f$ over $\mathcal{A}$ is defined by, $\|f\| = \sup_{\|A\|=1}\{|f(A)|\}$. A linear functional $\omega$ 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 $A = 0$. If $\omega_1$ and $\omega_2$ are two states then clearly,

$$\omega = \lambda\omega_1 + (1-\lambda)\omega_2,$$

is also a state for all $\lambda \in [0,1]$. The set of states is a convex subset of $\mathcal{A}^*$.

If $\omega$ is a positive linear functional over $\mathcal{A}$ then we can define a sesquilinear form, $\varrho(B, A) = \omega(B^*A)$, that is, $\varrho(\mu A, \lambda B) = \overline{\mu}\lambda\varrho(A, B)$, and $\varrho(A, B) = \overline{\varrho(B, A)}$. Since $\omega$ a positive linear functional we have,

$$\varrho(\lambda A - B, \lambda A - B) \geq 0.$$

On expanding it, we obtain,

$$|\lambda|^2\varrho(A, A) - \overline{\lambda}\varrho(A, B) - \lambda\varrho(B, A) + \varrho(B, B) \geq 0$$

By letting $\lambda = \varrho(A, B)/\varrho(A, A)$ we obtain, $0 \leq [|\varrho(A, B)|^2/\varrho(A, A)^2]\varrho(A, A) - 2[|\varrho(A, B)|^2/\varrho(A, A)] + \varrho(B, B)$. This gives us,

$$|\varrho(A, B)|^2 \leq B_\omega(A, A)B_\omega(B, B).$$

If $\omega$ is a positive linear functional then it satisfies the Cauchy-Schwarz inequality,

$$|\omega(A^*B)|^2 \leq \omega(A^*A)\omega(B^*B).$$

If $\omega_1$ and $\omega_2$ are two positive linear functionals we write $\omega_1 \geq \omega_2$ if $\omega_1 - \omega_2$ is positive. This gives an order on positive linear functionals. If $\omega_1$ and $\omega_2$ are two states over $\mathcal{A}$ and $0 < \lambda < 1$ then $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$ is also a state such that $\omega \geq \lambda\omega_1$ and $\omega \geq (1-\lambda)\omega_2$. The set of all states is a convex subset of $\mathcal{A}^*$ and we will denote it by $S(\mathcal{A})$. The extreme points of this convex set are called pure states. They are such that $\omega > \lambda\omega_1$ iff $\omega_1 = \omega$.

Given a closed two-sided ideal $\mathcal{J} \subseteq \mathcal{A}$, the quotient algebra is defined by,

$$\mathcal{A}_\mathcal{J} = \mathcal{A}/\mathcal{J} = \big\{[A] = A + J \;\mid\; J \in \mathcal{J}\big\}$$

with the norm, $\|[A]\| = \inf_{J\in\mathcal{J}}\{\|A + J\|\}$ the algebra $\mathcal{A}_\mathcal{J}$ is a $C^*$-algebra.

The Gelfand-Naimark-Segal theorem constructs for a given $C^*$-algebra $\mathcal{A}$ and a state $\omega$ a representation of the algebra of observables $\mathcal{A}$ on the set of bounded operators $\mathcal{B}(\mathcal{H})$ for some $\mathcal{H}$.

**Theorem 4.1. (Gelfand-Naimark-Segal)** *Let $\omega$ be a state on a unital $C^*$-algebra $\mathcal{A}$ then there exists a cyclic representation $(\mathcal{H}_\omega, \pi_\omega, |\Omega\rangle)$ of unit norm such that*

$$\omega(A) = \langle\Omega|\pi_\omega(A)\Omega\rangle_\omega$$

*for all $A \in \mathcal{A}$. The representation is unique in the sense that if $(\mathcal{H}, \pi, |\Omega_\varphi\rangle)$ is a cyclic representation such that, $\varphi(A) = \langle\Omega_\varphi|\pi(A)\Omega_\varphi\rangle_\omega$ then there exists a unique unitary operator $U : \mathcal{H} \to \mathcal{H}_\omega$, such that,*

$$\pi_\omega(A) = U\pi(A)U^*,$$

*and $U|\Omega_\varphi\rangle = [1]$.*

**Proof**

Given a state $\omega$ on $\mathcal{A}$ one considers the set in $\mathcal{A}$ defined by, $\mathcal{J}_\omega = \{A \mid \omega(A^*A) = 0\}$. By Cauchy-Schwarz inequality whenever $A \in \mathcal{J}_\omega$ for any $B \in \mathcal{A}$ we have,

$$|\omega((BA)^*BA)|^2 = |\omega(C^*A)|^2 \leq \omega(C^*C)\omega(A^*A) = 0,$$

where $C = B^*BA$. So, $BA \in \mathcal{J}_\omega$. So $\mathcal{J}_\omega$ is an ideal. Factorizing $\mathcal{A}$ by $\mathcal{J}_\omega$ an inner product is introduced on the quotient space $\mathcal{A}_{\mathcal{J}_\omega}$ defined by,

$$\big\langle [A] | [B] \big\rangle_\omega := \omega(A^*B).$$

where $[A]$ and $[B]$ denote the equivalence classes determined by $A$ and $B$ respectively. The new vector space is completed by adding all the Cauchy sequences and we denote the Hilbert space by $\mathcal{H}_\omega$. On this Hilbert space we have the representation of the algebra $\mathcal{A}$,

$$\pi_\omega : \mathcal{A} \to \mathcal{B}(\mathcal{H}_\omega),$$

defined by,

$$\pi_\omega(A)[B] \equiv [AB].$$

Let $[I] = [1]$. The expectation of any observable can then be written as,

$$\omega(A) = \langle [1] | A[1] \rangle_\omega.$$

A state on the algebra can be represented as a vector in some Hilbert space. A vector $\varphi \in \mathcal{H}$ is said to be cyclic for $\mathcal{A}$ if the closure of $\mathcal{A}\varphi$ is same as $\mathcal{H}$. A vector $\varphi$ is separating for $\mathcal{A}$ if $A\varphi = 0$ implies $A = 0$ for all $A \in \mathcal{A}$. The vector $[1]$ is cyclic for $\mathcal{A}$.

For every state $\omega$ on an algebra $\mathcal{A}$ there exists a cyclic representation $(\mathcal{H}_\omega, \pi_\omega, [1])$.

$$\omega(A) = \langle [1] | \pi_\omega(A)[1] \rangle_\omega, \quad \forall\, A \in \mathcal{A}.$$

If there is another cyclic representation $(\mathcal{H}, \pi, |\Omega\rangle)$ then define a map, $U\pi(A)|\Omega\rangle = \pi_\omega(A)[1]$. This is an isometry with an inverse, hence it extends to a unitary map. $\square$

The above construction of a representation for a $C^*$-algebra using the given state is called the GNS construction. We observe that the representation is faithful if the state is faithful.

**Theorem 4.2.** *Let $A \in \mathcal{A}$ be a self-adjoint element. Then there exists a cyclic representation $(\mathcal{H}, \pi, |\Omega_\varphi\rangle)$ of $\mathcal{A}$ such that*

$$\|\pi(A)\| = \|A\|$$

**Proof**

The norm of a self-adjoint operator is the same as its spectral radius,

$$\|A\| = \rho(A) = \sup_{\lambda \in \sigma(A)} \{|\lambda|\}$$

Let $\lambda$ be this maxima, using this we can define a functional on the algebra generated by $A$ and identity. Defined by,

$$\varphi_0 : \alpha A + \beta 1 \mapsto \alpha\lambda + \beta$$

It also maps $\varphi_0(1) = 1$. So the linear functional is also a state. Now by Hahn-Banach theorem this can be extended to a state $\varphi$ on $\mathcal{A}$ with $\varphi(1) = \varphi_0(1) = 1 = \|\varphi\|$. The GNS representation for this state satisfy,

$$\|A\| = |\varphi_0(A)| = |\varphi(A)| = |\langle\Omega_\varphi|\pi_\varphi(A)\Omega_\varphi\rangle| \leq \|\pi_\varphi(A)\|.$$

□

Now, to each element $A_i$, we have a representation such that $\|\pi_i(A_i)\| = A_i$. Using these representations we can form a direct sum representation. Let $\{A_i\}_{i\in I}$ be a dense set in $\mathcal{A}$, For each $i \in I$ we have a representation $(\mathcal{H}_i, \pi_i)_{i\in I}$ such that $\|\pi_i(A_i)\| = \|A_i\|$ because $\|\pi_i(A_i^* A_i)\| = \|A_i^* A_i\|$ and C$^*$ identity. Thus the direct sum will be such that,

$$\|\pi(A)\| = \|A\|$$

for all $A \in \mathcal{A}$. If $\mathcal{A}$ is separable, $I$ can be assumed to be countable set, and hence we can assume the representation $(\mathcal{H}, \pi)$ to be separable representation. Every separable $C^*$-algebra can be represented on a separable Hilbert space.

### 4.1.1 | Normal Representations

## 4.2 | Bounded Operators on a Hilbert Space

### 4.2.1 | Concrete $W^*$-Algebras

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

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

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

[10] S Sakai, $C^*$ and $W^*$-Algebras. Springer, 1971