universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Mathematical aspects of the Operator Algebraic Approach to Quantum Field Theory“

verfasst von / submitted by

Bharath Govind Ron

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna, 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Mathematik
Betreut von / Supervisor:	ao. Univ.-Prof. Mag. Günther Hörmann

Acknowledgements

I would like to extend my heartfelt gratitude to my parents for their unwavering support and love throughout my academic journey. They have been the source of hope whenever I felt lost in chaos.

I am profoundly thankful to my advisor, Dr. Günther Hörmann, whose unique approach allowed me to explore the subject matter in my own style. Their keen insights and gentle nudges were instrumental in keeping me on track, preventing me from delving too deep into rabbit holes. I believe their mentorship has contributed more to my mathematical development than all my previous years combined, instilling in me a newfound confidence in my own abilities.

Table of Contents

Table of Contents ii

1 Algebraic Quantum Theory 1
1.1 Generalised Probabilistic Theories . . . 1
1.1.1 Effects & Ensembles . . . 2
1.1.1.1 Locally Convex Embeddings . . . 3
1.1.2 Algebraic Quantum Theory . . . 5
1.1.2.1 Banach Spaces & Algebras . . . 5
1.1.2.2 †-Structure & W^*-Algebras . . . 10
1.2 Convexity in Locally Convex Spaces . . . 14
1.2.1 Conjugate Spaces & Conjugate Duality . . . 14
1.2.1.1 The Hahn-Banach Theorem . . . 14
1.2.1.2 Conjugate Duality . . . 17
1.2.2 Locally Convex Topologies . . . 21
1.2.2.1 Supporting Functions & Indicators . . . 21
1.2.2.2 Minkowski Functionals & Duality . . . 23
1.2.3 Conjugate Duality for Banach Space . . . 26
1.2.3.1 Bidual of Banach Spaces . . . 27
1.3 The σ-Weak Topology & Lattice Structures . . . 28
1.3.1 The Bidual of W^*-Algebras . . . 28
1.3.1.1 $\mathcal{A}$-Module Structure on $\mathcal{A}^*$ . . . 29
1.3.1.2 Arens Product on $\mathcal{A}^{**}$ . . . 29
1.3.2 The Conjugate Pairing $\langle \cdot \mid \cdot \rangle_{\mathcal{A}_*}$ . . . 31
1.3.2.1 Tensor Product of W^*-Algebras . . . 31
1.3.2.2 The Lattice of Projections . . . 33
1.3.2.3 Structure of the Dual Space . . . 35
1.3.2.4 The Predual: Sakai's Theorem . . . 36
1.4 Spectral Theory of W^*-Algebras . . . 37
1.4.1 What are Observables? . . . 38
1.4.1.1 Complete Boolean Algebras . . . 38
1.4.1.2 Stone Duality . . . 40
1.4.2 The Gelfand Spectrum & Stone Space . . . 44
1.4.2.1 The Gelfand Transform . . . 47
1.4.2.2 Spectral Theorem . . . 51
Appendices for the Chapter . . . 55
1.A Quantum Theory on Hilbert Spaces . . . 55
1.A.1 Necessity for Quantum Formalism . . . 55
1.A.1.1 Heisenberg's Approach . . . 57
1.A.2 The Space of Bounded Operators on $\mathcal{H}$ . . . 59
1.A.2.1 Von Neumann Subalgebras . . . 60
1.A.2.2 Lattice of Closed Subspaces . . . 60
1.A.3 Representation of W^*-Algebras . . . 62

2 Symmetries & The Spectrum Condition 66
2.1 Axioms of Local Quantum Physics . . . 66

2.2 The Space-Time Translation Group . . . 68
2.2.1 Analyticity of Fourier Transforms . . . 68
2.2.1.1 Paley-Wiener-Hörmander Theorem . . . 69
2.2.2 The Future Cone & Positivity of Energy . . . 71
2.3 The Spectrum Condition & Analyticity . . . 72
2.3.1 Representation of the Translation Group . . . 72
2.3.1.1 Stone's Theorem & Analyticity . . . 74
2.3.2 The Reeh-Schlieder Theorem . . . 75
2.4 Tomita-Takesaki Modular Theory . . . 79
2.4.1 σ-Finite von Neumann Algebras . . . 79
2.4.2 The Modular Group . . . 81
2.4.2.1 The Modular Relation . . . 84
2.4.2.2 Tomita's Theorem . . . 88
2.4.3 Borchers-Wiesbrock Fundamental Relations . . . 91
Appendices for the Chapter . . . 96
2.A The Haar Measure . . . 96

Summaries **103**

Bibliography **104**

104

1 | Algebraic Quantum Theory

The physical systems of interest in quantum theories are those whose observed phenomena can be described by the notion of an *observable*. The starting idea here is that the world can be described by sentences of the form *the observable R has the value R_i*. The next step is to articulate this idea in 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 or integrable functions from a manifold Ω, called the phase space, to the set of real numbers:

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

As we will discuss later, this model of an observable fails when the application domain is expanded to simultaneously include both discrete and continuous cases. Although quantum theory has the same idea, it is formulated in a different way that makes it compatible with the extended application domain. We now reconstruct quantum theory from ground up, and in doing so we end up exposing the unnecessary ideas of the classical formalism.

1.1 | Generalised Probabilistic Theories

The discussion here will be closest to the Ludwig school [27]-[29], with some stuff borrowed from the quantum logic literature and algebraic quantum theory. This can be thought of as a modern formulation of Heisenberg's original idea. Hans Primas [34] characterizes approaches to the formulation of quantum theory into three main categories: the quantum logic approach, algebraic approach, and the convex state-space approach. In this section we show the necessity of W^*-algebras for quantum theory by starting from simple operational postulates about how we observe the world around us.

The convex state-space approach was developed by Ludwig and his peers. This approach starts with measuring instruments and preparation instruments and is the most general approach among the three. A physical theory is in some sense interpreted from outside in terms of pre-theories not belonging to the theory in question itself. Usually when one tries to formulate quantum theory one starts with a pre-theory such as classical mechanics and then *quantizes* the theory. This makes the theory messy and the underlying physical ideas hidden and unclear. In Ludwig's approach the pre-theory is the theory of preparation and measuring instruments. The assumed structure in this instrumentalist view of physics can be motivated with simple thought experiments making it foundationally sound, and hence we will adopt this approach to the formulation of quantum theory. The construction and behavior of the instruments will not be of interest to us. Any changes occurring in the

instruments during *measurements* will be accepted as objective outcomes. In this point of view, the fundamental notions of quantum theory have to be defined operationally in terms of macroscopic instruments and prescriptions for their application. Quantum mechanics is then interpreted entirely in terms of such instruments and outcomes which are the changes occurring to instruments.

1.1.1 | Effects & Ensembles

From this instrumentalist or operational point of view, the notion of *state* can be defined in terms of the preparation procedure. A preparation procedure is characterized by the kind of system it prepares. The measuring instrument that is capable of undergoing changes when an experiment is performed for evaluating a collection of possible outcomes. The possible results of such an experiment are called outcomes of the experiment. The observable change in the instrument is called an effect.

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

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

The numbers $\mu(\rho|R_i)$ are called operational statistics.

Two completely different preparation procedures may give the same operational statistics, that is, they may give the same probabilities for every experiments. Such preparation procedures must be considered equivalent operationally. An equivalence class of preparation procedures yielding the same operational statistics for experiments is called an ensemble. Similarly, there may be two measuring instruments that have the same chances of undergoing a change for similarly prepared systems. Such measuring instruments must be operationally considered to be equivalent. An equivalence class of change for measuring instruments is called an effect. An effect is the equivalence class of all instruments that undergo a change for the same possible outcome. By considering equivalence classes we have obtained the structure of sets. We obtained the basic mathematical structure for modelling preparation procedures and measuring instruments in terms of the ensembles and effects they describe without introducing any new physical law. Since we are considering equivalence classes, the construction and behavior of the instruments are irrelevant to the discussion.

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 statistics

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

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

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

On a historic note, this is where we see theoretical physics branch off into two different groups. The first group does in the direction of understanding the structure of the space of ensembles, following Schrödinger, Dirac, Feynman, and others; since state spaces in classical theories are also closely related to topology and geometry this group has been able to make a lot of progress in developing the standard approach to quantum field theories behind the celebrated standard model by utilising the geometric intuition behind classical theories and by understanding and generalising gauge theoretic ideas of Maxwell and others. The other group branched out in direction of understanding the structure of the space of effects, following Heisenberg-Jordan-Born, von Neumann, Haag-Kastler-Araki, and others. This group has been able to make much less progress comparitively, possibly because the space of effects is very strange and unfamiliar compared to the more familiar and intuitive classical theories. This lack of pre-theories such as classical theory has left this group stranded in the middle of nowhere, and the ones who did pursue this new path are forced to build up starting from almost nothing. The second path is not popular possibly due to unfamiliarity. We will however follow the second path due to its stronger foundations and embrace the risk of losing the centuries old intuition, for classical theories.[1]

We expect every preparation procedure to be a part of the universe, and every state of the world can be thought of as being prepared by the universe itself. Hence in the abstract mathematical framework, *reality* can be identified with the set of all preparation procedures. Since we can only describe or understand reality by performing experiments on the world, the measuring instruments are our links to *object reality*.

1.1.1.1 | Locally Convex Embeddings

In order to be able to work with effects and ensembles we now embed them inside sets with mathematical structure that respects the expected operational relations. We now describe how the collection of effects and the collection of ensembles inherit mathematical structures. In our case the operational requirement is that we should be able to make sense of taking mixtures of effects and ensembles. Suppose we can prepare systems in different ensembles by varying settings of a preparation procedures, then by adjusting the settings we should be able to prepare systems in mixtures of ensembles. We expect the space of ensembles to have the structures to describe mixtures.

Let us assume that we can form such mixtures by changing the settings on the preparation procedures. If such adjustments are parametrized by real numbers, such as rotating a knob, then the notion of convexity needed for formalising comes from the space of real numbers.

[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.

The space of real numbers is a convenient space because its convexity structure is related to a lot of other structures, such as its topology, its vector space structure, and so on. As discussed in [16], such a notion of mixing corresponds mathematically to the notion of convex combination. Hence real vector space structure acts as a good starting point. Hence we may expect the set $\mathcal{S}$ to have the structures for making sense of convexity and we expect that 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 to make sense of linear 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 at least as large as the real numbers. It is important to note that preparation procedures 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 procedures and measuring instruments to ensembles and effects is a transition from the real physical world to the abstract mathematical world. It should also be noted that it does not make sense to *prepare* closed systems, one has to assume such systems start off in some state a priori.

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

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

where ρ_i are ensembles and α_i are in a field $\mathbb{K}$ and I is a finite set. Since the notion of convexity is inherited from the notion of convexity on the real line, we assume that the field $\mathbb{K}$ contains the field of real numbers. Any linear combination of such maps also belongs to $\mathcal{A}_*$, it follows that $\mathcal{A}_*$ is a $\mathbb{K}$-vector space. Denote by $\mathcal{A}$ the set of maps,

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

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

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

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

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

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

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

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

that coincides with μ for effects and ensembles, that is, $\langle E_{R_i}|\mu_\rho\rangle = \mu(R_i|\rho)$.

The deeper reason for the introduction of the complex number field is that it is algebraically closed and allows for factorisation. We will later see that the complex structure allows us mathematically to make sense of the physical idea that an instrument which undergoes a change for a collection of events can be replaced by a collection of instruments for those simultaneously measurable outcomes. Such a decomposition of a given effect into a collection of smaller, simultaneously measurable effects is made precise using complex structures by the spectral theory.

1.1.2 | Algebraic Quantum Theory

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

1.1.2.1 | Banach Spaces & Algebras

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

We assume the existence of the unit 1, which corresponds to the instrument which is always true for any preparation procedure the unit must be unique since any other such element must be operationally equivalent. Operationally, the unit corresponds to the *existence* of the system itself, since it must always be true. Similarly we assume the existence of the unique 0 element which outputs *no* for every preparation procedure, and can be thought of as a faulty instrument which never undergoes any changes.

Note that the less accurate instrument undergoes a change for more events than intended. A faulty instrument never undergoes any change, for any event. It has no possibility to be inaccurate, and hence correspond to the most accurate instruments.

Banach Space Structure

For any collection of instruments for a given outcome we can obtain a relation between such instruments by their accuracy. Consider two instruments which can undergo changes under some *similar* situations. This means that we may have some redundancies, or common inaccuracies, that is, they can measure some common possible outcomes, or alternatively, they have differing accuracy for measuring the common outcome. 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.

In our case, it is easier to model the inaccuracy. If changes in accuracy are modelled using an addition like operation, then the notion of inaccuracy must correspond to a function f such that,

$$f(E_R + E_S) \leqq f(E_R) + f(E_S).$$

Such functions are called sub-linear maps. The notion of inaccuracy of instruments naturally gives rise to triangle inequality like properties.

We can describe this *inaccuracy function* more conveniently as a supremum. Since inaccuracy represents the maximal likelihood of undergoing a change, it must correspond to the supremum of undergoing a change over all preparation procedures. That is to say, the accuracy in the real world can only depend on *reality* which we expect to be closely related to preparation procedures. This gives us a definition in terms of supremum over preparation procedures, and can be defined on all of $\mathcal{A}$,

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

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

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

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

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

$$\big\|\omega\big\|_{\mathcal{A}_*} \coloneqq \sup_{E\in\mathcal{E}} \Big[\big|\langle E|\omega\rangle\big|\Big],\ \ \omega \in \mathcal{A}_*$$

where the supremum is taken over the collection of effects which corresponds to the convex embedded subset $\mathcal{E}$ of $\mathcal{A}$. We can hence require both $\mathcal{A}$ and $\mathcal{A}_*$ to be a complex, normed spaces. Under the idealization that there exist instruments of every level of accuracy, $\mathcal{E}$ must be closed with respect to the norm $\|\cdot\|_{\mathcal{A}}$. Since $\mathcal{E}$ generates $\mathcal{A}$ we also assume $\mathcal{A}$ to be closed with respect to taking limits with respect to the norm $\|\cdot\|_{\mathcal{A}}$. Hence we may assume that the spaces $\mathcal{A}_*$ and $\mathcal{A}$ are Banach spaces. This allows us to talk about taking limits and to do mathematical analysis with ensembles and effects. In particular we now have access to Banach space tools such as the Hahn-Banach theorem.[2]

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

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

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

extends $\mu(\cdot|\cdot)$ while respecting the real-linearity properties. It follows that $\langle \mathcal{A}|\mathcal{A}_*\rangle$ is a dual pair. Since $\mathcal{A}$ and $\mathcal{A}_*$ separate each other, any bounded linear functional on $\mathcal{A}_*$ can be obtained as a limit of linear combinations of functionals corresponding to effects. By completeness assumptions on $\mathcal{A}$ we may assume that each bounded linear functional corresponds to an element of $\mathcal{A}$. Hence we must have

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

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

Complex Algebra Structure

Any measuring instrument can be applied to systems once it has been prepared following a preparation procedure. We can however also view the combined setup consisting of the measurement instrument applied to the preparation procedure as a preparation procedure itself. We can then perform measurements on this composite preparation procedure. Heuristically, when these concatenated measurements have inaccuracies which are independent of each other, we must expect the accuracy of the concatenation, to amplify the inaccuracies so the total inaccuracy would be that of the first instrument *and* the second instrument and hence must correspond to the product of inaccuracies. This can only be the case when the accuracies of the two instruments are independent. Hence the accuracy of the concatenation must be expected to be less than the product of individual accuracies. For any effects E and F in $\mathcal{A}$ we denote the above described concatenation of measurements by EF. This gives rise to a product structure on $\mathcal{A}$. Hence we will assume the vector space $\mathcal{A}$ to also possess an algebra structure (there is no need to assume that this algebra structure is commutative).

As soon as $\mathcal{A}$ possesses an algebra structure, we can make sense of polynomials in $\mathcal{A}$. Suppose a measuring instrument consists of a collection of mutually compatible instruments, that is, the measurements performed by such instruments do not affect the outcomes of the other instruments. Suppose the instrument undergoes a change if it undergoes a change for each of its component instruments or sub-instruments, then the probability of outcome for such an instrument should correspond to product of probabilities of its components. Let E be such an instrument and E_i its components. Then for any ensemble μ_ρ we must have,

$$\mu_\rho(E) = \prod_{i \in \mathcal{I}} \mu_\rho(E_i).$$

Since each effect corresponds to an equivalence class of measuring instruments, we may expect there to exist some measuring instrument which is the collection of compatible measuring instruments which cannot be further divided. Hence we may assume that it is always possible to decompose a given measuring instrument into its maximally divided sub-instruments.

We assume that it is always possible to decompose a measuring instrument as in terms of its component sub-instruments. To make sense of this we need structures that enable us to factorise polynomials in $\mathcal{A}$.

By the fundamental theorem of algebra, we can always factorise a polynomial in the field of complex numbers. Hence we will assume the pairing to be a real bilinear map to the field of complex numbers

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

For mathematical convenience, we will transfer this complex structure to the vector spaces $\mathcal{A}_*$ and $\mathcal{A}$ themselves, and assume that $\mathcal{A}_*$ and $\mathcal{A}$ are complex vector spaces. We require the scaling of both effects and ensembles by a unit length complex number to give the same pairing as the unscaled pairing. This requirement along with real linearity in each argument forces us to assume the pairing $\langle\cdot|\cdot\rangle$ to be sesquilinear.

Quantum mechanics can also be formulated in more complicated fields such as the field of quarternions. We will however avoid discussions on such choices since such a theory will not be universal, in the sense that the extra structure are not required by postulates about preparation procedures and measuring instruments.

Banach Algebra Structure

We expect varying accuracy of either of the measuring instruments corresponding to E or F to vary the accuracy of EF accordingly. Hence we expect the algebraic structure on $\mathcal{A}$ to respect the topological structure on $\mathcal{A}$ coming from the norm $\|\cdot\|_{\mathcal{A}}$. We now study the topological properties of the product following [24].

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

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

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

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

Similarly we have a natural right-action given by

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

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

$$\left\|\mathcal{R}_F\right\|_{\mathcal{B}(\mathcal{A})} = \sup_{E\in\mathcal{E}} \left[\left\|EF\right\|_{\mathcal{A}}\right].$$

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

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

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

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

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

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

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

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

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

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

$$\big\|EF\big\|_{\mathcal{A}} \leqq \big\|E\big\|_{\mathcal{A}} \big\|F\big\|_{\mathcal{A}}$$

Since $\|\mathcal{R}_1\|_{\mathcal{B}(\mathcal{A})} = 1$ we have, $\|\mathcal{R}_E(1)\|_{\mathcal{A}} = \|E\|_{\mathcal{A}} \leqq \|\mathcal{R}_E\|_{\mathcal{B}(\mathcal{A})} \leqq \|\mathcal{R}_1\|_{\mathcal{B}(\mathcal{A})} \, \|E\|_{\mathcal{A}} = \|E\|_{\mathcal{A}}$. That is,

$$\big\|R_E\big\|_{\mathcal{B}(\mathcal{A})} = \big\|E\big\|_{\mathcal{A}}.$$

Hence the embedding,

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

is an isometric isomorphism and $\mathcal{R}(\mathcal{A})$ is a Banach algebra. The space of states acts as continuous linear functionals on $\mathcal{R}(\mathcal{A})$, given by the map $\omega(R_A) = \big\langle A, \omega \big\rangle$, with $|\omega(R_A)| \leqq \|R_A\|_{\mathcal{B}(\mathcal{A})}$, hence the elements of $\mathcal{A}_*$ respect the topological properties of $\mathcal{R}(\mathcal{A})$.

At this stage, it is important to note what it means to measure for the same outcomes. For an outcome R, the preparation procedure concatenated by the measuring instrument corresponding to the effect E_R is itself a preparation procedure. To such preparation procedure, an application of the same measuring instrument again will not change the ensemble the procedure produces. Since the concatenation of measurements corresponds to the product structure of $\mathcal{R}(\mathcal{A})$, this operational requirement about repeated measurements corresponds to requiring that whenever E in $\mathcal{R}(\mathcal{A})$ corresponds to an effect it must correspond to a projection

$$E^2 = E.$$

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

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

1.1.2.2 | †-Structure & W^*-Algebras

We now consider operations on the space of effects that are physically meaningful. Suppose we perform a rearrangement or readjustments of all the measuring instruments. Then we can consider all the measuring instruments which have been turned to non-functioning instruments, these instruments never undergo any changes, and the re-arrangement or re-adjustment of instruments acts as a switch-off for those instruments. Assuming such manipulations keep the system within itself, it must keep the space of effects invariant. However the only way to find out what changes we have made to the system is by performing experiments on the system. Since the space of states is separating for the space of effects, any manipulation we do to the space of effects are *seen* by the space of states. Hence we consider the equivalence class of such manipulations on the space of effects, which affect the pairing $\langle\cdot|\cdot\rangle$ of $\mathcal{A}$ with $\mathcal{A}_*$ in the same way. Each such equivalence class of manipulation of the system corresponds to a linear map,

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

The manipulation T of the system can also be thought of as a manipulation of the preparation procedures. We can also consider the equivalence class of manipulations of the system giving rise to the same change in preparation procedures, and for every preparation procedure ρ this can be described in terms of operational statistics by the map

$$E \mapsto \mu_\rho(TE) \coloneqq \big(R_T\mu_\rho\big)(E).$$

This map of ensembles can be extended to $\mathcal{A}_*$ by linearity, such that for every ensembles ρ and σ and scalars λ and μ, $\alpha\mu_\rho + \beta\mu_\sigma \mapsto \alpha\mu\big(T(\cdot)|\rho\big) + \beta\mu\big(T(\cdot)|\sigma\big)$. Hence to each physically meaningful manipulation of the system there must exist a bounded linear map,

$$R_T : \mathcal{A}_* \to \mathcal{A}_*,$$

This map on $\mathcal{A}_*$ gives us a map,

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

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

$$\big\langle T^\dagger A|\omega\big\rangle = \big\langle A|R_T\omega\big\rangle$$

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

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

The operation $\dagger : \mathcal{R}(\mathcal{A}) \to \mathcal{R}(\mathcal{A})$ seems to be completely algebraic in nature, however it also possesses interesting topological properties, since the topologies on the space of effects and state space operationally correspond to the physical notion of accuracy. In order to describe the topological properties of the operation †, we now describe the notion of complementary effect and its relation to the notion of topological complementation on Banach spaces.

Suppose we have a measuring instrument which can measure an outcome R, the change occuring in the instrument corresponds to the effect for R. If the measuring instrument does not undergo any change, then we can infer that the outcome R had not occurred. Since the outcome R not occuring is also an outcome by itself, denoted by $\neg R$, and the

same measuring instrument also be used for measuring $\neg R$, the instrument not undergoing a change corresponds to the effect for $\neg R$. Hence we must have an operation on the space of effects which assigns to every effect E_R corresponding to the outcome R, the effect $E_{\neg R}$ corresponding to the outcome $\neg R$. We call $E_{\neg R}$ the complement of E_R and we denote it by

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

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

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

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

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

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

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

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

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

onto the closed complemented subspace $\mathcal{M}$ of $\mathcal{R}(\mathcal{A})$ is a continuous operator on $\mathcal{R}(\mathcal{A})$. Conversely, every continuous projection $E_{\mathcal{M}}$, corresponds to a unique closed complemented subspace $\mathcal{M}$. Since the space of effects is contained in $\mathcal{R}(\mathcal{A})$ and since each effect correspond to a continuous projection operator, it follows that every effect corresponds to a unique closed complemented subspace of $\mathcal{R}(\mathcal{A})$, hence the notation for the complement effect $E^{\perp}$ is closely related to the notion of topological complementation on $\mathcal{R}(\mathcal{A})$.[3]

Every laboratory can be thought of as a subsystem of the universe, and can equivalently be described in terms of the preparation procedures it can prepare, or by all the experiments which we can perform in the laboratory. By switching off the laboratory, we will switch off all the experiments we could perform in the laboratory, or equivalently switch of all the preparation procedures in the laboratory. The laboratory itself be thought of as a measuring instrument which is compatible with all the experiments it can perform, and for the laboratory itself, this must correspond to the identity element. Hence to each laboratory, there must exist a projection element, which commutes with every experiment that can be performed in the laboratory. To each preparation procedure in the laboratory, this projection element corresponds to the switching on of the procedure. For each preparation procedure, we have a collection of laboratories in which the preparation procedure can exist. For example, a

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

preparation procedure in a laboratory, is also a preparation procedure in a bigger laboratory which contains the smaller laboratory. The projection element corresponding to the smallest laboratory to which the preparation procedure can belong to, will be called the support of the preparation procedure. Hence we obtain a correspondence between projection elements and collection of preparation procedures for which the projection corresponds to the minimal laboratory. If μ_ρ is an ensemble corresponding to some preparation procedure, then we denote by E_ρ the projection element as described above. We must require that

$$\|\mu_\rho\|_{\mathcal{A}_*} = \|E_\rho\|_{\mathcal{A}}.$$

By linearity, the operator $(T^\dagger)^\dagger$ is determined by the pairing $\mu_\rho((T^\dagger)^\dagger(E)) = R_T\mu_\rho(E) = \mu_\rho(TE)$ for all effects E belonging to the support of μ_ρ, and hence also for $E_\rho^\perp$. By the uniqueness of the pairing of closed complemented subspaces with its complement, and since ensembles and effects separate each other, we must require that

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

For all physically relevant transformations T, the mapping $\dagger$ acts as an involution. The uniqueness of the topological complementation ensures that the operation $\dagger$ is an involution.

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

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

Such Banach algebras are called involutive Banach algebras, or Banach *-algebras.

However ensembles are themselves closely related to effects in an opposite way via the notion of support. Hence the composite manipulation of the system acts in an *independent* way on the space of effects and the state space. From the discussion about norms on $\mathcal{A}$ and $\mathcal{A}_*$, we know that the norms are closely related to the notion of accuracy of measuring instruments and preparation procedures. Since the manipulations introduced on the space of effects and state space are opposite of each other they are independent, and we should expect the accuracies of effects and ensembles to be multiplicative. The relation between support properties of ensembles and effects captures the complementary nature of such a composite manipulation $T^\dagger T$ of the system.

We now make this precise, via topological complementations and the expected relationship between subspaces of the state space and ideals of the space of effects. The norm of $T^\dagger T$ is given by

$$\|T^\dagger T\| = \sup_{\substack{E\in\mathcal{E}\\ \omega\in\mathcal{S}}} \Big[\big|\langle TE|R_T\omega\rangle\big|\Big].$$

For every ω we can split the state space as direct sum of complemented subspaces as $\mathcal{A}_* = TE_\omega\mathcal{A}_* \oplus TE_\omega^\perp\mathcal{A}_*$. Since the pairing vanishes if TE does not belong to the support of $R_T\omega$, we can assume that TE belongs to the ideal $TE_\omega^\perp\mathcal{A}$.

$$\|T^\dagger T\| = \sup_{\substack{E\in\mathcal{E}\\ \omega\in\mathcal{S}}} \Big[\big|\langle TE|R_T\omega\rangle\big|\Big] = \sup_{E\in\mathcal{E}} \Big[\|TE\|\Big] \sup_{\omega\in\mathcal{S}} \Big[\|R_T\omega\|\Big] = \|T\|^2.$$

Hence we have,

$$\|T^\dagger T\| = \|T\|^2.$$

Banach *-algebras satisfying this identity are called C^*-algebras.

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

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

A C^*-algebra $\mathcal{A}$ which is a dual Banach space of $\mathcal{A}_*$ is called a W^*-algebra, and in such a case the Banach space $\mathcal{A}_*$ is called the predual of $\mathcal{A}$.

Mathematical structures can be studied by studying function which respect such structures. 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 φ on $\mathcal{H}$ corresponds to a vector f_φ in $\mathcal{H}$ with $\|\varphi\|_{\mathcal{H}^*} = \|f_\varphi\|_{\mathcal{H}}$ such that,

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

Every Hilbert space, as 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. By definition of the dual Banach space the norm $\|\cdot\|_{\mathcal{A}}$ is given by

$$\big\|A\big\|_{\mathcal{A}} = \sup_{\|\omega\|_{\mathcal{A}_*} \leq 1} \Big[\big|\omega(A)\big|\Big].$$

In addition to being dual Banach spaces, W^*-algebras also come equipped with their own C^*-algebra structure. $\mathcal{A}$ consists of continuous linear functionals on $\mathcal{A}_*$.

Necessity of W^*-Algebras

We obtained the definition above of W^*-algebras starting from very simple operational postulates about measuring instruments and preparation procedures. We did assume some idealisations for mathematical convenience such as topological completeness, for which we did not give any justification. Hence, if we take this instrumentalist assumptions as the postulates of quantum theory then the theory of W^*-algebras acts as the appropriate mathematical theory for the articulation of physical ideas underlying quantum theory. The W^*-algebra corresponding to the system will be called the algebra of observables, and the projection elements of the W^*-algebra model effects of the system. However to show that the mathematical foundations of quantum theories are solid, we must show that the converse is also true, that is, the theory of W^*-algebras is sufficiently mathematically rich to accommodate the postulates of quantum theory about measuring instruments and preparation procedures. In the next sections we develop the theory of W^*-algebras, and prove theorems which reverse ideas we used here in showing the necessity of W^*-algebras.

Note that the notion of observables has not made its appearance yet. Hence the notion of observables is not the most primitive concept in quantum theory from an instrumentalist point of view, although it is already implicit. As we will see later, the notion of spectrum is closely related to the notion of observables. To respect the hierarchy of physical ideas we will not use the notion of spectrum to study W^*-algebras until we introduce the notion of observables.

1.2 | Convexity in Locally Convex Spaces

The linear structure on vector spaces allows us to study lines, and the notion of convexity is closely related. In §1.1.1.1, the notion of mixing of ensembles was formalised using the notion of convexity on real numbers. This was based on simple assumptions about how mixtures can be made in the *real* world via operations which can be parametrised by real numbers, such as rotating a bob to change the settings of a preparation procedure. The notion of convexity on reals and its relation to linear structure, topology, etc. was how we obtained the vector space structure for the space of effects and ensembles. Hence much of mathematical structures of quantum theories rely heavily on this arguably weak link to real numbers. To compensate this weakness we will approach the study of convexity using a more abstract language via conjugations and pairings of spaces, so that if need be, we can quickly generalise the discussion here, to the case where the notion of convexity does not depend on *reality*. We will use this section to isolate all the analytic aspects of the theory of W^*-algebras.

1.2.1 | Conjugate Spaces & Conjugate Duality

The linear structures of vector spaces is closely related to the notion of lines. A set K is said to be convex (respectively absolutely convex) if it contains the smallest line segments (respectively the smallest complex disc) joining its points, that is. We will use this line segment[4] description of convex sets for heuristics and intuition, and formalise this intuition in the abstract framework using dual spaces. The linear operations, such as vector addition, and scaling are continuous with respect to the topology induced by the norm. Every line passing through the origin in $\mathcal{X}$ corresponds to a real linear map from $\mathbb{R}$ to $\mathcal{X}$,

$$L : \mathbb{R} \to \mathcal{X}.$$

Since norms are necessarily non-vanishing, the composition,

$$\mathbb{R} \xrightarrow{L} \mathcal{X} \xrightarrow{\|\cdot\|_{\mathcal{X}}} \mathbb{R},$$

defines a non-trivial function from every line through the origin. Linear maps are transformation of lines passing through the origin to lines passing through the origin, and we need to generalise them to be not dependent on the origin. Affine maps are compositions of a translation map after a linear map, and take lines to lines. Heuristically, the norm on a Banach space sees every line on the space, and hence we may expect to characterise convex sets in terms of the properties of the norm.

1.2.1.1 | The Hahn-Banach Theorem

For any complex number λ, we denote its real part by $\Re(\lambda)$ and the imaginary part by $\Im(\lambda)$. We are interested in the two cases corresponding to convexity and absolute convexity. We will now prove the Hahn-Banach theorem for the real linearity case, and extend to the complex case as a corollary. The Hahn-Banach separation theorems make precise the heuristic relation between the notion of convexity and norm on the Banach space.

[4] Note that in the vector space context, convexity is closely related to the notion of lines. This notion of convexity is intuitive since it is very geometric. If needed this intuitive notion of convexity can be replaced with more general heuristics based on the needs. It might be possible to obtain an abstract categorical notion of convexity where we replace the order relation on reals by morphisms between objects of a category. A full subcategory of this category could then correspond to a *convex thing*. We will however not explore such abstract territory here since we do not see the need for it.

The proof of the Hahn-Banach theorem also works for sublinear maps on $\mathcal{X}$. A sublinear map is a real valued map f on $\mathcal{X}$ such that,

$$f(x+y) \leqq f(x) + f(y).$$

When $f(\lambda x) = |\lambda| f(x)$, it is called a semi-norm, and if $f(x)$ is zero only when x is the origin of $\mathcal{X}$ they correspond to norms. Hence the Hahn-Banach theorem would still be useful if the *accuracy* function for instruments does not correspond to a norm.

Note that we can write any complex number λ as $\Re(\lambda) + i\Im(\lambda)$. Hence we can write $\omega(x) = \Re(\omega(x)) + i\Im(\omega(x))$. Since ω is a complex functional, we also have $\omega(ix) = i\omega(x)$, and hence we have $\Re(i\omega(x)) = \Re(\omega(ix)) = -\Im(\omega(x))$. Therefore we must have,

$$\omega(x) = \Re(\omega(x)) - i\Re(\omega(ix)).$$

Hence the real part $\Re(\omega)$ already determines the functional ω.

Theorem 1.2.1. (Extension) *Let $\mathcal{X}$ be a vector space with a sublinear map f. Let $\mathcal{Y}$ be a subspace of $\mathcal{X}$. If φ is a linear functional on $\mathcal{Y}$ such that for all y in $\mathcal{Y}$*

$$\Re\big(\varphi(y)\big) \leqq f(y)$$

then there exists a linear functional $\widehat{\varphi}$ on $\mathcal{X}$ with $\widehat{\varphi}|_{\mathcal{Y}} = \varphi$ such that

$$\Re\big(\widehat{\varphi}(x)\big) \leqq f(x).$$

The Hahn-Banach theorem can now be used to study convexity properties of $\mathcal{X}$ since the sub-linear map relates the convexity structure on $\mathcal{X}$ with the convexity structure on $\mathbb{R}$. Given two disjoint convex sets, we can consider the collection of all lines joining the points of the two convex sets. We can describe these lines in terms of affine linear functionals, and describe when a line leaves a convex set using the values of the functional. This allows us to separate disjoint convex subsets using linear functionals. We now formalise this heuristics to obtain a characterisation of disjointness of convex sets in terms of the dual space.

A locally convex space $\mathcal{X}$ is a topological vector space (vector space operations are continuous with respect to its topology) such that its topology is generated by a family of convex neighborhoods of the origin together with the vector space operations. Let $\mathcal{N}_{\mathcal{X}}$ be such a neighborhood basis consisting of balanced, absorbent, convex neighborhoods of the origin. Then for each $\mathcal{N}$ in $\mathcal{N}_{\mathcal{X}}$, define a function $\mu_{\mathcal{N}}$ by

$$\mu_{\mathcal{N}}(x) = \inf_{x \in \lambda \mathcal{N}} \big[|\lambda|\big].$$

Here $\mu_{\mathcal{N}}(x)$ is the smallest amount descaling of x needed to bring it inside the neighborhood $\mathcal{N}$. By convention we assume the infimum of the empty set to be $-\infty$. It immediately follows from the definition that $\mu_{\mathcal{N}}(x) < 1$ only when x belongs to $\mathcal{N}$, and $\mu_{\mathcal{N}}(y) > 1$ if y is outside of $\mathcal{N}$. $\mu_{\mathcal{N}}$ is called the Minkowski functional for $\mathcal{N}$.

It follows that each $\mu_{\mathcal{N}}$ is a semi-norm on $\mathcal{X}$ and the collection $\{\mu_{\mathcal{N}}\}_{\mathcal{N} \in \mathcal{N}_{\mathcal{X}}}$ of semi-norms on $\mathcal{X}$ generate the topology on $\mathcal{X}$. The Minkowski functional $\mu_{\mathcal{N}}$ can distinguish points inside $\mathcal{N}$ from the points outside. For a set $\mathcal{M}$ in $\mathcal{X}$, we denote by $\langle\!\langle \mathcal{M} \rangle\!\rangle$ and $\langle\!\langle\!\langle \mathcal{M} \rangle\!\rangle\!\rangle$ the convex and absolutely convex closures respectively of $\mathcal{M}$.

Theorem 1.2.2. (Separation) *Let $\mathcal{X}$ be a locally convex space, and let $\mathcal{M}$ be a compact set with non-empty interior. Let $\mathcal{N}$ be such that $\langle\!\langle \mathcal{M} \rangle\!\rangle \cap \langle\!\langle \mathcal{N} \rangle\!\rangle$ is empty. Then there exists a functional φ in $\mathcal{X}^*$ such that*

$$\sup_{x\in\mathcal{M}} \Big[\Re\big(\varphi(x)\big)\Big] < \lambda_{[\mathcal{M}:\mathcal{N}]} < \inf_{y\in\mathcal{N}} \Big[\Re\big(\varphi(y)\big)\Big],$$

for some $\lambda_{[\mathcal{M}:\mathcal{N}]}$ in $\mathbb{R}$. In such a case, we say φ separates $\mathcal{M}$ and $\mathcal{N}$.

We can obtain complex versions of the Hahn-Banach extension and separation theorems as follows; Since $\varphi(x) = \Re(\varphi(x)) - i\Re(\varphi(ix))$ every complex functional is determined by its real part. Since $\Re(\varphi(x)) \leqq |\varphi(x)|$ it follows that whenever $|\varphi(y)| \leqq f(y)$, there exists an extension $\widehat{\varphi}$ of φ to $\mathcal{X}$ with

$$|\widehat{\varphi}(x)| \leqq f(x).$$

For the separation theorem we can adapt the separation theorem for the real case by replacing the use of extension theorem for the real case with the complex version, and hence if $\mathcal{M}\rangle\!\rangle\langle\!\langle\mathcal{N}$ there exist φ in $\mathcal{X}$ and

$$\sup_{x\in\mathcal{M}} \Big[\big|\varphi(x)\big|\Big] < \lambda_\varphi < \inf_{y\in\mathcal{N}} \Big[\big|\varphi(y)\big|\Big].$$

The set of all x in $\mathcal{X}$ with $\Re\big(\varphi(x)\big) = \lambda_\varphi$ defines a hyperplane in $\mathcal{X}$ such that $\langle\!\langle \mathcal{M} \rangle\!\rangle$ and $\langle\!\langle \mathcal{N} \rangle\!\rangle$ lie on its opposite sides. We say that the hyperplane $\Re\big(\varphi(x)\big) = \lambda_\varphi$ separates $\langle\!\langle \mathcal{M} \rangle\!\rangle$ and $\langle\!\langle \mathcal{N} \rangle\!\rangle$.

Weak & Weak* Topologies

Continuous functions relate the topologies of their domain with the topology of their range. By relating the topologies, continuous functions measure the nearness of points in the domain, via the nearness of points in the range. Let $\mathcal{X}$ and $\mathcal{Y}$ be two sets. If there exists a non-degenerate bilinear pairing

$$\big\langle \cdot \,|\, \cdot \big\rangle : \mathcal{X} \times \mathcal{Y} \to \mathbb{C}$$

then we say $\langle \mathcal{X} | \mathcal{Y} \rangle$ is a dual pair. The elements of $\mathcal{Y}$ define a topology on $\mathcal{X}$, and dually elements of $\mathcal{X}$ define a topology on $\mathcal{Y}$. Such topologies are called weak topologies, and are the universal topologies on the spaces making the pairing continuous.

If the topology on $\mathcal{X}$ is given to be τ, then the space of τ-continuous linear functionals on $\mathcal{X}$ is denoted by $(\mathcal{X}, \tau)^*$, and is the weakest topology for which $(\mathcal{X}, \tau)^*$ is a dual pair. If the topology on $\mathcal{X}$ is assumed then we denote by $\mathcal{X}^*$ the space of all continuous linear functionals on $\mathcal{X}$. The evaluation pairing is called the conjugate pairing and is denoted by,

$$\big\langle x | \varphi \big\rangle_{\mathcal{X}} \equiv \varphi(x).$$

It follows that

$$\big\langle \mathcal{X} | \mathcal{X}^* \big\rangle_{\mathcal{X}}$$

is a dual pair.

The weak*-topology on $\mathcal{X}^*$ is such that $\{\varphi \mapsto \varphi(x)\}_{x\in\mathcal{X}}$ is a continuous family of functions. The universal topology on $\mathcal{X}^*$ with this property is the one induced by the collection of seminorms,

$$\Big\{\varphi \mapsto \big|\varphi(x)\big|\Big\}_{x\in\mathcal{X}}$$

Since semi-norms are convex functions, the dual space $\mathcal{X}^*$ equipped with this topology is a locally convex space. Dually, the smallest topology on $\mathcal{X}$ making $x \mapsto \varphi(x)$ continuous for

all φ in $\mathcal{X}^*$ is called the weak topology on $\mathcal{X}$. The weak*-topology on $\mathcal{X}^*$ is also denoted by $\sigma(\mathcal{X}^*, \mathcal{X})$.

The bidual $\mathcal{X}^{**}$ of $\mathcal{X}$, is the space of all continuous linear functionals on $\mathcal{X}^*$ equipped with the strongest $\mathcal{X}$-compatible topology. In such a case, we have a canonical embedding of $\mathcal{X}$ inside $\mathcal{X}^{**}$,

$$\begin{aligned} \iota_{\mathcal{X}} \ : \ & \mathcal{X} \to \mathcal{X}^{**} \\ & x \mapsto \big\langle x \,|\, \cdot \,\big\rangle_{\mathcal{X}}. \end{aligned}$$

where $\big\langle x| \cdot \big\rangle_{\mathcal{X}} : \varphi \mapsto \langle x|\varphi\rangle_{\mathcal{X}}$ is a continuous linear functional on $\mathcal{X}^*$. We will denote $\langle \cdot | \cdot \rangle_{\mathcal{X}}$ by $\langle \cdot | \cdot \rangle$ whenever there is no confusion.

1.2.1.2 | Conjugate Duality

We now discuss some aspects of convexity for locally convex spaces. The discussion here is closely related to discussion in [5] and [22]. Let $\mathcal{X}$ be a locally convex space. We say a function is convex if it satisfies the Jensen inequality,

$$f(\lambda x + (1-\lambda)y) \leqq \lambda f(x) + (1-\lambda) f(y)$$

Heuristically this means that f maps straight line segments to line segments in $\mathbb{R}$.

If f is a convex function then the set $\mathcal{M}_\lambda$ of all x in $\mathcal{X}$ with $f(x) \leqq \lambda$ is a convex set for each real number λ in $\mathbb{R}$. Convex functions ca compress or stretch the line segments in $\mathcal{X}$. The interesting case is when the compression is not continuous, that is parts of line segments are collapsed completely. In such cases $\mathcal{M}_\lambda$ can contain full affine subspaces. The convexity properties contained in such subspaces cannot be seen by the convex function. Conjugate duality allows us to quantify the convexity unseen by a given convex function. One may study how much more convexity the set of affine functionals see compared to a given convex function. The total convexity seen by the linear functionals on $\mathcal{X}$ can be quantified by the taking supremum over all points of $\mathcal{X}$. We now make this precise by defining the conjugate function of a function.

Let $\overline{\mathbb{R}}$ denote the extended real line. For any function $f : \mathcal{X} \to \overline{\mathbb{R}}$, the conjugate function or the Legendre-Fenchel transform of f is the map $\mathcal{F}f : \mathcal{X}^* \to \overline{\mathbb{R}}$ defined by

$$\mathcal{F}f(\varphi) \equiv (f)^*(\varphi) \coloneqq \sup_{x \in \mathcal{X}} \Big[\Re \big\langle x|\varphi \big\rangle - f(x) \Big].$$

The value $f^*(\varphi)$ quantifies how much of the convexity of the space is lost by the function f when compared to the linear functional φ. If we take f to be $\Re\langle \cdot|\varphi\rangle$, then $f^*(\varphi)$ vanishes and f is just as convex as φ. If $g : \mathcal{X}^* \to \overline{\mathbb{R}}$ is a function then we define the coconjugate of g by,

$$\mathcal{G}g(x) \equiv \big(g\big)_*(x) = \sup_{\varphi \in \mathcal{X}^*} \Big[\Re \big\langle x|\varphi \big\rangle - g(\varphi) \Big].$$

As the supremum of affine linear functionals the Legendre transform is a convex function. Since f^* is the pointwise supremum of weak*-continuous affine functions, the conjugate function f^* is a weak*-continuous function on $\mathcal{X}^*$. The coconjugate g_* is a weakly continuous function on $\mathcal{X}$.

We immediately obtain the following properties of the Legendre transform;

$$\begin{aligned} &\mathcal{F}(f + c)(\chi) = \mathcal{F}f(\chi) - c \\ &\mathcal{F}f(\chi - x) = \mathcal{F}f(\chi) + \langle x, \chi \rangle \\ &\mathcal{F}f(t\chi) = \mathcal{F}f(\chi/t) \\ &\mathcal{F}(tf)(\chi) = t \cdot \mathcal{F}f(\chi/t). \end{aligned}$$

Similarly for $\mathcal{G}$. By definition of the Legendre-Fenchel transform, we must have

$$f(x) + f^*(\varphi) \geqq \Re\langle x|\varphi\rangle.$$

This inequality is called the Fenchel-Young inequality.

We can also define the absolute Legendre transform to focus on absolute convexity instead of convexity. In such a case, the Legendre transform of interest can be defined to be,

$$f^*(\varphi) \equiv \sup_{x\in\mathcal{X}} \Big[\big|\langle x|\varphi\rangle\big| - f(x)\Big].$$

We will use either of these Legendre transforms depending on whether the context is about convexity or absolute convexity, and only explicitly indicate which one we are using if there would be confusion.

Properness & Lower Semi-Continuity

Topology and the vector space structure on the real line are closely related to each other. If a convex function compresses certain line segments in the space, then it should be possible to notice it from the topological properties of the real line. Hence the compression issues of convex functions can be studied by their continuity properties. This is possible in locally convex spaces since the topology is generated by neighborhood bases of convex sets, and hence the convexity and topological properties of locally convex spaces can be related with convexity and topological properties on the real line.

We must also ensure that f does not stretch line segments to infinity. A function f is said to be proper if it is not always ∞ and is never $-\infty$. The set of all x with $f(x) < \infty$ is called the domain of f and denoted by $\mathcal{D}_f$. For a proper function f on $\mathcal{X}$ we define the conjugate function as,

$$\mathcal{F}f(\varphi) \equiv f^*(\varphi) \coloneqq \sup_{x\in\mathcal{D}_f} \Big[\Re\langle x|\varphi\rangle_{\mathcal{X}} - f(x)\Big].$$

Similarly if g is a proper function on $\mathcal{X}^*$ then we define the coconjugate of g as

$$\mathcal{G}g(x) \equiv g_*(x) = \sup_{\varphi\in\mathcal{D}_g} \Big[\Re\langle x|\varphi\rangle - g(\varphi)\Big].$$

A function f on $\mathcal{X}$ is said to be lower semi-continuous if for every x with $\lambda < f(x)$ for some λ in $\mathbb{R}$, there exists an open neighborhood N_x of x such that for every y in N_x we have $\lambda < f(y)$.

As a supremum over affine functions, for any function f on $\mathcal{X}$, the conjugate function f^* is lower semi-continuous function on $\mathcal{X}^*$ with respect to the weak*-topology. We denote the collection of all lower semi-continuous convex functions on $\mathcal{X}$ by $\Gamma(\mathcal{X})$. We will denote the collection of all lower semi-continuous proper convex functions on $\mathcal{X}$ denoted by $\overline{\Gamma}(\mathcal{X})$. It follows that $\overline{\Gamma}(\mathcal{X})$ is a vector space.

Lemma 1.2.3. *If f is in $\overline{\Gamma}(\mathcal{X})$ then f^* is in $\overline{\Gamma}(\mathcal{X}^*)$ and the Legendre-Fenchel conjugate transform is an injective mapping;*

$$\mathcal{F} : \overline{\Gamma}(\mathcal{X}) \rightarrowtail \overline{\Gamma}(\mathcal{X}^*)$$

In particular, then there exists φ_f in $\mathcal{X}^*$ such that for all x and y in $\mathcal{X}$,

$$\begin{aligned} f(x) - f(y) &\geqq \Re\big[\langle x - y|\varphi_f\rangle\big]; \\ f^*(\varphi_f) &\leqq \Re\langle y|\varphi_f\rangle - f(y). \end{aligned}$$

Proof

Since f is lower semi-continuous and convex, the set of all (x, λ) with $f(x) \leqq \lambda$ is a closed convex subset of $\mathcal{X} \oplus \mathbb{R}$ denoted by $\mathcal{M}_f$. Any (y, μ) with $\mu < f(y)$ lies outside, and hence by the Hahn-Banach separation theorem, there exists a continuous linear functional φ on $\mathcal{X} \oplus \mathbb{R}$ such that,

$$\sup_{\mathcal{M}_f} \Big[\Re \big\langle (x, \lambda) | \varphi \big\rangle_{\mathcal{X} \oplus \mathbb{R}} \Big] < \Big[\Re \big\langle (y, \mu) | \varphi \big\rangle_{\mathcal{X} \oplus \mathbb{R}} \Big].$$

The functional φ can be decomposed as an extension of the linear functional φ on $\mathcal{X}$ with, $\varphi(x, \lambda) = \varphi(x, 0) + \lambda \alpha_\varphi$ where $\alpha_\varphi \equiv \varphi(0, 1)$, for simplicity we denote $\Re(\alpha_\varphi) = \alpha_\varphi$ and denote $\varphi(x, 0)$ by $\varphi_0(x)$. Hence we must have

$$\sup_{f(x) \leqq \lambda} \big[\Re \big\langle x | \varphi_0 \big\rangle + \lambda \alpha_\varphi \big] < \Re \big\langle y | \varphi_0 \big\rangle + \mu \alpha_\varphi.$$

If f is proper it must be bounded below. By taking $-\mu$ large enough, we infer that α_φ is negative. Hence by taking $\lambda = f(x)$ and $\mu = f(y) - \epsilon$ we must have; $\Re[\langle x | \varphi_0 \rangle] + \alpha_\varphi f(x) \leqq \Re[\langle y | \varphi_0 \rangle] + \alpha_\varphi f(y)$. By rescaling φ_0, and considering the functional $\varphi_f \equiv -\lambda_\varphi^{-1} \varphi_0$, it follows that

$$f(x) - f(y) \geqq \Re \big[\big\langle x - y | \varphi_f \big\rangle \big]$$

The evaluation of the conjugate function f^* at φ_f is

$$f^*(\varphi_f) = \sup_{x \in \mathcal{D}_f} \Big[\Re \big\langle x | \varphi_f \big\rangle - f(x) \Big] \leqq \Re \big\langle y | \varphi_f \big\rangle - f(y).$$

Hence f^* is a proper function. This completes the proof of the lemma. □

We now show that the Legendre-Fenchel transform is an isomorphism of vector spaces by showing $\mathcal{G} \circ \mathcal{F}$ is the identity map on $\overline{\Gamma}(\mathcal{X})$. Note that the symmetric role of the pairing with respect to weak topology on $\mathcal{X}$ and weak*-topology on $\mathcal{X}^*$ guarantees the isomorphism. $\mathcal{G} \circ \mathcal{F} f$ is called the biconjugate of f and is given by

$$(\mathcal{G} \circ \mathcal{F} f)(x) \equiv (f)^*_*(x) \equiv \sup_{\varphi \in \mathcal{D}_{f^*}} \Big[\Re \big\langle x | \varphi \big\rangle - f^*(\varphi) \Big].$$

Clearly, $(f)^*_*$ is a proper lower semi-continuous convex function that is, $(\mathcal{G} \circ \mathcal{F} f)$ is in $\overline{\Gamma}(\mathcal{X})$. We obtain the Fenchel-Young inequality for the biconjugate from the above expression;

$$(f)^*_*(x) + f^*(\chi) \geqq \Re \langle x | \chi \rangle.$$

We now prove the duality theorem due to Fenchel-Moreau-Rockafellar;

Theorem 1.2.4. (Conjugate Duality)

$$\overline{\Gamma}(\mathcal{X}) \xrightarrow[\cong]{\mathcal{F}} \overline{\Gamma}(\mathcal{X}^*)$$

Proof

We show that $\mathcal{G}$ is a left inverse of $\mathcal{F}$. That is, for every f in $\overline{\Gamma}(\mathcal{X})$, we now show that $(f)^*_*(x) = f(x)$ for all x in $\mathcal{X}$. By definition of the Legendre-Fenchel transform, we must have

$$f(x) + f^*(\varphi) \geqq \Re \big\langle x | \varphi \big\rangle.$$

Combining this with the definition of the biconjugate we obtain,

$$f(x) \geqq \sup_{\varphi \in \mathcal{D}_{\mathcal{F}f}} \Big[\Re\langle x|\varphi\rangle - f^*(\varphi) \Big] \equiv (f)^*_*(x).$$

From the discussion before the previous lemma, we have that whenever f is a lower semi-continuous proper convex function, there exists some φ_f such that $f(x) - f(y) \geqq \Re\langle x|\varphi_f\rangle - \Re\langle y|\varphi_f\rangle$ and $f(y) \leqq \Re\big\langle y|\varphi_f\big\rangle - f^*(\varphi_f)$. By definition of the biconjugate as supremum over all such linear functionals, we must have, $\Re\langle x|\varphi_f\rangle - f^*(\varphi_f) \leqq (f)^*_*(x)$ for all x in its domain. Hence we have,

$$f(x) \leqq (f)^*_*(x).$$

Combining with the above inequality we have $(f)^*_*(x) = f(x)$ for all x in $\mathcal{D}_f$.

We now show that if $f(x)$ is infinite, then $(f)^*_*(x)$ must also be infinite. If φ_f is as above, it must belong to the domain of f^*. For any other functional φ in its domain consider the functional $\chi = \varphi + \lambda\varphi_f$. By definition of the conjugate function we have,

$$\begin{aligned} f^*(\chi) &= \sup_{y \in \mathcal{D}_f} \Big[\Re\big\langle y|\varphi\big\rangle + \lambda \Re\big\langle y|\varphi_f\big\rangle - f(y) \Big] \\ &\leqq \underbrace{\sup_{y \in \mathcal{D}_f} \Big[\Re\big\langle y|\varphi\big\rangle - f(y) \Big]}_{f^*(\varphi)} + \lambda \sup_{y \in \mathcal{D}_f} \Big[\Re\big\langle y|\varphi_f\big\rangle \Big]. \end{aligned}$$

Hence we have;

$$f^*(\chi) - f^*(\varphi) \leqq \sup_{y \in \mathcal{D}_f} \Big[\Re\big[\big\langle y|\chi - \varphi\big\rangle\big] \Big] = \lambda \sup_{y \in \mathcal{D}_f} \Big[\Re\big\langle y|\varphi_f\big\rangle \Big].$$

That is, $\chi = \varphi + \lambda\varphi_f$ is in the domain of f^*. By the Fenchel-Young inequality for biconjugate function we have, $(f)^*_*(x) + f^*(\chi) \geqq \Re\big\langle x|\chi\big\rangle$. Substituting for $f^*(\chi)$ we obtain,

$$(f)^*_*(x) \geqq \Big[\Re\langle x|\varphi\rangle_{\mathcal{X}} - f^*(\varphi) \Big] + \lambda \Big[\inf_{y \in \mathcal{D}_f} \big[\Re\langle x - y|\varphi_f\rangle \big] \Big].$$

Since $f(x) - f(y) \geqq \Re\langle x - y|\varphi_f\rangle$ it follows that

$$(f)^*_*(x) \geqq \lambda_{x,\varphi} + \lambda\big[f(x) - f(y)\big]$$

where $\lambda_{x,\varphi} = \Re\langle x|\varphi\rangle - f^*(\varphi)$ is a finite. Since f is a proper function, there must exist some y for which $f(y)$ is finite. Since λ can be arbitrarily large it follows that $(f)^*_*(x) \geqq \infty$, whenever $f(x) = \infty$. This completes the proof. □

Since there exists a natural embedding $\iota_{\mathcal{X}}$ of $\mathcal{X}$ into $\mathcal{X}^{**}$, for every function h on $\mathcal{X}^{**}$, we obtain a function on $\mathcal{X}$ given by the composition of the mappings; $h \circ \iota_{\mathcal{X}} : \mathcal{X} \to \mathbb{R}$. Since $\mathcal{G} \circ \mathcal{F}$ is an isomorphism on $\overline{\Gamma}(\mathcal{X})$ it follows that for any f in $\overline{\Gamma}(\mathcal{X})$, we must have, $f(x) = (f)^*_*(x) = (f^*)^*(\iota_{\mathcal{X}}(x))$.

The conjugate function f^* of f is lower semi-continuous regardless of the continuity properties of f. Hence $(f)^*_*$ acts as weakly lower semi-continuous approximate of f. If the second conjugate $(f)^*_*$ is identical to f, then it immediately follows that the function is a lower semi-continuous function.

1.2.2 | Locally Convex Topologies

Since every locally convex space has absolutely convex neighborhoods of the origin which generates the topology, the Minkowski functionals associated to such neighborhoods are very closely related to the topology and convexity properties of the locally convex space. Hence we can expect to study the topological and convexity properties in terms of the Minkowski functionals. We now make this duality precise by studying how to transfer the topological and convexity properties from the locally convex space to the functions on the space via Minkowski functionals, and back.

A function is said to be on affine functional on a convex set $\mathcal{K}$ if it maps line segments joining points in $\mathcal{K}$ to line segments in the target space. The restriction of an affine functional on $\mathcal{X}$ to $\mathcal{K}$ is an affine functional on $\mathcal{K}$, and conversely, every affine functional can be extended by linearity, and using the Hahn-Banach theorem to an affine functional on $\mathcal{X}$. The space of affine functionals on $\mathcal{K}$ is denoted by $\mathcal{A}(\mathcal{K})$. We have,

$$\mathcal{A}(\mathcal{X}) \cong \mathcal{X}^* \oplus \mathbb{R}.$$

If f is in $\overline{\Gamma}(\mathcal{X})$, then by conjugate duality,

$$(f)^*_*(x) = \sup_{\varphi \in \mathcal{D}_{f^*}} \Big[\Re\langle x|\varphi\rangle - f^*(\varphi) \Big] = f(x).$$

Hence f can be written as a supremum of affine continuous functions. Since each affine continuous functional is a proper lower semi-continuous conex function, the supremum of a collection of affine continuous functionals must also be proper lower semi-continuous. Hence we have proved the following lemma;

Lemma 1.2.5. *A function f is in $\overline{\Gamma}(\mathcal{X})$ if and only if*

$$f(x) = \sup_{\alpha \leq f} \big[\alpha(x)\big],$$

where α are continuous affine functionals on the domain of f.

Note that $(f)^*_*$ is weakly lower semi-continuous. The function $(f)^*_*$ must always be supremum of affine functionals, and hence must possess nice convexity properties. We also know by conjugate duality that the function $(f)^*_*$ is identical to f whenever the proper function f is lower semi-continuous convex function. Hence conjugate duality provides a duality between convexity properties and topological properties.

1.2.2.1 | Supporting Functions & Indicators

Since convex functions on $\mathcal{X}$ are closely related to convex subsets of $\mathcal{X}$, the conjugate duality also establishes a duality between convex subsets of $\mathcal{X}$ and $\mathcal{X}^*$ with appropriate closure properties. To make the relation between convex sets and convex functions precise, we define the notion of supporting functions. We will assume that $\mathcal{X}$ is a locally convex space, and $\mathcal{N}_{\mathcal{X}}$ is an convex neighborhood basis of the origin of $\mathcal{X}$.

For any convex function f on a locally convex space $\mathcal{X}^*$ and λ in $\mathbb{R}$, the set of all χ such that $f(\chi) \leqq \lambda$ is a convex set. A lower semi-continuous proper convex function f is said to be positively homogeneous if for every t in $\mathbb{R}_+$ the function satisfies,

$$f(\lambda\varphi) = \lambda f(\varphi).$$

Positively homogeneous convex functions are well-behaved around the origin.

Since locally convex topologies are closely related to the convex neighborhoods of the origin, we may expect to describe the semi-norms in terms of positively homogeneous functions in $\overline{\Gamma}(\mathcal{X})$. Suppose a function f in $\overline{\Gamma}(\mathcal{X})$ is positively homogeneous, then $f(\lambda x) = \lambda f(x)$ for all λ in $\mathbb{R}^+$ and we must have,

$$\begin{aligned} f^*(\varphi) &= \sup_{x\in\mathcal{X}} \Big[\big| \langle x|\varphi\rangle_{\mathcal{X}} \big| - f(x) \Big] \\ &= \sup_{x\in\mathcal{X}} \Big[\big| \langle \lambda x|\varphi\rangle_{\mathcal{X}} \big| - f(\lambda x) \Big] = \lambda f^*(\varphi). \end{aligned}$$

Hence f^* can only take the values 0 or ∞. Since f^* is a convex function its domain $\mathcal{D}_{f^*}$ is a closed convex set which contains origin. Hence we must have,

$$(f)^*_*(x) = \sup_{\varphi\in\mathcal{D}_{f^*}} \Big[\big| \langle x|\varphi\rangle \big| \Big] = f(x).$$

We denote the positively homogeneous functions in $\overline{\Gamma}(\mathcal{X})$ by $\overline{\Gamma}_+(\mathcal{X})$. Every function in $\overline{\Gamma}_+(\mathcal{X})$ is the supremum of linear functionals in the domain of f^*. Hence we have proved the following lemma;

Lemma 1.2.6. *A function f is in $\overline{\Gamma}_+(\mathcal{X})$ if and only if*

$$f(x) = \sup_{\varphi\in\mathcal{K}} \Big[\big| \langle x|\varphi\rangle_{\mathcal{X}} \big| \Big],$$

for a convex neighborhood $\mathcal{K}$ of the origin in $\mathcal{X}^*$.

The semi-norms of a locally convex topological space $\mathcal{X}$ correspond to positively homogeneous continuous proper convex functions with respect to the topology. By the above lemma, every semi-norm must be a supremum over some convex neighborhood of the origin in $\mathcal{X}^*$. Hence the dual space allows us to describe the topology in terms of its neighborhoods of the origin. For any subset $\mathcal{N}$ of $\mathcal{X}$ the supporting function of $\mathcal{N}$ is defined by,

$$H_{\mathcal{N}}(\varphi) = \sup_{x\in\mathcal{N}} \Big[\big| \langle x|\varphi\rangle \big| \Big].$$

By linearity of each φ, it follows immediately that

$$H_{\mathcal{N}} \equiv H_{\langle\langle\langle\mathcal{N}\rangle\rangle\rangle}.$$

Clearly, supporting functions are positively homogeneous. As supremum over continuous functionals, supporting functions of non-empty sets are lower semi-continuous proper convex functions. Hence we have, $H_{\mathcal{N}} = (H_{\mathcal{N}})^*_*$. The Legendre transform of $H_{\mathcal{N}}$ is called the indicator of $\mathcal{N}$ and is denoted by $I_{\mathcal{N}}$. We have,

$$I_{\mathcal{N}} = (H_{\mathcal{N}})^* \; ; \; H_{\mathcal{N}} = (I_{\mathcal{N}})_*.$$

Upon expanding the definition of Legendre-Fenchel transform for $H_{\mathcal{N}}$ we obtain,

$$I_{\mathcal{N}}(x) \equiv \sup_{\varphi\in\mathcal{X}^*} \inf_{y\in\mathcal{N}} \Big[\big| \langle x|\varphi\rangle \big| - \big| \langle y|\varphi\rangle \big| \Big].$$

Hence the indicator $I_{\mathcal{N}}$ must vanish inside $\mathcal{N}$, and be infinite outside.

1.2.2.2 | Minkowski Functionals & Duality

The Minkowski functional associated with the elements from a basis of neighborhoods of the origin consisting of convex sets of a locally convex space generate its topology, and they are convenient to study since we can apply conjugate duality to Minkowski functionals. However we usually do not have the neighborhood basis to work with. Hence it is useful to have tools that allow us to construct such absolutely convex neighborhoods of the origin starting from functions on the space.

We can use this property to construct absolutely convex neighborhoods around origin of $\mathcal{X}^*$, using the supporting functions. Let $\mathcal{N}$ be any set in $\mathcal{X}$, and let $H_{\mathcal{N}}$ be its supporting function. Then for any λ, let $\mathcal{N}^\lambda$ denote the absolutely convex set of all χ in $\mathcal{X}^*$ such that $H_{\mathcal{N}}(\chi) \leqq \lambda$. We can represent $\mathcal{N}^\lambda$ more conveniently in terms of its indicator as;

$$I_{\mathcal{N}^\lambda} \equiv I_{[H_{\mathcal{N}} \leqq \lambda]},$$

where $[H_{\mathcal{N}} \leqq \lambda]$ is the set of all x such that $H_{\mathcal{N}}(x) \leqq \lambda$. Since $H_{\mathcal{N}}$ is positively homogeneous it follows that K^1 is a closed absolutely convex neighborhood of the origin. $\mathcal{N}^1$ is called the polar of $\mathcal{N}$ and is denoted by $\mathcal{N}^\circ$. We have,

$$I_{\mathcal{N}^\circ} \equiv I_{[H_{\mathcal{N}} \leqq 1]}.$$

Similarly, for any subset $\mathcal{M}$ of $\mathcal{X}^*$, we define the polar $\mathcal{M}_\circ$ of $\mathcal{M}$ in $\mathcal{X}$ to be the absolutely convex neighborhood of the origin given by,

$$I_{\mathcal{M}_\circ} \equiv I_{[H_{\mathcal{M}} \leqq 1]}.$$

Hence we can obtain absolutely convex neighborhoods of the origin using the subsets of the dual space and their supporting functions.

Relation to Supporting Functions

We denote by $\mathcal{N}_{\mathcal{X}}$ and $\mathcal{N}_{\mathcal{X}^*}$ the balanced, absorbing, absolutely convex neighborhood basis of the origin which generate the locally convex topologies $\mathcal{X}$ and $\mathcal{X}^*$ respectively. By taking the interior of closed sets, or by taking closures of open sets, we can study the topologies by either by their closed sets or open sets.

Recall that the Minkowski functional $\mu_{\mathcal{N}}$ corresponding to an absolutely convex neighborhood $\mathcal{N}$ of the origin is corresponds to be the function which assigns to each x, the minimal descaling of x needed to bring it inside $\mathcal{N}$. It is hence defined by,

$$\mu_{\mathcal{N}}(x) = \inf_{x \in \lambda \mathcal{N}} \big[|\lambda|\big].$$

The Minkowski functional is a semi-norm, and for every λ in $\mathbb{R}^+$,

$$\mu_{\mathcal{N}}(\lambda x) = \lambda \mu_{\mathcal{N}}(x).$$

The Minkowski functional is always greater than 1 outside of $\mathcal{N}$ and less than 1 inside of $\mathcal{N}$. Hence, the Minkowski functional can also be expressed as an infimum over all indicators for neighborhoods $\lambda\mathcal{N}$, and for every λ and hence we have,

$$\mu_{\mathcal{N}}(x) = \inf_{x \in \lambda \mathcal{N}} \big[|\lambda|\big] = \inf_{\lambda \in \mathbb{R}^+} \big[I_{\lambda \mathcal{N}}(x) + \lambda\big].$$

This is justified because if x belongs to $\lambda\mathcal{N}$ then $I_{\lambda\mathcal{N}}$ is zero, and if λ is too small and x lies outside of $\lambda\mathcal{N}$ then $I_{\lambda\mathcal{N}}$ is infinite. Hence the infimum is attained for the minimal λ for which x lies inside $\lambda\mathcal{N}$. By the definition of the conjugate function of $\mu_{\mathcal{N}}$ is then given by,

$$\begin{aligned}\big(\mu_{\mathcal{N}}\big)^*(\chi) &= \sup_{\lambda\in\mathbb{R}^+}\big[H_{\lambda\mathcal{N}}(\chi)-\lambda\big]\\ &= \sup_{\lambda\in\mathbb{R}^+}\big[\lambda(H_{\mathcal{N}}(\chi)-1)\big].\end{aligned}$$

For the last step; $\lambda(H_{\mathcal{N}}(\chi)-1)$ is positive whenever χ is outside of $\mathcal{N}^\circ$, in which case the above supremum will be infinite. When χ is inside $\mathcal{N}^\circ$, $H_{\mathcal{N}}(\chi)-1$ is negative, and hence the supremum must be zero. Hence we must have,

$$\big(\mu_{\mathcal{N}}\big)^* = I_{\mathcal{N}^\circ}.$$

Similarly, if $\mathcal{M}$ is in $\mathcal{N}_{\mathcal{X}^*}$, then we must have

$$\big(\mu_{\mathcal{M}}\big)_* = I_{\mathcal{M}_\circ}.$$

Since the polar sets are absolutely convex neighborhoods of the origin, we can define their Minkowski functionals, and by the above discussion for any subset $\mathcal{K}$ in $\mathcal{X}^*$ we must have,

$$\mu_{\mathcal{K}_\circ} = \big(I_{\mathcal{K}}\big)_*.$$

Similarly, for any subset $\mathcal{L}$ of $\mathcal{X}$ we must have,

$$\mu_{\mathcal{L}^\circ} \equiv \big(I_{\mathcal{L}}\big)^*.$$

We have, $I_{(\mathcal{N}^\circ)_\circ} = (\mu_{\mathcal{N}^\circ})_* = (I_{\mathcal{N}})^*_* = I_{\mathcal{N}}$, and since $H_{\mathcal{N}} = H_{\langle\!\langle\mathcal{N}\rangle\!\rangle}$ whenever $\mathcal{N}$ contains the origin, we must have $I_{(\mathcal{N}^\circ)_\circ} = I_{\langle\!\langle\mathcal{N}\rangle\!\rangle}$. If $\langle\!\langle\!\langle\mathcal{M}\rangle\!\rangle\!\rangle^*$ is the absolutely convex weak*-closure of $\mathcal{M}$. Then we should have $I_{(\mathcal{M}_\circ)^\circ} = I_{\langle\!\langle\!\langle\mathcal{M}\rangle\!\rangle\!\rangle^*}$ whenever $\mathcal{M}$ contains the origin of $\mathcal{X}^*$. Hence we have proved the bipolar theorem;

Theorem 1.2.7. (Bipolar Theorem) *If $\mathcal{N}\subseteq\mathcal{X}, \mathcal{M}\subseteq\mathcal{X}^*$ contains the origin, then*

$$\begin{aligned}(\mathcal{N}^\circ)_\circ &= \langle\!\langle\!\langle\mathcal{N}\rangle\!\rangle\!\rangle,\\ (\mathcal{M}_\circ)^\circ &= \langle\!\langle\!\langle\mathcal{M}\rangle\!\rangle\!\rangle^*\end{aligned}$$

For any non-empty absolutely convex neighborhood $\mathcal{N}$ of the origin the polar $\mathcal{N}^\circ$ is an absolutely convex neighborhood of the origin for the dual space $\mathcal{X}^*$. For any x by definition of the Minkowski functional we have $x\in\mu_{\mathcal{N}}(x)\mathcal{N}$, and for any small ϵ, $\lambda_\epsilon = [\mu_{\mathcal{N}}(x)+\epsilon]$ is such that $x\in\lambda_\epsilon\mathcal{N}$. Similarly, by taking $\lambda_{\epsilon^*} = [\mu_{\mathcal{N}^\circ}(\chi)+\epsilon]$ we must have, $\chi\in\lambda_{\epsilon^*}\mathcal{N}^\circ$. Hence by the definition of polar sets we must have

$$\big\langle\lambda_\epsilon^{-1}x|\lambda_{\epsilon^*}^{-1}\chi\big\rangle \leqq 1.$$

Since $\mu_{\mathcal{N}} = H_{\mathcal{N}^\circ}$, and $\mu_{\mathcal{N}^\circ} = H_{(\mathcal{N}^\circ)_\circ} = H_{\mathcal{N}}$. Hence we must have,

$$\big\langle x|\chi\big\rangle \leqq \big[H_{\mathcal{N}}(\chi)+\epsilon\big]\big[H_{\mathcal{N}^\circ}(x)+\epsilon\big].$$

Here ϵ can be arbitrarily small. This proves a Cauchy-Schwarz type inequality for the conjugate pairing in terms of the neighborhood bases of the origin and their polars.

Lemma 1.2.8. (Cauchy-Schwarz)

$$\langle x|\chi\rangle \leqq H_{\mathcal{N}}(\chi)H_{\mathcal{N}^\circ}(x),$$

whenever $\mathcal{N}$ is an absolutely convex neighborhood of the origin.

The space of all continuous functions from $\mathcal{X}$ to $\mathbb{K}$ is a subset of the product space $\mathbb{K}^{\mathcal{X}} = \prod_{x\in\mathcal{X}} \mathbb{K}$, equipped with the product topology. The space $\mathcal{X}^*$ of continuous linear functionals on $\mathcal{X}$ is a subset of $\mathbb{K}^{\mathcal{X}}$. Given a subset $\mathcal{U}$ of $\mathcal{X}^*$, we denote by $\langle \cdot \,|\, \mathcal{U}\rangle$ the image of $\mathcal{U}$ under the map,

$$\varphi \mapsto \langle \,\cdot\,|\,\varphi\rangle \in \mathbb{K}^{\mathcal{X}}.$$

Since linearity, and topology are related to each other on $\mathbb{K}$, and since the product topology on $\mathbb{K}^{\mathcal{X}}$ corresponds to pointwise evaluations, it follows that $\mathcal{X}^*$ is a closed subset of $\mathbb{K}^{\mathcal{X}}$. The product topology on $\mathbb{K}^{\mathcal{X}}$ restricted to continuous functionals coincides with the weak* topology on $\mathcal{X}^*$.

Let $\mathcal{X}$ be a locally convex space and $\mathcal{N}_{\mathcal{X}}$ be a neighborhood basis of the origin of absolutely convex subsets. For each element $\mathcal{N}$ of $\mathcal{N}_{\mathcal{X}}$ the associated Minkowski functional $\mu_{\mathcal{N}}$ possesses useful convexity property, and continuity properties. We now exploit these properties, to study topological properties of subsets of the dual space $\mathcal{X}^*$.

Theorem 1.2.9. (Alaoglu-Bourbaki) *If $\mathcal{N}$ is in $\mathcal{N}_{\mathcal{X}}$ then $\mathcal{N}^\circ$ is weak*-compact.*

The supporting functions can 'see sets only up to their absolutely convex closures'. Hence, we can isolate the relation between convexity and topology by constructing topologies using supporting functions. Since supporting function are proper, lower semi-continuous with respect to the weak topologies we can make use of conjugate duality and related theorems for studying the properties. The topology on $\mathcal{X}$ generated by the collection of semi-norms,

$$\Big\{ x \mapsto \underbrace{\sup_{\varphi\in\mathcal{K}} \Big[\big|\langle x|\varphi\rangle\big| \Big]}_{H_{\mathcal{K}}(x)} \Big\}_{\mathcal{K}\in\mathcal{X}^*},$$

where each $\mathcal{K}$ is weak*-compact subset of $\mathcal{X}^*$, is called the Mackey topology on $\mathcal{X}$, denoted by $\tau(\mathcal{X},\mathcal{X}^*)$. The topology on $\mathcal{X}^*$ generated by the collection of semi-norms,

$$\Big\{ \varphi \mapsto \underbrace{\sup_{x\in\mathcal{N}} \Big[\big|\langle x|\varphi\rangle\big| \Big]}_{H_{\mathcal{N}}(\varphi)} \Big\}_{\mathcal{N}\in\mathcal{X}},$$

where each $\mathcal{N}$ is weakly compact subset of $\mathcal{X}$ is called the Mackey topology on $\mathcal{X}^*$ denoted by $\tau(\mathcal{X}^*,\mathcal{X})$. The Mackey topologies isolate the convexity aspects of a space with respect to the pairing space, and hence we may expect the Mackey topologies to have some universal property for locally convex topologies.

Since $\sigma(\mathcal{X},\mathcal{X}^*)$ is the weakest topology such that the maps $x \mapsto \langle x|\varphi\rangle_{\mathcal{X}}$ are continuous for all φ in $\mathcal{X}^*$, and since $(\mathcal{X},\tau)^* = \mathcal{X}^*$ we must have, $\sigma(\mathcal{X},\mathcal{X}^*) \subseteq \tau$.

We must now prove that for any τ-continuous linear functional φ the map $|\varphi|$ is one of the semi-norm generating the topology $\tau(\mathcal{X},\mathcal{X}^*)$, that is, there exists some weak*-compact subset $\mathcal{K}_\varphi$ of $\mathcal{X}^*$ such that $|\varphi| = H_{K_\varphi}$. Since $(\mathcal{X},\tau)^* = \mathcal{X}^*$, the functional φ must be in $\mathcal{X}^*$.

The polar $\{\varphi\}_\circ$ is a weakly-closed absolutely convex neighborhood of the origin in $\mathcal{X}$. By the Alaoglu-Bourbaki theorem $(\{\varphi\}_\circ)^\circ$ must be weak*-compact. Let

$$\mathcal{K}_\varphi \coloneqq \big(\{\varphi\}_\circ\big)^\circ.$$

By the bipolar theorem it is the closed absolutely convex hull generated by φ and the origin in $\mathcal{X}^*$. Hence the map $\chi \mapsto |\langle x|\chi\rangle|$ attains suprema at φ in the set $\mathcal{K}_\varphi$. Hence we have,

$$\big|\varphi(x)\big| = \sup_{\chi\in\mathcal{K}_\varphi}\Big[\big|\langle x|\chi\rangle\big|\Big] \equiv H_{\mathcal{K}_\varphi}(x).$$

Hence φ must be $\tau(\mathcal{X},\mathcal{X}^*)$-continuous, since it is one of the semi-norms generating $\tau(\mathcal{X},\mathcal{X}^*)$. Hence we have $\tau \subseteq \tau(\mathcal{X},\mathcal{X}^*)$. The following theorem due to Arens & Mackey says that $\tau(\mathcal{X},\mathcal{X}^*)$ is also compatible with the pairing.

Theorem 1.2.10. (Mackey-Arens [47])

$$(\mathcal{X},\tau)^* = \mathcal{X}^* \Leftrightarrow \sigma(\mathcal{X},\mathcal{X}^*) \subseteq \tau \subseteq \tau(\mathcal{X},\mathcal{X}^*).$$

Let $\mathcal{N}^\circ_\mathcal{X}$ denote the set of all polars $\mathcal{N}^\circ$ of neighborhoods $\mathcal{N}$ in $\mathcal{N}_\mathcal{X}$. Let $\mathcal{X}^{*\circ}$ denote the vector space of all functionals on $\mathcal{X}^*$ which are weak*-continuous on each subset $\mathcal{N}^\circ$ of $\mathcal{X}^*$ for every $\mathcal{N}$ in $\mathcal{N}_\mathcal{X}$. Since restriction of weak*-continuous functionals are automatically weak*-continuous, we must have

$$\mathcal{X} \subseteq \mathcal{X}^{*\circ}$$

Every element x of $\mathcal{X}$ acts as a linear functional on $\mathcal{X}^*$, given by $\langle x|\cdot\rangle_\mathcal{X} : \varphi \mapsto \langle x|\varphi\rangle_\mathcal{X}$. We denote the vector space of all such functionals on $\mathcal{X}^*$ by $\langle\mathcal{X}|\cdot\rangle_\mathcal{X}$.

Theorem 1.2.11. (Grothendieck [1]) *$\mathcal{X}$ is complete if and only if $\mathcal{X}^{*\circ} \cong \big\langle\mathcal{X}|\cdot\big\rangle_\mathcal{X}$.*

A locally convex space, whose topology is induced by a translation invariant metric and is complete with respect to the metric is called a Fréchet space. By the translation invariance we can describe the metric in terms of the convex function $\|\cdot\|$.

Theorem 1.2.12. (Krein-Smulian [1]) *Let $\mathcal{X}$ be a Fréchet space. A linear functional on $\mathcal{X}^*$ is $\sigma(\mathcal{X}^*,\mathcal{X})$-continuous iff its restriction to $\mathcal{N}^\circ$ is $\sigma(\mathcal{X}^*,\mathcal{X})$-continuous, $\forall\mathcal{N}\in\mathcal{N}_\mathcal{X}$.*

1.2.3 | Conjugate Duality for Banach Space

Let $\mathcal{X}$ be a Banach space. By linearity, the unit ball $B_\mathcal{X}$ in $\mathcal{X}$ consisting of all x in $\mathcal{X}$ with $\|x\|_\mathcal{X} \leqq 1$ contains the data required to generate all lines of $\mathcal{X}$. The supporting function of $B_\mathcal{X}$ is denoted by,

$$\big\|\varphi\big\|_{\mathcal{X}^*} \coloneqq \sup_{\|x\|_\mathcal{X}\leqq 1}\Big[\big|\langle x|\varphi\rangle_\mathcal{X}\big|\Big] \equiv H_{B_\mathcal{X}}(\varphi).$$

As a supremum of weak* continuous functionals $\|\cdot\|_{\mathcal{X}^*}$ is in $\overline{\Gamma}_+(\mathcal{X}^*)$. $\|\cdot\|_{\mathcal{X}^*}$ defines a norm on $\mathcal{X}^*$, and $\mathcal{X}^*$ equipped with the norm $\|\cdot\|_{\mathcal{X}^*}$ is a Banach space.

Any subspace $\mathcal{Y}$ of $\mathcal{X}$ inherits the norm $\|\cdot\|_\mathcal{X}$ by restriction, and hence we can describe the dual space $\mathcal{Y}^*$ as $\|\cdot\|_\mathcal{X}$-bounded functionals φ with the norm, $\|\varphi\|_{\mathcal{Y}^*} \equiv H_{B_\mathcal{Y}}(\varphi)$, where $B_\mathcal{Y}$ is the collection of all y in $\mathcal{Y}$ with $\|y\|_\mathcal{X} \leqq 1$. The map

$$f^\varphi_\mathcal{X}(x) = \big\|\varphi\big\|_{\mathcal{Y}^*}\big\|x\big\|_\mathcal{X}$$

is a sub-linear map since the norm $\|\cdot\|_{\mathcal{X}}$ on $\mathcal{X}$ is sub-additive. Hence by the Hahn-Banach theorem, there exists an extension $\widehat{\varphi}$ of φ such that $|\langle x|\widehat{\varphi}\rangle_{\mathcal{X}}| \leqq f^{\varphi}_{\mathcal{X}}(x) = \|\varphi\|_{\mathcal{Y}^*}\|x\|_{\mathcal{X}}$. Since $\|\cdot\|_{\mathcal{X}^*}$ is the supremum over all such functionals, it follows that $\|\widehat{\varphi}\|_{\mathcal{X}^*}$ is less than $\|\varphi\|_{\mathcal{Y}^*}$. Since $\mathcal{X}$ already contains $\mathcal{Y}$ we have

$$\|\widehat{\varphi}\|_{\mathcal{X}^*} = \|\varphi\|_{\mathcal{Y}^*}.$$

For any x in $\mathcal{X}$, we can consider the subspace $\mathbb{K}x$ generated by x, and define a linear functional on $\mathbb{K}x$ by, $\lambda x \mapsto \lambda\|x\|^2_{\mathcal{X}}$. This is clearly a bounded linear functional on $\mathbb{K}x$, and by Hanh-Banach theorem, can be thought of as a restriction of the linear functional φ_x with the same norm. The norm of the linear functional φ_x is given by,

$$\big\|\varphi_x\big\|_{\mathcal{X}^*} = \sup_{\|\lambda x\|\leqq 1}\Big[\big|\big\langle \lambda x\big|\varphi_x\big\rangle_{\mathcal{X}}\big|\Big] = \big\|x\big\|_{\mathcal{X}},$$

with

$$\big\langle x|\varphi_x\big\rangle_{\mathcal{X}} = \|x\|^2_{\mathcal{X}}.$$

Since $\|\cdot\|_{\mathcal{X}^*} = H_{B_{\mathcal{X}}}$, by the Cauchy-Schwarz inequality we must have $|\langle x|\varphi\rangle_{\mathcal{X}}| \leqq \|x\|_{\mathcal{X}}\|\varphi\|_{\mathcal{X}^*}$. The set $B_{\mathcal{X}^*}$ consists of all φ which are majorised by $\|\cdot\|_{\mathcal{X}}$. Hence it follows that $\|\cdot\|_{\mathcal{X}}$ is a supremum of weak*-continuous functionals, $\|\cdot\|_{\mathcal{X}}$ must be weak* lower semi-continuous and positively homoegneous. Hence it must belong to $\overline{\Gamma}_+(\mathcal{X})$. Hence we have proved the following theorem;

Theorem 1.2.13. *If $\|\cdot\|_{\mathcal{X}}$ is a norm on $\mathcal{X}$, then $\|\cdot\|_{\mathcal{X}}$ is in $\overline{\Gamma}_+(\mathcal{X})$, with*

$$\big\|x\big\|_{\mathcal{X}} = \sup_{\varphi\leq\|\cdot\|_{\mathcal{X}}}\Big[\big|\big\langle x\big|\varphi\big\rangle_{\mathcal{X}}\big|\Big].$$

Heuristically, the norm already contains data about all line segments on the space, and relates each line segment to a line-segment on the real line. Since the topology and vector space structure on $\mathbb{R}$ are related to each other, we may expect the norm to introduce such a relation on the vector space.

1.2.3.1 | Bidual of Banach Spaces

Let $\mathcal{X}$ be a Banach space with respect to a norm $\|\cdot\|_{\mathcal{X}}$. We may expect topological closure properties to be closely related to closure with respect to convexity structure, as is the case for real numbers. The bidual of $\mathcal{X}$ also inherits the operator norm from the norm $\|\cdot\|_{\mathcal{X}^*}$, defined by $\|\cdot\|_{\mathcal{X}^{**}} \coloneqq H_{B_{\mathcal{X}^*}}$ and hence

$$\big\|X\big\|_{\mathcal{X}^{**}} = \sup_{\varphi\in B_{\mathcal{X}^*}}\Big[\big|\big\langle \varphi\big|X\big\rangle_{\mathcal{X}^*}\big|\Big].$$

The conjugate transform of $\|\cdot\|_{\mathcal{X}}$ is $(\|\cdot\|_{\mathcal{X}})^*(\varphi) = \sup_{\lambda\in\mathbb{R}^+}\lambda\sup_{x\in B_{\mathcal{X}}}\big[|\langle x|\varphi\rangle_{\mathcal{X}}| - \|x\|_{\mathcal{X}}\big]$. Whenever φ is not in $B_{\mathcal{X}^*}$, then we have $\|\varphi\|_{\mathcal{X}^*} > 1$, then there exists some x in $B_{\mathcal{X}}$ such that $\langle x|\varphi\rangle > 1$. Hence $(\|\cdot\|_{\mathcal{X}})^*(\varphi)$ must be infinite. When φ is in $B_{\mathcal{X}^*}$, by the Cauchy-Schwarz inequality we must have, $(\|\cdot\|_{\mathcal{X}})^*(\varphi) = \sup_{x\in B_{\mathcal{X}}}\big[|\langle x|\varphi\rangle_{\mathcal{X}}| - \|x\|_{\mathcal{X}}\big] \leqq \sup_{x\in B_{\mathcal{X}}}\big[\|x\|_{\mathcal{X}}[\|\varphi\|_{\mathcal{X}^*} - 1]\big] \leqq 0$ which implies that $(\|\cdot\|_{\mathcal{X}})^*(\varphi)$ must be zero. Hence it follows that

$$(\|\cdot\|_{\mathcal{X}})^* = I_{B_{\mathcal{X}^*}}.$$

For every f in $\overline{\Gamma}(\mathcal{X})$ we must have $(\mathcal{F} \circ \mathcal{F}f)(\iota_{\mathcal{X}} x) = f(x)$. Hence we must have $\|x\|_{\mathcal{X}} = \|\iota_{\mathcal{X}}(x)\|_{\mathcal{X}^{**}}$. Hence the canonical embedding of $\mathcal{X}$ into $\mathcal{X}^{**}$ is an isometric embedding. In terms of the polar sets, we have $(B_{\mathcal{X}})^{\circ} = B_{\mathcal{X}^*}$. Hence we have

$$(B_{\mathcal{X}}{}^{\circ})^{\circ} = (B_{\mathcal{X}^*})^{\circ} = B_{\mathcal{X}^{**}}.$$

From the relation between indicators and Minkowski functionals we must have,

$$I_{B_{\mathcal{X}^{**}}} = I_{((B_{\mathcal{X}})^{\circ})^{\circ}} = (\mu_{(B_{\mathcal{X}})^{\circ}})^* = (I_{B_{\mathcal{X}}})^{**}.$$

Since $f^{**}(\iota_{\mathcal{X}}(x)) = f(x)$ for every function f in $\overline{\Gamma}(\mathcal{X})$, it follows that $B_{\mathcal{X}^{**}}$ is the weak** closure of the set $\iota_{\mathcal{X}}(B_{\mathcal{X}})$. Hence we have proved the following theorem;

Theorem 1.2.14. (Goldstine) *$\iota_{\mathcal{X}}(B_{\mathcal{X}})$ is weak* dense in $B_{\mathcal{X}^{**}}$.*

1.3 | The σ-Weak Topology & Lattice Structures

The pairings of mathematical objects allows the objects to share their properties and their properties work synergistically with the pairing objects. Heuristically speaking the pairing reduce the burden on each of the objects. As we had discussed in the formulation of quantum theory, the space of effects and the state space are related to each other, complementarily. This complementary relation is closely related to the support properties of the states viewed as linear functionals. The support of an ensemble is related to the largest effect for which the pairing of the ensemble and the effect do not vanish. We now make this complementary relation precise in the context of abstract W^*-algebras and establish a relation between the predual and projections of the W^*-algebra. This leads to the uniqueness of predual, due to Sakai [36]. Sakai's approach makes use of the notion of spectrum and related concepts. From our physically motivated point of view, the notion of spectrum is closely related to the physical notion of observables. Since we have not yet introduced the notion of observable we will refrain from using the tools of spectral theory.

Sakai's [36] definition of W^*-algebras is in the spirit of how we formulated quantum theory. Hence we follow a mixed approach to studying the relation between a W^*-algebra and its predual, with definitions from Sakai [36] and utilising the methods of Takesaki & Tomiyama [39]-[43], which makes the role of the pairing clear.

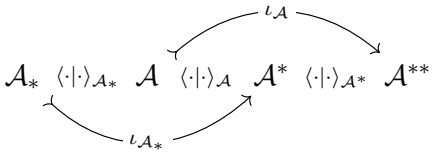
1.3.1 | The Bidual of W^*-Algebras

Let $\mathcal{A}$ be a W^*-algebra and let $\mathcal{A}_*$ be its predual. The weak* topology on $\mathcal{A}$ coming from the predual $\mathcal{A}_*$ is called the σ-weak topology. We have an isometric embedding of $\mathcal{A}$ inside the bidual $\mathcal{A}^{**}$,

$$\iota_{\mathcal{A}} : \mathcal{A} \to \mathcal{A}^{**}.$$

The pairing of $\mathcal{A}^*$ with $\mathcal{A}^{**}$ is denoted by $\langle\cdot|\cdot\rangle_{\mathcal{A}^*}$.

The W^*-algebra $\mathcal{A}$ is itself the dual space of $\mathcal{A}_*$, with the pairing denoted by $\langle\cdot|\cdot\rangle_{\mathcal{A}_*}$. $\mathcal{A}^*$ is the bidual for $\mathcal{A}_*$. Hence we obtain an isometric embedding $\iota_{\mathcal{A}_*}$ of $\mathcal{A}_*$ inside the bidual $\mathcal{A}^*$. The setup of our structural analysis is the following pair of isometric embeddings,



with

$$\langle \omega | A \rangle_{\mathcal{A}_*} = \langle A | \iota_{\mathcal{A}_*}(\omega) \rangle_{\mathcal{A}} = \langle \iota_{\mathcal{A}_*}(\omega) | \iota_{\mathcal{A}}(A) \rangle_{\mathcal{A}^*}$$

for all A in $\mathcal{A}$ and ω in $\mathcal{A}_*$. This allows us to transfer the properties of any of these Banach spaces to their bidual when such a property also respects the pairing. We begin by studying the embedding of the W^*-algebra inside its second conjugate.

1.3.1.1 | $\mathcal{A}$-Module Structure on $\mathcal{A}^*$

The product structure on $\mathcal{A}$ allows us to precompose a continuous linear functional with an element of $\mathcal{A}$, and similarly we can also describe a post composition of continuous linear functionals with elements of $\mathcal{A}$. This introduces algebraic actions of $\mathcal{A}$ on $\mathcal{A}^*$. Since $\mathcal{A}$ is a C^*-algebra, we must have, $\|AB\| \leqq \|A\|\|B\|$. Hence for any φ in $\mathcal{A}^*$, we have

$$|\varphi(AB)| \leqq \|AB\|_{\mathcal{A}} \|\varphi\|_{\mathcal{A}^*} \leqq \|A\|_{\mathcal{A}} \|B\|_{\mathcal{A}} \|\varphi\|_{\mathcal{A}^*}.$$

Hence the functional $B \mapsto \varphi(AB)$ defines a bounded linear functional on $\mathcal{A}$. Similarly the mapping $B \mapsto \varphi(BA)$ is also a bounded linear functional on $\mathcal{A}$. Hence the pairing allows us to introduce an $\mathcal{A}$-module structure on $\mathcal{A}^*$. Hence we define, $A\varphi$ and φA in $\mathcal{A}^*$ by

$$\begin{aligned} \langle A(\varphi) | B \rangle_{\mathcal{A}} &\coloneqq \langle \varphi | BA \rangle_{\mathcal{A}} \\ \langle (\varphi)A | B \rangle_{\mathcal{A}} &\coloneqq \langle \varphi | AB \rangle_{\mathcal{A}}. \end{aligned}$$

Similarly, we define,

$$\langle \varphi^\dagger | B \rangle_{\mathcal{A}} \coloneqq \overline{\langle \varphi | B^\dagger \rangle}_{\mathcal{A}}$$

for all B in $\mathcal{A}$. It immediately follows that $(\varphi A)^\dagger = A^\dagger \varphi^\dagger$.

1.3.1.2 | Arens Product on $\mathcal{A}^{**}$

The product structures on C^*-algebras respect the norm, and linearity, and hence the convexity properties are respected. The dual space $\mathcal{A}^*$ comes equipped with an $\mathcal{A}$-bimodule structure via the pairing $\langle \cdot | \cdot \rangle_{\mathcal{A}}$. We now use this structure to extend the product structure on $\mathcal{A}$ to obtain a product structure on $\mathcal{A}^{**}$ using the module structure on $\mathcal{A}^*$, and allowing it to *carry* the product structure via the pairing.

Since $A\varphi$ and φA are continuous linear functionals in $\mathcal{A}^*$, every element X in $\mathcal{A}^{**}$ can act on these elements, we define the action of $\mathcal{A}^{**}$ on $\mathcal{A}^*$ by

$$\langle A | X\varphi \rangle_{\mathcal{A}} \equiv \langle A\varphi | X \rangle_{\mathcal{A}^*}, \quad \langle A | \varphi X \rangle_{\mathcal{A}} \equiv \langle \varphi A | X \rangle_{\mathcal{A}^*}.$$

The functionals $X\varphi$ and φX are bounded since the norm on $\mathcal{A}^*$ can be expressed both as operator norm using the norm on $\mathcal{A}$, as well as the norm on $\mathcal{A}^{**}$. Then for any Y in $\mathcal{A}^{**}$ we define,

$$\begin{aligned} \langle \varphi | X \rtimes Y \rangle_{\mathcal{A}^*} &\equiv \langle \varphi X | Y \rangle_{\mathcal{A}^*}, \\ \langle \varphi | X \ltimes Y \rangle_{\mathcal{A}^*} &\equiv \langle Y\varphi | X \rangle_{\mathcal{A}^*}. \end{aligned}$$

By Goldstine's theorem, since the unit ball of $\mathcal{A}$ is weak* dense in the unit ball of $\mathcal{A}^{**}$, the Krein-Smulian theorem says that weak* continuity is equivalent to weak* continuity on bounded sets. After rescaling if necessary, we can approximate the elements X and Y by elements of $\mathcal{A}$, and hence there exist nets $\{A_\alpha\}_{\alpha \in \mathcal{I}}$ and $\{B_\beta\}_{\beta \in \mathcal{J}}$ such that

$$X = \lim_{\alpha} A_\alpha, \quad Y = \lim_{\beta} B_\beta$$

in the sense of weak* convergence. We abused notation and wrote A_α instead of $\iota_{\mathcal{A}}(A_\alpha)$. We can interchange the weak*-limits and the pairings by continuity, and hence we have;

$$\begin{aligned}\big\langle X\varphi|Y\big\rangle_{\mathcal{A}^*} &= \Big\langle X\varphi\big|\lim_\beta B_\beta\Big\rangle_{\mathcal{A}^*} = \lim_{\beta\in\mathcal{J}}\Big[\big\langle B_\beta|X\varphi\big\rangle_{\mathcal{A}}\Big]\\ &= \lim_{\beta\in\mathcal{J}}\Big[\big\langle B_\beta\varphi\big|X\big\rangle_{\mathcal{A}^*}\Big] = \lim_{\beta\in\mathcal{J}}\Big[\Big\langle B_\beta\varphi\big|\lim_\alpha A_\alpha\Big\rangle_{\mathcal{A}^*}\Big] = \lim_{\beta\in\mathcal{J}}\lim_{\alpha\in\mathcal{I}}\Big[\big\langle B_\beta A_\alpha|\varphi\big\rangle_{\mathcal{A}}\Big].\end{aligned}$$

Similarly, for the other product. Hence we have

$$\begin{aligned}\big\langle\varphi|X\rtimes Y\big\rangle_{\mathcal{A}^*} &= \lim_\beta\lim_\alpha\Big[\big\langle B_\beta A_\alpha|\varphi\big\rangle_{\mathcal{A}}\Big],\\ \big\langle\varphi|X\ltimes Y\big\rangle_{\mathcal{A}^*} &= \lim_\alpha\lim_\beta\Big[\big\langle B_\beta A_\alpha|\varphi\big\rangle_{\mathcal{A}}\Big].\end{aligned}$$

Since the norm on $\mathcal{A}^{**}$ is given by operator norm, it follows that $\mathcal{A}^{**}$ is a Banach algebra with either of the products, by the isometric embedding of $\mathcal{A}$ in $\mathcal{A}^{**}$ together with Goldstine's theorem the norm on $\mathcal{A}^{**}$ is determined. $\mathcal{A}$ is said to be Arens regular when the two products coincide, that is, it does not matter which side of $\mathcal{A}^*$ was used to carry the product structure.

For C^*-algebras, the †-operation is an isometric involution on $\mathcal{A}$, and reverses the order of products. It follows that $X^\dagger\varphi^\dagger = (\varphi X)^\dagger$. This allows us to conveniently manipulate:

$$\big\langle\varphi|X\rtimes Y\big\rangle_{\mathcal{A}^*} = \big\langle\varphi X|Y\big\rangle_{\mathcal{A}^*} = \overline{\big\langle(\varphi X)^\dagger|Y^\dagger\big\rangle}_{\mathcal{A}^*} = \overline{\big\langle X^\dagger\varphi^\dagger|Y^\dagger\big\rangle}_{\mathcal{A}^*}. = \overline{\big\langle\varphi^\dagger|Y^\dagger\ltimes X^\dagger\big\rangle}_{\mathcal{A}^*}.$$

Similarly for $X\ltimes Y$. Since $A_\alpha^\dagger B_\beta^\dagger = (B_\beta A_\alpha)^\dagger$, and since †-operation is an isometry it follows by Goldstine's theorem that the order of taking limits does not matter. We have,

$$\begin{aligned}\big\langle\varphi^\dagger|Y^\dagger\ltimes X^\dagger\big\rangle_{\mathcal{A}^*} &= \lim_\alpha\lim_\beta\Big[\big\langle B_\beta^\dagger A_\alpha^\dagger|\varphi^\dagger\big\rangle_{\mathcal{A}}\Big]\\ &= \lim_\beta\lim_\alpha\Big[\big\langle B_\beta^\dagger A_\alpha^\dagger|\varphi^\dagger\big\rangle_{\mathcal{A}}\Big] = \big\langle\varphi^\dagger|Y^\dagger\rtimes X^\dagger\big\rangle_{\mathcal{A}^*}\end{aligned}$$

Hence it follows that

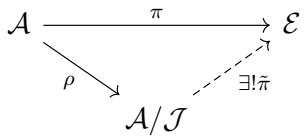
$$\big\langle\varphi|X\rtimes Y\big\rangle_{\mathcal{A}^*} = \big\langle\varphi|X\ltimes Y\big\rangle_{\mathcal{A}^*}.$$

Hence it follows that the bidual of a C^*-algebra is Arens regular. It follows by Goldstine's theorem together with Krein-Smulian theorem that the bidual of a C^*-algebra is also a C^*-algebra. Since $\mathcal{A}^*$ is the predual of $\mathcal{A}^{**}$, it follows that $\mathcal{A}^{**}$ is a W^*-algebra.

A *-homomorphism is a mapping between C^*-algebras is an algebra homomorphism which commutes with the †-operation on the two spaces. If π is a *-homomorphism from $\mathcal{A}$ to a W^*-algebra $\mathcal{E}$, then $\mathcal{J} = \ker(\pi)$ is a two-sided closed ideal of $\mathcal{A}$, and we can make sense of the quotient C^*-algebra $\mathcal{A}/\mathcal{J}$, see §I.8 [40]. Equipped with the norm,

$$\|A+\mathcal{J}\|_\pi = \inf_{J\in\mathcal{J}}\big\{\|A+J\|\big\}.$$

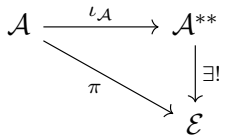
It is a C^*-algebra. The map



$\tilde{\pi}$ is injective, and hence every ω in $\mathcal{E}^*$ acts on $\mathcal{A}/\mathcal{J}$ and the predual of $\mathcal{E}$ acts as continuous functionals on $\mathcal{A}/\mathcal{J}$ with

$$\big\langle\omega|\pi(A)\big\rangle_{\mathcal{E}_*} = \big\langle\iota_{\mathcal{A}_*}\big(\tilde{\pi}^*(\omega)\big)|\iota_{\mathcal{A}}(A)\big\rangle_{\mathcal{A}^*}.$$

Using density of $\iota_{\mathcal{A}} : \mathcal{A} \to \mathcal{A}^{**}$, the composition $\iota_{\mathcal{A}}$ with the transpose of the map $\tilde{\pi}^*$ gives us a *-homomorphism from $\mathcal{A}^{**}$ to $\mathcal{E}$. Hence we have,



For every C^*-algebra, the bidual equipped with the Arens product is the unique W^*-algebra satisfying the above universal property. $\mathcal{A}^{**}$ is called the universal enveloping W^*-algebra of $\mathcal{A}$. Sherman-Takeda theorem, [37]-[38] establishes the above universal property in the framework of Hilbert spaces.

1.3.2 | The Conjugate Pairing $\langle \cdot | \cdot \rangle_{\mathcal{A}_*}$

We now study the additional relation between conjugate pairings using the additional structures on W^*-algebra. We assume that the conjugate pairing is a bilinear pairing for the sake of simplicity. From the point of view of physics, the existence of identity element of the algebra of observables is equivalent to the existence of the system itself, since the identity element corresponds to the effect which is always true for every preparation procedure. Hence we will assume all physically relevant algebras of observables to contain the identity element. Note that the existence of identity for W^*-algebras can be showed using approximate identities. With approximate identities, we can tranfer the approximate identity to the other side of the pairing, and take the limit on the more comfortable side. However, in this thesis we will assume the existence of identity.

Let $\mathcal{A}$ be a W^*-algebra with a predual $\mathcal{A}_*$. The weak*-topology $\sigma(\mathcal{A}, \mathcal{A}_*)$ on $\mathcal{A}$ is called the σ-weak topology. σ-weakly continuous functional on $\mathcal{A}$ are called normal functionals on $\mathcal{A}$. By definition of the σ-weak topology, a linear functional φ is normal if and only if there exists some finite subset F of $\mathcal{A}_*$ and a constant λ_φ such that,

$$\big|\varphi(A)\big| \leqq \lambda_\varphi \sup_{\omega \in F} \Big[\big|\langle \omega | A \rangle_{\mathcal{A}_*}\big|\Big].$$

In the above case we say φ is majorised by the collection F. Note that $\mathbb{C}$ is a W^*-algebra with itself as the predual.

1.3.2.1 | Tensor Product of W^*-Algebras

A mapping $T : \mathcal{A} \to \mathcal{B}$ of W^*-algebras is said to be σ-weakly continuous if it is the adjoint of a continuous linear mapping $T_* : \mathcal{B}_* \to \mathcal{A}_*$ on the preduals. Equivalently, for every φ in $\mathcal{B}_*$, the functional $\varphi \circ T$ belongs in $\mathcal{A}_*$ with

$$\big\langle \varphi \circ T | A \big\rangle_{\mathcal{A}_*} = \big\langle \varphi | TA \big\rangle_{\mathcal{B}_*}$$

for every A in $\mathcal{A}$ and φ in $\mathcal{B}_*$. Hence if a net $\{A_\alpha\}$ converges σ-weakly to A in $\mathcal{A}$, then the net $\{TA_\alpha\}$ also converges σ-weakly to TA in $\mathcal{B}$.

Let $\mathcal{A}$ and $\mathcal{B}$ be two W^*-algebras and $\mathcal{A}_*$ and $\mathcal{B}_*$ be their preduals respectively. Denote by $\mathcal{A}_* \odot \mathcal{B}_*$ the algebraic tensor product of $\mathcal{A}_*$ and $\mathcal{B}_*$ equipped with the norm $\|\omega \odot \varphi\|_{\mathcal{A}_* \otimes \mathcal{B}_*} = \|\omega\|_{\mathcal{A}_*} \|\varphi\|_{\mathcal{B}_*}$. The completion of $\mathcal{A}_* \odot \mathcal{B}_*$ with this norm, is denoted by $\mathcal{A}_* \otimes \mathcal{B}_*$ and is a Banach space. We denote the algebraic tensor product of $\mathcal{A}$ with $\mathcal{B}$ by $\mathcal{A} \odot \mathcal{B}$. The elements of $\mathcal{A} \odot \mathcal{B}$ are of the form $\sum_{i \in F} (A_i \odot B_i)$ where A_i are in $\mathcal{A}$ and B_i are in $\mathcal{B}$, where F is a

finite set. Their action on the elements of $\mathcal{A}_* \otimes \mathcal{B}_*$ is defined pointwise by,

$$\Big\langle \omega \otimes \varphi \big| \sum_{i\in F} \big[A_i \odot B_i\big] \Big\rangle_{(\mathcal{A}\otimes\mathcal{B})_*} := \sum_{i\in F} \big\langle \omega | A_i \big\rangle_{\mathcal{A}_*} \big\langle \varphi | B_i \big\rangle_{\mathcal{B}_*}.$$

where F is a finite set.

We denote the completion of the algebraic tensor product $\mathcal{A} \odot \mathcal{B}$ equipped with the operator norm coming from $\|\cdot\|_{\mathcal{A}_*\otimes\mathcal{B}_*}$ by $\mathcal{A} \otimes \mathcal{B}$, and we call it the tensor product of the W^*-algebra $\mathcal{A}$ with $\mathcal{B}$. Note that $\mathcal{A} \otimes \mathcal{B}$ has a natural C^*-algebra structure, with respect to pointwise operations.

For every ω in $\mathcal{A}_*$ and X in $\mathcal{A} \otimes \mathcal{B}$, the functional $f_{\omega:X} : \varphi \mapsto \big\langle \omega \otimes \varphi | X \big\rangle_{(\mathcal{A}\otimes\mathcal{B})_*}$ is a bounded linear functional on $\mathcal{B}_*$. Since $(\mathcal{B}_*)^* \cong \mathcal{B}$, it follows that there exists an element $R_\omega(X)$ in $\mathcal{B}$ such that

$$f_{\omega:X}(\varphi) = \big\langle \varphi | R_\omega(X) \big\rangle_{\mathcal{B}_*}.$$

Hence for each ω in $\mathcal{A}_*$ we obtain a mapping

$$\begin{aligned} R_\omega : \mathcal{A} \otimes \mathcal{B} &\to \mathcal{B} \\ X &\mapsto R_\omega(X). \end{aligned}$$

such that

$$\big\langle \omega \otimes \varphi | X \big\rangle_{(\mathcal{A}\otimes\mathcal{B})_*} = \big\langle \varphi | R_\omega(X) \big\rangle_{\mathcal{B}_*}.$$

Hence R_ω is the adjoint of the bounded mapping $(R_\omega)_* : \varphi \mapsto \omega \otimes \varphi$ and hence it follows that R_ω is σ-weakly continuous mapping from $\mathcal{A} \otimes \mathcal{B}$ to $\mathcal{B}$.

For any element $\sum_{i\in F}(A_i \otimes B_i)$ in $\mathcal{A}\otimes\mathcal{B}$ for a finite set F, we must have, $\big\langle \omega\otimes\varphi \big| \sum_{i\in F} \big[A_i \otimes B_i\big] \big\rangle_{(\mathcal{A}\otimes\mathcal{B})_*} = \sum_{i\in F} \langle \omega | A_i \rangle_{\mathcal{A}_*} \langle \varphi | B_i \rangle_{\mathcal{B}_*}$. Hence it follows that,

$$R_\omega\Big[\sum_{i\in F} \big[A_i \otimes B_i\big] \Big] = \sum_{i\in F} \big\langle \omega | A_i \big\rangle_{\mathcal{A}_*} B_i$$

From these relations we get, $R_\omega\big[[1 \otimes A]\big[\sum_{i\in F}[A_i \otimes B_i]\big][1 \otimes B]\big] = R_\omega\big[\sum_{i\in F} A_i \otimes [AB_iB]\big] = \sum_{i\in F} \langle \omega | A_i \rangle_{\mathcal{A}_*} [AB_iB]$. Since $\mathcal{A} \odot \mathcal{B}$ is σ-weakly dense in $\mathcal{A} \otimes \mathcal{B}$, see [45], and hence it follows that

$$R_\omega\Big[\big[1 \otimes A\big] X \big[1 \otimes B\big]\Big] = A R_\omega(X) B.$$

Similarly for each φ in $\mathcal{B}_*$, we get the σ-weakly continuous mapping, $L_\varphi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{A}$ with $L_\varphi\big[[A \otimes 1]X[B \otimes 1]\big] = A L_\varphi(X) B$. We have,

$$\big\langle \varphi | R_\omega(X) \big\rangle_{\mathcal{B}_*} = \big\langle \omega | L_\varphi(X) \big\rangle_{\mathcal{A}_*}.$$

If $R_\omega(X) = 0$ for all ω in $\mathcal{A}_*$, then $\big\langle \omega\otimes\varphi | X \big\rangle = \big\langle \varphi | R_\omega(X) \big\rangle = 0$. Hence $\langle \big[\sum_{i\in F} \omega_i \otimes \varphi_i\big] | X \rangle = 0$, and hence it follows that $X = 0$. Hence the families of maps $\{R_\omega\}_{\omega\in\mathcal{A}_*}$ and $\{L_\varphi\}_{\varphi\in\mathcal{B}_*}$ are total in $\mathcal{A} \otimes \mathcal{B}$.

Lemma 1.3.1. *The maps $A \mapsto AB$, $A \mapsto BA$ and $A \mapsto A^\dagger$ are σ-weakly continuous.*

Proof

Consider the map $\omega_B : A \mapsto \omega(BA)$. This is a bounded linear functional on $\mathcal{A}$ and hence belongs to $\mathcal{A}^*$ and hence is in the predual of $\mathcal{A}^{**}$. The functional $R_{\omega_B} : \mathcal{A}^{**} \otimes \mathcal{A} \to \mathcal{A}$ is a σ-weakly continuous map given by

$$R_{\omega_B}\big[X \otimes A'\big] \mapsto \Big[\varphi \mapsto \big\langle \iota_{\mathcal{A}_*}(\omega) | \iota_{\mathcal{A}}(B) X \big\rangle \big\langle \varphi | A' \big\rangle\Big]_{\varphi\in\mathcal{A}_*}.$$

By Goldstine's theorem, X can be written as a weak* limit of elements of $\mathcal{A}$. Since the embedding map $\iota_{\mathcal{A}} : \mathcal{A} \to \mathcal{A}^{**}$ is weak* dense on bounded sets it follows that the map $A \mapsto BA$ is a composition of σ-weakly continuous maps, with

$$\omega(BA) = \left\langle \varphi | R_{\omega_B}\left[\iota_{\mathcal{A}}(A) \otimes 1\right] \right\rangle_{\mathcal{A}_*}.$$

for some φ with $\langle \varphi | 1 \rangle = 1$, which can be chosen by appropriate scaling. Hence the map $A \mapsto BA$ must be σ-weakly continuous. It follows similarly that the maps $A \mapsto AB$ and $A \mapsto A^\dagger$ are σ-weakly continuous. □

Note that this line of proofs using the bidual to prove topological properties is due to Tomiyama, [44].

1.3.2.2 | The Lattice of Projections

A subspace $\mathcal{M}$ of $\mathcal{A}_*$ is said to be left-invariant with respect to $\mathcal{A}$ if $\mathcal{A}(\mathcal{M}) \subset \mathcal{M}$. Similarly, we say $\mathcal{M}$ is right-invariant if $(\mathcal{M})\mathcal{A} \subset \mathcal{M}$. We now consider subsets which are invariant with respect to the algebraic and topological structures. Note that the †-operation exchanges left, and right actions.

Let $\mathcal{L}(\mathcal{A}_*)$ be the collection of all closed left-invariant subspaces of $\mathcal{A}_*$. Let $\mathcal{M}$ be a closed left-invariant subspace of $\mathcal{A}_*$. Since every element of $\mathcal{A}_*$ is a bounded linear functional on $\mathcal{A}$, we have,

$$\left\langle \varphi | BA \right\rangle_{\mathcal{A}_*} = \left\langle A(\varphi) | B \right\rangle_{\mathcal{A}_*} = \left\langle \varphi' | B \right\rangle_{\mathcal{A}_*}.$$

where φ' also belongs to $\mathcal{M}$, since $\mathcal{A}(\mathcal{M}) \subset \mathcal{M}$. The polar $\mathcal{M}^\circ$ is the largest subset of $\mathcal{A}$ such that

$$\left\langle \omega | M \right\rangle_{\mathcal{A}_*} \leqq 1,$$

for all ω in $\mathcal{M}$ and M in $\mathcal{M}^\circ$. If B belongs to $\mathcal{M}^\circ$, it follows by the above equality that, BA is also in $\mathcal{M}^\circ$. Hence, it follows that $\mathcal{M}^\circ$ is a right-ideal of $\mathcal{A}$. Since polars of sets are already weak*-closed subsets, it follows that, $\mathcal{M}^\circ$ is a σ-weakly closed right-invariant subset of $\mathcal{A}$. Let $\mathcal{J}$ be a σ-weakly closed right-ideal of $\mathcal{A}$. For any φ in $\mathcal{A}_*$ we have,

$$\left\langle \varphi A | J \right\rangle_{\mathcal{A}_*} = \left\langle \varphi | AJ \right\rangle_{\mathcal{A}_*}.$$

When φ belongs to the polar, it follows by the above equality that φA also belongs to the polar, and $\mathcal{J}_\circ$ is a closed left-invariant subset of $\mathcal{A}_*$. By considering the closed subspaces spanned by $\mathcal{J}_\circ$ we obtain a correspondence between objects in $\mathcal{A}_*$ and $\mathcal{A}$ which are algebraically, and topologically invariant under the action of $\mathcal{A}$.

We denote the collection of all σ-weakly closed right-ideals of $\mathcal{A}$ by $\mathcal{R}_\sigma(\mathcal{A})$. With respect to inclusions, $\mathcal{R}_\sigma(\mathcal{A})$ is a lattice. Let the collection of all closed left-invariant subspaces of $\mathcal{A}_*$ be denoted by $\mathcal{L}(\mathcal{A}_*)$. With respect to subspace inclusions, $\mathcal{L}(\mathcal{A}_*)$ is also a lattice.

Theorem 1.3.2. (Takesaki [39])

$$\begin{aligned} \circ : \mathcal{L}(\mathcal{A}_*) &\to \mathcal{R}_\sigma(\mathcal{A}) \\ \mathcal{M} &\mapsto \mathcal{M}^\circ \end{aligned}$$

is an isomorphism of lattices.

Similarly if the lattice of all closed right-invariant subspaces of $\mathcal{A}_*$ is denoted by $\mathcal{R}(\mathcal{A}_*)$, and the lattice of all σ-weakly closed left-ideals of $\mathcal{A}$ by $\mathcal{L}_\sigma(\mathcal{A})$, then the two lattices are isomorphic. We now exploit the Banach space structure of $\mathcal{A}_*$.

The involution $\dagger$ on $\mathcal{A}$ is particularly interesting, for any A and B in $\mathcal{A}$, we have $A^\dagger B = (B^\dagger A)^\dagger$ and hence it reverses the product structure on $\mathcal{A}$. This relates the left and right multiplication operations on $\mathcal{A}$. If $\mathcal{L}$ is a right-ideal, then

$$A\mathcal{L}^\dagger = (\mathcal{L}A^\dagger)^\dagger = \mathcal{L}^\dagger$$

Hence $\mathcal{L}^\dagger$ is a left-ideal of $\mathcal{A}$. The intersection,

$$\mathcal{A}_\mathcal{L} \equiv \mathcal{L} \cap \mathcal{L}^\dagger$$

is the largest C^*-subalgebra containing the ideal $\mathcal{L}$, and is a two-sided ideal of $\mathcal{A}$. Since the intersection of closed sets is closed, it follows that if $\mathcal{L}$ is σ-weakly closed, then so is $\mathcal{A}_\mathcal{L}$. In such a case, let $\mathcal{A}_\mathcal{L}^\perp$ be the largest set of normal functionals such that $\langle \mathcal{A}_\mathcal{L}^\perp | \mathcal{A}_\mathcal{L} \rangle_{\mathcal{A}_*} = 0$. By the bilinearity of the pairing $\langle \cdot | \cdot \rangle_{\mathcal{A}_*}$, it follows that this collection is a subspace of $\mathcal{A}_*$, and by the continuity of the pairing $\mathcal{A}_\mathcal{L}^\perp$ is also a closed subspace. Consider the quotient space,

$$(\mathcal{A}_\mathcal{L})_* \equiv \mathcal{A}_* / \mathcal{A}_\mathcal{L}^\perp$$

which consists of equivalence classes $[\varphi] = \varphi + \mathcal{A}_\mathcal{L}^\perp$. We can equip it with the norm,

$$\big\|[\varphi]\big\|_{(\mathcal{A}_\mathcal{L})_*} = \inf_{\psi \in [\varphi]} \Big[\|\psi\| \Big],$$

$(\mathcal{A}_\mathcal{L})_*$ is a Banach space. $\langle [\varphi] | A \rangle_{\mathcal{A}_\mathcal{L}} \equiv \langle \varphi | A \rangle_{\mathcal{A}_*}$ defines a well-defined pairing of $(\mathcal{A}_\mathcal{L})_*$ and the C^*-subalgebra $\mathcal{A}_\mathcal{L}$. Since $\mathcal{A}_*$ separates $\mathcal{A}$, it follows that $(\mathcal{A}_\mathcal{L})_*$ is a predual of $\mathcal{A}_\mathcal{L}$ and hence $\mathcal{A}_\mathcal{L}$ is a W^*-algebra.

Let $E_\mathcal{L}$ be the identity element of the W^*-subalgebra $\mathcal{A}_\mathcal{L}$. Hence we must have $E_\mathcal{L}^2 = E_\mathcal{L} = E_\mathcal{L}^\dagger$. Since $\mathcal{A}_\mathcal{L}$ is a right-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 $\mathcal{A}_\mathcal{L}$. Since $E_\mathcal{L}$ is the identity element of $\mathcal{A}_\mathcal{L}$ 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 the identity element, $E_\mathcal{L}$ is uniquely determined. Hence every σ-weakly 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 σ-weakly 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 σ-weakly 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 intersections 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 structure from the σ-weakly 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 1.3.3. *If $\mathcal{A}$ is a W^*-algebra then $\mathcal{P}(\mathcal{A})$ is a complete lattice.*

A family of projections $\{E_\alpha\}_{\alpha\in\mathcal{I}}$ is said to be orthogonal if $E_\alpha \wedge E_\beta$ is zero whenever α is not equal to β. In such a case, we write

$$\vee_{\alpha\in\mathcal{I}} E_\alpha = \sum_{\alpha\in\mathcal{I}} E_\alpha.$$

The effects in quantum theory correspond to *yes-no* questions, and mathematically correspond to projection elements of the W^*-algebra modeling the system. For physical systems, we expect every question about the system can be decomposed into *yes-no* questions. Hence, we must also expect the W^*-algebra to have sufficiently many projection elements to describe all possible questions about the system.

1.3.2.3 | Structure of the Dual Space

The enveloping W^*-algebra $\mathcal{A}^{**}$ of the W^*-algebra $\mathcal{A}$ is itself a W^*-algebra with the dual space $\mathcal{A}^*$ as its predual. Hence we can study the conjugacy relationship between closed invariant subspaces of $\mathcal{A}^*$ and σ-weakly closed ideals of $\mathcal{A}^{**}$. By the linearity of the action, the action of any element of $\mathcal{A}^{**}$ can be studied by any scaling of it, and rescaling back.

By the Krein-Smulian theorem it is sufficient to understand how the elements of the closed unit ball $B_{\mathcal{A}^{**}}$ of $\mathcal{A}^{**}$ act on $\mathcal{A}^*$. By Goldstine's theorem, $\iota_{\mathcal{A}}(B_{\mathcal{A}})$ is weak*-dense in $B_{\mathcal{A}^{**}}$, hence by continuity of the pairing, that is, by the Krein-Smulian theorem, the action of elements in $B_{\mathcal{A}^{**}}$ is determined by how elements in $B_{\mathcal{A}}$ act. Hence it follows that $\mathcal{A}_*$ is invariant under the action of $\mathcal{A}^{**}$. Since the predual $\mathcal{A}_*$ of $\mathcal{A}$ is itself a Banach subspace of $\mathcal{A}^*$, it is a closed two-sided invariant subspace of $\mathcal{A}^*$. Hence there must exist a central projection $E_{\mathcal{A}_*}$ of $\mathcal{A}^{**}$ such that

$$\left(\iota_{\mathcal{A}_*}(\mathcal{A}_*)\right)^\circ \equiv \mathcal{A}^{**}(1 - E_{\mathcal{A}_*}).$$

Hence we can write

$$\mathcal{A}_* \equiv (1 - E_{\mathcal{A}_*})\mathcal{A}^*.$$

We denote $E_{\mathcal{A}_*}\mathcal{A}^*$ by $\mathcal{A}_*^\perp$. Note that $E_{\mathcal{A}_*}$ is unique, by the universal property of $\mathcal{A}^{**}$, that any *-homomorphism of a C^* algebra to a W^*-algebra must factor through $\mathcal{A}^{**}$, and the uniqueness of the identity element.

Theorem 1.3.4. (Takesaki [39]) *If $\mathcal{M}$ is an invariant closed subspace of $\mathcal{A}^*$, then for any φ in $\mathcal{M}$ there exist a unique φ_σ in $(1 - E_{\mathcal{A}_*})\mathcal{A}^*$ and $\varphi_\sigma^\perp$ in $E_{\mathcal{A}_*}(\mathcal{A}^*)$ such that*

$$\|\varphi\|_{\mathcal{A}^*} = \|\varphi_\sigma\|_{\mathcal{A}^*} + \|\varphi_\sigma^\perp\|_{\mathcal{A}^*}.$$

More compactly, we write

$$\mathcal{M} \equiv \big[\mathcal{M} \cap A_*\big] \oplus_{L^1} \big[\mathcal{M} \cap \mathcal{A}_*^\perp\big].$$

Proof

For any φ in $\mathcal{M}$ define the linear functionals, $\varphi_\sigma^\perp := E_{\mathcal{A}_*}\varphi$, and $\varphi_\sigma := (1 - E_{\mathcal{A}_*})\varphi$. Clearly, the functional $\varphi_\sigma^\perp$ belongs to $\mathcal{A}_*^\perp$, and similarly φ_σ belongs to $\mathcal{A}_*$. This decomposition is unique because $E_{\mathcal{A}_*}$ is uniquely determined, since $\mathcal{A}^{**}$ is the universal enveloping W^*-algebra.

We have,

$$\begin{aligned}\|\varphi\|_{\mathcal{A}^*} &= \sup_{\|X\|\leqq 1}\Big[\big|\langle\varphi \mid X\rangle\big|\Big] = \sup_{\|X\|\leqq 1}\Big[\big|\langle\varphi_\sigma + \varphi_\sigma^\perp|X\rangle\big|\Big] \\ &= \sup_{\|X\|\leqq 1}\Big[\big|\langle\varphi_\sigma|X\rangle\big|\Big] + \sup_{\|X\|\leqq 1}\Big[\big|\langle\varphi_\sigma^\perp|X\rangle\big|\Big] = \|\varphi_\sigma\|_{\mathcal{A}^*} + \|\varphi_\sigma^\perp\|_{\mathcal{A}^*}.\end{aligned}$$

For the third equality we used the fact that the identity element of $\mathcal{A}^{**}$ is given by $1_{\mathcal{A}^{**}} = E_{\mathcal{A}_*} + (1 - E_{\mathcal{A}^*})$. This proves the theorem. □

The linear functionals belonging to $\mathcal{A}_*^\perp$ are called singular linear functionals on $\mathcal{A}$. Any linear functional can be decomposed into singular, and normal functionals since $\mathcal{A}^*$ is itself an invariant subspace.

1.3.2.4 | The Predual: Sakai's Theorem

Using the above decomposition theorem we now obtain a characterisation of normal functionals. Since the norm on $\mathcal{A}$ coincides with the operator norm coming from the norm on its predual $\mathcal{A}_*$. For all A in $\mathcal{A}$ and ω in $\mathcal{A}_*$,

$$\Big\langle \omega \,\Big|\, \mathcal{A}\big[\lim_{\mathcal{J}} E_{\mathcal{J}}\big]\Big\rangle = \lim_{\mathcal{J}}\langle\omega|AE_{\mathcal{J}}\rangle = \lim_{\mathcal{J}}\langle E_{\mathcal{J}}\omega|A\rangle = \Big\langle \big[\lim_{\mathcal{J}} E_{\mathcal{J}}\big]\omega \,\Big|\, \mathcal{A}\Big\rangle.$$

where the limit is taken in the σ-weak topology.

Using this we obtain an alternative characterisation of normality in terms of properties of the lattice of projections of the W^*-algebra. A continuous linear functional ω is said to be completely additive if

$$\omega\Big[\sum_{\alpha\in\mathcal{I}} E_\alpha\Big] = \sum_{\alpha\in\mathcal{I}}\omega(E_\alpha)$$

for any family $\{E_\alpha\}_{\alpha\in\mathcal{I}}$ of orthogonal projections in $\mathcal{P}(\mathcal{A})$.

If ω is a normal functional, for any finite subset $\mathcal{J}$ of $\mathcal{I}$ let $E_{\mathcal{J}} \equiv \sum_{\alpha\in\mathcal{J}} E_\alpha$. Since $\mathcal{P}(\mathcal{A})$ is a complete lattice, the net $\{E_{\mathcal{J}}\}$ converges to $\sum_{\alpha\in\mathcal{I}} E_\alpha$, σ-weakly. Hence we have,

$$\begin{aligned}\omega\Big[\sum_{\alpha\in\mathcal{I}} E_\alpha\Big] &= \omega\Big[\lim_{\mathcal{J}}\sum_{\alpha\in\mathcal{J}} E_\alpha\Big] \\ &= \lim_{\mathcal{J}}\omega\Big[\sum_{\alpha\in\mathcal{J}} E_\alpha\Big] = \sum_{\alpha\in\mathcal{I}}\omega(E_\alpha).\end{aligned}$$

Hence every normal functional is completely additive.

Suppose ω is a completely additive functional, then by Takesaki's theorem above we can decompose it into its normal and singular part $\omega = \omega_\sigma + \omega_\sigma^\perp$. In particular $\omega - \omega_\sigma^\perp$ must also be completely positive. Note that $\omega - \omega_\sigma$ is a singular functional.

Let $\mathrm{supp}(\omega - \omega_\sigma)$ be the smallest closed invariant subspace of $\mathcal{A}^*$ called the support of $\omega - \omega_\sigma$ and let E be the corresponding projection. E cannot be in $\mathcal{A}$, because otherwise $E(\omega - \omega_\sigma)$ would be a normal functional in $\mathcal{A}^*$, contradicting the uniqueness of decomposition. Hence if ω is completely additive then ω must also be normal. Hence we have proved the following characterisation of normality;

Theorem 1.3.5. *ω is completely additive if and only if it is normal.*

The lattice of projections of a W^*-algebra is closely related to closed subspaces of its predual. Hence the notion of normality of functionals is closely related to the projection lattice structure of a W^*-algebra.

Let $\mathcal{B}$ be a Banach space such that $\mathcal{B}^* \equiv \mathcal{A}$. The elements of $\mathcal{B}$ correspond to a functionals on $\mathcal{A}$, and since $\mathcal{B}$ is an closed invariant subspace of $\mathcal{A}^*$, each κ, can be decomposed uniquely as $\kappa = \kappa_\sigma + \kappa_\sigma^\perp$ with

$$\|\kappa\|_{\mathcal{B}} \equiv \|\kappa_\sigma\|_{\mathcal{A}^*} + \|\kappa_\sigma^\perp\|.$$

Since $\mathcal{B}^* \equiv \mathcal{A}$, we also have $\mathcal{B}^\circ \equiv \mathcal{A}$. Hence $(\mathcal{B}^\circ)_\circ = \mathcal{A}_*$. Hence κ in $\mathcal{B}$ must correspond to a normal functional, $\kappa_\sigma^\perp$ must be identically zero. Hence κ_σ corresponds to an element of $\mathcal{A}_*$ with

$$\|\kappa\|_{\mathcal{B}} = \|\kappa_\sigma\|_{\mathcal{A}_*}$$

Heuristically, we *need* all of $\mathcal{A}$ to separate closed convex subsets of $\mathcal{B}$.

Similarly, we can think of each element of $\mathcal{A}_*$ as an element of $\mathcal{B}$. Hence the mapping, $\kappa \mapsto \kappa_\sigma$ is an isometric isomorphism of Banach spaces $\mathcal{B}$ and $\mathcal{A}_*$, and hence the predual is unique. This establishes the following uniqueness theorem of the predual for W^*-algebras due to Sakai;

Theorem 1.3.6. (Sakai) *Upto isometric isomorphism, the predual of a W^*-algebra is unique.*

Universal properties usually endow the object with some uniqueness properties. In our case, the pairings also relate the objects to their pairing spaces. We obtained the uniqueness of the predual via pairings, and a universal property of its *third* dual space, mainly because the norm on the third dual space was already fixed by the predual.

1.4 | Spectral Theory of W^*-Algebras

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

1.4.1 | What are Observables?

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

1.4.1.1 | Complete Boolean Algebras

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Every Boolean algebra can also be thought of as a lattice and vice-versa with the order give by,

$$R_i \leqq R_j \Leftrightarrow R_i \vee R_j = R_j$$

The notion of completeness can be described in terms of the lattice order and in terms of existence of least upper bounds, and this lattice theoretic notion of completeness is closer to the completeness of $\mathbb{R}$ in terms of order coming from the notion of positivity on real numbers. The space $\mathcal{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$ is a σ-complete Boolean algebra. If we assume every measuring instrument has its values labelled by real numbers, $\mathcal{B}(\mathbb{R})$ acts as an appropriate Boolean algebra for modeling observables.

The Notion of Observables

The notion of observable is an idealisation, and describes collections of events measured by the most accurate measuring instrument, which is capable of measuring all the smallest outcomes to the highest accuracy. From the discussion above the space of effects of a system may be modeled using a W^*-algebra $\mathcal{A}$. Since observables correspond to collections of simultaneously measurable outcomes the notion of an observable must be closely related to complete Boolean algebras. Hence the collection of effects corresponding to such measuring instrument corresponding to Boolean subalgebras of the space of effects.

We now make this precise by introducing the notion of Boolean maps. Boolean maps are those which preserve Boolean algebra structures. Such maps allow us to identify Boolean subalgebras of the space of effects via maps into the space of effects. If Σ and Σ' are Boolean algebras, a mapping

$$E : \Sigma \to \Sigma'$$

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

Since observables correspond to a complete Boolean algebra, we define an observable in algebraic quantum theory as a pair (Σ, E_Σ), where Σ is a complete Boolean algebra and

$$E_\Sigma : \Sigma \to \mathcal{A}.$$

is a Boolean map to its image $E_\Sigma(\Sigma) \subseteq \mathcal{A}$. In particular, $E_\Sigma(\Sigma)$ will have to be a commutative subalgebra of $\mathcal{A}$. Hence, the commutative subalgebras of the algebra of observables $\mathcal{A}$ will be of special interest to us.

Remark: Generalisation of Observables?

Döring and Isham, in [12], revisit the definition of observables, by questioning the Boolean algebraic nature of simultaneously measurable outcomes. They model observables at an even higher level of generality by replacing the Boolean algebra structure on Σ_R with the so called Heyting algebras. If Boolean algebra operations corresponds to the classical set theoretic logical operations $\vee$, $\wedge$ and $\neg$, the Heyting algebra operations correspond to the intuitionistic logical operations, which does not assume the law of excluded middles, or the idempotency of the $\neg$-operation. The Heyting algebras can be modeled categorically in terms of toposes, similar to how Boolean algebras can be modelled by sets. A topos, as a category is equivalent to the category of sheaves on a space, and hence each object of a topos corresponds to some

local property of a space. This model takes into account all categories where the notion of *locality* makes sense, this allows us to talk about *where* the values of the observable lies from a very abstract and general point of view. This approach to the formulation of quantum theory, starting with remodeling the structure expected of simultaneously measurable outcomes is known as topos quantum theory.

Topos quantum theory however continues with the assumption that the algebra of observables corresponds to a W^*-algebra. Note that in our case, we obtained the W^*-algebraic structure based on operational assumptions about instruments. However we *secretly* used the idempotency of $\neg$ via the uniqueness of topological complementation, which was essential for the operation $\dagger$ to be an involution. In the absense of $\neg$, the existence of the involution $\dagger$ may have to be assumed. Which would then send us back to the drawing board, and defeating all the progress we have made so far. Hence we will not discuss topos quantum theory here and refer an interested reader to [12]-[13].

1.4.1.2 | Stone Duality

Boolean algebra structures are very abundant in mathematics. Given any set K, the powerset $\mathcal{P}(K)$ of the set K, equipped with union, intersection and complementation as the operations is a Boolean algebra. This Boolean algebra can also be thought of as a Boolean algebra of the topology on K equipped with the discrete topology, where every element of the Boolean algebra corresponds to a simultaneously open and closed set of K. Note that this simultaneous openness and closedness of subsets is essential to make sense of Boolean algebraic structures topologically since we expect the complement of the set to also correspond to an element of the Boolean algebra. The topologies of interest to us are not usually discrete topology. The topology on a set carries relevant information about additional structures on the set. Given a topological space K, the simultaneously open and closed subsets are called clopen sets. Equipped with union, intersection and complementation as the operations, the collection of clopen subsets Σ_K of K is a Boolean algebra.

Stone duality says that the converse is also true, that is, any complete Boolean algebra Σ can be thought of as the Boolean algebra of clopen sets of a topological space, called the Stone space K_Σ, such that the minimal elements of Σ correspond to points of K_Σ.

The Space of Maximal Ideals of Σ

The idea of Stone duality is to use the elements of the Boolean algebra as regions. The order of elements of the Boolean algebra provides a natural relation between the regions of the space, and using these regions as basis we obtain a topology on the Stone space. We start by describing minimal elements algebraically in terms of the Boolean algebraic operations.

Suppose we have a collection of outcomes which cannot further be divided as a combination of more accurate outcomes, then any further refining the collection into more accurate outcomes should correspond to the same collection. This notion of refining allows for an algebraic description of the notion of collection of outcomes of highest accuracy. One can refine a given collection $\mathcal{I}$ of simultaneously measurable outcomes by wedging the collection with all the other simultaneously measurable outcomes. If the collection is already maximally refined, then we must have,

$$\mathcal{I} \wedge \Sigma \equiv \mathcal{I}.$$

Furthermore, if we also require simultaneous measurability, that is, it must be closed under $\vee$, we can describe the idealisation of observables algebraically in terms of the notion of ideals of an algebra. The maximal ideals can only leave out a single minimal outcome. For the case

of complete Boolean algebras, the maximal ideals can be used to describe the collection of outcome using a set. The set K_Σ of all maximal ideals is called the Stone space of Σ.

Outcomes can be isolated by considering maps into the trivial Boolean algebra describing a single *yes-no* outcome. The Boolean algebra corresponding to a single *yes-no* outcome is $\{0,1\}$ where 0 corresponds to a *no* and 1 corresponds to *yes*, Since such maps are expected to contain the data about the outcome, it must respect the Boolean algebra structure. Hence we consider Boolean maps,

$$f : \Sigma \to \mathbf{2} = \{0,1\}.$$

Let $\mathrm{Hom}(\Sigma, \mathbf{2})$ be the collection of all such Boolean maps to the trivial Boolean algebra $\{0,1\}$. $\mathrm{Hom}(\Sigma, \mathbf{2})$ is itself a Boolean algebra with pointwise operations. Each element R of the Boolean algebra can be studied using the sets,

$$[R]_\Sigma = \{f \mid f(R) = 1\},$$

where each $[R]_\Sigma$ is a subset of $\mathrm{Hom}(\Sigma, \mathbf{2})$. If f is a Boolean map we must have $f(R \vee S) = f(R) \vee f(S)$ and $f(R \wedge S) = f(R) \wedge f(S)$, it follows that $[R \wedge S]_\Sigma = [R]_\Sigma \cap [S]_\Sigma$ and $[R \vee S]_\Sigma = [R]_\Sigma \cup [S]_\Sigma$. This gives us a Boolean map

$$\Gamma_\Sigma : R \mapsto [R]_\Sigma$$

Γ_Σ is called the Stone map. The set $[R]_\Sigma$ consists of all functions which have the value 1 for R, and must also be 1 for all outcomes S greater than R. $[R]_\Sigma$ is the collection of all events *greater* than R, and hence we can think about the outcome R in terms of the set $[R]_\Sigma$.

Since the Boolean algebra $\mathbf{2} = \{0,1\}$ does not contain any proper ideal, the kernels of the Boolean maps from Σ to $\mathbf{2}$ must be maximal ideals of Σ. Hence we have, $\mathrm{Hom}(\Sigma, \mathbf{2}) \subseteq K_\Sigma$. On the other hand, for every R in Σ there exists some maximal ideal $\mathcal{I}_{\neg R}$, such that $\neg R$ belongs to $\mathcal{I}_{\neg R}$.[5] Using this maximal ideal we can construct a Boolean map,

$$f_R : \Sigma \to \mathbf{2}$$

which sends each S in Σ to 1 if S belongs to $\mathcal{I}_{\neg R}$ and to 0 otherwise. Hence for every R in Σ, there exists a Boolean map $f_R : \Sigma \to \mathbf{2}$, such that $f_R(R) = 1$. Hence as a set we must have,

$$\mathrm{Hom}(\Sigma, \mathbf{2}) \equiv K_\Sigma.$$

Since $\mathbf{2}$ with the discrete topology is a compact Hausdorff space, the product topological space $\mathbf{2}^\Sigma$ is also a compact Hausdorff space by Tychonoff's theorem.

Since our goal is to think of every element of Σ as an element of the topology on Σ, we construct the topology on K_Σ to be the weakest topology which turns $[R]_\Sigma$ into clopen sets.

[5] A proper-ideal of Σ is a set $\mathcal{I} \subsetneq \Sigma$ containing the bottom element of Σ, and closed under the operation $\vee$, such that $\mathcal{I} \wedge \Sigma = \mathcal{I}$. By considering the set $R \wedge \Sigma$, we observe that every non-unit element R of Σ belongs to a proper-ideal. A maximal ideal is a proper-ideal not contained in any other proper-ideal. Let K_Σ be the collection of all maximal ideals in Σ. By Zorn's lemma, every proper-ideal belongs to a maximal ideal. For every $R \in \Sigma$, a maximal ideal cannot contain both R and $\neg R$, because that would mean the ideal also contains the identity of the Boolean algebra, which contradicts that it is a proper-ideal. Each maximal ideal describes a smallest outcome, and it corresponds to the outcome it does *not* contain.

Conversely, if a set $\mathcal{I}$ contains only one of R or $\neg R$ for every R in Σ, it must be a maximal ideal, because if it was a subset of a proper-ideal $\mathcal{J}$ then by considering $S \in \mathcal{J} \backslash \mathcal{I}$, then $S \wedge \Sigma \subsetneq \mathcal{J}$ is a proper-ideal and must contain both S and $\neg S$ since $\mathcal{I}$ is already assumed to contain all S for which this is not the case. But this implies that $S \wedge \Sigma$ is a proper ideal, since $S \vee \neg S = 1 \in S \wedge \Sigma$. Hence every maximal ideals is characterised by the fact that for every element R in Σ only of R or $\neg R$ must belong to it.

In this case, for every R in Σ, let $\chi(R)$ be the evaluation map on $\mathbf{2}^\Sigma$. This corresponds to the projection map

$$\begin{aligned}\chi_R : \mathbf{2}^\Sigma &\to \mathbf{2} \\ f &\mapsto f(R).\end{aligned}$$

If we require this map to be continuous then by definition of product topology the sets $\chi_R^{-1}(1)$ and $\chi_R^{-1}(0) \equiv \mathbf{2}^\Sigma \backslash \chi_R^{-1}(1)$ are open in $\mathbf{2}^\Sigma$. Hence $\mathbf{2}^\Sigma$ is totally disconnected, that is, it is a Hausdorff space for which every open set is the union of clopen sets, or equivalently, the topology is generated by clopen sets. The collection of all clopen sets of K is denoted by $\mathrm{Cl}(K)$.

The maps $f \mapsto f(\neg R)$ and $f \mapsto \neg f(R)$ are both continuous functions. Hence the set of all f such that $f(\neg R) = \neg f(R)$ is a closed set in $\mathbf{2}^\Sigma$, Similarly, for any collection $\{R_\alpha\}$, we can consider the maps $f \mapsto f(\vee_\alpha R_\alpha)$ and $f \mapsto \vee_\alpha f(R_\alpha)$, and similarly for $\wedge$.

The intersection of these closed sets must be a closed set, and the intersection precisely corresponds to the Stone space $K_\Sigma \equiv \mathrm{Hom}(\Sigma, \mathbf{2})$.[6] Hence, as a closed subset of a compact Hausdorff totally disconnected space, the Stone space K_Σ is also a compact Hausdorff and totally disconnected. Hence we have proved Stone duality, also known as the Stone representation theorem;

Theorem 1.4.1. (Stone Duality)

$$\begin{aligned}\Gamma_\Sigma : \Sigma &\rightarrowtail \mathit{Cl}(K_\Sigma) \\ R &\mapsto [R]_\Sigma\end{aligned}$$

is a Boolean embedding where

$$K_\Sigma \equiv \mathrm{Hom}(\Sigma, \mathbf{2})$$

is a compact Hausdorff and totally disconnected consisting of maximal ideals of Σ. K_Σ *is called the Stone space of* Σ.

For every complete Boolean algebra Σ the Stone space K_Σ is a compact Hausdorff and is totally disconnected. Conversely, for every compact Hausdorff space K which is totally disconnected the collection of clopen sets Σ_K is a complete Boolean algebra with clopen sets corresponding to points of K as minimal elements of Σ_K. Hence it is convenient to isolate such topological spaces.

Lattice Structure on $\mathcal{C}_\mathbb{R}(K)$

An extremely disconnected space is a compact Hausdorff space such that every open set is a union of clopen sets. We describe a convenient alternative characterisation of extremely disconnected spaces, which will be useful later in the abstract characterisation of observables in quantum theory. Function spaces inherit structures from the target space. If the functions are also expected to respect the structures on the initial space, then the function space relates and carries the data about such structures.

We now use the Stone space of $\mathbb{R}$ as the target space to characterise extremely disconnected spaces in terms of an appropriate space of functions. Let K be an extremely disconnected space, and let $\mathcal{C}_\mathbb{R}(K)$ be the space of real valued continuous functions on K. The continuity

[6] Here we used the fact that the set $\{x \mid f(x) = g(x)\}$ for continuous functions f and g is always closed.

of functions relate the topologies on the space K and $\mathbb{R}$. We can use the lattice structure on $\mathbb{R}$ to introduce a lattice structure on $\mathcal{C}_{\mathbb{R}}(K)$.

For any two functions f_α and f_β, we say, $f_\beta \leq f_\alpha$ if and only if $f_\beta(x) \leqq f_\alpha(x)$ for every x in K. Together with this order, $\mathcal{C}_{\mathbb{R}}(K)$ is a lattice. An indexed collection $\{f_\alpha\}_{\alpha\in\mathcal{I}}$ of functions in $\mathcal{C}_{\mathbb{R}}(K)$ is said to be uniformly bounded, if there exists k in $\mathbb{R}$ such that

$$|f_\alpha(x)| \leqq k, \quad \forall\alpha, \forall x \in K.$$

Let $\mathcal{I}$ be a directed set. We say a uniformly bounded collection $\{f_\alpha\}_{\alpha\in\mathcal{I}}$ in $\mathcal{C}_{\mathbb{R}}(K)$ is a directed collection if

$$\alpha \leq \beta \Rightarrow f_\beta \leq f_\alpha.$$

We say the lattice $\mathcal{C}_{\mathbb{R}}(K)$ is conditionally complete if every uniformly bounded directed collection $\{f_\alpha\}$ has a least upper bound in $\mathcal{C}_{\mathbb{R}}(K)$, that is, there exists a function f, denoted by

$$f = \sup_\alpha\{f_\alpha\}$$

such that for any g with $f_\alpha \leq g$, we must also have $f \leq g$. The topology on $\mathbb{R}$ is closely related to the order on $\mathbb{R}$, and hence we can expect the completion properties of Boolean algebra corresponding to K to be closely related to the lattice structure on $\mathcal{C}_{\mathbb{R}}(K)$.

For any f in $\mathcal{C}_{\mathbb{R}}(K)$, by the compactness of the space K we may assume that the range of f is contained in a compact interval K_f. For any n, we decompose the interval K_f into n equal length subintervals, and we enumerate them by K_f^i where $i \leqq n$. Let χ_i be the characteristic functions for $f^{-1}(K_f^i)$. Let $f_i = f(x_i)$ for some $x \in f^{-1}(K_f^i)$. The function f in $\mathcal{C}_{\mathbb{R}}(K)$ can now be approximated as,

$$f(x) \sim \textstyle\sum_i f_i\chi_i(x).$$

By dividing the interval K_f into smaller and smaller pieces, that is, by increasing n, we obtain closer and closer approximation of the function f as finite combinations of simpler characteristic functions on K. Hence we have,

$$\textstyle\sum_i f_i\chi_i(x) \to f(x)$$

uniformly. Whether we can make sense of such limits from within the space of continuous functions depends on whether the limits of characteristic functions remain continuous. Hence, the completeness of Boolean algebra associated with K is closely related the continuity of limits of characteristic functions on K.

From the discussion above, it is clear that a compact Hausdorff space K is an extremely disconnected space whenever each characteristic function of K is continuous since in such a case the characteristic function of an open set coincides with the characteristic function of its closure. The criterion of continuity of characteristic functions allows us to characterise of an extremely disconnected space K in terms of the lattice properties of $\mathcal{C}_{\mathbb{R}}(K)$.

Theorem 1.4.2. *A compact Hausdorff space K is stonean whenever $\mathcal{C}_{\mathbb{R}}(K)$ is a conditionally complete lattice.*

Proof

From the discussion above it is sufficient to prove that every characteristic function on open sets of K is continuous, which would also mean that it is the characteristic function for the

closure of the set. We now combine the lattice properties of $\mathcal{C}_{\mathbb{R}}(K)$ with the topological properties of K. As a compact Hausdorff space K is a normal space. Hence we can use the Urysohn's lemma for normal spaces, which says that, the space of continuous functions on K separates the closed subsets of K. That is, for every pair of disjoint closed subsets there exists a continuous function with value 1 on one closed subset, and 0 on the other.

By considering the directed collection of all open sets N_α containing the closure of N, we obtain a collection of continuous functions $\{f_\alpha\}$ with $f_\alpha(x) \in [0,1]$ such that $f_\alpha|_{N_\alpha^c} \equiv 0$, and $f_\alpha|_N \equiv 1$. Hence characteristic functions χ_U can be approximated by continuous functions $\{f_\alpha\}$, such that

$$\chi_N(x) = \sup_\alpha f_\alpha(x),$$

for all x in K. By assumption, any uniformly bounded collection $\{f_\alpha\} \subset \mathcal{C}_{\mathbb{R}}(K)$ possesses a supremum, that is, there exists a function f in $\mathcal{C}_{\mathbb{R}}(K)$ such that

$$f = \sup_\alpha \{f_\alpha\}.$$

The function f as defined above majorises the characteristic function on N, and by continuity of f it also majorizes the characteristic function on the closure $\overline{N}$. We now prove that f is the same as χ_N, and hence χ_N must be a continuous function. Suppose f is not the characteristic function χ_N, then there must exist some y outside of N such that $f(y)$ is non-zero.

Since every singleton set of a Hausdorff space is closed, by Urysohn's lemma there exists continuous function g on K vanishing at y and $g(x) = 1$ for all x in N. Hence we have $g(y) < f(y)$, and hence we have $g \leq f$. This is a contradiction since f was assumed to be the supremum of the collection $\{f_\alpha\}$. □

The above approximation always makes sense in the context of measurable functions since every characteristic function of a measurable set is already measurable. The convergence in the space of measurable functions is weaker in the sense that meaures on the space K do not *see* the peculiarities of the topological data of K.

1.4.2 | The Gelfand Spectrum & Stone Space

From the discussion above, commutative subalgebras of the algebra of observables are of special interest to us. Complete boolean algebras are closely related to topological spaces, as is made precise by the Stone duality. Stone duality allows us to study a complete Boolean algebra as a Boolean algebra of a suitable collection of subsets of a topological space, called the Stone space. In our context the Boolean algebraic structures are inherited from the algebraic properties of the space of effects, which is a W^*-algebras. The space of effects already comes equipped with a notion of convergence coming from its state space, and hence we must expect the topological data of the Stone space for Boolean algebras corresponding to simultaneously measurable outcomes to be closely related to its built in topology. Stone spaces of interest to us in the context of quantum theory are closely related to the operator algebraic notion of the Gelfand spectrum. The notion of spectrum provides us with analytic tools for the construction and study of Stone spaces. We now discuss the spectral theory of commutative C^*-algebras with identity. The need for a complex structure in the formulation of quantum theory came from the need to make sense of the minimal effects, and hence we may expect to use the tools for complex analysis to be useful for the analysis of spectrum in the study of maximal ideals of commutative C^*-algebras.

The involution in C^*-algebras possesses properties similar to complex conjugation. While the set of complex numbers only has a unique non-invertible element, C^*-algebras need not

have this uniqueness. Such elements would belong to some ideals, and hence we can describe maximal ideals in terms of non-invertibility. The collection of such non-invertible elements acts as the starting point.

Let $\mathcal{A}$ be a commutative C^*-algebra with identity. Let the group of invertible elements of $\mathcal{A}$ be denoted by $\mathcal{G}(\mathcal{A})$, which is an open set in $\mathcal{A}$.[7] 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 $\mathcal{A}$. 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 $\mathcal{A}$. For every element A in $\mathcal{A}$ define a map R_A from $R(A)$ to $\mathcal{A}$ by

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

The map R_A is called the resolvent of A, and it introduces complex analytic tools. $R(A)$ is called the resolvent of A. Since $\mathcal{G}(\mathcal{A})$ is open, it follows that $R(A)$ is open. The spectrum of A in $\mathcal{A}$ is defined to be the closed set

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

By rearranging we can check that for all $\lambda, \mu \in R(A)$, λ not equal to μ,

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

Hence R_A defines an analytic function with $R_A'(\lambda) = R_A(\lambda)^2$ from the open subset $R(A)$ 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 $\mathbb{C}$ with $\|R_A(\lambda)\| = |\lambda|^{-1} \|(\lambda^{-1} A - 1)^{-1}\|$. Hence it follows that $\|R_A(\lambda)\| \to 0$ as λ tends to infinity.

For any bounded linear functional ω by the continuity and linearity we have

$$\begin{aligned}(\omega \circ R_A)'(\lambda) &= \lim_{\lambda \to \mu} \left[\tfrac{\omega(R_A(\lambda)) - \omega(R_A(\mu))}{\lambda - \mu} \right] \\ &= \omega \Big[\lim_{\lambda \to \mu} \tfrac{(R_A(\lambda) - R_A(\mu))}{(\lambda - \mu)} \Big] = \omega\big(R_A(\lambda)^2\big).\end{aligned}$$

The composition $\omega \circ R_A$ 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 $\omega \circ R_A \equiv 0$ which is absurd since ω was arbitrarily chosen. Hence $\sigma(A)$ must be non-empty for every $A \in \mathcal{A}$, and we must have $|\lambda| > \|A\|$ whenever λ belongs to $\sigma(A)$.

Suppose every non-zero element of $\mathcal{A}$ is invertible, and $\mathcal{A}$ 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 1.4.3. (Gelfand-Mazur) *If $\mathcal{A}$ is a commutative C^*-algebra with unit and $\mathcal{G}(\mathcal{A}) = \mathcal{A} \backslash \{0\}$ then $\mathcal{A} \cong \mathbb{C}$.*

Although the notion of spectrum may be defined in general algebra, the import of complex analytic tools is possible due to the topology on the Banach algebra. The spectrum depends

[7] If $A \in B_1(I)$, that is, $\|I - A\| < 1$ then we can make sense of the sum $B = \sum_{\mathbb{N}} (A - I)^i$ since $\sum_{\mathbb{N}} \|A - I\|^i$ 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 the product with the neighborhood we obtain an open neighborhood of A of invertible elements $AB_\epsilon(I) \subset \mathcal{G}(\mathcal{A})$. Which proves that $\mathcal{G}(\mathcal{A})$ is open.

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^*-algebra, $(A-\lambda)^{-1}$ belongs to the C^*-algebra generated by the elements $A, A^\dagger$ and 1. Hence the notion of spectrum in the C^*-algebra context is intrinsic to the elements and does not depend on the ambient C^*-algebra.

The map $p(z) \mapsto p(A)$ is an algebra homomorphism from the space of polynomials $\mathbb{C}[z]$ to the algebra $\mathbb{C}[A]$ generated by 1 and A. For any fixed complex number λ, by the fundamental theorem of algebra, the polynomial can be decomposed as a product of linear factors, hence we have

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

where λ_i are the roots of the polynomial. 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 $(A^\dagger - \overline{\lambda}) = (A-\lambda)^\dagger$ is invertible. Where $\overline{\lambda}$ denotes the complex conjugate of λ. Hence we have proved the following theorem:

Theorem 1.4.4. (Spectral Mapping for Polynomials)

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

We now relate the norm $\|A\|$ of an element A to the *size* of its spectrum. For every λ 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

$$0 \leqq \rho(A) \coloneqq \sup_{\lambda \in \sigma(A)} \{|\lambda|\} \subseteq \big[-\|A\|, \|A\|\big].$$

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

By the spectral mapping theorem, whenever λ is $\sigma(A)$, it follows that λ^n is in $\sigma(A^n)$. Hence we have $\rho(A) \leqq \underline{\lim}_k \|A^k\|^{1/k}$. For a bounded linear functional ω and $|w| > \rho(A)$, the function,

$$f(w) = w(\omega \circ R_A)(w),$$

defines an analytic function on the open disc $B_{1/\rho(A)}(0)$. Expanding in terms of its power series at $w = z^{-1}$ we obtain,

$$f(z) = \sum_{i \in \mathbb{N}} \omega(A^i) z^i.$$

for z in $B_{1/\|A\|}(0) \subset B_{1/\rho(A)}(0)$. It follows that the same series must converge on the disc $B_{1/\rho(A)}(0)$ since f is analytic on the bigger disc as well. Hence it must be that the sequence $\omega(A^i)\lambda^i = \omega((\lambda A)^i)$ is a bounded sequence for every bounded linear functional ω, hence $(\lambda A)^i$ is bounded. By the Banach-Steinhaus theorem or uniform boundedness principle the set $\{\|(\lambda A)^i\|\}$ is bounded, by say M. By letting $|\lambda|$ tend to $\rho(A)^{-1}$ we obtain,

$$\overline{\lim_{k \in \mathbb{N}}} \big\|A^k\big\|^{\frac{1}{k}} \leqq \rho(A) \leqq \underline{\lim_{k \in \mathbb{N}}} \big\|A^k\big\|^{\frac{1}{k}}.$$

Since $\overline{\lim}$ must always greater than $\underline{\lim}$, the above inequalities must be an equality. This gives us the following formula for spectral radius;

Theorem 1.4.5. (Spectral Radius Formula)

$$\rho(A) = \lim_{n\in\mathbb{N}} \left\|A^n\right\|^{\frac{1}{n}}, \quad \forall A \in \mathcal{A}.$$

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

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

Since $\sigma(A^\dagger) = \overline{\sigma(A)}$ the spectrum of a self-adjoint element consists of real numbers.

A self-adjoint element A of a C^*-algebra is called positive if its spectrum consists of non-negative real numbers, and is denoted by $A \geq 0$. The relation $\geq$ gives rise to an order on C^*-algebras, see §I.6 [40].

1.4.2.1 | 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^*-algebras to the complex numbers. We are interested in studying the case when all the C^*-algebraic structures are preserved. Since the product structure on $\mathbb{C}$ is commutative, the product can only be preserved if $\mathcal{A}$ is a commutative C^*-algebra. We will assume $\mathcal{A}$ is a commutative C^*-algebra, and all results in this section will be for commutative C^*-algebras.

Let $\mathcal{A}$ be a commutative unital C^*-algebra, a linear functionals on $\mathcal{A}$ is called a character if it is also a *-algebra homomorphism and non-zero. The collection of all characters on $\mathcal{A}$ is called the Gelfand spectrum of $\mathcal{A}$ denoted by $K_\mathcal{A}$. If φ is a character then *-algebra homomorphism requirements can be summarised by

$$\chi(AB + \lambda C^\dagger) = \chi(A)\chi(B) + \lambda\overline{\big(\chi(C)\big)}.$$

This definition immediately implies that $\chi(1) = \chi(1^2) = \chi(1)^2$ and hence $\chi(1) = 1$. Hence $\chi(A)$ is non-zero whenever A is invertible. Hence the existence of inverse of $(A - \lambda)$ implies that $\chi(A) - \lambda$ is non-zero. Hence $\chi(A) \in \sigma(A)$. For self-adjoint elements, we have

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

it follows that

$$|\chi(A)| \leqq \|A\|, \ \ \forall A \in \mathcal{A}.$$

Hence every character χ is a bounded linear functionals with $\|\chi\| = \sup_{\|A\|\leqq 1} |\chi(A)| = 1$. $K_\mathcal{A}$ is a closed subset of the unit ball in the space of continuous linear functionals with respect to the weak* topology.[8] By Alaoglu's theorem, $K_\mathcal{A}$ is a compact Hausdorff space.

Theorem 1.4.6. *$K_\mathcal{A}$ is a compact Hausdorff space.*

The space of continuous functions on $K_\mathcal{A}$, denoted by $\mathcal{C}(K_\mathcal{A})$ contains the topological data of the space $K_\mathcal{A}$, and equipped with the operator norm it is a commutative C^*-algebra.

[8] χ is near to κ if $\chi(A)$ is near to $\kappa(A)$ as complex numbers for every A in $\mathcal{A}$.

Due to the weak*-continuity of characters, the map $\chi \mapsto \chi(A)$ is a continuous function on $K_{\mathcal{A}}$ for each $A \in \mathcal{A}$. Since every character χ is a C^*-homomorphism it follows that

$$\begin{aligned} \Gamma : \mathcal{A} &\to \mathcal{C}(K_{\mathcal{A}}) \\ A &\mapsto \Gamma_A \end{aligned} \qquad \text{(Gelfand transform)}$$

where Γ_A is a continuous function on $K_{\mathcal{A}}$ given by $\chi \to \chi(A)$, is a C^*-homomorphism. Γ is called the Gelfand transform. We now relate $\mathcal{A}$ and $\mathcal{C}(K_{\mathcal{A}})$ topologically, proving the Gelfand-Naimark theorem.

From the discussion above, we know that for each A in the commutative C^*-algebra $\mathcal{A}$, its Gelfand transform Γ_A is a continuous function. Since the norm on the space of continuous functions is the uniform norm, we have to make sure there exists a point in $K_{\mathcal{A}}$, that is, there exists a character χ on $\mathcal{A}$, such that

$$\big|\chi(A)\big| = \|A\|_{\mathcal{A}}.$$

The spectral radius formula relates the spectrum of A with the norm $\|A\|_{\mathcal{A}}$. We now use the spectrum to construct the required character. This would ensure the Gelfand transform is an isometry between $\mathcal{A}$ and $\mathcal{C}(K_{\mathcal{A}})$.

Theorem 1.4.7. (Gelfand-Naimark) *Γ is an isometric C^*-isomorphism.*

Proof

If λ is in the range of the continuous function Γ_A there must exist a character χ such that $\Gamma_A(\chi) = \lambda$, and hence by linearity of χ, we have $\chi(A - \lambda) = 0$. Since characters can never be zero for invertible elements $A - \lambda$ must not be invertible. Hence

$$\sigma(A) \subseteq \big\{\Gamma_A(\chi)\big\}_{\chi \in K_{\mathcal{A}}}.$$

If λ 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

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

π_λ is a Banach*-algebra homomorphism. $\mathcal{A}/\mathcal{I}_\lambda$ cannot have any non-invertible elements, because otherwise it would contain 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 $\chi_\lambda = \psi_\lambda \circ \pi_\lambda$ will also be a C^*-homomorphism. Hence χ_λ is a character. By construction the image of $A - \lambda$ under π_λ must be zero, and hence we have $\chi_\lambda(A - \lambda) = 0$. Hence we have

$$\{\Gamma_A(\chi)\}_{\chi \in K_{\mathcal{A}}} \subseteq \sigma(A),$$

Hence we have,

$$\sigma(A) \equiv \big\{\Gamma_A(\chi)\big\}_{\chi \in K_{\mathcal{A}}}$$

This relates the range of Γ_A with the spectral radius;

$$\big\|\Gamma_A\big\|_{\sup} = \sup_{\chi\in K_{\mathcal{A}}}\Big\{\big|\Gamma_A(\chi)\big|\Big\} = \big\|A\big\|_{\sigma}.$$

Since $\mathcal{A}$ is a C^*-algebra we have for every self-adjoint element $\|A^\dagger\|_{\mathcal{A}} = \|A\|_{\mathcal{A}}$, and hence we can relate the spectral radius with the norm. To reduce the problem to the case of self-adjoint elements, for any $B \in \mathcal{A}$ we consider the element $A = B^\dagger B$ which is always self-adjoint. Hence we have,

$$\big\|A^{2^k}\big\|_{\mathcal{A}} = \big\|(A^{2^k-1})^\dagger(A^{2^k-1})\big\|_{\mathcal{A}} = \big\|A^{2^k-1}\big\|^2_{\mathcal{A}}.$$

By repeatedly doing the same, it follows that $\|A^{2^k}\|^{\frac{1}{2^k}}_{\mathcal{A}} = \|A\|_{\mathcal{A}}$, and when combined with the spectral radius formula we have $\rho(A) = \|A\|_{\mathcal{A}}$. Hence we have,

$$\big\|\Gamma_A\big\|_{\sup} = \big\|A\big\|_{\mathcal{A}}.$$

The Gelfand transformation is an isometry.

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

$$\Gamma_A(\chi) = \chi(A) \neq \psi(A) = \Gamma_A(\psi).$$

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

The Gelfand-Naimark theory relates properties of a commutative C^*-algebra with the properties of its Gelfand spectrum. Since the Gelfand-Naimark theorem only used the equality of the spectral radius and the norm for self-adjoint elements, the theorem is also applicable to symmetric commutative Banach *-algebras, for which the norm is invariant under involutions operation. The Gelfand-Naimark theory is also called the spectral theory of C^*-algebras.

For the non-unital case, the notion of the Gelfand spectrum becomes more complex, and the Gelfand spectrum corresponds to a locally compact Hausdorff space instead of a compact Hausdorff space. The construction remains the same as for the unital case, however in the non-unital case we embed the C^*-algebra inside the space of automorphisms of the C^*-algebra, and use the identity of this space of automorphisms as the identity. The Gelfand spectrum of this *unitised* C^*-algebra is the one-point compactification of the Gelfand spectrum of the non-unital C^*-algebra we started with.

Continuous Functional Calculus

Let K be a locally compact Hausdorff space, and $\mathcal{C}_c(K)$ be the space of all compactly supported continuous functions on K. The Riesz-Markov representation theorem says that every bounded linear functional $\widehat{\mu}$ on the space $\mathcal{C}_c(K)$ corresponds to a complex Radon measures μ on K such that,

$$\widehat{\mu}(f) = \int_K f(\lambda)d\mu(\lambda), \quad \forall f \in \mathcal{C}_c(K).$$

Hence the dual Banach space $\mathcal{C}_c(K)^*$ is identified with the space of Radon measures $\mathcal{M}(K)$. Since the Gelfand spectrum $K_{\mathcal{A}}$ for a unital commutative C^*-algebra $\mathcal{A}$ is a compact Hausdorff space, continuous linear functionals on $\mathcal{A}$ correspond to measures on $K_{\mathcal{A}}$, the Gelfand transform allows us to import many of the tools for continuous functions to the study of commutative C^*-algebras.

The Gelfand-Naimark theorem allows us to introduce tools for continuous functions on real numbers to the study of self-adjoint elements, for any continuous function f on the complex numbers the composition $f \circ \Gamma_A$ is a continuous function on $K_{\mathcal{A}}$. Hence for every bounded linear functional ω on $\mathcal{A}$ corresponds to a Radon measure μ_ω on $K_{\mathcal{A}}$, such that

$$\omega(f(A)) = \int_{K_{\mathcal{A}}} (f \circ \Gamma_A)(\lambda) d\mu_\omega(\lambda), \quad \forall A \in \mathcal{A}.$$

Hence for any continuous function f we can make sense of the element $f(A)$ defined by,

$$\Gamma_{f(A)} \equiv f \circ \Gamma_A.$$

Since the spectrum is identical to the range of the Gelfand transform, it follows that for self-adjoint elements Γ_A has real range. Since $f(A)^* = \overline{f(A)} = f(A)$ it follows that $f(A)$ corresponds to a self-adjoint element for real valued functions.

If $\mathcal{A}$ is generated by a self-adjoint element A and the identity, the continuous function Γ_A acts as identity on the spectrum $\sigma(A)$ and is zero outside, that is, $K_{\mathcal{A}} \backslash \sigma(A)$. Hence for any bounded linear functional ω there exists a Radon measure μ_ω such that

$$\omega(A) = \int_{\sigma(A)} \lambda d\mu_\omega(\lambda), \quad \forall A \in \mathcal{A}.$$

We cannot however make sense of this integral without reference to the state ω since the measure μ_ω depends on ω.

The Spectrum & Stone Spaces

It is the relation between Boolean algebra structures and measuring instruments which made us study the structure of commutative C^*-algebras. We now make this relation precise, by constructing Boolean algebras from the Gelfand spectrum of commutative C^*-algebra. Given a commutative C^*-algebra $\mathcal{A}$ with identity, the Gelfand spectrum $K_{\mathcal{A}}$ equipped with the weak* topology is a compact Hausdorff space.

Let E be a projection element of the commutative C^*-algebra $\mathcal{A}$, then its Gelfand transform Γ_E corresponds to a continuous function on $K_{\mathcal{A}}$. Since

$$\Gamma_E = \Gamma_{E^2} = \Gamma_E^2$$

and since 0 and 1 are the only complex complex numbers with this property, it follows that Γ_E must be the characteristic function on the closure of the set of all characters χ with $\Gamma_E(\chi) = 1$. If the set of all such characters if denoted by U_E, then since $U_E^c = \Gamma_E^{-1}(0)$, it follows that U_E is both closed and open. Every projection element of $\mathcal{A}$ corresponds to a continuous characteristic function on a clopen set,

$$\chi_{U_E} = \Gamma_E.$$

We can construct a Boolean algebra using the clopen sets of the topological space $K_{\mathcal{A}}$. By Stone duality, the clopen sets of $K_{\mathcal{A}}$ together with intersection, union and complementation is a Boolean algebra, and by the above, such clopen subsets can be identified with continuous characteristic functions.

If the spectrum $K_{\mathcal{A}}$ of a commutative C^*-algebra $\mathcal{A}$ corresponds to the Stone space of a complete Boolean algebra of simultaneously measurable outcomes, then the points of $K_{\mathcal{A}}$ correspond to the smallest outcomes. The points of the space can be described as

limits of characteristic functions of regions converging to a point as described in the above limit. However such a limit need not make sense from within the C^*-algebra, since not every characteristic function is continuous. For such a Boolean algebra to be complete, we should be able to make sense of taking limits of clopen sets, that is, each point of the Gelfand spectrum must correspond to an element of the Boolean algebra, or that each point in the Gelfand spectrum must be a clopen set. By compactness of $K_{\mathcal{A}}$ we can assume that the range of a continuous f on $K_{\mathcal{A}}$ is contained in the compact square. For any n, we decompose this square into n^2 smaller squares, denote these smaller squares by $K_{i,j}$ where $i, j \leqq n$. Let $\chi_{i,j}$ be the characteristic functions for $f^{-1}(K_{i,j})$. Let $f_{i,j} = f(k_{i,j})$ for some $k_{i,j} \in f^{-1}(K_{i,j})$. Then we have,

$$\sum_{i,j} f_{i,j}\chi_{i,j}(x) \to f(x)$$

uniformly. Hence every continuous function in $\mathcal{C}(K_{\mathcal{A}})$ can be uniformly approximated by finite linear combinations of projections. Every characteristic function is measurable, and Radon measures on a space can usually only see the regions upto their boundary. However, if there exist a collection of measures which can *see* the regions precisely, then we may be able to remedy the issue of limits by transfering some of the burden to such measures.

1.4.2.2 | Spectral Theorem

Every abstract stonean space gives rise to a complete Boolean algebra of clopen subsets, and conversely, by Stone duality, every complete Boolean algebra can be thought of as the complete Boolean algebra of clopen subsets of the stonean space of maximal ideals. We now obtain a characterisation of complete Boolean maps, in the context of arbitrary W^*-algebras, and since observables are Boolean maps from a complete Boolean algebra to the algebra of observables we also obtain a characterisation of the notion of observables in quantum theory.

In the context of the Gelfand-Naimark theorem, the topology on the Gelfand spectrum is closely related to the norm of the C^*-algebra, and from the discussion above this topology is too strong, and does not allow us to make sense of limits of characteristic functions. For the case of quantum theory, every ensemble must give rise to a probability measure for the collection of outcomes of an observables. The Banach space generated by ensembles is the predual of the algebra of observables. We now study how the pairing of a W^*-algebra with its predual reduces the *burden* on the W^*-algebra, and discuss how the σ-weak topology on W^*-algebras allows us to take limits of characteristc functions. The topology on the Gelfand spectrum is closely related to spectral order on C^*-algebras. This is similar to how the topological structure on real numbers is closely related to the order structure on real numbers. For the Gelfand spectrum to be an abstract stonean space it must be extremely disconnected, or alternatively, every point must be simultaneously open and closed, or equivalently the characteristic functions for every open set should be continuous. We say a C^*-subalgebra is maximally commutative if it is a commutative C^*-subalgebra, which is not contained in any other commutative C^*-subalgebra. We now develop the necessary background to it starting with the characterisation of the spectrum, via theorem 1.4.2.

Lemma 1.4.8. *The spectrum of any maximally commutative C^*-subalgebra of $\mathcal{A}$ is Stonean.*

Proof

We must prove that any bounded net $\{A_\alpha\}$ of positive elements in a maximally commutative C^*-algebra $\mathcal{C}$ has a supremum. Note that $\sup_\alpha A_\alpha$ makes sense, as an element in $\mathcal{A}$. Since $\mathcal{C}$ is maximal, it is sufficient to show that $\sup_\alpha A_\alpha$ commutes with every element of $\mathcal{C}$.

To further simplify, we will show that $\sup_\alpha\{A_\alpha\}$ commutes with every unitary element of $\mathcal{C}$, and since we can obtain every element of $\mathcal{C}$ can be obtained as finite linear combination of unitary elements, it would follow that $\sup_\alpha\{A_\alpha\}$ also belongs to $\mathcal{C}$.

For unitary elements, we have $U = U^\dagger = U^{-1}$. We have,

$$\sup_\alpha \{U^\dagger A_\alpha U\} = U^\dagger \big[\sup_\alpha\{A_\alpha\}\big] U$$

Hence

$$U^\dagger \big[\sup_\alpha\{A_\alpha\}\big] U = \sup_\alpha\{A_\alpha\}$$

This proves the lemma. □

Recall that by the Riesz-Markov representation theorem says that the dual Banach space of the space of compactly supported continuous functions $C_c(K)$ of a locally compact Hausdorff space K is the space of Radon measures on K, and every continuous linear functional ω on $C_c(K)$ is of the form,

$$\omega(f) = \int_K f(x) d\mu_\omega(x), \quad \forall f \in C_c(K).$$

Let $\mathcal{A}$ be a W^*-algebra and $\mathcal{A}_*$ be its predual. We now prove that W^*-algebras have properties that allow for an intrinsic meaning for the above integral expression. We now decompose a self-adjoint element as a collection of supports.

A spectral measure E_A is a Boolean map from the σ-complete Boolean algebra $\mathcal{B}(\mathbb{R})$ generated by intervals into the lattice of projections of a W^*-algebra. That is,

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

such that

$$E_A(\varnothing) = 0\,;\; E_A(\mathbb{R}) = 1,$$

and $E_A(\mathbb{R}\backslash U) = 1 - E_A(U)$, and for every sequence $\{U_i\}$ of bounded increasing sequence in $\mathcal{B}(\mathbb{R})$ converging to a set U, we must have $E_A(\vee_{i\in\mathbb{N}} U_i) = \vee_{i\in\mathbb{N}}\big(E_A(U_i)\big)$. Equivalently, for any countable collection of disjoint sets in $\mathcal{B}(\mathbb{R})$, we must have,

$$E_A(\vee_i U_i) = \textstyle\sum_i E_A(U_i),$$

whenever $U_i \cap U_j$ is empty for i different from j.

Since $\mathcal{P}(\mathcal{A})$ is a complete lattice, the collection of characteristic functions $\{\chi_U\}_{x\in U}$ where U belongs to some neighborhood base at x, has a limit, and corresponds to the characteristic function of the point x. For each self-adjoint element A, we obtain a spectral measure,

$$\begin{aligned} E_A : \mathcal{B}(\mathbb{R}) &\to \mathcal{P}(\mathcal{A}) \\ U &\mapsto E_A(U) \equiv \Gamma^{-1}\big(\chi_U\big). \end{aligned}$$

where U is a Borel set in the spectrum $\sigma(A)$, and by complete additivity of normal functionals, it follows that $\mu_{\omega,A} \equiv \omega \circ E_A$ is a σ-finite measure. Combining with the fact that $\mathcal{A}_*$ separates $\mathcal{A}$, obtain the spectral theorem as a corollary, from §1.11 [36];

Theorem 1.4.9. (Spectral Theorem) *If $\mathcal{A}$ is a W^*-algebra, then for every self-adjoint element A in $\mathcal{A}$, there exists a unique spectral measure*

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

such that

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

such that for every ω in $\mathcal{A}_*$ we have

$$\Big\langle \omega \Big| \Big[\int_{\mathbb{R}} \lambda dE_A(\lambda) \Big] \Big\rangle_{\mathcal{A}_*} = \int_{\mathbb{R}} \lambda d\langle \omega | E_A(\lambda) \rangle.$$

Let (Ω, μ) be a measure space, then $L^\infty(\Omega, \mu)$ of essentially bounded μ-measurable functions on Ω is a C^*-algebra, and since its the dual space of $L^1(\Omega, \mu)$ it is a W^*-algebra. A direct sum of finite measure spaces is called a localizable space. We now show that the converse also holds, that is, every commutative W^*-algebra can be thought of as $L^\infty(\Gamma, \nu)$ for some localizable measure space (Γ, ν).

Theorem 1.4.10. *Let $\mathcal{C}$ be a commutative W^*-algebra. Then $\mathcal{C}$ is isomorphic to $L^\infty(\Gamma, \nu)$ for some localizable measure space (Γ, ν).*

Proof

Let $K_{\mathcal{C}}$ be the Gelfand spectrum of $\mathcal{C}$. By Riesz-markov representation theorem any normal state ω, defines a unique Radon measure on $K_{\mathcal{C}}$ such that,

$$\omega(A) = \int_{K_{\mathcal{C}}} \Gamma_A(x) d\mu_\omega(x) \ \forall A \in \mathcal{C}.$$

Hence we must have $C(K_{\mathcal{C}}) \subseteq L^\infty(K_{\mathcal{C}}, \mu_\omega)$. The Gelfand transform gives us a map,

$$\begin{aligned} \Gamma_\omega : \mathcal{C} &\to L^\infty(K_{\mathcal{C}}, \mu_\omega) \\ A &\mapsto \Gamma_\omega(A). \end{aligned}$$

Since the Gelfand transform is a *-algebra homomorphism, it also follows that Γ_ω is a normal homomorphism. $\Gamma_\omega(\mathcal{C}) \subseteq L^\infty(K_{\mathcal{C}}, \mu_\omega)$. The kernel $\ker(\Gamma_\omega)$ is a closed two-sided ideal in $\mathcal{C}$, and hence there exists a projection E_ω with $\ker(\Gamma_\omega) = \mathcal{A}(1 - E_\omega)$. Choose a maximal family of normal states $\{\omega_\alpha\}$ whose supports are orthogonal to each other. Then we have,

$$\begin{aligned} \mathcal{C} = \sum_\alpha{}^{\oplus} \big(\mathcal{C}E_{\omega_\alpha}\big) &= \sum_\alpha{}^{\oplus} L^\infty\big(K_{\mathcal{C}_\alpha}, \mu_{\omega_\alpha}\big) \\ &= L^\infty\Big(\bigcup_\alpha K_{\mathcal{C}_\alpha}, \textstyle\sum_\alpha^{\oplus} \mu_{\omega_\alpha}\Big). \end{aligned}$$

Here $K_{\mathcal{C}_\alpha}$ is the image of $K_{\mathcal{C}}$ under $\Gamma_{E_{\omega_\alpha}}$. Hence $\mathcal{C}$ is isomorphic to $L^\infty\big(\cup_\alpha K_{\mathcal{C}_\alpha}, \sum_\alpha^{\oplus} \mu_{\omega_\alpha}\big)$. $\square$

A positive Radon measure μ on a stonean space K is called normal if for any directed collection $\{f_\alpha\} \subset \mathcal{C}_{\mathbb{R}}(K)$ of uniformly bounded functions, with supremum $f = \sup_\alpha\{f_\alpha\}$, we have

$$\mu(f) = \sup_\alpha\{\mu(f_\alpha)\}.$$

An abstract stonean space K is said to be hyperstonean if it admits sufficiently many positive normal measures, that is, if for any non-zero function f in $\mathcal{C}_{\mathbb{R}}(K)$ there exists a normal measure μ_f such that

$$\mu_f(f) \neq 0.$$

That is, the space of real valued continuous functions can be separated by normal measures. Then $K_{\mathcal{A}}$ is hyperstonean if and only if $\mathcal{A}$ is a W^*-algebra.

Appendices for the Chapter

1.A | Quantum Theory on Hilbert Spaces

W^*-algebras are special case of C^*-algebras analogous to how Hilbert spaces are special as Banach spaces. By viewing operator algebras as subalgebras of bounded operators on some Hilbert space, we can use the tools for the study of Hilbert spaces. In this section, we quickly study the concrete operator algebras, that is, operator algebras which are subalgebras of bounded operators on some Hilbert space.

1.A.1 | Necessity for Quantum Formalism

By the end of the nineteenth century, it was clear that certain elementary processes obeyed 'discontinuous' laws. That is to say, there exist observables whose collection of effects forms discrete sets, and also observables whose collection of effects forms a continuous set.

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

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

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

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

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

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

where $\mu_\rho^Q(Q_i)$ is the probability the observable has a value Q_i when the state is μ_ρ. In particular, the sum should make sense. The expected value of the observable R for the state μ_ρ is given by,

$$\langle Q\rangle_\rho = \textstyle\sum_{i\in\mathbb{N}} Q_i \mu_\rho^Q(Q_i).$$

On the other hand if the collection of effects of an observable can be labeled by a continuous set, that is, the measurements are labelled by countable unions of open intervals of $\mathbb{R}$, such an observable corresponds to a map, which assigns to intervals R_i of $\mathbb{R}$ the effect E_{R_i} where

R_i is the interval in which the value of the observable R lies. Each ensemble μ_ρ corresponds to a function, $\mu_\rho : \mathcal{A} \to \mathbb{R}$. The gives us a composite map,

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

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

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

In particular the integral should make sense. Each state corresponds to the measure $\mu_\rho^R d\lambda$. The expected value of the observable R for the state μ_ρ is given by,

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

Classical theories model observables using integrable functions on a manifold. It is sufficient to assume integrability of the functions for this model. If the manifold is Ω, observables in classical physics correspond to integrable functions of the form

$$R : \Omega \to \mathbb{R}.$$

Denote by R^* the pre-image map $\mathcal{B}(\mathbb{R}) \to \mathcal{B}(\Omega)$. The state in classical physics corresponding to the ensemble ρ is modeled by a positive Radon measure μ_ρ on Ω, and corresponds to probability measures on Ω. The probability measure on Ω the state μ_ρ associates with the observable R is

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

This is the composite map,

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

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

$$\int_\Omega d\mu_\rho(\omega) = 1.$$

That is, μ_ρ^R is the image measure of μ_ρ under R. The expected value of the observable R for an ensemble ρ is given by

$$\langle R \rangle_\rho = \int_\Omega R(\omega) d\mu_\rho(\omega).$$

The collection of all characteristic functions on Ω is itself an observable and since every integrable function can be approximated as combinations of simple function, every observable in the classical model coexists with each other.

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

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

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

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

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

such that,

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

where $\mu_\rho^Q(Q_i) = \mu_\rho(Q^{-1}(Q_i))$. Such functions correspond to summable sequences denoted by $l^1(\Omega)$. The coexistence is not possible because the space of integrable functions $L^1(\Omega)$ is not isomorphic to the space of summable sequences $l^1(\Omega)$.[9]

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

1.A.1.1 | Heisenberg's Approach

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

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

[9] 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.

Von Neumann & Hilbert Spaces

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 a complete orthonormal system and an element of the Hilbert space can be uniquely specified by its coordinates with respect to a complete orthonormal system. A separable Hilbert space is a Hilbert space which has a countable complete orthonormal system, and any two separable Hilbert spaces are isomorphic.

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

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

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

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

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

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

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

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

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.A.1. (Riesz-Fischer) *$L^2(\mathbb{R}) \cong l^2(\mathbb{N})$ as Hilbert spaces, and are separable Hilbert spaces.*

This isomorphism provides us with a state space that allows for coexistence of discrete and continuous observables. See §1.4 [48] 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 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.

1.A.2 | The Space of Bounded Operators on $\mathcal{H}$

Since our goal is to give a quick overview, we will skip the details about Hilbert spaces, and refer an interested reader to [32]. Let $\mathcal{H}$ be a complex Hilbert space with a sesquilinear inner product $\langle\cdot|\cdot\rangle_{\mathcal{H}}$. Equipped with the operator norm, the algebra of all bounded linear operators on $\mathcal{H}$ is denoted by $\mathcal{B}(\mathcal{H})$. We will assume $\mathcal{H}$ is a separable, that is, its bases are countable. For any A in $\mathcal{B}(\mathcal{H})$, by the Riesz-Frechet representation theorem, the map

$$\nu \mapsto \big\langle \nu | A\eta \big\rangle_{\mathcal{H}}$$

corresponds to a unique vector $A^\dagger\nu$ in $\mathcal{H}$ such that $\langle\eta|A^\dagger\nu\rangle_{\mathcal{H}} = \langle A\eta|\nu\rangle_{\mathcal{H}}$, and hence for each operator A in $\mathcal{B}(\mathcal{H})$ we obtain an operator $A^\dagger$ in $\mathcal{B}(\mathcal{H})$, and since $|\langle\nu|A^\dagger A\nu\rangle_{\mathcal{H}}| = |\langle A\nu|A\nu\rangle_{\mathcal{H}}| = \|A\nu\|^2$. $\mathcal{B}(\mathcal{H})$ is a C^*-algebra, see §2 [32].

Hilbert spaces are their own dual spaces. We now relate the operators on $\mathcal{H}$ and the vectors in $\mathcal{H}$. This allows us to use Hilbert spaces to study operator subalgebras of $\mathcal{B}(\mathcal{H})$. Let $E_{\mathcal{H}}$ be a complete orthonormal system of $\mathcal{H}$, then a non-negative self-adjoint operator ρ in $\mathcal{B}(\mathcal{H})$ is trace class if

$$\mathrm{Tr}(\rho) = \sum_{\nu\in E_{\mathcal{H}}} \big\langle \nu|\rho\nu\big\rangle_{\mathcal{H}}$$

is finite. The vector space generated by trace class operators on $\mathcal{H}$ is denoted by $L^1(\mathcal{H})$. Since trace can be shown to be independent of the choice of basis it follows that $L^1(\mathcal{H})$ equipped with the norm $\|\cdot\|_{L^1} = \mathrm{Tr}(|\cdot|)$ is a normed space. $L^1(\mathcal{H})$ is also a Banach space, see [32].

The space of bounded operators $\mathcal{B}(\mathcal{H})$ acts as the dual space of $L^1(\mathcal{H})$, and each element A in $\mathcal{B}(\mathcal{H})$ corresponds to a linear functional on $L^1(\mathcal{H})$ given by

$$\rho \mapsto \mathrm{Tr}(\rho A).$$

Hence $\mathcal{B}(\mathcal{H})$ is a W^*-algebra with $L^1(\mathcal{H})$ as its predual. The space $L^1(\mathcal{H})$ of all trace-class operators is generated by positive trace-class operators of norm 1. Such operators are called density operators.

We now introduce a convenient manipulation tool via resolution of identities using a complete orthonormal system. Let P_η be the orthogonal projection to the subspace spanned by η. If $E_{\mathcal{H}} = \{\nu\}$ is a complete orthonormal system, then we have, $I = \sum_{\mu\in E_{\mathcal{H}}} P_\nu$ pointwise, in the sense that we can write any vector η in $\mathcal{H}$ as

$$\eta = \sum_{\nu\in E_{\mathcal{H}}} P_\nu\eta = \sum_{\nu\in E_{\mathcal{H}}} \big\langle\nu\big|\eta\big\rangle_{\mathcal{H}}\nu \equiv \Big[\sum_{\nu\in E_{\mathcal{H}}} |\nu\rangle\langle\nu|\Big]\eta.$$

where $\langle\nu|\eta = \langle\nu|\eta\rangle_{\mathcal{H}}$. Since $E_{\mathcal{H}}$ is a complete orthonormal system we must have, $\langle\nu|\eta\rangle_{\mathcal{H}} = \delta_{\nu,\eta}$. We can now manipulate this by using $\langle\nu|\eta\rangle_{\mathcal{H}} = \delta_{\nu,\eta}$, and we obtain,

$$\big|\,\mathrm{Tr}(\rho A)\big| = \sum_{\nu\in E_{\mathcal{H}}}\sum_{\eta\in E_{\mathcal{H}}} \big|\big\langle\eta\big|\rho\eta\big\rangle_{\mathcal{H}}\big|\big|\big\langle\nu\big|A\nu\big\rangle_{\mathcal{H}}\big| \leqq \big\|A\big\|\,\mathrm{Tr}(\rho).$$

Every orthogonal projection in $\mathcal{H}$ is a trace class operator. If A is non-zero, then there exists some η and ν such that $\langle\eta|A\nu\rangle$ is non-zero. By taking $\rho = P_\eta$ we must have

$$\mathrm{Tr}(P_\eta A) = \sum_{\nu\in E_{\mathcal{H}}} \langle\nu|P_\eta A\nu\rangle_{\mathcal{H}} = \langle\eta|A\nu\rangle_{\mathcal{H}}$$

The σ-weak topology on $\mathcal{B}(\mathcal{H})$ is induced by the trace-class operators with the topology induced by the collection of semi-norms $\{A \mapsto |\langle\nu|A\eta\rangle|\}_{\nu,\eta\in\mathcal{H}}$.

1.A.2.1 | Von Neumann Subalgebras

A W^*-algebra $\mathcal{A}$ which is a subalgebra of $\mathcal{B}(\mathcal{H})$ is called a von Neumann subalgebra of $\mathcal{B}(\mathcal{H})$. Every A in $\mathcal{A}$ is a linear functional on $L^1(\mathcal{H})$. Denote by $\mathcal{A}_*^\perp$ the subspace of $L^1(\mathcal{H})$ which is annihilated by $\mathcal{A}$. This annihilator subspace is a closed subspace, and hence the quotient space $L^1(\mathcal{H})/\mathcal{A}_*^\perp$ is a Banach space. Hence by the uniqueness of the predual we can identify the predual $\mathcal{A}_*$ of $\mathcal{A}$ with $L^1(\mathcal{H})/\mathcal{A}_*^\perp$. Hence it is useful to study the structure of the lattice of projections of $\mathcal{B}(\mathcal{H})$.

The von Neumann subalgebras on Hilbert spaces have a convenient algebraic characterisation due to von Neumann. For any subset S of $\mathcal{B}(\mathcal{H})$, the collection of all operators which commute with the elements of S is called the commutant of S and denoted by S'. Then clearly, we must have,

$$S \subseteq S''.$$

Clearly, S' contains the identity operator.

Let $\mathcal{A}$ be a C^*-algebra. If $AB = BA$ for all A in $\mathcal{A}$, then $A^\dagger B = BA^\dagger$. In such a case we have $B^\dagger A = AB^\dagger$. Hence commutants of C^*-algebras are C^*-algebras. Von Neumann's double commutant theorem characterises W^*-algebraic in terms of the commutatnts.

Theorem 1.A.2. (Double Commutant) *If $\mathcal{A}$ is a W^*-algebra on $\mathcal{H}$ then $\mathcal{A}'' = \mathcal{A}$.*

That is, $\mathcal{A}_*$ acts as the predual of $\mathcal{A}''$ as well. For a proof of the double commutant theorem, see [32]. The double commutant of any C^*-algebra acts as its universal enveloping von Neumann algebra $\mathcal{A}''$, and we must have an isomorphism between $\mathcal{A}''$ and the bidual $\mathcal{A}^{**}$. This is known as the Sherman-Takeda theorem [37]-[38].

1.A.2.2 | Lattice of Closed Subspaces

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

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

of the form,

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

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

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

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

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

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

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

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

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

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

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

Theorem 1.A.3. (Gleason) *If $\mathcal{H}$ is a Hilbert space of dimension greater than 2 then every non-commutative probability measure ω is of the form*

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

where ρ is a positive element in $L^1(\mathcal{H})$ with $\mathrm{Tr}(\rho) = 1$.

The operator algebraic approach via trace-class operators, Hilbert-Schmidt operators, and bounded operators, is simpler if one knows more sophisticated mathematics. We refer an interested reader to the discussion on Hilbert-Schmidt operators and trace-class operators in [32].

The extreme points of the state space are called pure states and are of the form, $\rho = \rho^2$ and corresponds to some vector $|\varphi\rangle$ in the Hilbert space $\mathcal{H}$ and ρ is the projection onto the subspace generated by $|\varphi\rangle$. Such a state is denoted by $|\varphi\rangle\langle\varphi|$. For an observable with an

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

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

The expectation value of the observable will be,

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

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

1.A.3 | Representation of W^*-Algebras

Recall that $\mathcal{G}(\mathcal{A})$ denotes the collection of invertible elements of $\mathcal{A}$. If π is a C^*-homomorphism from $\mathcal{A}$ to $\mathcal{B}$ then we must have, $\mathcal{G}(\mathcal{A}) \subseteq \mathcal{G}(\mathcal{B})$. Since the spectrum of an element A is defined to the all λ such that $A - \lambda$ is not invertible, it follows that $\sigma(\pi(A)) \subset \sigma(A)$. For C^*-algebras, the norm of elements is the same as the spectral radius for self-adjoint elements. By combining with the C^*-identity we must have

$$\big\|\pi(A)\big\|_{\mathcal{B}} \leqq \big\|A\big\|_{\mathcal{A}}.$$

Hence every C^*-homomorphism is automatically norm-continuous. If π is an isomorphism of the C^*-algebras $\mathcal{A}$ and $\mathcal{B}$ then we must have $\mathcal{G}(\mathcal{A}) = \mathcal{G}(\mathcal{B})$ and $\|A\|_{\mathcal{A}} = \|\pi(A)\|_{\mathcal{B}}$. A representation of $\mathcal{A}$ on the Hilbert space $\mathcal{H}$, is a C^*-homomorphism,

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

A representation π is called a faithful representation if π is injective, or $\mathcal{A}$ is isomorphic to its image $\pi(\mathcal{A})$. In such a case $\mathcal{A}$ can be studied as operators on $\mathcal{H}$. A normal representation π_ω of a W^*-algebra, is a *-homomorphism into a W^*-subalgebra of $\mathcal{B}(\mathcal{H}_\omega)$. That is, for every A in $\mathcal{A}$ the composition of $\pi_\omega(A)$ acts as a bounded functional on predual of the target W^*-algebra. By von Neumann's double commutant theorem in this case we must have $\pi_\omega(\mathcal{A})'' = \pi_\omega(\mathcal{A})$.

Two representations $(\mathcal{H}, \pi_\mathcal{H})$ and $(\mathcal{K}, \pi_\mathcal{K})$ of an algebra $\mathcal{A}$ are said to be unitarily equivalent if there exists a unitary operator $U : \mathcal{H} \to \mathcal{K}$ such that

$$\pi_\mathcal{H} \circ U = U \circ \pi_\mathcal{K}.$$

Since conjugation with unitary maps preserve the norm, equivalent representations contain the same algebraic and topological data about $\mathcal{A}$.

Let $(\mathcal{H}, \pi)$ be a representation of $\mathcal{A}$, let $\mathcal{M}$ be a subspace of $\mathcal{H}$ which is invariant under π, that is, $\pi(A)\mathcal{M} \subseteq \mathcal{M}$ for all $A \in \mathcal{A}$. Suppose $\mathcal{M}$ is closed and let $P_\mathcal{M}$ be the projection with range $\mathcal{M}$ then the invariance implies,

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

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

$$\begin{aligned} \pi(A)P_\mathcal{M} &= (P_\mathcal{M} \pi(A^\dagger) P_\mathcal{M})^\dagger \\ &= (\pi(A^\dagger)P_\mathcal{M})^\dagger = P_\mathcal{M}\pi(A). \end{aligned}$$

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

$$\pi_\mathcal{M}(A) = P_\mathcal{M} \pi(A) P_\mathcal{M},$$

then $(\mathcal{M}, \pi_{\mathcal{M}})$ 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 π. If $\mathcal{M}$ is invariant under π then so is $\mathcal{M}^{\perp}$. Hence we obtain another subrepresentation. The original Hilbert space $\mathcal{H}$ can be written as a direct sum, $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$ and each operator $\pi(A)$ then decomposes as a direct sum

$$\pi(A) = \pi_{\mathcal{M}}(A) \oplus \pi_{\mathcal{M}^{\perp}}(A).$$

So the representation can be written as $(\mathcal{H}, \pi) = (\mathcal{M}, \pi_{\mathcal{M}}) \oplus (\mathcal{M}^{\perp}, \pi_{\mathcal{M}^{\perp}})$.

Let $(\mathcal{H}_\alpha, \pi_\alpha)_\alpha$ be a family of representations of $\mathcal{A}$. The vector space $\sum^{\oplus} \mathcal{H}_\alpha$ together with the inner products $\langle (x_\alpha) | (y_\alpha) \rangle \coloneqq \sum_\alpha \langle x_\alpha | y_\alpha \rangle_\alpha$ is a Hilbert space. We can construct a representation of $\mathcal{A}$ on the Hilbert space $\sum^{\oplus} \mathcal{H}_\alpha$ using the representations π_α by

$$\begin{aligned} \pi \equiv \big[\textstyle\sum^{\oplus} \pi_\alpha \big] : \mathcal{A} &\to \mathcal{B}\big(\textstyle\sum^{\oplus} \mathcal{H}_\alpha \big) \\ A &\mapsto \textstyle\sum^{\oplus} \pi_\alpha(A). \end{aligned}$$

Since each $\|\pi_\alpha(A)\| \leqq \|A\|$ it follows that $\|\pi(A)\| \leqq \|A\|$. Hence the map π is a representation of $\mathcal{A}$ on $\mathcal{H}$. This is called the direct sum of representation of the collection $\{(\mathcal{H}_\alpha, \pi_\alpha)\}$ of representations and is denoted by,

$$\sum_{\alpha \in I}^{\oplus} \Big[\big(\mathcal{H}_\alpha, \pi_\alpha \big) \Big].$$

The trivial part of a representation $(\mathcal{H}, \pi)$ is the largest subspace $\mathcal{K}$ of $\mathcal{H}$ annihilated by the action of $\mathcal{A}$. Equivalently, it corresponds to the subspace spanned by the maximal element of the lattice of all projections for which $P\pi P$ vanishes. A representation $(\mathcal{H}, \pi)$ is non-degenerate if it has no trivial part.

Let $(\mathcal{H}, \pi)$ be a non-degenerate representation of $\mathcal{A}$. Take a maximal family of non-zero vectors $\{\nu_\alpha\}$ in $\mathcal{H}$ such that whenever α and β are different, we have $\pi(\mathcal{A})\nu_\alpha \perp \pi(\mathcal{A})\nu_\beta$. In such a case let $\mathcal{H}_\alpha$ be the Hilbert space generated by the completion of $\pi(\mathcal{A})\nu_\alpha$ with respect to the inner product $\langle \cdot | \cdot \rangle_{\mathcal{H}}$. By construction, this is an invariant subspace of $\mathcal{H}$. If P_α denotes the orthogonal projection of $\mathcal{H}$ to the subspace $\mathcal{H}_\alpha$, then we obtain a representation of $\mathcal{A}$ on the Hilbert space $\mathcal{H}_\alpha$ given by

$$\pi_\alpha = P_\alpha \pi P_\alpha.$$

By construction $\{\mathcal{H}_\alpha\}$ is a pairwise orthogonal family of closed subspaces of $\mathcal{H}$. Hence every representation $(\mathcal{H}, \pi)$ can be decomposed as a direct sum of cyclic representations;

$$(\mathcal{H}, \pi) = \sum_{\alpha \in I}^{\oplus} \Big[\big(\mathcal{H}_\alpha, \pi_\alpha \big) \Big].$$

So every non-degenerate representation can be written as a direct sum of a family of cyclic subrepresentations. If there are no invariant subspaces the representation $(\mathcal{H}, \pi)$ of $\mathcal{A}$ is called irreducible.

The Gelfand-Naimark-Segal Construction

The pairing of a locally convex space with its conjugate space allows us to study the topological properties of the locally convex space in terms of its convexity properties. Hilbert spaces are their own dual spaces, and hence all the topological and convexity data is contained within. By representing a C^*-algebra on a Hilbert space, we can study the conjugacy pairing using the inner product on the Hilbert space. We now define inner products on C^*-algebras using their continuous linear functionals and by taking the completion with respect to such

inner products, we obtain representations of C^*-algebras on Hilbert spaces. Throughout this part, we assume $\mathcal{A}$ contains the identity element.

A continuous linear functional φ on $\mathcal{A}$ is said to be positive if it preserves the notion of positivity on $\mathcal{A}$, coming from the spectrum of elements. That is, it maps positive elements, to positive reals. Hence, if φ is a positive linear functional, we must have

$$\varphi(A^\dagger A) \geq 0, \quad \forall A \in \mathcal{A}.$$

The positive linear functional φ is called a state if $\|\varphi\|_{\mathcal{A}^*} = 1$. We say a state φ is faithful if $\varphi(A^\dagger A) = 0$ if and only if A is zero. The set of all states on $\mathcal{A}$ is denoted by $S_{\mathcal{A}^*}$. It follows that $\varphi(A)$ is a real for every self-adjoint element A.

If φ is a positive linear functional then by expanding $\varphi((A+\lambda B)^\dagger(A+\lambda B)) \geqq 0$ we obtain the Cauchy-Schwarz inequality for positive linear functionals;

$$\left|\varphi(A^\dagger B)\right|^2 \leqq \varphi(A^\dagger A)\varphi(B^\dagger B).$$

Hence we can expect to define inner products using positive linear functionals. By the Cauchy-Schwarz inequality above, we must have $|\varphi(A)|^2 \leqq \varphi(1)\varphi(A^\dagger A)$. However, since inner products should be non-degenerate, we should quotient out such degenerate elements of $\mathcal{A}$. Consider the subspace of $\mathcal{A}$ defined by,

$$\mathcal{I}_\varphi \equiv \big\{A \mid \varphi(A^\dagger A) = 0\big\}.$$

By the Cauchy-Schwarz inequality above, $\varphi(AB)$ must vanish whenever A is in $\mathcal{I}_\varphi$.

Hence $\mathcal{I}_\varphi$ is a left-ideal of $\mathcal{A}$ and will be called the Gelfand ideal for φ. Using the Gelfand ideal for φ we can define the quotient space;

$$\begin{aligned} \Xi_\varphi : \mathcal{A} &\to \mathcal{A}/\mathcal{I}_\varphi \\ A &\mapsto [A]. \end{aligned}$$

We denote $\mathcal{A}/\mathcal{I}_\varphi$ by $\mathcal{A}_\varphi$.

Theorem 1.A.4. (Gelfand-Naimark-Segal) *If φ is a state on $\mathcal{A}$, then*

$$\big\langle [A] | [B] \big\rangle_\varphi = \varphi(B^\dagger A)$$

is an inner product on the quotient space $\mathcal{A}_\varphi$. The completion $\mathcal{H}_\varphi$ of $\mathcal{A}_\varphi$ with respect to $\langle\cdot|\cdot\rangle_\varphi$ has a representation of $\mathcal{A}$;

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

where $\pi_\varphi(A) : [B] \mapsto [AB]$. The vector $\Omega_\varphi = [1]$ is cyclic and separating for $\mathcal{A}$, and is called the Gelfand vector.

If $\mathcal{I}_\varphi$ is a two-sided ideal, and if $\alpha : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ is a cyclic representation with $\varphi(A) = \langle x|\alpha(A)x\rangle$, with the cyclic vector x, then α is unitarily equivalent to π_φ.

Sketch of Proof

Let Ξ_φ be the quotient map, $A \mapsto [A]$ from $\mathcal{A}$ to $\mathcal{A}_\varphi$. Since φ is a positive linear functional, we obtain an inner product on $\mathcal{A}_\varphi$, given by

$$\big\langle [A] | [B] \big\rangle_\varphi \coloneqq \varphi(A^\dagger B).$$

By construction this is well-defined, since we have removed all the degenerate elements by taking quotient with the Gelfand ideal. The algebraic product on $\mathcal{A}$ naturally gives us an action of $\mathcal{A}$ on $\mathcal{A}_\varphi$, given by

$$\begin{aligned}\tilde{\pi}_\varphi(A) : \mathcal{A}_\varphi &\to \mathcal{A}_\varphi \\ [B] &\mapsto [AB].\end{aligned}$$

Since $\|\pi_\varphi(A)\|^2 = |\varphi(A^\dagger A)| \leqq \|A\|^2$, it is a bounded operator on $\mathcal{A}_\varphi$. By taking the completion of $\mathcal{A}_\varphi$ with respect to the inner product $\langle\cdot|\cdot\rangle_\varphi$, we obtain the Hilbert space $\mathcal{H}_\varphi$. By construction $\mathcal{A}_\varphi$ is dense in $\mathcal{H}_\varphi$ and hence by continuity the mapping $\tilde{\pi}_\varphi(A)$ gives us bounded linear map $\pi_\varphi(A) : \mathcal{H}_\varphi \to \mathcal{H}_\varphi$. Hence we obtain a norm-continuous representation of $\mathcal{A}$ on the Hilbert space $\mathcal{H}_\varphi$

$$\pi_\varphi : \mathcal{A} \to \mathcal{B}(\mathcal{H}_\varphi).$$

By letting $\Omega_\varphi \equiv [I]$ for the unit element I or using an approximate identity we have,

$$\varphi(A) = \big\langle \Omega_\varphi | \pi_\varphi(A)\Omega_\varphi \big\rangle_\varphi.$$

By construction $\pi_\varphi(\mathcal{A})\Omega_\varphi$ is dense in $\mathcal{H}_\varphi$, since $\mathcal{H}_\varphi$ is completion of $\mathcal{A}_\varphi$ whose elements are of the form $[A] = \pi_\varphi(A)[I] \equiv \pi_\varphi(A)\Omega_\varphi$. Hence Ω_φ is a cyclic vector for $\mathcal{A}$.

If there is any other cyclic representation of $\mathcal{A}$ on $\mathcal{H}$ with a cyclic vector ν, then let $U_{\varphi,\nu}\Omega_\varphi = \nu$ then $U_{\varphi,\nu}$ is an invertible isometry, and is hence a unitary operator. It follows that $\pi_\varphi \circ U_{\varphi,\nu} = U_{\varphi,\nu} \circ \pi_\varphi$. Hence all cyclic representations are unitarily equivalent. □

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

2 | Symmetries & The Spectrum Condition

The formulation of quantum theory takes into account only the mathematical structures expected of measuring instruments and preparation procedures. We did not need any additional pre-theories in the formulation of algebraic quantum theory. The generality of algebraic quantum theory makes it applicable to very general systems, this however keeps its predictivity low. To increase its predictivity about the physical world, we must impose additional constraints coming from the ideas physicists have come up with about how the physical world works. Since classical theories already possess a high level of predictive power for various phenomena, it is natural to attempt to reuse these classical ideas to constrain quantum theories. The starting point of classical theories is the description of a collection of transformations that keep the system invariant. This collection of transformations acts as the symmetry group of the system. These symmetry groups constrain the structure of classical theories, making them more predictive. The state-space of classical theories are modeled using Borel measures on a manifold and the group of isometries of the manifold correspond to the symmetries of such systems. Algebraic quantum field theory (AQFT) attempts to describe these classical constraints on the objects of quantum theory consisting of von Neumann algebras and normal states. In this chapter, we express the ideas of (special) relativity in terms of the objects of algebraic quantum theory.

In this chapter, we focus on Minkowski space-time, and the group of symmetries corresponds to the proper orthochronous Poincaré group. The main objective is to prove the Reeh-Schlieder theorem. Much of the discussion here is based on the works of Borchers.

2.1 | Axioms of Local Quantum Physics

Such a description of relativistic ideas in the quantum framework is made possible by the assumption that every measurement is performed in a finite region of space, and in a finite time interval. This approach is due to Haag-Kastler [18] and Araki [2], and is called local quantum physics or AQFT. Since the measuring instruments are modeled in terms of von Neumann algebras, we can assign to each region of space-time a collection of effects corresponding to measurements in that region. This gives us an association,

$$\mathcal{O} \mapsto \mathcal{A}(\mathcal{O})$$

of von Neumann algebras to each region of space-time, which we will assume to be the Minkowski space. This association is called a net of observable algebras.

If an observable can be measured in a space-time region $\mathcal{O}_\alpha$, then it can be measured in any space-time region $\mathcal{O}_\beta$ containing $\mathcal{O}_\alpha$. If $\mathcal{O}_\alpha \subset \mathcal{O}_\beta$ we have an associated co-restriction map, such that, for every observable $A \in \mathcal{A}(\mathcal{O}_\alpha)$ we have the co-restriction $A|^{\mathcal{O}_\beta} \in \mathcal{A}(\mathcal{O}_\beta)$.

The notation $|^{\mathcal{O}_\beta}$ is the co-restriction to $\mathcal{O}_\beta$. The map, $A \mapsto A|^{\mathcal{O}_\beta}$ is a map from $\mathcal{A}(\mathcal{O}_\alpha)$ to $\mathcal{A}(\mathcal{O}_\beta)$. If $\mathcal{O}_\alpha \subset \mathcal{O}_\beta \subset \mathcal{O}_\gamma$ are nested open sets then the co-restriction is transitive.

$$\left(A|^{\mathcal{O}_\beta}\right)\Big|^{\mathcal{O}_\gamma} = A|^{\mathcal{O}_\gamma}.$$

If **vN** is the collection of all von Neumann algebras, related to each other by normal maps, the assignment of $\mathcal{A}(\mathcal{O})$ to the open set $\mathcal{O}$ is a mapping (functor),

$$\mathcal{A} : \mathcal{O}(\mathcal{M}) \to \mathbf{vN},$$

where $\mathcal{O}(\mathcal{M})$ is the collection of all open sets of $\mathcal{M}$ which get mapped to von Neumann algebras, and the inclusions maps $\mathcal{O}_\alpha \subset \mathcal{O}_\beta$ get mapped to normal injective maps of von Neumann algebras. To each such inclusion in $\mathcal{O}(\mathcal{M})$ we get a restriction morphism in **vN**, $\{\mathcal{O}_\alpha \subset \mathcal{O}_\beta\} \mapsto \{\mathcal{A}(\mathcal{O}_\alpha) \to \mathcal{A}(\mathcal{O}_\beta)\}$, the co-restriction maps corresponds to subalgebra inclusions.

Postulate. (Monotone Property) *If $\mathcal{O}_\alpha \subset \mathcal{O}_\beta$ then $\mathcal{A}(\mathcal{O}_\alpha) \subset \mathcal{A}(\mathcal{O}_\beta)$.*

The self-adjoint elements of the von Neumann $\mathcal{A}(\mathcal{O})$ correspond to the observables which can be measured in the region $\mathcal{O}$. This collection of von Neumann algebras is a partially ordered collection with the order given by von Neumann subalgebra inclusions. Naturally, every chain is bounded by the von Neumann algebra of all observables. This maximal von Neumann algebra is called the quasi-local algebra, denoted by $\mathcal{A}$.

Suppose we have a Hilbert space $\mathcal{H}$ on which each $\mathcal{A}(\mathcal{O})$ can be represented then for any open cover $\{\mathcal{O}_\alpha\}$ of the Minkowski space, we must have

$$\Big[\bigcup_\alpha \mathcal{A}(\mathcal{O}_\alpha)\Big]'' = \mathcal{A}.$$

The von Neumann subalgebras $\mathcal{A}(\mathcal{O})$ of the quasi-local algebra $\mathcal{A}$ are called local algebras. The quasi-local algebra $\mathcal{A}$ is the starting point of AQFT.

Einstein-Weyl Causality

Pre-relativity physics is a collection of physical theories which treat the notion of time as the primary concept. The mathematical language used to articulate the physical ideas was in terms of a manifold called space-time, and time is described geometrically as the time coordinate of space-time. Maxwell's theory of electrodynamics expected special properties out of the notion of light, and these expected properties were also experimentally verified. This resulted in a revisiting of the foundations of the notion of time. As a result, time lost its special status. Relativity on the other hand is a theory that treats the notion of light as the primary concept. Relativity continues to use the geometric notions of space-time and differential geometry to articulate the notion of light and its properties. Relativity then becomes the study of manifolds with constraints imposed by light. The properties of light correspond to the geometric properties of the space-time. In the case of special relativity, the speed of light is not influenced by anything and hence can be described by a flat manifold space-time.

The starting point for the reconsideration of the foundations of space-time is that the Galilean transformations of space-time change the speed of light and hence cannot be treated as symmetries of space-time. The Galilean group needs to be replaced with a suitable group,

and this corresponds to the proper orthochronous Poincaré group $\mathcal{P}_+^\dagger$, consisting of all transformations preserving the properties of light and keeping the direction of time fixed.

If an observable in a region $\mathcal{O}$ is transformed by a proper orthochronous Poincaré transformation g we obtain an observable in the region $g\mathcal{O}$. So we should expect each such causal automorphism g an algebra automorphism α_g such that,

$$\alpha_g \mathcal{A}(\mathcal{O}) \subset \mathcal{A}(g\mathcal{O}).$$

But we expect the same for g^{-1}. So we have the reverse inclusion as well.

$$\alpha_{g^{-1}} \mathcal{A}(g\mathcal{O}) \subset \mathcal{A}(g^{-1}g\mathcal{O}) = \mathcal{A}(\mathcal{O}).$$

This amounts to the physical idea that one can repeat experiments at different places and at different times.

Postulate. (Covariance) *$\forall g \in \mathcal{P}_+^\dagger$ we have $\alpha_g \mathcal{A}(\mathcal{O}) = \mathcal{A}(g\mathcal{O})$.*

This gives us the action of the Poincaré group on the algebra of observables. This however does not capture the Einstein-Weyl causality completely, which requires that no interaction can be faster than the speed of light. Geometrically, the Einstein-Weyl causality says that events in space-like separated regions cannot influence each other. From the point of view of local quantum physics the Einstein-Weyl causality corresponds to the requirement that effects belonging to space-like separated regions should be independent, that is, the effects in these regions should commute with each other.

If two regions $\mathcal{U}$ and $\mathcal{V}$ are spacelike separated means that quantum events in the region $\mathcal{U}$, that is, updating the state with operators belonging to $\mathcal{A}(\mathcal{U})$ and quantum events in the region $\mathcal{V}$ do not interfere with each other. The measurement values can be known jointly. So, the operators in $\mathcal{A}(\mathcal{U})$ must commute with operators in $\mathcal{A}(\mathcal{V})$. Following Borchers, we will use the notation $\rangle\langle$ for space-like separation.

Postulate. (Locality) *If $\mathcal{U}, \mathcal{V} \subseteq \mathcal{M}$ with $\mathcal{U} \rangle\langle \mathcal{V}$ then $\big[\mathcal{A}(\mathcal{U}), \mathcal{A}(\mathcal{V})\big] = 0$.*

The locality postulate hints at the existence of a relation between the geometric properties of the underlying space-time and the sub-algebra relations of the quasi-local algebra. Our goal is to investigate this relation between sub-algebra relations and geometry, focusing on flat space-time.

2.2 | The Space-Time Translation Group

From the classical point of view, the symmetries of the system correspond to the Poincaré group of transformations of space-time. The abelian group of space-time translations is of special interest to us. The dual group of the translation group corresponds to the classical notion of energy-momentum in terms of which laws of classical theories are formulated. To implement classical ideas on local quantum physics it is necessary to understand the translation group and its dual group.

2.2.1 | Analyticity of Fourier Transforms

The space of functions on a group inherits two product structures, one from the group structure on the underlying space, the other from the product structure on the target space, in

our case this corresponds to the translations on $\mathbb{R}^n$, and the product of complex numbers. Fourier analysis relates these two product structures. The Paley-Wiener theorems relate the support properties of a function or a distribution with the analyticity properties of its Fourier transform.

We now prove a generalization of the Paley-Wiener theorem for closed convex sets following Hörmander [22], and formulate it in the language of supporting functions of the corresponding closed convex subset.

2.2.1.1 | Paley-Wiener-Hörmander Theorem

Let $\mathcal{S}(\mathbb{R}^n)$ be the space of all Schwarz functions on $\mathbb{R}^n$. Let f be a lower semi-continuous proper convex function on $\mathcal{X}$, then we define the set $\mathcal{S}_f(\mathbb{R}^n)$ to be the set of all Schwartz functions φ such that for every multi-indices α and β there exists a constant $\lambda^f_{\alpha,\beta}$ such that

$$\sup_{x\in\mathbb{R}^n}\Big[e^{f(x)}\Big|x^\alpha\partial^\beta\varphi(x)\Big|\Big]\equiv\lambda^f_{\alpha,\beta}<\infty$$

For any multi-index β with $|x|^{n+1}<|x^\beta|$ outside the closed unit cube I centered at the origin in $\mathbb{R}^n$, we must have, $\big[|x|^{n+1}|\varphi(x)|\big]\,e^{f(x)}<\big[|x^\beta||\varphi(x)|\big]\,e^{f(x)}\leqq\lambda^f_{0,\beta}$. Equivalently, we must have

$$e^{f(x)}\big[|\varphi(x)|\big]<\lambda^f_{0,\beta}/|x|^{n+1}.$$

The choice of such a β ensures integrability since $1/|x|^{n+1}$ is an integrable function on $\mathbb{R}^n\backslash I$. Combining with the Fenchel-Young inequality, $\langle x|\chi\rangle-f^*(\chi)\leqq f(x)$, we obtain,

$$e^{\langle x|\chi\rangle}e^{-f^*(\chi)}\big[|\varphi(x)|\big]<\lambda^f_{0,\beta}/|x|^{n+1}.$$

whenever χ belongs to the domain $\mathcal{D}_{f^*}$. We can write $\mathbb{R}^n$ as the disjoint union of I and I^c and by definition, the Fourier-Laplace transform of φ is given by

$$\mathcal{L}\varphi(\kappa+i\chi)=\Big[\int_I+\int_{I^c}\Big]\Big[e^{-i\langle x|\kappa+i\chi\rangle}\varphi(x)\Big]d\lambda(x)$$

Since I has unit measure, and since the function $e^{i\langle x|\kappa+i\chi\rangle}\varphi(x)$ is continuous the integral of $\big[e^{i\langle x|\kappa+i\chi\rangle}\varphi(x)\big]$ on the closed unit cube I must be finite. For the integral of $\big[e^{i\langle x|\kappa+i\chi\rangle}\varphi(x)\big]$ on I^c we have the bound

$$\int_{I^c}\Big|e^{-i\langle x|\kappa+i\chi\rangle}\varphi(x)\Big|d\lambda(x)\leqq\lambda^f_{0,\beta}\big|e^{(f^*)(\chi)}\big|\int_{I^c}\underbrace{\big|e^{i\langle x|\kappa\rangle}\big|}_{=1}\big[1/|x|^{n+1}\big]\,d\lambda(x).$$

Since $\int_{I^c}\big[1/|x|^{n+1}\big]\,d\lambda(x)$ is finite, it follows that the above integral must be finite whenever χ belongs to the domain $\mathcal{D}_{f^*}$. Hence $\widehat{\varphi}(\kappa+i\chi)$ is well-defined whenever χ is in $\mathcal{D}_{f^*}$. By differentiating under the integral sign we observe that

$$\overline{\partial}_i\big(\widehat{\varphi}\big)=\int_{\mathbb{R}^n}\underbrace{\overline{\partial}_i\Big[e^{-i\langle x|\chi+i\kappa\rangle}\Big]}_{\equiv0}\varphi(x)d\lambda(x).$$

where the underbraced term is zero by the Cauchy-Riemann equations for the exponential map. Hence the Fourier-Laplace transform of φ is analytic on the interior of $\mathcal{D}_{f^*}$. We shall ignore the uninteresting case when the interior of $\mathcal{D}_{f^*}$ is empty. Since $(i\partial_\kappa)^\alpha(\kappa+i\chi)\widehat{\varphi}(\kappa+i\chi)$ is the Fourier transform of $x^\alpha i^{|\beta|}\partial^\beta\varphi$, we must have

$$\sup_\chi\big|e^{-f^*(\chi)}(\kappa+i\chi)^\beta\partial^\alpha_\kappa\widehat{\varphi}(\kappa+i\chi)\big|<\infty$$

Since f^* is lower semi-continuous, $-f^*$ must be upper semi-continuous, and hence the map

$$\chi \mapsto \left|e^{-f^*(\chi)}(\kappa+i\chi)^\beta \partial_\kappa^\alpha \widehat{\varphi}(\kappa+i\chi)\right|$$

must be upper semi-continuous, and bounded.

On the other hand, if $\widehat{\varphi}$ is an analytic function on the tube $\mathbb{R}^n + i\big(\mathrm{int}\big(\mathcal{D}_{f^*}\big)\big)$, then by analyticity it must be infinitely differentiable, and the product with any polynomial must also be analytic. Hence we can take the Fourier inverse of the function, as a Schwartz function. When $\chi \in \mathrm{int}\big(\mathcal{D}_{f^*}\big)$ it follows that

$$\varphi_\chi(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\langle x|\kappa+i\chi\rangle}\widehat{\varphi}(\kappa+i\chi)d\lambda(\kappa)$$

is defined and since the Fourier inverse is an isomorphism of the Schwartz space, it follows that $e^{\langle\cdot|\chi\rangle}\varphi_\chi$ is in $\mathcal{S}(\mathbb{R}^n)$. If f is such that $\left|e^{-f^*(\chi)}(\kappa+i\chi)^\beta\partial_\kappa^\alpha\widehat{\varphi}(\kappa+i\chi)\right|$ is upper semi-continuous, and bounded then

$$e^{\langle x|\chi\rangle - f^*(\chi)}|x^\beta\partial^\alpha\varphi_\chi| = e^{-f^*(\chi)}|x^\beta\partial^\alpha e^{\langle x|\chi\rangle}\varphi_\chi| < \infty.$$

By the Cauchy integral formula, it follows that the function

$$\int_\Delta e^{i\langle x|\chi+\tau(\kappa_1-\kappa_2)+i\kappa_j\rangle}\widehat{\varphi}(\chi+\tau(\kappa_1-\kappa_2)+i\kappa_j)d\tau$$

has the same value for both k_1 and k_2, where the integral is taken over a polydisc Δ. Hence it follows that the functions φ_κ correspond to the same function, say φ, and it must be a Schwartz function. By the conjugate duality theorem there must exist some χ_φ such that $f(x)+f^*(\chi_\varphi) = \langle x|\chi_\varphi\rangle$, and hence we must have,

$$\sup_{x\in\mathbb{R}^n}\Big[e^{f(x)}\Big|x^\alpha\partial^\beta\varphi(x)\Big|\Big] \equiv \lambda^f_{\alpha,\beta} < \infty.$$

Hence if $\widehat{\varphi}$ is analytic on $\mathbb{R}^n + i\big(\mathrm{int}(\mathcal{D}_{f^*})\big)$, then there exists a Schwartz function in $\mathcal{S}_f(\mathbb{R}^n)$ such that $\widehat{\varphi}$ is its Fourier-Laplace transform.

The conditions on $\mathcal{S}_f(\mathbb{R}^n)$ can be transferred over to the tempered distributions via the pairing between Schwartz functions and tempered distributions. Let τ be a distribution on $\mathbb{R}^n$ such that for every multi-indices α and β there exists some constant $\lambda^f_{\alpha,\beta}$ with

$$\big|\big\langle\varphi\big|\tau\big\rangle\big| \leqq \lambda^f_{\alpha,\beta}\sup_{x\in\mathbb{R}^n}\Big[e^{-f(x)}\big|x^\beta\partial^\alpha\varphi(x)\big|\Big].$$

Let $\tau_\chi = e^{\langle\cdot|\chi\rangle}\tau$, since $\langle x|\chi\rangle \leqq f(x)+f^*(\chi)$, we must have,

$$\begin{aligned}
\langle\varphi|\tau_\chi\rangle &\leqq \lambda^f_{\alpha,\beta}\sup_{x\in\mathbb{R}^n}\Big[e^{-f(x)}x^\beta\partial^\alpha\big[e^{\langle x|\chi\rangle}\varphi(x)\big]\Big]\\
&\leqq \lambda^f_{\alpha,\beta}\sup_{x\in\mathbb{R}^n}\Big[e^{-f(x)}x^\alpha e^{\langle x|\chi\rangle}\big[\chi^\alpha\varphi(x)+\cdots+\partial^\alpha\varphi(x)\big]\Big]\\
&\leqq \lambda'\sup_{x\in\mathbb{R}^n}\Big[e^{f^*(\chi)}\big|x^\beta\partial^\alpha\varphi(x)\big|\Big]/(1+|\chi|)^{|\alpha|+|\beta|}.
\end{aligned}$$

Hence τ_χ must be a tempered distribution whenever $f^*(\chi)$ is finite. The space of all such distributions is denoted by $\mathcal{S}'_f(\mathbb{R}^n)$, in which case, we can make sense of the Fourier transform.

The character χ maybe expressed as a linear combination $\sum_i \lambda_i \chi_i$ of affine independent characters χ_i. The function,

$$\psi_i(x) = e^{\langle x|\chi\rangle} / \big[\sum_i e^{\langle x|\chi_i\rangle} \big]$$

is in a bounded subset of $\mathcal{S}(\mathbb{R}^n)$. Since $\tau e^{\langle\cdot|\chi\rangle} = \psi_i \sum_i \tau_{\chi_i} e^{\langle\cdot|\chi_i\rangle}$, it follows that $\widehat{\tau}_\chi$ is smooth function with

$$\widehat{\tau}_\kappa(\kappa) = \sum_i \big\langle \psi_\chi e^{-i\langle\cdot|\kappa\rangle} | \tau e^{\langle\cdot|\chi_i\rangle} \big\rangle.$$

The function $\kappa + i\chi \mapsto \psi_\chi e^{\langle\cdot|\kappa\rangle}$ is an $\mathcal{S}(\mathbb{R}^n)$ valued analytic function, it follows that $\widehat{\tau}_\chi(\kappa)$ is an analytic function whenever χ is in the interior of the domain $\mathcal{D}_{f^*}$. We denote the interior of $\mathcal{D}_{f^*}$ by $\mathcal{M}_f$.

Theorem 2.2.1. ([22]) *Let $f \in \overline{\Gamma}(\mathbb{R}^n)$. Then the Fourier-Laplace transform maps $\mathcal{S}'_f(\mathbb{R}^n)$ to the space of analytic functions on $\mathbb{R}^n + i\mathcal{M}_f$, such that,*

i) For every τ in $\mathcal{S}'_f$ and χ in a compact subset K of $\mathcal{M}_f$, there exist constants λ and N such that

$$\widehat{\tau}(\kappa + i\chi) \leqq \lambda_K (1 + |\chi|)^N.$$

ii) The map, $\chi \mapsto \widehat{\tau}_\chi = \widehat{\tau}(\cdot + i\chi)$ extends to a continuous function on $\mathcal{D}_{f^*}$.

iii) $e^{-f^*(\chi)} \widehat{\tau}_\chi (1 + |\chi|)^N$ is bounded in $\mathcal{S}'(\mathbb{R}^n)$.

Conversely, for any analytic function h satisfying the above conditions, $e^{\langle\cdot|\chi\rangle} h(\cdot + i\chi)$ is independent of χ and

$$e^{\langle\cdot|\chi\rangle - f^*(\chi)} h / (1 + |\chi|)^N$$

is a bounded tempered distribution, whenever χ belongs to $\mathcal{M}_f$.

Proof

It remains to prove that $e^{\langle\cdot|\chi\rangle} h(\cdot + i\chi)$ is independent of χ. This follows by the Cauchy-Riemann equation for analytic functions, which implies that the derivative with respect to χ is zero. □

2.2.2 | The Future Cone & Positivity of Energy

The Einstein-Weyl causality requires the speed of light to be special in the sense that no signal can travel faster than it. In the framework of local quantum physics, this corresponds to requiring measuring instruments in two different regions to be related to each other only if they can be reached by light. This can be described in terms of the translation group. Since the translation group is abelian, we will denote a translation of a region $\mathcal{O}$ by (t, x) by $\mathcal{O} + (t, x)$. For a given region $\mathcal{O}$, the region that can be reached by light can be described in terms of the action of the translation group as

$$\mathcal{O} + F_+ \equiv \bigcup_{(x,t)\in F_+} \big(\mathcal{O} + (t, x)\big)$$

where the closed convex cone $F_+ = \big\{(t, x) \,|\, ct \geq |x|\big\} \subset \mathbb{R}^n$ is called the forward light-cone. One can study convex sets by convex functions, and in particular a closed convex set can be studied in terms of its supporting function.

We can study the forward cone F_+ by its supporting function. For convenience, we will write the time coordinate separately. The supporting function of the forward cone is

$$H_{F_+}(\tau,\chi) = \sup_{F_+}\Big[t\tau + \langle x|\chi\rangle\Big] = \sup_{t\in\mathbb{R}_+}\Big[t\big(\tau + c|\chi|\big)\Big].$$

As a supremum of affine linear maps, the supporting function is a convex function. The characters (τ,χ) of the translation group correspond to energy-momentum. The support of H_{F_+} is a closed convex set consisting of the vectors (τ,χ) for which $H_{F_+}(\tau,\chi)$ is finite, and must be such that $\tau \leqq -c|\chi|$. We will denote this cone by E_+.

2.3 | The Spectrum Condition & Analyticity

Any laboratory located in a finite region can only contain finitely many massive particles and possibly infinitely many massless particles with finite energy. The state with no particles, known as the vacuum, corresponds to the lowest energy situation. Hence we need to make sense of representations of the translation group such that the energy of the system is positive. Such situations are called finite particle representations. The spectrum condition is an essential condition for quantum field theories and implements the classical ideas about energy-momentum in quantum theory, consisting of the quasi-local algebra and normal states on it.

2.3.1 | Representation of the Translation Group

On the space-time, there is a partial order coming from causality. That is, $x \leq y$ if x can influence y. The set of all one-to-one transformations f of the space-time such that both f and f^{-1} preserve the causal order is a group and is called causality group. It was shown by [50] that the causal group is generated by the orthochronous Lorentz transformation, translations of the space-time, and the dilations of the space-time. This implies that the concept of causality is intricately linked to the metric of Minkowski space-time. The group of all linear transformations preserving the Minkowski metric, and the direction of time is called the orthochronous Poincaré group.

The Lorentz group itself remains invariant under conjugation with any space-time translation and hence is a normal subgroup of the Poincaré group. Hence the Poincaré group can be decomposed as a semi-direct product,

$$\mathcal{P} \equiv \mathcal{L} \rtimes_\alpha \mathcal{T}.$$

Physically meaningful symmetric transformations must preserve the predual, and must correspond to inner automorphisms. Hence the map α_g must be an inner automorphism of the von Neumann algebra for each g in $\mathcal{P}$. A group of automorphisms of $\mathcal{A}$ is called σ-weakly continuous if for every A, and for every normal linear functional $\omega \in \mathcal{A}_*$, the map,

$$g \mapsto \omega\big(\alpha_g(A)\big)$$

is continuous.

Hilbert spaces come equipped with a lot of useful structures and the analysis of these structures is simpler compared to the much more abstract structures on an abstract von Neumann algebras. It is convenient to find a physically meaningful representation and utilise the nicer structures of Hilbert spaces for the study of quasi-local algebra. If we have a physically meaningful state on the quasi-local algebra, then such a representation is always

possible. Recall that given any state on a C^*-algebra, one can construct a representation of the C^*-algebra on a Hilbert space. If the state is also faithful, then the representation is faithful.

Let ω be a faithful normal state on $\mathcal{A}$ and $\mathcal{H}$ is the corresponding GNS Hilbert space. The representations of the Poincaré group on the quasi-local algebra can be studied in terms of their representations on Hilbert spaces. Let $\alpha, g \mapsto \alpha_g$ be a σ-weakly continuous group of automorphisms of von Neumann algebra $\mathcal{A}$. The map,

$$g \mapsto \Big\langle \Omega | \pi_\omega\big(\alpha_g(A)\big)\Omega \Big\rangle_\omega .$$

must be continuous whenever ω is a normal state.

Every locally compact group $\mathcal{G}$ comes equipped with a left Haar measure λ, which is invariant under the left action of the group, that is, every region of the group related to each other by a left action by an element of the group has the same left Haar volume. Using the left Haar measure, one obtains the notion of integrability of functions on $\mathcal{G}$. We say a complex-valued measurable function f on $\mathcal{G}$ is integrable if

$$\|f\|_{L^1(\mathcal{G})} = \int_{\mathcal{G}} \big|f(x)\big| d\lambda(x) < \infty$$

The representation theory of the group $\mathcal{G}$ is itself closely related to the representations of the Banach*-algebra $L^1(\mathcal{G})$. The group product and the product structure of complex numbers induce product structures on the space $L^1(\mathcal{G})$ of integrable functions, given by,

$$(f * g)(x) = \int_{\mathcal{G}} f(y) g(y^{-1}x) d\lambda(y).$$

$$(f \cdot g)(x) = f(x)g(x).$$

The complex conjugation introduces an involution operation by

$$f^*(x) = \overline{f(x^{-1})}.$$

Together with these operations $L^1(\mathcal{G})$ is a Banach*-algebra. For the translation group $\mathbb{R}^n$, the group operation is the abelian group operation of addition, and hence it follows that $f * g(x) = g * f(x)$ and $f^*(x) = \overline{f(-x)}$.

Suppose we have a strongly continuous representation $\pi_\mathcal{G}$ of $\mathcal{G}$ as unitary operators on a Hilbert space $\mathcal{H}$. Each function on $\mathcal{G}$ can be thought of as providing a weight for points on $\mathcal{G}$, and hence each function f on $\mathcal{G}$ gives rise to a bounded operator $\pi_\mathcal{G}(f)$ on the Hilbert space $\mathcal{H}$ defined by

$$\pi_\mathcal{G}(f) = \int_{\mathcal{G}} f(x) \pi_\mathcal{G}(x) d\lambda(x).$$

Heuristically, this can be thought of as $f(x)$ weighted combination of unitaries $\pi_\mathcal{G}(x)$.[1] The properties of the Haar measure ensures that, $\pi_\mathcal{G}(f * g) = \pi_\mathcal{G}(f)\pi_\mathcal{G}(g)$, $\pi_\mathcal{G}(f^*) = \pi_\mathcal{G}(f)^\dagger$ and for every x in $\mathcal{G}$, we have,

$$\pi_\mathcal{G}(L_x f) = \pi_\mathcal{G}(x)\pi_\mathcal{G}(f).$$

[1] More precisely, the convergence is in the sense of strong operator topology on $\mathcal{B}(\mathcal{H})$, the integral is to be viewed in terms of the inner product on $\mathcal{H}$ as,

$$\langle \nu | f_\omega \eta \rangle_\omega = \int_{\mathcal{G}} f(x) \langle \nu | \pi_\mathcal{G}(x) \eta \rangle_\omega d\lambda(x).$$

This integral is well-defined since $\pi_\mathcal{G}(x)$ is a unitary operator, which makes $\langle \nu | \pi_\mathcal{G}(x)\eta\rangle_\omega$ bounded, and since f is an integrable function, we can make sense of the integral. This satisfies $\big|\langle \nu | \pi_\mathcal{G}(f)\eta\rangle_\omega\big| \leqq \|f\|_{L^1}\|\nu\|_\omega\|\eta\|_\omega$. Hence $\pi_\mathcal{G}(f)$ defines a bounded linear operator on $\mathcal{H}$, with norm $\|\pi_\mathcal{G}(f)\| \leqq \|f\|_{L^1}$.

Hence we obtain a *-representation of $L^1(\mathcal{G})$ on the Hilbert space $\mathcal{H}$. The operators $\pi_{\mathcal{G}}(x)$ correspond to unitary operators and $\pi_{\mathcal{G}}(f)$ correspond to bounded operators on $\mathcal{H}$.

Every integrable function can be approximated by compactly supported continuous functions and can be approximated by characteristic functions. The characteristic functions correspond to projection operators on $\mathcal{H}$. Similarly, any $\pi_{\mathcal{G}}(x)$ can be described using an approximate identity as a limit

$$\pi_{\mathcal{G}}(x) = \lim_{i\in\mathbb{N}} \pi_{\mathcal{G}}\big(L_x\varphi_i\big),$$

where φ_i is an approximate identity, where support of φ_i shrinks to the identity of $\mathcal{G}$. Hence it follows that the von Neumann algebras generated by $\pi_{\mathcal{G}}(\mathcal{G})$ and $\pi_{\mathcal{G}}(L^1(\mathcal{G}))$ are identical. Hence we can now utilise the theory of C^*-algebras and their representations. The details of the representation theory of locally compact groups can be found in [15].

2.3.1.1 | Stone's Theorem & Analyticity

The translation subgroup $\mathcal{T} \equiv \mathbb{R}^n$ is a locally compact abelian group and hence is especially convenient to study. The unitary representations of the locally compact abelian groups are closely related to Fourier transforms on such groups. $L^1(\mathbb{R}^n)$ is a commutative Banach *-algebra where the Haar measure λ is the Lebesgue measure on $\mathbb{R}^n$. The Fourier transform is given by

$$\mathcal{F} : L^1(\mathbb{R}^n) \to \mathcal{C}_0((\mathbb{R}^n)')$$

such that for every χ in $(\mathbb{R}^n)'$ we have,

$$\mathcal{F}f(\chi) = \int_{\mathbb{R}^n} e^{-i\langle x|\chi\rangle} f(x) d\lambda(x).$$

If we have a strongly continuous unitary representation $\pi_{\mathcal{T}}$ of the group $\mathcal{T}$ on the Hilbert space $\mathcal{H}$, we must also have a representation of $L^1(\mathcal{T})$ as bounded operators on $\mathcal{H}$. We will denote this subalgebra of bounded operators by $L^1(\mathcal{T}, \mathcal{H})$.

In such a case, by the spectral theorem, there exists a unique spectral measure $E_{\mathcal{T}} : \mathcal{B}((\mathbb{R}^n)') \to \mathcal{P}(\mathcal{H})$ such that the operators in $L^1(\mathcal{G}, \mathcal{H})$ are given by,

$$\pi_{\mathcal{T}}(f) = \int_{(\mathbb{R}^n)'} \mathcal{F}f(\chi) dE_{\mathcal{T}}(\chi).$$

Since the representation of $\mathbb{R}^n$ also gives rise to a representation of $L^1(\mathbb{R}^n)$ as bounded operators, we have a representation of Schwartz functions as bounded operators on $\mathcal{H}$.

The unitary element corresponding to x is then given by, $U_{\mathcal{T}}(x) = \lim_{U\to\{1\}} \pi_{\mathcal{G}}(L_x\varphi_U)$, hence corresponds to the representation of translates of the delta-distribution on the Hilbert space $\mathcal{H}$. We obtain,

$$U_{\mathcal{T}}(x) = \int_{(\mathbb{R}^n)'} e^{-i\langle x|\chi\rangle} dE_{\mathcal{T}}(\chi) = e^{-ixA}.$$

where $A = \int_{(\mathbb{R}^n)'} \chi dE_{\mathcal{T}}(\chi)$. Every strongly continuous unitary representation of a locally compact abelian group is of this form. This is called Stone's theorem, and it allows us to describe representations of the translation group in terms of the dual group of translations via Fourier transforms.

Assume that we have a representation of the quasi-local algebra $\mathcal{A}$ on the Hilbert space $\mathcal{H}$. Then the group of weakly continuous automorphisms of the translation group corresponds to a strongly continuous unitary representation,

$$g \mapsto U_{\mathcal{T}}(g)$$

on the Hilbert space $\mathcal{H}$ such that, $\pi_\omega\big(\alpha_g(A)\big) = U_{\mathcal{T}}(g)\pi_\omega(A)U_{\mathcal{T}}(g)^\dagger$. By Stone's theorem there exists a self-adjoint operator A (possibly unbounded) on $\mathcal{H}$ with a spectral decomposition

$$A = \int_{(\mathbb{R}^n)'} \chi dE_{\mathcal{T}}(\chi)$$

such that

$$U_{\mathcal{T}}(x) = \int_{(\mathbb{R}^n)'} e^{-i\langle x|\chi\rangle} dE_{\mathcal{T}}(\chi).$$

Let the forward light-cone be denoted by F_+, and let F_+° be its interior. Let I_{F_+} be the indicator function for the forward light-cone. The support of H_{F_+} is called the dual cone of F_+ and is denoted by E_+. Recall that the indicator function of F_+ and the supporting function of F_+ will be each others' Legendre transforms. By the Fenchel-Young inequality, we must have,

$$\langle (t,x)|(\tau,\chi)\rangle \leqq H_{F_+}(\tau,\chi) + I_{F_+}(t,x).$$

The domain of I_{F_+} is the forward light cone F_+.

A normal state ω is said to be a positive energy state if its support is $E_{\mathcal{T}}(\mathcal{K}_\omega)$ for some compact subset of the positive energy cone. The representation $U_{\mathcal{T}}$ of $\mathbb{R}^n$ is called a representation with spectrum condition if the spectrum of $U_{\mathcal{T}}$ is contained in the forward light-cone. Since $U_{\mathcal{T}}$ is a strongly continuous representation, for each A and B in $\mathcal{A}$ since $A \mapsto BA$ is σ-weakly continuous it follows that the map,

$$x \mapsto \omega(B\alpha_x(A)) = \Big\langle [B^\dagger] | U_{\mathcal{T}}(x)[A] \Big\rangle_\omega$$

is a continuous function on $\mathbb{R}^n$, with domain in F_+. The unitary operator $U_{\mathcal{T}}(x)$ is the Fourier transform of the operator valued Dirac delta distribution at x on $\mathbb{R}^n$ with support in the forward cone F_+.

By 2.2.1, its Fourier-Laplace transform must be an analytic function on $i\mathbb{R}^n + \big(\mathrm{int}\big(\mathcal{D}_{I_{F_+}}\big)\big) = i\mathbb{R}^n + E_+$. Note that the pairing in Stone's theorem is $i\langle x|\chi\rangle$, so, we consider first a map $\chi \mapsto i\chi$ before applying the Paley-Wiener-Hörmander theorem. For more details on the relation between spectrum condition and analyticity, see [6].

2.3.2 | The Reeh-Schlieder Theorem

The Reeh-Schlieder theorem is a very foundational and rigorous result for the analysis of quantum field theories. Many of the other rigorous results of quantum field theories rely on the Reeh-Schlieder theorem. Here we will prove the theorem following a slightly more general approach due to Borchers [7].

We will assume the following setting before stating and proving the theorem and its corollaries. The Reeh-Schlieder theorem shows that the Gelfand vector corresponding to the vacuum representation is cyclic and separating for $\mathcal{A}(\mathcal{O})$ for every open set $\mathcal{O}$ of the Minkowski space. More abstractly we consider a von Neumann algebra $\mathcal{A}$ acting on a Hilbert space $\mathcal{H}$ equipped with a strongly continuous unitary representation of the abelian group $\mathbb{R}^n$

$$x \mapsto U_{\mathcal{T}}(x) \equiv \int_{E_+ \subset (\mathbb{R}^n)'} e^{-i\langle x|\chi\rangle} dE_{\mathcal{T}}(\chi).$$

where E_+ is a proper closed convex cone of $(\mathbb{R}^n)'$. Let $\mathcal{B} \subset \mathcal{A}$ and let $\mathcal{B}(x) = U_{\mathcal{T}}(x)\mathcal{B}U_{\mathcal{T}}(x)^\dagger$. We say that $\mathcal{B}$ is weakly additive for $\mathbb{R}^n$ as a von Neumann subalgebra of $\mathcal{A}$ if

$$\Big[\bigcup_{x\in\mathbb{R}^n} \mathcal{B}(x)\Big]'' = \mathcal{A} \qquad \text{(weak additivity)}$$

For the remainder of this section, we will use the following notation;

$$\Big[\bigcup_{x\in\mathcal{N}}\mathcal{B}(x)\Big]'' = \mathcal{B}(\mathcal{N}).$$

Theorem 2.3.1. (Reeh-Schlieder-Borchers) *Let $x \mapsto U_{\mathcal{T}}(x)$ be as above, and $\mathcal{B} \subset \mathcal{A}$ such that*

$$\Big[\bigcup_{x\in\mathbb{R}^n}\mathcal{B}(x)\Big]'' = \mathcal{A}$$

If $\nu \in \mathcal{H}$ is analytic for $\mathbb{R}^n$, then for any open neighborhood $\mathcal{N} \subseteq \mathbb{R}^n$ of the origin,

$$\overline{\mathcal{B}(\mathcal{N})\nu} = \overline{\mathcal{A}\nu}.$$

Proof

The idea is to use the analyticity to find an appropriate complex neighborhood C of the origin, then use the commutativity and continuity of the action of the group $\mathbb{R}^d$ to show that the function must also be analytic on the tube $C + i\mathbb{R}^n$. Then by Bochner's tube theorem, the function can be extended to an analytic function on the convex hull of $C + i\mathbb{R}^n$. We then use analytic continuation to show density properties.

Let ν be an analytic vector for the abelian group $\mathbb{R}^n$. This means that the Hilbert space valued function $f : \mathbb{R}^n \to \mathcal{H}$ given by

$$f(x) = U_{\mathcal{T}}(x)\nu$$

has an analytic extension to the neighborhood of origin $0 \in C \subset \mathbb{C}^n$. Since $\mathbb{R}^n$ is an abelian group we have, $U_{\mathcal{T}}(x + y) = U_{\mathcal{T}}(y)U_{\mathcal{T}}(x)$ and by the strong continuity of the unitary representation it follows that,

$$f : x \mapsto U_{\mathcal{T}}(x + y)\nu$$

is analytic for all $y \in \mathbb{R}^n$, hence $f(x)$ is analytic on a tube. By Bochner's tube theorem, we can assume that C is convex and $f(x)$ is analytic on the tube $C + \mathbb{R}^n$.

By assumption, the spectrum of $U_{\mathcal{T}}(x)$ lies in a proper convex cone $E_+ \subset (\mathbb{R}^n)' \equiv \mathbb{R}^n$, it follows that $U_{\mathcal{T}}(x)$ for $x \in \mathbb{R}^n$ is the boundary value of an analytic operator-valued function, analytic on the tube

$$T_+ \equiv \Big\{x + iy \,|\, y \in F_+\Big\},$$

where F_+ is the dual cone of E_+. Hence for any finite set F, and a collection of bounded operators B_i on $\mathcal{H}$ with i in F, the product, $(x_i)_{i\in F} \mapsto [\prod_{i\in F} U_{\mathcal{T}}(x_i)B_i]$ is an operator-valued function, analytic on the tube, $\oplus_{i\in F}T_+$.

Consider the subspace $\mathcal{A}\nu$ of $\mathcal{H}$, obtained by the action of $\mathcal{A}$ on the vector ν. By weak additivity, any element $A \in \mathcal{A}$ can be written as a limit of elements of the form

$$A \equiv \prod_{i\in F} B_i(x_i) \in \Big[\bigcup_{x\in\mathbb{R}^n}\mathcal{B}(x)\Big]'' = \mathcal{A},$$

where F is a finite set. The elements B_i act as bounded operators on $\mathcal{H}$. For any η in $\mathcal{H}$ consider the function

$$(x_i)_{i\in F} \mapsto \Big\langle \prod_{i\in F} B_i(x_i)\nu \,|\, \eta\Big\rangle_\omega$$

By rewriting $B_i(x_i) = U_{\mathcal{T}}(x_i)B_iU_{\mathcal{T}}(x_i)^\dagger$, we obtain

$$\prod_{i\in F} B_i(x_i) = \Big[\prod_{i\in F} U_{\mathcal{T}}(x_i)U_{\mathcal{T}}(x_{i-1})^\dagger B_i\Big]U_{\mathcal{T}}(x_n)^\dagger,$$

where $d = |F|$. We know that the function $(y_i)_{i\in F} \mapsto [\prod_{i\in F} U_{\mathcal{T}}(y_i)B_i]$, is an analytic operator valued function on the tube, $y_i \in T_+$. We also know that the function $f(x) = U(x)\nu$ is analytic on the convex open neighborhood $T = C + i\mathbb{R}^n$. Hence by changing the coordinates from x_i to $x_i - x_{i-1}$, the function

$$(x_i)_{i\in F} \mapsto \Big\langle \Big[\prod_{i\in F} U_{\mathcal{T}}(x_i)B_i\Big]U_{\mathcal{T}}(x_n)\nu \,|\, \eta\Big\rangle_\omega.$$

is analytic on the tube, $T_F \equiv [T_+ \oplus \cdots \oplus T_+] \oplus [-T \cup T_+] \subset \mathbb{C}^{n\cdot d}$. If η is a vector such that

$$\Big\langle \prod_{i\in F} B_i(x_i)\nu \,|\, \eta\Big\rangle_\omega \equiv 0$$

on T_F, then by analytic extension, it must be identically zero everywhere. Hence we have,

$$\overline{\Big[\bigcup_{x\in N} \mathcal{B}(x)\Big]''\nu} = \overline{\mathcal{A}\nu}.$$

□

The interpretation of the Reeh-Schlieder theorem can be quite tricky. When we apply the Reeh-Schlieder theorem to the quantum field theories setting we obtain that for any open set $\mathcal{O}$ of the Minkowski space the subset $\mathcal{A}(\mathcal{O})\Omega$ of $\mathcal{H}$ is also dense in $\mathcal{H}$. That is,

$$\overline{\mathcal{A}(\mathcal{O})\Omega} = \overline{\mathcal{A}\Omega} = \mathcal{H},$$

where Ω is the Gelfand vector corresponding to the vacuum state. Here by the vacuum state we mean a translation invariant state on the quasi-local algebra which has the lowest energy.

For the proof take an open set $\mathcal{O}_\epsilon$ such that $\mathcal{O} = \bigcup_{x\in N_\epsilon}\{\mathcal{O}_\epsilon + x\}$ where $N_\epsilon = \{x \mid \|x\| \leqq \epsilon\}$. In such a case the von Neumann subalgebra $\mathcal{A}(\mathcal{O}_\epsilon) \subset \mathcal{A}$ satisfies the assumptions of the Reeh-Schlieder theorem and hence we must have

$$\underbrace{\overline{\Big[\bigcup_{x\in N_\epsilon} \mathcal{A}(\mathcal{O}_\epsilon + x)\Big]''\Omega}}_{\mathcal{A}(\mathcal{O})} = \mathcal{H}.$$

It appears as if the observables in a small region of space-time contain data about all the observables. This is however not the case. The analyticity of the vacuum is a very strong condition on the vacuum. This analyticity already assumes the state is well-behaved under all translations. So, even though $\mathcal{A}(\mathcal{O})$ describes only a part of the total data of the algebra of observables the remaining data is contained in the state ω, this is known as 'entanglement' in quantum foundations and the quantum information literature.

Clearly, quantum field theories satisfy weak additivity for every $\mathcal{A}(\mathcal{O})$,

$$\Big[\bigcup_{x\in\mathbb{R}^n} \mathcal{A}(\mathcal{O} + x)\Big]'' = \mathcal{A}.$$

The weak additivity says that the data in $\mathcal{A}$ is equivalent to the combined data in $\mathcal{A}(\mathcal{O})$ and the abelian group $\mathbb{R}^n$. This means that if we are allowed to translate to any region of space-time then any measuring instrument in the world works similar to how measuring instruments work in a given region. Heuristically, this says that the assumption about the analyticity of Ω with respect to $\mathbb{R}^n$ can already describe the extra data coming from the action of $\mathbb{R}^n$. This extra data is put into the state when it is prepared. The subset $\mathcal{A}(\mathcal{O})\Omega$ already contains the combined data, and the analyticity of the state tells us that $\mathcal{A}(\mathcal{O})\Omega$ is spread out 'correctly' in $\mathcal{H}$.

If the spectrum of $U_{\mathcal{T}}(x)$ is bounded, then every vector ν is analytic for $U_{\mathcal{T}}(x)$. Hence for every ν in $\mathcal{H}$, the projection onto the subset $\overline{\mathcal{B}(\mathcal{N})\nu}$ is an element of $\mathcal{A}'$. Hence it follows that $\mathcal{A}' = \mathcal{B}(\mathcal{N})'$ or equivalently $\mathcal{B}(\mathcal{N}) = \mathcal{A}$.

Corollary 2.3.2. *Let $\mathcal{N} \subsetneq \mathbb{R}^n$, and suppose*

$$\Big[\bigcup_{x\in\mathcal{N}} \mathcal{B}(x)\Big]'' \subsetneq \mathcal{A}$$

then the spectrum of $U_{\mathcal{T}}(x)$ must be unbounded.

Heuristically, the condition of the above corollary says that if there exist observables outside of a given region, then how the regions are related to each other must be non-trivial data, because otherwise, every state would contain such data, which cannot be the case if the algebra of observables in the smaller region are expected to be strictly contained in the quasi-local algebra.

Theorem 2.3.3. *Let $\mathcal{A}$ be a von Neumann algebra on $\mathcal{H}$, and*

$$U_{\mathcal{T}}(x) = \int_{E_+\subset(\mathbb{R}^n)'} e^{-i\langle x|\chi\rangle} dE_{\mathcal{T}}(\chi).$$

be a unitary representation of translations. If Ω is an analytic vector for $U_{\mathcal{T}}$ such that $\overline{\mathcal{B}\Omega} = \overline{\mathcal{A}\Omega}$, for a von Neumann subalgebra $\mathcal{B} \subset \mathcal{A}$, then

$$\Big[\bigcup_{x\in\mathbb{R}^n} \mathcal{B}(x)\Big]'' \equiv \mathcal{A}.$$

Proof

The spectrum of $U_{\mathcal{T}}$ is contained in a closed proper cone. Let $\mathcal{A}$ be any von Neumann subalgebra which is invariant under $U_{\mathcal{T}}$, that is $U_{\mathcal{T}}(x)\mathcal{A}U_{\mathcal{T}}(x)^\dagger = \mathcal{A}$. In such a case, the projection lattice structure of $\mathcal{A}$ must be preserved, and hence $U_{\mathcal{T}}$ acts as an inner automorphism of $\mathcal{A}$. We can hence assume that $U_{\mathcal{T}}(x)$ belongs to $\mathcal{A}$. The von Neumann subalgebra

$$\Big[\bigcup_{x\in\mathbb{R}^n} \mathcal{B}(x)\Big]''$$

is automatically invariant under $U_{\mathcal{T}}(x)$ since we have constructed it to be such a von Neumann algebra by taking conjugations with all $U_{\mathcal{T}}(x)$s. Hence $U_{\mathcal{T}}$ act as automorphisms on the projection lattice of $[\cup_{x\in\mathbb{R}^n}\mathcal{B}(x)]''$ also.

To complete the proof we must show that the projection lattices for the two von Neumann algebras are the same. Since $\overline{\mathcal{B}\Omega} = \overline{\mathcal{A}\Omega}$, it also follows that the projections in $[\cup_{x\in\mathbb{R}^n}\mathcal{B}(x)]''$ also separate the predual. Hence we must have,

$$\Big[\bigcup_{x\in\mathbb{R}^n}\mathcal{B}(x)\Big]'' \equiv \mathcal{A}$$

which proves the theorem. □

2.4 | Tomita-Takesaki Modular Theory

Quantum theory requires the notion of von Neumann algebras, and normal states on them. Relativity on the other hand uses semi-Riemannian geometric notions to formulate its ideas. We however expect quantum theory to be a universal framework capable of modeling and describing all observed phenomena. This makes it important to formulate the geometric ideas of relativity from within the operator algebraic framework of quantum theory. We need a way to represent the geometric ideas of relativity on the operator algebraic framework of quantum theory. Hence it becomes important to find a correspondence between geometric objects, and operator algebraic concepts while being compatible with the physical ideas. The compatibility with physical ideas is an essential requirement. If we remove this requirement, the relation we may obtain between geometric objects and operator algebraic concepts may not be able to articulate the physical ideas. For example, we already have many relations between geometric objects and operator algebras, such as the Gelfand-Naimark theory. This relation however does not respect the physical ideas of relativity, such as causality, locality, etc. as described before. So, although it may be of interest to generalize Gelfand-Naimark theory for commutative C^*-algebras to the non-commutative case from a purely mathematical point of view, it may not be the right path from the point of view of physics.

In this section, we describe the various attempts to recover the infinitesimal geometric data, assuming that the vacuum state on the quasi-local algebra is invariant under such transformations. This turns the study of geometric data into the study of operator algebraic data. In this sense, part of the geometric data can also be described from within the operator algebraic setting using states on the system. This is similar to how in algebraic geometry where one studies geometric data about spaces by studying commutative algebras. The main tool we use is the Tomita-Takesaki modular theory for von Neumann algebras.

Von Neumann algebras have algebraic and topological data. To be able to study a von Neumann algebra by its states we must understand how the state carries both the algebraic and topological data. Faithful states 'see' every element of the von Neumann algebra. Normal states respect the topological data of the von Neumann algebra. So we start with describing faithful normal states.

2.4.1 | σ-Finite von Neumann Algebras

Since we expect the values of every single measurement instrument to be modeled using real numbers we expect the collection of effects corresponding to the single measuring instrument to inherit these properties of real numbers. In this case, the effects correspond to whether a measurement value lies in an interval. Since rational numbers are dense in real numbers, every interval has a rational number. Hence we expect any collection of mutually disjoint intervals to be of at most countable cardinality.

A von Neumann algebra is said to be σ-finite, if every collection of mutually orthogonal projections has at most a countable cardinality. For the reason discussed above, we will

assume all von Neumann algebras of interest to us to be σ-finite.

Suppose $\mathcal{A}$ is a σ-finite von Neumann algebra on a Hilbert space $\mathcal{H}$ with inner product $\langle\cdot|\cdot\rangle$. Suppose $\{\kappa_i\}_{i\in\mathcal{I}} \subseteq \mathcal{H}$ be a maximal family of non-zero vectors such that $P_i \perp P_j$ whenever i and j are different. Here P_i is the projection to the subspace spanned by $\mathcal{A}'\kappa_i$. Since $\mathcal{A}$ is a von Neumann algebra, by von Neumann's density theorem we have $\mathcal{A}'' = \mathcal{A}$, and hence it follows that the projection P_i lies in $\mathcal{A}$. By maximality of $\{\kappa_i\}_{i\in\mathcal{I}}$ we have

$$\sum_{i\in\mathcal{I}} P_i = 1.$$

Hence $\{\kappa_i\}_{i\in\mathcal{I}}$ is cyclic for $\mathcal{A}'$, and equivalently separating for $\mathcal{A}$.[2] This is a mutually orthogonal collection of projection operators, and since $\mathcal{A}$ is σ-finite it follows that $\{P_i\}$ is countable, and hence the collection $\{\kappa_i\}_{\mathcal{I}}$ is countable. We will hence assume $\mathcal{I} \equiv \mathbb{N}$.

Normalise the collection of vectors $\{\kappa_i\}$ such that $\sum_{\mathbb{N}} \|\kappa_i\|^2 = 1$. Using this collection we can construct a faithful normal state by

$$\omega(A) = \sum_{i\in\mathbb{N}} \langle \kappa_i | A\kappa_i \rangle.$$

Hence we have proved the following theorem,

Theorem 2.4.1. (Existence) *If $\mathcal{A}$ is σ-finite then $\mathcal{A}$ has faithful normal states.*

σ-finite von Neumann algebras are convenient to study because they have faithful normal states and one can use the Hilbert space structure to study them. Since this assumption is physically justified as well we will assume all von Neumann algebras appearing in the future to be σ-finite.

Let $\mathcal{A}$ be a von Neumann algebra together with a faithful normal state ω. Since ω is faithful, the Gelfand ideal is trivial. The state ω defines an inner product on the von Neumann algebra $\mathcal{A}$ given by,

$$\big\langle [A]|[B] \big\rangle_\omega \coloneqq \omega(A^\dagger B).$$

where $[A]$ denotes the equivalence class determined by A, in this case, each element of the von Neumann algebra determines an equivalence class uniquely. The Gelfand vector [1] will be denoted by Ω. If $\mathcal{H}$ is the completion with respect to the inner product, then the left multiplication operator uniquely determines a representation π_ω of $\mathcal{A}$ on the Hilbert space $\mathcal{H}$, that is,

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

By normality of the state ω, $\pi_\omega(\mathcal{A})$ is a von Neumann subalgebra of $\mathcal{B}(\mathcal{H})$ and hence

$$\pi_\omega(\mathcal{A})'' = \pi_\omega(\mathcal{A}).$$

Since ω is faithful $\omega(A^\dagger A) = \langle [A]|[A]\rangle_\omega$ is non-zero whenever A is non-zero, and hence it follows that $[A]$ is non-zero for every $A \in \mathcal{A}$. The Gelfand vector Ω is separating for $\pi_\omega(\mathcal{A})$ hence is cyclic for the commutant $\pi_\omega(\mathcal{A})'$.

By construction $\pi_\omega(\mathcal{A})\Omega$ is dense in $\mathcal{H}$, or Ω is cyclic for $\pi_\omega(\mathcal{A})$, or separating for the commutant $\pi_\omega(\mathcal{A})'$. Ω is cyclic and separating for both $\pi_\omega(\mathcal{A})$ and $\pi_\omega(\mathcal{A})'$. Hence we have proved the following lemma,

[2] A collection of vectors $\mathcal{K} \subset \mathcal{H}$ is separating for $\mathcal{A}$ if for every non-zero $A \in \mathcal{A}$ there exists some $\kappa \in \mathcal{K}$ such that $A\kappa$ is nonzero. $\mathcal{K} \subset \mathcal{H}$ is cyclic for $\mathcal{A}$ if $\mathcal{A}\mathcal{K}$ is dense in $\mathcal{H}$. It can be shown that $\mathcal{K}$ is separating for $\mathcal{A}$ if and only if it is cyclic for the commutant $\mathcal{A}'$.

Lemma 2.4.2. *The vector Ω is cyclic and separating for both $\pi_\omega(\mathcal{A})$ and $\pi_\omega(\mathcal{A})'$.*

The data about the von Neumann algebra $\mathcal{A}$, as 'seen' by the state ω is contained in the dense subset $\pi_\omega(\mathcal{A})\Omega$, and similarly, the data about the commutant $\mathcal{A}'$ is contained in $\pi_\omega(\mathcal{A})'\Omega$. Modular theory studies the relation between a von Neumann algebra and its commutant by relating $\pi_\omega(\mathcal{A})\Omega$ and $\pi_\omega(\mathcal{A})'\Omega$.

2.4.2 | The Modular Group

We will assume that $\mathcal{A}$ is a von Neumann subalgebra on $\mathcal{H}$ with a cyclic and separating vector Ω, and denote $A\Omega$ by $[A]$. The approach of modular theory is to study the relation between $\mathcal{A}$ and $\mathcal{A}'$ by studying the relation between the dense subsets $\mathcal{A}\Omega$ and $\mathcal{A}'\Omega$ using the inner product. The idea is to use the inner product $\langle\cdot|\cdot\rangle_\omega$ on $\mathcal{H}$ and exploit the fact that for every closed linear operator T on $\mathcal{H}$ we must have,

$$\big\langle \kappa | T\nu \big\rangle_\omega = \big\langle T^\dagger \kappa | \nu \big\rangle_\omega.$$

for every ν in the domain of T and κ in the domain of $T^\dagger$. Similarly, for a closed anti-linear operator S we must have

$$\big\langle \kappa | S\nu \big\rangle_\omega = \big\langle \nu | S^\dagger \kappa \big\rangle_\omega.$$

If one of the arguments is from $\mathcal{A}\Omega$ and the other is from $\mathcal{A}'\Omega$, and if there exists an operator of this type, we would have related the two dense subsets. The goal is to try to find such a map S on $\mathcal{H}$. Observe that for any $A \in \mathcal{A}$ and $B \in \mathcal{A}'$ we have

$$\big\langle [B^\dagger] | [A] \big\rangle_\omega = \omega(BA) = \omega(AB) = \big\langle [A^\dagger] | [B] \big\rangle_\omega.$$

So, we should expect the map T or S to be related to the involutions $[A] \mapsto [A^\dagger]$ for $A \in \mathcal{A}$ and $[B] \mapsto [B^\dagger]$ for $B \in \mathcal{A}'$. Hence the starting point is to study the two maps,

$$\underline{S}_{\mathcal{A}} : [A] \mapsto [A^\dagger], \quad \underline{F}_{\mathcal{A}} : [B] \mapsto [B^\dagger],$$

for all $A \in \mathcal{A}$ and $B \in \mathcal{A}'$, with domains $\mathcal{A}\Omega$ and $\mathcal{A}'\Omega$ respectively.

Algebraic Properties

By construction and the assumption that ω is faithful and normal, both these sets are dense in $\mathcal{H}$. Hence $\underline{S}_{\mathcal{A}}$ and $\underline{F}_{\mathcal{A}}$ are densely defined anti-linear operators on $\mathcal{H}$. Let $\underline{S}^\dagger_{\mathcal{A}}$ and $\underline{F}^\dagger_{\mathcal{A}}$ be the adjoints of $\underline{S}_{\mathcal{A}}$ and $\underline{F}_{\mathcal{A}}$ respectively. For any $B \in \mathcal{A}'$ and for any $A \in \mathcal{A}$ we have $\langle \underline{F}_{\mathcal{A}}[B] | [A] \rangle_\omega = \langle [B^\dagger] | [A] \rangle_\omega = \langle [A^\dagger] | [B] \rangle_\omega = \langle \underline{S}_{\mathcal{A}}[A] | [B] \rangle_\omega$. By the anti-linearity of $\underline{S}_{\mathcal{A}}$ we have $\big\langle \underline{S}_{\mathcal{A}}[A] | [B] \big\rangle_\omega = \big\langle \underline{S}^\dagger_{\mathcal{A}}[B] | [A] \big\rangle_\omega$. Hence we get

$$\big\langle \underline{F}_{\mathcal{A}}[B] | [A] \big\rangle_\omega = \big\langle \underline{S}^\dagger_{\mathcal{A}}[B] | [A] \big\rangle_\omega.$$

So, whenever $[B]$ is in the domain of $\underline{F}_{\mathcal{A}}$ it is also in the domain of $\underline{S}^\dagger_{\mathcal{A}}$. Hence we have

$$\underline{F}_{\mathcal{A}} \subseteq \underline{S}^\dagger_{\mathcal{A}}.$$

Similarly we have $\underline{S}_{\mathcal{A}} \subseteq \underline{F}^\dagger_{\mathcal{A}}$. In particular, they are densely defined. An operator is closeable if and only if it has a densely defined adjoint. Hence $\underline{S}_{\mathcal{A}}$ and $\underline{F}_{\mathcal{A}}$ are closable with closures $S_{\mathcal{A}}$ and $F_{\mathcal{A}}$ respectively. Since adjoints of densely defined operators are closed we have,

$$\underline{F}_{\mathcal{A}} \subseteq F_{\mathcal{A}} \subseteq \underline{S}^\dagger_{\mathcal{A}},$$

and similarly we have $\underline{S}_{\mathcal{A}} \subseteq S_{\mathcal{A}} \subseteq \underline{F}_{\mathcal{A}}^{\dagger}$.

We can immediately describe some of the algebraic properties of $S_{\mathcal{A}}$. Let $S_{\mathcal{A}}^{\dagger}$ be the adjoint of $S_{\mathcal{A}}$. Since $\underline{S}_{\mathcal{A}} = \underline{S}_{\mathcal{A}}^{-1}$, $S_{\mathcal{A}}$ is injective and we have

$$S_{\mathcal{A}}^2 = 1$$

as an unbounded operator. Similarly, we have $(S_{\mathcal{A}}^{\dagger})^2 = 1$. Both $S_{\mathcal{A}}$ and $S_{\mathcal{A}}^{\dagger}$ are involutions on $\mathcal{H}$ with dense ranges. They possess unique polar decompositions. Let the polar decomposition of $S_{\mathcal{A}}$ be given by

$$S_{\mathcal{A}} = J_{\mathcal{A}} \Delta_{\mathcal{A}}^{\frac{1}{2}}$$

where $\Delta_{\mathcal{A}} = S_{\mathcal{A}}^{\dagger} S_{\mathcal{A}}$ and $J_{\mathcal{A}}$ is an anti-unitary operator on $\mathcal{H}$, that is $J_{\mathcal{A}} J_{\mathcal{A}}^{\dagger} = J_{\mathcal{A}}^{\dagger} J_{\mathcal{A}} = 1$.

Since $S_{\mathcal{A}}$ is an involution it follows that $\Delta_{\mathcal{A}}$ is non-singular, positive, and hence self-adjoint. Using the fact that $S_{\mathcal{A}}^2 = 1$ and $(S_{\mathcal{A}}^{\dagger})^2 = 1$ we have $S_{\mathcal{A}}^{\dagger} S_{\mathcal{A}} S_{\mathcal{A}} S_{\mathcal{A}}^{\dagger} = 1$. Hence we have, $S_{\mathcal{A}} S_{\mathcal{A}}^{\dagger} = \Delta_{\mathcal{A}}^{-1}$, and we get,

$$J_{\mathcal{A}} \Delta_{\mathcal{A}}^{\frac{1}{2}} J_{\mathcal{A}} = \Delta_{\mathcal{A}}^{-\frac{1}{2}}.$$

Now we also have, $\Delta_{\mathcal{A}}^{-1/2} = J_{\mathcal{A}}^2 J_{\mathcal{A}}^{\dagger} \Delta_{\mathcal{A}}^{1/2} J_{\mathcal{A}}$. Since $J_{\mathcal{A}}^2$ is a unitary operator and $J_{\mathcal{A}}^{\dagger} \Delta_{\mathcal{A}}^{1/2} J_{\mathcal{A}}$ is a positive self-adjoint operator. By the uniqueness of polar decomposition, it follows that $J_{\mathcal{A}}^2 = 1$ hence we have

$$J_{\mathcal{A}} = J_{\mathcal{A}}^{\dagger}.$$

Taking the adjoint of the polar decomposition we get,

$$S_{\mathcal{A}}^{\dagger} = \left(J_{\mathcal{A}} \Delta_{\mathcal{A}}^{\frac{1}{2}}\right)^{\dagger} = \Delta_{\mathcal{A}}^{\frac{1}{2}} J_{\mathcal{A}} = J_{\mathcal{A}} \Delta_{\mathcal{A}}^{-\frac{1}{2}}.$$

To study the properties of $\Delta_{\mathcal{A}}$, it is convenient to consider its spectral decomposition. Let $E_{\mathcal{A}} : \mathcal{B}(\mathbb{R}) \to \mathcal{P}(\mathcal{H})$ be the spectral measure of $\Delta_{\mathcal{A}}$, that is,

$$\left\langle \eta \,|\Delta_{\mathcal{A}} \nu \right\rangle_{\omega} = \int_{\mathbb{R}^+} \lambda d\left\langle \eta \,|E_{\mathcal{A}}(\lambda)\nu \right\rangle_{\omega}.$$

for every ν in the domain of $\Delta_{\mathcal{A}}$. Note that it is sufficient to consider the integral over positive numbers because $\Delta_{\mathcal{A}}$ is positive, and hence support of the spectral measure $E_{\mathcal{A}}$ lies in the positive reals. The spectral measure for $\Delta_{\mathcal{A}}^{-1}$ corresponds to $J_{\mathcal{A}} E_{\mathcal{A}} J_{\mathcal{A}}$, and hence for any bounded Borel function f on $\mathbb{C}$, we have,

$$\begin{aligned}
\left\langle f(\Delta_{\mathcal{A}}^{-1})\nu | \nu \right\rangle_{\omega} &= \int_{\mathbb{R}} f(\lambda) d\left\langle J_{\mathcal{A}} E_{\mathcal{A}}(\lambda) J_{\mathcal{A}} \nu \,|\nu \right\rangle_{\omega} \\
&= \int_{\mathbb{R}} f(\lambda) d\left\langle J_{\mathcal{A}} \nu \,|E_{\mathcal{A}}(\lambda) J_{\mathcal{A}} \nu \right\rangle_{\omega} \\
&= \left\langle f(\Delta_{\mathcal{A}}) J_{\mathcal{A}} \nu \,|J_{\mathcal{A}} \nu \right\rangle_{\omega} = \left\langle J_{\mathcal{A}} \nu \,|\tilde{f}(\Delta_{\mathcal{A}}) J_{\mathcal{A}} \nu \right\rangle_{\omega}.
\end{aligned}$$

where $\tilde{f}(x) = \overline{f(x)}$. By the anti-unitarity of $J_{\mathcal{A}}$ we have,

$$\left\langle f(\Delta_{\mathcal{A}}^{-1})\nu \,|\, \nu \right\rangle_{\omega} = \left\langle J_{\mathcal{A}} \tilde{f}(\Delta_{\mathcal{A}}) J_{\mathcal{A}} \nu \,|\, \nu \right\rangle_{\omega}.$$

Hence it follows that, $f(\Delta_{\mathcal{A}}^{-1}) = J_{\mathcal{A}} \tilde{f}(\Delta_{\mathcal{A}}) J_{\mathcal{A}}$. By taking $f(x) = x^{-it}$ we obtain,

$$J_{\mathcal{A}} \Delta_{\mathcal{A}}^{it} = \Delta_{\mathcal{A}}^{it} J_{\mathcal{A}}.$$

$J_{\mathcal{A}}$ is called the modular conjugation and $\Delta_{\mathcal{A}}$ is called the modular operator of ω.

Topological Properties

We now prove that it is indeed possible to understand the properties of the von Neumann algebra $\mathcal{A}$ and its commutant $\mathcal{A}'$ as seen by the state ω, by studying the dense subset $\mathcal{A}\Omega$ and $\mathcal{A}'\Omega$. The difficulty lies in understanding how to go back from the dense subsets $\mathcal{A}\Omega$ to $\mathcal{A}$ and recover the von Neumann algebraic structures. The elements of $\mathcal{A}$ and $\mathcal{A}'$ act as bounded operators on the Hilbert space $\mathcal{H}$. We know that

$$S_{\mathcal{A}}[A] = A^{\dagger}\Omega$$

Hence we can think of the image of $\mathcal{A}\Omega$ under $S_{\mathcal{A}}$ as the image of Ω under the action of the von Neumann algebra $\mathcal{A}$, and similarly for the commutant $\mathcal{A}'$. Hence the operator $S_{\mathcal{A}}$ allows us to go between $\mathcal{A}\Omega$ and $\mathcal{A}$.

On the other hand, for $A \in \mathcal{A}$ and $B \in \mathcal{A}'$, and suppose additionally if we have $[A] \in \mathcal{A}'\Omega$ and $[B] \in \mathcal{A}\Omega$, then we have $B[A] = [BA] = [AB] = A[B]$. The goal of the discussion below is to prove that such vectors are dense in $\mathcal{H}$ and can be described in terms of domains of $S_{\mathcal{A}}$ and $F_{\mathcal{A}}$ respectively. Such vector as continuous actions of either $\mathcal{A}$ or $\mathcal{A}'$

$$\begin{array}{ccc} [B] & & [A] \\ \downarrow & & \downarrow \\ A[B] & = & B[A] \end{array}$$

Suppose ν is in the domain of $F_{\mathcal{A}}$, then consider the map

$$B_{\nu}, [A] \mapsto A\nu$$

If B_{ν} is continuous, then $AB_{\nu}[C] = AC\nu = B_{\nu}[AC] = B_{\nu}A[C]$ for all $C \in \mathcal{A}$. Hence we have $AB_{\nu}\nu = B_{\nu}A\nu$. Since ν is in the domain of $F_{\mathcal{A}}$, it also follows that

$$\langle [C]|B_{\nu}^{\dagger}[A]\rangle_{\omega} = \langle B_{\nu}[C]||[A]|\rangle_{\omega} = \langle C\nu|[A]\rangle_{\omega} = \langle \nu|[C^{\dagger}A]\rangle_{\omega}.$$

Since ν is already in the domain of $F_{\mathcal{A}}$ and using the fact that $\underline{F}_{\mathcal{A}} \subset \underline{S}_{\mathcal{A}}^{\dagger}$ we have $\langle \nu|[C^{\dagger}A]\rangle_{\omega} = \langle \nu|\underline{S}_{\mathcal{A}}[A^{\dagger}C]\rangle_{\omega} = \langle [A^{\dagger}C]|F_{\mathcal{A}}\nu\rangle_{\omega} = \langle [C]|A(F_{\mathcal{A}}\nu)\rangle_{\omega}$. Hence we showed that

$$A(F_{\mathcal{A}}\nu) = B_{\nu}^{\dagger}[A].$$

Hence $B_{\nu} \in \mathcal{A}'$. The domain of $S_{\mathcal{A}}^{\dagger}$ consists of all the vectors ν in $\mathcal{H}$ such that

$$\eta \mapsto \langle \nu|S_{\mathcal{A}}\eta\rangle_{\omega}$$

is continuous for all $\eta \in \mathcal{D}(S_{\mathcal{A}})$. We consider the case when η is of the form $[A]$. Clearly $[A]$ belongs to the domain of $S_{\mathcal{A}}$ by definition of $\underline{S}_{\mathcal{A}}$. Since $\langle \nu|S_{\mathcal{A}}[A]\rangle_{\omega} = \langle A\nu|\Omega\rangle_{\omega}$ it reduces to the case when $[A] \mapsto A\nu$ is continuous.

By the discussion above, we know that there exist operators B_{ν} such that $B_{\nu}A = AB_{\nu}$ for all ν as above. Every vector in the domain of the closure $S_{\mathcal{A}}$ can be written as a sequence of vectors $\{\nu_{\alpha}\}$ trending to ν, and must correspond to an operator B in the commutant $\mathcal{A}'$. The domain of $F_{\mathcal{A}}$, as the closure of the operator $\underline{F}_{\mathcal{A}}$ consists of all vectors ν which are limits of vectors of the form $[B]$ for $B \in \mathcal{A}'$. Since $\mathcal{A}'\Omega$ is dense in $\mathcal{H}$, ν can be approximated by vectors in the domain of $\underline{F}_{\mathcal{A}}$, hence ν is in the domain of its closure $F_{\mathcal{A}}$, and B_{ν} is affiliated to $\mathcal{A}'$. Hence we have $S_{\mathcal{A}}^{\dagger} \subset F_{\mathcal{A}}$. We have proved the following lemma,

Lemma 2.4.3. (Takesaki)

$$S_{\mathcal{A}}^{\dagger} \equiv F_{\mathcal{A}}.$$

The support of the spectral measure $E_{\mathcal{A}}$ of $\Delta_{\mathcal{A}}$ belongs in $\mathbb{R}^{+}$. Since spectral measures resolve the identity, it follows that the projections $E_{\mathcal{A}}((0,t))$ converge strongly to the identity on $\mathcal{H}$, as t tends to infinity. Hence the union of their domains is dense in $\mathcal{H}$.

By Takesaki's lemma above, the domain of $F_{\mathcal{A}}$ is the same as that of $\Delta^{-1/2}$, which is again the same as that of $\Delta_{\mathcal{A}}$ by the positivity of the spectral measure. It follows that $\cup_t \mathcal{D}(E_{\mathcal{A}}(t)) \subseteq \mathcal{D}(S_{\mathcal{A}}) \cap \mathcal{D}(F_{\mathcal{A}})$. It follows that $\mathcal{D}(S_{\mathcal{A}}) \cap \mathcal{D}(F_{\mathcal{A}})$ is dense in $\mathcal{H}$. Hence we have proved the following;

Lemma 2.4.4.

$$\overline{\mathcal{D}\big(\Delta_{\mathcal{A}}^{\frac{1}{2}}\big) \cap \mathcal{D}\big(\Delta_{\mathcal{A}}^{-\frac{1}{2}}\big)} = \mathcal{H}.$$

This dense subset acts as a convenient subset to study the behavior of operators, since it contains the properties both of the von Neumann algebra $\mathcal{A}$ and its commutant $\mathcal{A}'$.

2.4.2.1 | The Modular Relation

We now relate the von Neumann algebra $\mathcal{A}$ and its commutant $\mathcal{A}'$ in terms of the modular relations. We will then solve this relation in terms of the modular operator and the modular conjugation. The goal is to use the Hilbert space $\mathcal{H}$ to relate the von Neumann algebra $\mathcal{A}$ and its commutant $\mathcal{A}'$, and then remove any terms involving the Hilbert space directly. We will follow [23], since they treat both $\mathcal{A}$ and $\mathcal{A}'$ on an equal footing.

For any $A, B, C \in \mathcal{A}$, we have $BCA^{\dagger} = (AC^{\dagger}B^{\dagger})^{\dagger}$. Since we can view the algebra $\mathcal{A}$ in terms of the dense subset $\mathcal{A}\Omega$, viewing C as the vector $[C]$ in $\mathcal{A}\Omega$, we have,

$$BS_{\mathcal{A}}AS_{\mathcal{A}}[C] = S_{\mathcal{A}}AS_{\mathcal{A}}B[C].$$

So, when viewed from the subset $\mathcal{A}\Omega$, the operator $S_{\mathcal{A}}AS_{\mathcal{A}}$ acts as if it is in the commutant of $\mathcal{A}$. We use this heuristic relation as motivation to construct an element of $\mathcal{A}'$ for each element in the von Neumann algebra $\mathcal{A}$.

Sakai-Radon-Nikodym Theorem

We prove a useful technical lemma, due to Sakai, which will provide us with a link between the von Neumann algebra $\mathcal{A}$ and its commutant $\mathcal{A}'$ on the Hilbert space $\mathcal{H}$. If ω is a faithful normal state on $\mathcal{A}$, it contains non-trivial data about every operator in $\mathcal{A}$.

Now, suppose $\varphi \ll \omega$. This by definition means that for every operator $C \in \mathcal{A}$,

$$(\omega - \varphi)\big(C^{\dagger}C\big) \geq 0.$$

Heuristically this means that the state ω contains more data about $\mathcal{A}$ than the data contained in φ about $\mathcal{A}$ or that φ has 'less' data about the von Neumann algebra $\mathcal{A}$ than the data contained in ω. Quantifying 'how much?' usually gives rise to a Radon-Nikodym type relation between φ and ω. We will describe below such a Radon-Nikodym type relation due to Sakai.

Let φ be a state on $\mathcal{A}$. Then, $\widehat{\varphi}(C,B) = \varphi(C^{\dagger}B)$ defines a sesquilinear form. Since φ a positive linear functional we have, $\widehat{\varphi}\big(\mu C - B, \mu C - B\big) \geq 0$. Upon expanding and taking

$$\mu = \widehat{\varphi}(C,B)/\widehat{\varphi}(C,C)$$

we obtain a Cauchy-Schwarz like inequality,

$$\big|\varphi(C^\dagger B)\big| \leq \big|\varphi(C^\dagger C)\big|^{\frac{1}{2}} \big|\varphi(B^\dagger B)\big|^{\frac{1}{2}}.$$

If $\varphi \ll \omega$, we have, $\varphi(C^\dagger C) \leq \omega(C^\dagger C)$. Hence we must have

$$\big|\varphi(C^\dagger B)\big| \leq \big|\omega(C^\dagger C)\big|^{\frac{1}{2}} \big|\omega(B^\dagger B)\big|^{\frac{1}{2}}.$$

We now use the fact that for a von Neumann algebra $\mathcal{A}$, the collection of all normal states $\mathcal{A}_*$ is the predual of $\mathcal{A}$, and hence every continuous linear functional on $\mathcal{A}_*$ can be thought of as an element of $\mathcal{A}$. The idea is to exploit this fact about von Neumann algebras, along with the Hahn-Banach separation theorem to show that φ can be constructed using ω and some operator A. Heuristically, A quantifies the extra data contained in ω.

Theorem 2.4.5. (Sakai) *Let φ be such that*

$$\big|\varphi(C^\dagger B)\big| \leq \big|\omega(C^\dagger C)\big|^{\frac{1}{2}} \big|\omega(B^\dagger B)\big|^{\frac{1}{2}}$$

then for any fixed $\lambda \in \mathbb{R}_+$, there exists some $A \in \mathcal{A}$, with $\|A\| \leq 1$ such that

$$\varphi(X) = \omega\big(\lambda AX + \lambda^{-1} XA\big).$$

Proof

Fix $\lambda \in \mathbb{R}_+$. Construct a parametrized collection of linear functionals on $\mathcal{A}$ given by,

$$\omega_{\lambda,A}, X \mapsto \omega(\lambda AX + \lambda^{-1} XA).$$

for every $A \in \mathcal{A}$. The idea is to show that if φ is not of the type $\omega_{\lambda,A}$ then φ can be separated by a continuous linear functional using the Hahn-Banach separation theorem. Since continuous linear functionals on von Neumann algebras are more constrained, we end up arriving at a contradiction.

The role of λ here is to absorb some of the wildness and to ensure A is well-behaved. Since $\omega_{\lambda,A}$ is constructed out of ω, it clearly has 'less data' about the von Neumann algebra $\mathcal{A}$ than ω itself. If A was the identity, then $\omega_{\lambda,I}$ is only a scaling of ω, or would have the same data once normalized.

By normality of the state ω it follows that whenever A is near to C in $\mathcal{A}$, $\omega_{\lambda,A}(X)$ must be near to $\omega_{\lambda,C}(X)$ as complex numbers. By linearity of ω we also have, $\omega_{\lambda,(A+C)}(X) = \omega_{\lambda,A}(X) + \omega_{\lambda,C}(X)$. So the mapping

$$A \mapsto \omega_{\lambda,A}$$

defines a continuous linear map from $\mathcal{A}$ to $\mathcal{A}_*$. Hence the set $\omega_{\lambda,1} \equiv \{\omega_{\lambda,A} \mid \|A\| \leq 1\}$ is a convex subset of $\mathcal{A}_*$, since it is the image of a closed set it is weakly compact in $\mathcal{A}_*$. We now prove that φ belongs to this set, which would prove the lemma.

Suppose there exists no A such that $\varphi = \omega_{\lambda,A}$. Then by the Hahn-Banach separation theorem, there exists a continuous real linear functional κ on $\mathcal{A}_*$ which separates the compact convex subset ω_λ in $\mathcal{A}_*$ from the point φ in $\mathcal{A}_*$. That is, the interval, $\kappa(\omega_{\lambda,A})$ for $\|A\| \leq 1$ and $\kappa(\varphi)$ are disjoint. Assume without loss of generality that $\kappa(\varphi) > \kappa(\omega_{\lambda,A})$ for every $\|A\| \leq 1$.

Since $\mathcal{A}$ is a von Neumann algebra, $\mathcal{A}_*$ is its predual, and hence κ corresponds to an element H in $\mathcal{A}$ and we must have,

$$\mathrm{Re}(\varphi(H)) > \big|\omega_{\lambda,1}(H)\big| \equiv \sup\Big\{\mathrm{Re}(\omega_{\lambda,A}(H)) \mid \omega_{\lambda,A} \in \omega_{\lambda,1}\Big\}.$$

Consider the polar decomposition, $H = U|H| = |H^\dagger|U$ where U is a unitary operator and both U and $|H|$ are in $\mathcal{A}$. We can now use U and $|H|$ to arrive at a contradiction. We have,

$$\omega_{\lambda,\frac{U}{2}^\dagger}(H) = \tfrac{1}{2}\lambda\omega\big(U^\dagger H\big) + \tfrac{1}{2}\lambda^{-1}\omega\big(HU^\dagger\big) = \tfrac{1}{2}\lambda\omega\big(|H|\big) + \tfrac{1}{2}\lambda^{-1}\omega\big(|H^\dagger|\big).$$

By unitarity of U we have $\|U\|^2 = \|U^\dagger U\| = 1$, and hence $\omega_{\lambda,\frac{U}{2}^\dagger} \in \omega_{\lambda,1}$. Hence we have

$$\mathrm{Re}\Big(\omega_{\lambda,\frac{U}{2}^\dagger}(H)\Big) \leq \big|\omega_\lambda(H)\big| < \mathrm{Re}(\varphi(H))$$

Combining this with the AM-GM inequality we have,

$$\begin{aligned}\mathrm{Re}(\varphi(H)) > |\omega_{\lambda,1}(H)| &\geq \tfrac{1}{2}\big(\lambda\omega\big(|H|\big) + \lambda^{-1}\omega\big(|H^\dagger|\big)\big)\\ &\geq \omega\big(|H|\big)^{\frac{1}{2}}\omega\big(|H^\dagger|\big)^{\frac{1}{2}}.\end{aligned}$$

Since we assumed that $|\varphi(C^\dagger B)| \leq |\omega(C^\dagger C)|^{1/2}|\omega(B^\dagger B)|^{1/2}$, for every $C, B \in \mathcal{A}$, we have,

$$\begin{aligned}\big|\mathrm{Re}(\varphi(H))\big| \leq |\varphi(H)| &= \big|\varphi(U|H|^{\frac{1}{2}}|H|^{\frac{1}{2}})\big|\\ &\leq \big|\omega(|H|)\big|^{\frac{1}{2}}\big|\omega(U|H|U)\big|^{\frac{1}{2}} = \big|\omega(|H|)\big|^{\frac{1}{2}}\big|\omega(|H^\dagger|)\big|^{\frac{1}{2}}.\end{aligned}$$

Which implies that $\omega(|H|)^{1/2}\omega(|H^\dagger|)^{1/2} > \omega(|H|)^{1/2}\omega(|H^\dagger|)^{1/2}$ which is absurd and we arrive at a contradiction. Hence there must exist some A such that $\varphi = \omega_{\lambda,A}$. □

Relation between $\mathcal{A}$ and $\mathcal{A}'$

We now use the Sakai-Radon-Nikodym theorem to relate the von Neumann algebra $\mathcal{A}$ with its commutant $\mathcal{A}'$. By construction, the Hilbert space $\mathcal{H}$ describes all the data about the von Neumann algebra $\mathcal{A}$ contained in the state ω. We can now ask how much of this data remains after the action of $\mathcal{A}'$. Instead of the cyclic and separating vector Ω, we consider the action of $\mathcal{A}$ on elements of $\mathcal{A}'\Omega$ and what we can say about this action using the state ω. So, we are interested in data of the form, $\omega(XB)$ for $X \in \mathcal{A}$ and $B \in \mathcal{A}'$.

To study such data consider the linear functionals of the form

$$\omega_B(X) = \big\langle [X]|[B]\big\rangle_\omega,$$

for $B \in \mathcal{A}'$ and note that $\omega(X) = \omega_I(X)$. Since $B^\dagger B \leq \|B\|^2$, for every $B \in \mathcal{A}'$ and $C \in \mathcal{A}$ we have $B^\dagger C^\dagger CB = C^\dagger B^\dagger BC \leq C^\dagger\|B\|^2 C$ and by the positivity of the state ω we have $\omega(B^\dagger C^\dagger CB) \leq \|B\|^2\omega(C^\dagger C)$. Hence we obtain,

$$\begin{aligned}\omega\big(C^\dagger BC\big) &\leq \big|\omega((B^\dagger C)^\dagger B^\dagger C)\big|^{\frac{1}{2}}\big|\omega(C^\dagger C)\big|^{\frac{1}{2}}\\ &= \big|\omega(C^\dagger BB^\dagger C)\big|^{\frac{1}{2}}\big|\omega(C^\dagger C)\big|^{\frac{1}{2}}\\ &\leq \big|\omega(C^\dagger\|B\|^2 C)\big|^{\frac{1}{2}}\big|\omega(C^\dagger C)\big|^{\frac{1}{2}} = \|B\|\,\omega(C^\dagger C).\end{aligned}$$

Hence we have, $\omega_B(C^\dagger C) \leq \|B\|\omega(C^\dagger C)$ or equivalently $\omega_B \ll \omega$. The assumptions of the theorem of Sakai above are satisfied. by ω_B. For every $\lambda \in \mathbb{R}_+$ there exists some A in $\mathcal{A}$ such that, $\omega_B(X) = \omega_{\lambda,A}(X)$

$$\omega_B(X) = \omega(\lambda XA + \lambda^{-1}AX).$$

We now use this application of Sakai's theorem to obtain a relation between $\mathcal{A}$ and $\mathcal{A}'$.

Theorem 2.4.6. (The Modular Relation) *Let $\lambda \in \mathbb{R}_+$. $\forall B \in \mathcal{A}'$, $\exists A \in \mathcal{A}$ with*

$$\lambda\big\langle \Delta_{\mathcal{A}}^{\frac{1}{2}}\nu | A\Delta_{\mathcal{A}}^{-\frac{1}{2}}\eta\big\rangle_\omega + \lambda^{-1}\big\langle \Delta_{\mathcal{A}}^{-\frac{1}{2}}\nu | A\Delta_{\mathcal{A}}^{\frac{1}{2}}\eta\big\rangle_\omega = \big\langle \nu | J_{\mathcal{A}}BJ_{\mathcal{A}}\eta\big\rangle_\omega,$$

$\forall \eta, \nu \in \mathcal{D}\big(\Delta_{\mathcal{A}}^{-1/2}\big) \cap \mathcal{D}\big(\Delta_{\mathcal{A}}^{1/2}\big)$.

Proof

Using the fact that $\omega(A) = \langle\Omega|[A]\rangle_\omega$ we can express the above relation in terms of the inner product on $\mathcal{H}$, $\langle[B]|[X]\rangle_\omega = \lambda\langle[AX]|\Omega\rangle_\omega + \lambda^{-1}\langle[XA]|\Omega\rangle_\omega$. Hence we obtain a relation between the elements of $\mathcal{A}$ and $\mathcal{A}'$,

$$\big\langle [X]|[B]\big\rangle_\omega = \lambda\big\langle [X]|[A^\dagger]\big\rangle_\omega + \lambda^{-1}\big\langle [X^\dagger]|[A]\big\rangle_\omega,$$

upon rearranging we get, $\langle[X^\dagger]|[A]\rangle_\omega = \lambda\langle[X]|[B] - \lambda[A^\dagger]\rangle_\omega$. Since $A, X \in \mathcal{A}$ we observe that $[A], [X]$ are in the domain of $S_{\mathcal{A}}$. Since the right-hand side is well-defined we conclude that $[A] \in \mathcal{D}(S_{\mathcal{A}}^\dagger)$.

Hence we can express the vector $[B]$ in the set $\mathcal{A}'\Omega$ in terms of a vector $[A]$ in the set $\mathcal{A}\Omega$ as $[B] = (\lambda S_{\mathcal{A}} + \lambda^{-1}F_{\mathcal{A}})[A]$ or equivalently,

$$[B] = \Big[\underbrace{\lambda J_{\mathcal{A}}\Delta_{\mathcal{A}}^{\frac{1}{2}} + \lambda^{-1}J_{\mathcal{A}}\Delta_{\mathcal{A}}^{-\frac{1}{2}}}_{G_\lambda}\Big][A].$$

Let $\nu \in \mathcal{D}(\Delta_{\mathcal{A}}^{1/2}) \cap \mathcal{D}(\Delta_{\mathcal{A}}^{-1/2})$. Set

$$G_1\nu = \eta.$$

If $\eta \in \mathcal{D}(F_{\mathcal{A}})$ then so does $F_{\mathcal{A}}\eta$. Since every vector η in the domain of $F_{\mathcal{A}}$ corresponds to an operator B_η affiliated with $\mathcal{A}'$ such that $B_\eta\Omega = \eta$. Hence there exists a sequence $\{B_i\} \subset \mathcal{A}'$ with $[B_i] \to \eta$. For $\lambda = 1$ we obtain for each $B_i \in \mathcal{A}'$ there is an element $A_i \in \mathcal{A}$ such that

$$G_1[A_i] = [B_i].$$

Since the spectrum of $(\Delta_{\mathcal{A}}^{1/2} + \Delta_{\mathcal{A}}^{-1/2})$ does not contain 0 it follows that G_1 has a bounded inverse. By letting $\nu_i = G_1^{-1}[B_i]$ we have, $[A_i] = G_1^{-1}[B_i] = \nu_i \to \nu$.

So we can view any $\nu, \eta \in \mathcal{D}(F_{\mathcal{A}})$ similar to vectors in $\mathcal{A}\Omega$ and let $A_\nu, A_\eta \in \mathcal{A}$ be such vectors, we have

$$\begin{aligned}\big\langle \eta | S_{\mathcal{A}}AS_{\mathcal{A}}\nu\big\rangle &= \big\langle [A_\eta]|S_{\mathcal{A}}AS_{\mathcal{A}}[A_\nu]\big\rangle \\ &= \big\langle [A_\eta]|[(AA_\nu^\dagger)^\dagger]\big\rangle = \big\langle S_{\mathcal{A}}[A]|S_{\mathcal{A}}^\dagger[A_\nu^\dagger A_\eta]\big\rangle\end{aligned}$$

With a similar manipulation, we obtain

$$\big\langle \eta | F_{\mathcal{A}}AF_{\mathcal{A}}\nu\big\rangle = \big\langle [A_\nu^\dagger A_\eta]|[A^\dagger]\big\rangle.$$

By taking $[X] = [A_\nu^\dagger A_\eta]$ and rearranging we obtain

$$\lambda\big\langle \nu | S_{\mathcal{A}} A S_{\mathcal{A}} \eta \big\rangle_\omega + \lambda^{-1} \big\langle \nu | F_{\mathcal{A}} A F_{\mathcal{A}} \eta \big\rangle_\omega = \big\langle \nu | B \eta \big\rangle_\omega,$$

We now take $J_{\mathcal{A}}\nu$ and $J_{\mathcal{A}}\eta$ instead of ν and η. Substituting $S_{\mathcal{A}} = J_{\mathcal{A}}\Delta_{\mathcal{A}}^{1/2}$ and $F_{\mathcal{A}} = J_{\mathcal{A}}\Delta_{\mathcal{A}}^{-1/2}$, and using the fact, $J_{\mathcal{A}}^2 = 1$, it follows that,

$$\begin{aligned}\big\langle J_{\mathcal{A}}\nu | S_{\mathcal{A}} A S_{\mathcal{A}} J_{\mathcal{A}}\eta \big\rangle_\omega &= \big\langle J_{\mathcal{A}}\nu | J_{\mathcal{A}}\Delta_{\mathcal{A}}^{\frac{1}{2}} A J_{\mathcal{A}}\Delta_{\mathcal{A}}^{\frac{1}{2}} J_{\mathcal{A}}\eta \big\rangle_\omega \\ &= \big\langle \Delta_{\mathcal{A}}^{\frac{1}{2}} J_{\mathcal{A}}^2 \nu | A \Delta_{\mathcal{A}}^{-\frac{1}{2}} J_{\mathcal{A}}^2 \eta \big\rangle_\omega = \big\langle \Delta_{\mathcal{A}}^{\frac{1}{2}} \nu | A \Delta_{\mathcal{A}}^{-\frac{1}{2}} \eta \big\rangle_\omega.\end{aligned}$$

Similarly, we have,

$$\big\langle J_{\mathcal{A}}\nu | F_{\mathcal{A}} A F_{\mathcal{A}} J_{\mathcal{A}}\eta \big\rangle_\omega = \big\langle \Delta_{\mathcal{A}}^{-\frac{1}{2}} \nu | A \Delta_{\mathcal{A}}^{\frac{1}{2}} \eta \big\rangle_\omega.$$

Hence we have,

$$\lambda\big\langle \Delta_{\mathcal{A}}^{\frac{1}{2}} \nu | A \Delta_{\mathcal{A}}^{-\frac{1}{2}} \eta \big\rangle_\omega + \lambda^{-1} \big\langle \Delta_{\mathcal{A}}^{-\frac{1}{2}} \nu | A \Delta_{\mathcal{A}}^{\frac{1}{2}} \eta \big\rangle_\omega = \big\langle \nu | J_{\mathcal{A}} B J_{\mathcal{A}} \eta \big\rangle_\omega.$$

whenever ν and η are in the domain of $\Delta_{\mathcal{A}}^{1/2}$ and $\Delta_{\mathcal{A}}^{-1/2}$. □

2.4.2.2 | Tomita's Theorem

The modular relation gives rise to a relation between $\mathcal{A}$ and $\mathcal{A}'$ for each $\lambda \in \mathbb{R}_+$. We now remove the dependence on λ. The idea is to think of the relation as the Fourier transform of a complex, rapidly decreasing function. The parameter λ can then be viewed in terms of characters. We need to develop certain tools to be able to do this.

Let f be a function bounded on the strip $S \equiv \{-\frac{1}{2} \leq \mathrm{Re}(z) \leq \frac{1}{2}\}$. Then we can define a meromorphic function by,

$$g(z) = \pi f(z)(\sin(\pi z))^{-1}.$$

This is a meromorphic function with a pole at $z = 0$. By the power series expansion of sin, we have $\sin(\pi z) = \pi z - \pi^3 z^3/3! + \cdots$, it follows that it is a simple pole. Hence, the residue is given by,

$$\mathrm{res}_g(0) = \lim_{z\to 0} \pi f(z)\big(\pi - \tfrac{\pi^3 z^2}{3!} + \cdots\big)^{-1} = f(0).$$

By the residue theorem, for any loop $\gamma \subseteq \mathcal{H}(S\backslash E)$, we have

$$\sum_{x\in E} \mathrm{res}_g(x) n(\gamma, x) = (2\pi i)^{-1} \int_\gamma g(z) dz,$$

where $n(\gamma, x)$ is the winding number for the loop γ around the point x, which corresponds to the number of times the loop 'winds' around the point. As z tends to infinity, since the function f is assumed to be bounded, the meromorphic function g rapidly tends to zero for $|\chi| < \pi$. Since there is only one pole for g, and at 0, for any loop γ not passing through 0 we have,

$$f(0) n(\gamma, 0) = (2\pi i)^{-1} \Big[\int g(\gamma(t))\gamma'(t) dt \Big].$$

Now, choose the loop to be the one along the infinite rectangle, then we can split the curve into four parts, two of which are along the lines, $z = \pm\frac{1}{2}$, given by, $\gamma_\pm(t) = \pm it \pm \frac{1}{2}$, the other two curves are at infinity, and since g rapidly dies at infinity, they will not have any contribution. Hence we have,

$$f(0) = (2\pi i)^{-1} \int_{\mathbb{R}} \Big[ig\big(it + \tfrac{1}{2}\big) - ig\big(it - \tfrac{1}{2}\big) \Big] dt$$

Substituting $g(z) = \pi f(z)(\sin(\pi z))^{-1}$ we obtain,

$$f(0) = \tfrac{1}{2}\int_{\mathbb{R}} \Big[\Big[\frac{f(it+\frac{1}{2})}{\sin\big(i\pi t+\frac{\pi}{2}\big)}\Big] - \Big[\frac{f(it-\frac{1}{2})}{\sin\big(i\pi t-\frac{\pi}{2}\big)}\Big]\Big]dt.$$

Using the fact that $\sin(i\pi t+\pi/2) = \cos(i\pi t) = \cosh(\pi t)$, and $\sin(i\pi t-\pi/2) = -\cos(i\pi t) = -\cosh(\pi t)$, and by substituting $\cosh(\pi t) = (e^{\pi t}+e^{-\pi t})/2$, we obtain,

$$f(0) = \int_{\mathbb{R}} \Big[\frac{f\big(it+\frac{1}{2}\big)+f\big(it-\frac{1}{2}\big)}{e^{\pi t}+e^{-\pi t}}\Big]dt.$$

We now use this formula to obtain an operator equation using the modular relation;

Lemma 2.4.7. (Operator Equation) *If $A \in \mathcal{A}, B \in \mathcal{A}'$ are such that*

$$\lambda\big\langle \Delta_{\mathcal{A}}^{\frac{1}{2}}\nu | A\Delta_{\mathcal{A}}^{-\frac{1}{2}}\eta\big\rangle_\omega + \lambda^{-1}\big\langle \Delta_{\mathcal{A}}^{-\frac{1}{2}}\nu | A\Delta_{\mathcal{A}}^{\frac{1}{2}}\eta\big\rangle_\omega = \big\langle \nu | J_{\mathcal{A}}BJ_{\mathcal{A}}\eta\big\rangle_\omega$$

for any $\eta,\nu \in \mathcal{D}(\Delta_{\mathcal{A}}^{-1/2}) \cap \mathcal{D}(\Delta_{\mathcal{A}}^{1/2})$, then,

$$A = \int_{\mathbb{R}} \lambda^{2it}\Big[\frac{\Delta_{\mathcal{A}}^{it}J_{\mathcal{A}}BJ_{\mathcal{A}}\Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]dt.$$

Proof

For every $\nu,\eta \in \mathcal{D}(\Delta_{\mathcal{A}}^{-1/2}) \cap \mathcal{D}(\Delta_{\mathcal{A}}^{1/2})$, consider the complex-valued function

$$f_{\nu,\eta}(z) = \lambda^{2z}\Big\langle \Delta_{\mathcal{A}}^{\overline{z}}\nu \,\Big|\, A\Delta_{\mathcal{A}}^{-z}\eta\Big\rangle_\omega.$$

By strong continuity of $\Delta_{\mathcal{A}}^z$, normality of the action of $\mathcal{A}$ on $\mathcal{H}$, and by skew-linearity with respect to the first argument in the inner product, $f_{\nu,\eta}$ is an analytic function, and we have,

$$\begin{aligned} f_{\nu,\eta}\big(it+\tfrac{1}{2}\big) &= \lambda^{2(it+\frac{1}{2})}\Big\langle \Delta_{\mathcal{A}}^{\overline{it+\frac{1}{2}}}\nu \,\Big|\, A\Delta_{\mathcal{A}}^{-it-\frac{1}{2}}\eta\Big\rangle_\omega \\ &= \lambda^{2(it+\frac{1}{2})}\Big\langle \Delta_{\mathcal{A}}^{\frac{1}{2}}\Delta_{\mathcal{A}}^{-it}\nu \,\Big|\, A\Delta_{\mathcal{A}}^{-\frac{1}{2}}\Delta_{\mathcal{A}}^{-it}\eta\Big\rangle_\omega. \end{aligned}$$

Similarly we have,

$$\begin{aligned} f_{\nu,\eta}\big(it-\tfrac{1}{2}\big) &= \lambda^{2(it-\frac{1}{2})}\Big\langle \Delta_{\mathcal{A}}^{\overline{it-\frac{1}{2}}}\nu \,\Big|\, A\Delta_{\mathcal{A}}^{-it+\frac{1}{2}}\eta\Big\rangle_\omega \\ &= \lambda^{2(it-\frac{1}{2})}\Big\langle \Delta_{\mathcal{A}}^{-\frac{1}{2}}\Delta_{\mathcal{A}}^{-it}\nu \,\Big|\, A\Delta_{\mathcal{A}}^{\frac{1}{2}}\Delta_{\mathcal{A}}^{-it}\eta\Big\rangle_\omega. \end{aligned}$$

$f_{\nu,\eta}$ is also bounded on the strip $S = \{-\frac{1}{2} \le \mathrm{Re}(z) \le \frac{1}{2}\}$, so $g(z) = \pi f_{\nu,\mu}(z)(\sin(\pi z))^{-1}$ decreases rapidly as z tends to infinity. Hence by the calculation prior to this lemma, we have,

$$f_{\nu,\eta}(0) = \int_{\mathbb{R}} \Big[\frac{f_{\nu,\eta}(it+\frac{1}{2})+f_{\nu,\eta}(it-\frac{1}{2})}{e^{\pi t}+e^{-\pi t}}\Big]dt.$$

Since $f_{\nu,\eta}(0) = \big\langle \nu | A\eta\big\rangle_\omega$, we have,

$$\big\langle \nu | A\eta\big\rangle_\omega = \int_{\mathbb{R}} \lambda^{2(it+\frac{1}{2})}\Big[\frac{\langle \Delta_{\mathcal{A}}^{\frac{1}{2}}\Delta_{\mathcal{A}}^{-it}\nu | A\Delta_{\mathcal{A}}^{-\frac{1}{2}}\Delta_{\mathcal{A}}^{-it}\eta\rangle_\omega}{e^{\pi t}+e^{-\pi t}}\Big]dt + \int_{\mathbb{R}} \lambda^{2(it-\frac{1}{2})}\Big[\frac{\langle \Delta_{\mathcal{A}}^{-\frac{1}{2}}\Delta_{\mathcal{A}}^{-it}\nu | A\Delta_{\mathcal{A}}^{\frac{1}{2}}\Delta_{\mathcal{A}}^{-it}\eta\rangle_\omega}{e^{\pi t}+e^{-\pi t}}\Big]dt.$$

Note that $\Delta_{\mathcal{A}}^{-it}\nu, \Delta_{\mathcal{A}}^{-it}\eta \in \mathcal{D}(\Delta_{\mathcal{A}}^{-1/2}) \cap \mathcal{D}(\Delta_{\mathcal{A}}^{1/2})$. Hence using the modular relation;

$$\lambda\langle \Delta_{\mathcal{A}}^{\frac{1}{2}}\nu | A\Delta_{\mathcal{A}}^{-\frac{1}{2}}\eta\rangle_\omega + \lambda^{-1}\langle \Delta_{\mathcal{A}}^{-\frac{1}{2}}\nu | A\Delta_{\mathcal{A}}^{\frac{1}{2}}\eta\rangle_\omega = \langle \nu | J_{\mathcal{A}}BJ_{\mathcal{A}}\eta\rangle_\omega,$$

we get,

$$\begin{aligned}\langle \nu | A\eta\rangle_\omega &= \int_{\mathbb{R}} \lambda^{2it}\Big\langle \Delta_{\mathcal{A}}^{-it}\nu \,\Big|\, \Big[\frac{J_{\mathcal{A}}BJ_{\mathcal{A}}}{e^{\pi t}+e^{-\pi t}}\Big]\Delta_{\mathcal{A}}^{-it}\eta\Big\rangle_\omega dt \\ &= \Big\langle \nu \,\Big|\, \Big[\int_{\mathbb{R}} \lambda^{2it}\Big[\frac{\Delta_{\mathcal{A}}^{it}J_{\mathcal{A}}BJ_{\mathcal{A}}\Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]dt\Big]\eta\Big\rangle_\omega.\end{aligned}$$

Since the choice of ν, η was arbitrary it follows that,

$$A = \int_{\mathbb{R}} \lambda^{2it}\Big[\frac{\Delta_{\mathcal{A}}^{it}J_{\mathcal{A}}BJ_{\mathcal{A}}\Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]dt$$

on the dense subset $\mathcal{D}(\Delta_{\mathcal{A}}^{-1/2}) \cap \mathcal{D}(\Delta_{\mathcal{A}}^{1/2})$, and hence it holds on $\mathcal{H}$. □

With this lemma, we got rid of ν and η and showed that A and B can be related directly by an operator equation. We now have all the ingredients to prove Tomita's theorem, which establishes a relation between von Neumann algebras and their commutants in terms of the modular operator and the modular conjugation.

For every $B \in \mathcal{A}'$,

$$\int_{\mathbb{R}} \lambda^{2it}\Big[\frac{\Delta_{\mathcal{A}}^{it}J_{\mathcal{A}}BJ_{\mathcal{A}}\Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]dt \in \mathcal{A}.$$

Consider the function,

$$f_{\nu,\eta}(t) = \Big\langle \eta \,\Big|\, \Big[\frac{\Delta_{\mathcal{A}}^{it}J_{\mathcal{A}}BJ_{\mathcal{A}}\Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]\nu\Big\rangle_\omega.$$

The damping factor $1/(e^{\pi t}+e^{-\pi t})$ ensures that this function decreases faster than any polynomial in t, and allows us to do Fourier analysis. We now use the positivity of λ. Since $\lambda \in \mathbb{R}_+$ there exists χ such that $\lambda = e^{\chi/2}$. The Fourier transform of $f_{\nu,\eta}(t)$ is

$$\begin{aligned}\mathcal{F}f_{\nu,\eta}(\chi) &= \int_{\mathbb{R}} \lambda^{2it}\Big\langle \eta \,\Big|\, \Big[\frac{\Delta_{\mathcal{A}}^{it}J_{\mathcal{A}}BJ_{\mathcal{A}}\Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]\nu\Big\rangle_\omega dt \\ &= \Big\langle \eta \,\Big|\, \Big[\int_{\mathbb{R}} \lambda^{2it}\Big[\frac{\Delta_{\mathcal{A}}^{it}J_{\mathcal{A}}BJ_{\mathcal{A}}\Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]dt\Big]\nu\Big\rangle_\omega.\end{aligned}$$

We now prove Tomita's theorem via injectivity of Fourier transforms.

Theorem 2.4.8. (Tomita) *If $J_{\mathcal{A}}$ and $\Delta_{\mathcal{A}}$ are as before, then*

$$J_{\mathcal{A}}\mathcal{A}J_{\mathcal{A}} = \mathcal{A}',\ \Delta_{\mathcal{A}}^{it}\mathcal{A}\Delta_{\mathcal{A}}^{-it} = \mathcal{A},\ \forall t \in \mathbb{R}.$$

Proof

Similar to $f_{\nu,\eta}$, consider the function,

$$g_{\nu,\eta}(t) = \Big\langle \eta \,\Big|\, U\Big[\frac{\Delta_{\mathcal{A}}^{it}J_{\mathcal{A}}BJ_{\mathcal{A}}\Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]U^\dagger\nu\Big\rangle_\omega,$$

where U is a unitary operator in $\mathcal{A}'$. The Fourier transform of $g_{\nu,\eta}$ is given by,

$$\begin{aligned}\mathcal{F}g_{\nu,\eta}(\chi) &= \int_{\mathbb{R}} \lambda^{2it}\Big\langle \eta \,\Big|\, U\Big[\frac{\Delta_{\mathcal{A}}^{it} J_{\mathcal{A}} B J_{\mathcal{A}} \Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]U^{\dagger}\nu\Big\rangle_{\omega} dt \\ &= \Big\langle U^{\dagger}\eta \,\Big|\Big[\int_{\mathbb{R}} \lambda^{2it}\Big[\frac{\Delta_{\mathcal{A}}^{it} J_{\mathcal{A}} B J_{\mathcal{A}} \Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]dt\Big]U^{\dagger}\nu\Big\rangle_{\omega}.\end{aligned}$$

Since $\langle U\nu|U\eta\rangle_{\omega} = \langle\nu|\eta\rangle_{\omega}$ it follows that,

$$\mathcal{F}g_{\nu,\eta}(\chi) = \Big\langle \eta \,\Big|\Big[\int_{\mathbb{R}} \lambda^{2it}\Big[\frac{\Delta_{\mathcal{A}}^{it} J_{\mathcal{A}} B J_{\mathcal{A}} \Delta_{\mathcal{A}}^{-it}}{e^{\pi t}+e^{-\pi t}}\Big]dt\Big]\nu\Big\rangle_{\omega} = \mathcal{F}f_{\nu,\eta}(\chi).$$

By injectivity of Fourier transformation it follows that $f_{\nu,\eta} = g_{\nu,\eta}$. Since the choice of ν and η was also arbitrary it follows that $U(\Delta_{\mathcal{A}}^{it} J_{\mathcal{A}} B J_{\mathcal{A}} \Delta_{\mathcal{A}}^{-it}) = (\Delta_{\mathcal{A}}^{it} J_{\mathcal{A}} B J_{\mathcal{A}} \Delta_{\mathcal{A}}^{-it})U$. Since the choice of the unitary $U \in \mathcal{A}'$ was arbitrary we conclude that

$$\Delta_{\mathcal{A}}^{it} J_{\mathcal{A}} B J_{\mathcal{A}} \Delta_{\mathcal{A}}^{-it} \in \mathcal{A} \;\; \forall t \in \mathbb{R}.$$

By taking $t = 0$ we have, $J_{\mathcal{A}}\mathcal{A}'J_{\mathcal{A}} \subseteq \mathcal{A}$ and by symmetry we also have, $J_{\mathcal{A}}\mathcal{A}J_{\mathcal{A}} \subseteq \mathcal{A}'$. Hence we have,

$$J_{\mathcal{A}}\mathcal{A}J_{\mathcal{A}} = \mathcal{A}'.$$

Hence every $A \in \mathcal{A}$ is of the form, $A = J_{\mathcal{A}} B J_{\mathcal{A}}$ for some $B \in \mathcal{A}'$. We have, $\Delta_{\mathcal{A}}^{it} A \Delta_{\mathcal{A}}^{-it} = \Delta_{\mathcal{A}}^{it} J_{\mathcal{A}} B J_{\mathcal{A}} \Delta_{\mathcal{A}}^{-it} \in \mathcal{A}$. Hence it follows that,

$$\Delta_{\mathcal{A}}^{it} \mathcal{A} \Delta_{\mathcal{A}}^{-it} \in \mathcal{A}$$

this completes the proof. □

2.4.3 | Borchers-Wiesbrock Fundamental Relations

If $(\mathcal{A}, U, \Omega)$ is a causal Borchers triple, Borchers' theorem relates the one-parameter groups of the Borchers triples and the modular automorphisms corresponding to the state ω. We will prove the theorem following [14].

Theorem 2.4.9. (Borchers) *Let $(\mathcal{A}, U, \Omega)$ be a Borchers triple. Then*

$$\Delta_{\mathcal{A}}^{it} U(s) \Delta_{\mathcal{A}}^{-it} = U(e^{-2\pi t}s),$$
$$J_{\mathcal{A}} U(s) J_{\mathcal{A}} = U(s)^{\dagger}$$

for all $s, t \in \mathbb{R}$.

Proof

Let $\nu \in \mathcal{A}\Omega$ and $\eta \in \mathcal{A}'\Omega$, consider the function

$$f_{\nu,\eta}^{+}(z) = \Big\langle \Big[\Delta_{\mathcal{A}}^{iz} U(e^{2\pi z}s)\Delta_{\mathcal{A}}^{-iz}\Big]\nu \,\Big|\, \eta\Big\rangle_{\omega}, \quad s \in \mathbb{R}_{+}.$$

This is a bounded function whenever $\mathrm{Im}(e^{2\pi z}s)$ is positive,[3] which is the case on the sets

$$S(k, k+\tfrac{1}{2}) \equiv \big\{x + iy \mid y \in \big[k, k+\tfrac{1}{2}\big]\big\}$$

[3]The problematic part is $U(e^{2\pi z}s)$, by Stone's theorem, we can write $U(s) = e^{isP}$, and hence we have $\|U(e^{2\pi z}s)\| = \big\|e^{i\mathrm{Re}(e^{2\pi z}s)P}e^{-\mathrm{Im}(e^{2\pi z}s)P}\big\| \leq \big\|e^{-\mathrm{Im}(e^{2\pi z}s)P}\big\|$. Here we used the fact that the first term in the product is a unitary operator whose norm must be 1. The second term is a bounded operator whenever $\mathrm{Im}(e^{2\pi z}s)$ is positive.

and hence $U(e^{2\pi z}s)$ is a bounded operator for all $z \in S(k, k+\frac{1}{2})$.

Since $U(s)\mathcal{A}U(s)^\dagger \subseteq \mathcal{A}$ for all $s \in \mathbb{R}_+$, it follows that for any $A \in \mathcal{A}$ and $B \in \mathcal{A}'$, $[U(s)^\dagger BU(s), A] = U(s)^\dagger[B, U(s)AU(s)^\dagger]U(s) = 0$. So we have $U(s)^\dagger \mathcal{A}' U(s) \subseteq \mathcal{A}'$, for all $s \in \mathbb{R}_+$. If we define

$$V(s) = J_{\mathcal{A}} U(s)^\dagger J_{\mathcal{A}}$$

By Tomita's theorem, $J_{\mathcal{A}}\mathcal{A}J_{\mathcal{A}} = \mathcal{A}'$, it follows that $V(s)\mathcal{A}V(s)^\dagger \subseteq \mathcal{A}$, for all $s \in \mathbb{R}_+$. So, $V(s)$ has similar properties as U. For similar reasons as above, the function,

$$f^-_{\nu,\eta}(z) = \left\langle \Delta^{iz}_{\mathcal{A}} V(e^{2\pi z}s)\Delta^{-iz}_{\mathcal{A}}\nu \,\middle|\, \eta \right\rangle_\omega,$$

is the bounded operator on the sets, $S(k-\frac{1}{2}, k) = \left\{x + iy \mid y \in \left[k-\frac{1}{2}, k\right]\right\}$ for all $s \in \mathbb{R}_+$.

We now show that we can define a well-defined analytic function on the complex plain by taking these functions on small patches and then gluing them up together. The idea is to make use of Schwarz's reflection principle from complex analysis, which states that, if an analytic function is defined on the upper half-plane and has well-defined (non-singular) real values on the real axis, then it can be extended to the conjugate function on the lower half-plane.

To show that $f^+_{\nu,\eta}$ and $f^-_{\nu,\eta}$ glue-up to give an analytic function we must verify that they agree on the common domains, that is, we have to verify that $f^+_{\nu,\eta}(t+i/2) = f^-_{\nu,\eta}(t)$. We have

$$\begin{aligned}
f^+_{\nu,\eta}\big(t+\tfrac{i}{2}\big) &= \left\langle \Delta^{it}_{\mathcal{A}}\Delta^{-\frac{1}{2}}_{\mathcal{A}} U(e^{2\pi t + i\pi}s)\Delta^{-it}_{\mathcal{A}}\Delta^{\frac{1}{2}}_{\mathcal{A}}\nu \,\middle|\, \eta \right\rangle_\omega \\
&= \left\langle \Delta^{it}_{\mathcal{A}}\Delta^{-\frac{1}{2}}_{\mathcal{A}} J^2_{\mathcal{A}} U(-e^{2\pi t}s)\Delta^{-it}_{\mathcal{A}} J^2_\omega \Delta^{\frac{1}{2}}_{\mathcal{A}}\nu \,\middle|\, \eta \right\rangle_\omega \\
&= \left\langle \Delta^{it}_{\mathcal{A}} J_{\mathcal{A}}\Delta^{\frac{1}{2}}_{\mathcal{A}} V(e^{2\pi t}s)\Delta^{-it}_{\mathcal{A}} J_{\mathcal{A}}\Delta^{\frac{1}{2}}_{\mathcal{A}}\nu \,\middle|\, \eta \right\rangle_\omega.
\end{aligned}$$

In the second step we used the fact that $J^2_\omega = I$, and in the third step we used $\Delta^{-1/2}_{\mathcal{A}} J_{\mathcal{A}} = J_{\mathcal{A}}\Delta^{1/2}_{\mathcal{A}}$, and $J_{\mathcal{A}}\Delta^{it}_{\mathcal{A}} = \Delta^{it}_{\mathcal{A}} J_{\mathcal{A}}$, and substituted $J_{\mathcal{A}}U(s)^\dagger J_{\mathcal{A}} = V(s)$.

By Tomita's theorem, we have, $\Delta^{it}_{\mathcal{A}} A \Delta^{-it}_{\mathcal{A}} \in \mathcal{A}$ for every $A \in \mathcal{A}$. Hence we must have $V(e^{2\pi t}s)\Delta^{it}_{\mathcal{A}} A^\dagger \Delta^{-it}_{\mathcal{A}} V(e^{2\pi t}s)^\dagger \in \mathcal{A}$ and. Since Ω is invariant under the action of V and $\Delta_{\mathcal{A}}$, we have, $\Delta^{it}_{\mathcal{A}} V(e^{2\pi t}s)^\dagger \Omega = \Omega$. Plugging these into the above equation, we obtain,

$$\begin{aligned}
f^+_{\nu,\eta}\big(t+\tfrac{i}{2}\big) &= \left\langle \Delta^{it}_{\mathcal{A}} J_{\mathcal{A}}\Delta^{\frac{1}{2}}_{\mathcal{A}} V(e^{2\pi t}s)\Delta^{it}_{\mathcal{A}} J_{\mathcal{A}}\Delta^{\frac{1}{2}}_{\mathcal{A}}[A] \,\middle|\, \eta \right\rangle_\omega \\
&= \left\langle \Delta^{it}_{\mathcal{A}} J_{\mathcal{A}}\Delta^{\frac{1}{2}}_{\mathcal{A}} \big(V(e^{2\pi t}s)\Delta^{it}_{\mathcal{A}} A^\dagger\big)\Omega \,\middle|\, \eta \right\rangle_\omega \\
&= \left\langle \Delta^{it}_{\mathcal{A}} J_{\mathcal{A}}\Delta^{\frac{1}{2}}_{\mathcal{A}} \big(V(e^{2\pi t}s)\Delta^{it}_{\mathcal{A}} A^\dagger \Delta^{it}_{\mathcal{A}} V(e^{2\pi t}s)^\dagger\big)\Omega \,\middle|\, \eta \right\rangle_\omega \\
&= \left\langle \Delta^{it}_{\mathcal{A}} V(e^{2\pi t}s)\Delta^{-it}_{\mathcal{A}} A\Omega \,\middle|\, \eta \right\rangle_\omega = f^-_{\nu,\eta}(t).
\end{aligned}$$

Similarly, we also have,

$$f^-_{\nu,\eta}\big(t+\tfrac{i}{2}\big) = f^+_{\nu,\eta}(t).$$

Hence the function,

$$F_{\nu,\eta}(x+iy) = \begin{cases} f^+_{\nu,\eta}(x+iy) & \forall y \in \left[n, n+\frac{1}{2}\right] \\ f^-_{\nu,\eta}(x+iy) & \forall y \in \left[n-\frac{1}{2}, n\right] \end{cases}$$

is a bounded, analytic function on $\mathbb{C}$. By Liouville's theorem, $F_{\nu,\eta}$ must be a constant function. Hence we have, $F_{\nu,\eta}(0) = F_{\nu,\eta}(t)$. Since ν, η are arbitrary, and are dense inside $\mathcal{H}$, it follows that

$$\Delta_{\mathcal{A}}^{it} U(e^{2\pi t} s) \Delta_{\mathcal{A}}^{-it} = U(s)$$

or equivalently,

$$\Delta_{\mathcal{A}}^{it} U(s) \Delta_{\mathcal{A}}^{-it} = U(e^{-2\pi t} s)$$

Similarly, we must have $F_{\nu,\eta}(0) = F_{\nu,\eta}\big(\frac{i}{2}\big)$, which yields,

$$U(s) = J_{\mathcal{A}} U(-s) J_{\mathcal{A}}.$$

□

We now prove a convenient generalization of the previous theorem. Let $\mathcal{B} \subset \mathcal{A}$ be a von Neumann subalgebra, ω is a faithful normal state for both $\mathcal{A}$ and $\mathcal{B}$, equivalently the Hilbert space $\mathcal{H}$ has a common cyclic and separating vector Ω, that is

$$\overline{\mathcal{B}\Omega} = \overline{\mathcal{A}\Omega} = \mathcal{H}.$$

We let $\Delta_{\mathcal{A}}$, $\Delta_{\mathcal{B}}$ and $J_{\mathcal{A}}$, $J_{\mathcal{B}}$ denote the modular operator and modular conjugation corresponding to $\mathcal{A}$ and $\mathcal{B}$ respectively. All unitary groups we will encounter will be assumed to be strongly continuous.

Theorem 2.4.10. (Borchers-Wiesbrock) *Let $U(s)$ be such that $U(s)\Omega = \Omega \ \forall s \in \mathbb{R}$ be a unitary group with analytic extension on $S(0, \frac{1}{2})$, continuous on its boundary. If*

$$U(s)\mathcal{B}U(s)^{\dagger} \subset \mathcal{A} \qquad \forall s \in \mathbb{R}$$
$$U(s + \tfrac{i}{2})\mathcal{B}'U(s + \tfrac{i}{2})^{\dagger} \subset \mathcal{A}' \quad \forall s \in \mathbb{R}$$

then

$$\Delta_{\mathcal{A}}^{it} U(s) \Delta_{\mathcal{B}}^{-it} = U(s - t) \quad \forall s, t \in \mathbb{R}$$
$$J_{\mathcal{A}} U(s) J_{\mathcal{B}} = U(s + \tfrac{i}{2}) \qquad \forall s \in \mathbb{R}.$$

Proof

The idea of the proof is the same as the proof of Borchers' theorem in the last section, with a lot more tedious manipulations. We will only sketch some of these manipulations and give the main ideas as to what we are trying to achieve with those manipulations. Let $\nu \in \mathcal{B}\Omega$ and $\eta \in \mathcal{A}'\Omega$. For a fixed $s \in \mathbb{R}$ define the two functions in the variable t by,

$$f^{+}_{\nu,\eta}(t) = \Big\langle \big[\Delta_{\mathcal{A}}^{it} U(s+t) \Delta_{\mathcal{B}}^{-it}\big]\nu \,\Big|\, \eta \Big\rangle_{\omega},$$
$$f^{-}_{\nu,\eta}(t) = \Big\langle \big[\Delta_{\mathcal{B}}^{it} U(s+t)^{\dagger} \Delta_{\mathcal{A}}^{-it}\big]\eta \,\Big|\, \nu \Big\rangle_{\omega}.$$

By assumption $\nu = [B] \in \mathcal{B}\Omega$ and $\eta = [A'^{\dagger}] \in \mathcal{A}'\Omega$ $B \in \mathcal{B}$ and $A' \in \mathcal{A}'$. So we have,

$$f^{+}_{\nu,\eta}(t) = \Big\langle \Big[\Delta_{\mathcal{A}}^{it} U(s+t) \Delta_{\mathcal{B}}^{-it}[B] \,\Big|\, [A'^{\dagger}]\Big\rangle_{\omega} = \Big\langle A' \Big[\underbrace{\Delta_{\mathcal{A}}^{it} \Big[U(s+t) \big[\Delta_{\mathcal{B}}^{-it} B \Delta_{\mathcal{B}}^{it}\big] U(s+t) \Big] \Delta_{\mathcal{A}}^{-it}}_{\in \mathcal{A}} \Big] \Omega \,\Big|\, \Omega \Big\rangle_{\omega}$$

Here we used the assumption that $U(s)\mathcal{B}U(s)^\dagger \subset \mathcal{A}$ for all $s \in \mathbb{R}$. This allows us to exchange elements of $\mathcal{A}$ and $\mathcal{A}'$ which ensures that

$$f^+_{\nu,\eta}(t) = f^-_{\nu,\eta}(t).$$

Since $U(s)$ extends to an analytic function on the strip $S(0, \frac{1}{2})$, the function $f^+_{\nu,\eta}(t)$ posses bounded extensions into the strip $S(0, \frac{1}{2})$, and $f^-_{\nu,\eta}(t)$ into the strip $S(-\frac{1}{2}, 0)$.

$$\begin{aligned} f^+_{\nu,\eta}(t+\tfrac{i}{2}) &= \Big\langle \Big[\Delta_{\mathcal{A}}^{i(t+\frac{i}{2})} U\big(s+t+\tfrac{i}{2}\big)\Delta_{\mathcal{B}}^{-i(t+\frac{i}{2})}\Big][B] \,\Big|\, [A'^\dagger]\Big\rangle_\omega \\ &= \Big\langle \Big[\Delta_{\mathcal{A}}^{-\frac{1}{2}}\Delta_{\mathcal{A}}^{it} U\big(s+t+\tfrac{i}{2}\big)\Delta_{\mathcal{B}}^{-it}\Delta_{\mathcal{B}}^{\frac{1}{2}}\Big][B] \,\Big|\, [A'^\dagger]\Big\rangle_\omega \\ &= \Big\langle \Big[J_{\mathcal{A}}\Delta_{\mathcal{A}}^{it} U\big(s+t+\tfrac{i}{2}\big)\Delta_{\mathcal{B}}^{-it} J_{\mathcal{B}}\Big][B^\dagger] \,\Big|\, [A']\Big\rangle_\omega \\ &= \Big\langle \Big[A'^\dagger J_{\mathcal{A}}\Delta_{\mathcal{A}}^{it} U\big(s+t+\tfrac{i}{2}\big)\Delta_{\mathcal{B}}^{-it} J_{\mathcal{B}} B^\dagger\Big]\Omega \,\Big|\, \Omega\Big\rangle_\omega. \end{aligned}$$

Using $U(s+t+\frac{i}{2})^\dagger\Omega = \Omega$ and $\Delta_{\mathcal{A}}\Omega = \Delta_{\mathcal{B}}\Omega = \Omega$ we obtain that,

$$f^+_{\nu,\eta}(t+\tfrac{i}{2}) = \Big\langle \underbrace{[\,\Delta_{\mathcal{A}}^{-it}(\underbrace{J_{\mathcal{A}}A'^\dagger J_{\mathcal{A}}}_{\in\mathcal{A}})\Delta_{\mathcal{A}}^{it}\,]}_{\in\mathcal{A}} \underbrace{\Big[\, U\big(s+t+\tfrac{i}{2}\big)\,[\Delta_{\mathcal{B}}^{-it}(\underbrace{J_{\mathcal{B}}B^\dagger J_{\mathcal{B}}}_{\in\mathcal{B}'})\Delta_{\mathcal{B}}^{it}]U\big(s+t+\tfrac{i}{2}\big)\,\Big]}_{\in\mathcal{A}'}\Omega|\Omega\Big\rangle.$$

Here we used the assumption that $U(s+\frac{i}{2})\mathcal{B}'U(s+\frac{i}{2})^\dagger \subset \mathcal{A}'$. This allows us to exchange the two underbraced terms, which ensures that

$$f^+_{\nu,\eta}(t+\tfrac{i}{2}) = f^-_{\nu,\eta}(t-\tfrac{i}{2}).$$

Hence by Schwarz reflection principle, we obtain an entire analytic function,

$$F_{\nu,\eta}(x+iy) = \begin{cases} f^-_{\nu,\eta}(x+iy) & \forall y \in \big[n-\frac{1}{2}, n\big] \\ f^+_{\nu,\eta}(x+iy) & \forall y \in \big[n, n+\frac{1}{2}\big] \end{cases}$$

Since the function is bounded by $\sup\{\|[B^\dagger]\|\|[A]\|, \|[A^\dagger]\|\|[B]\}$ by Liouville's theorem it must correspond to a constant function. Hence by evaluating $F_{\nu,\eta}$ using both $f^+_{\nu,\eta}$ and $f^-_{\nu,\eta}$ at $t \in \mathbb{R}$ and using the density of the choices ν and η it follows that

$$\Delta_{\mathcal{A}}^{it} U(s) \Delta_{\mathcal{B}}^{-it} = U(s-t), \;\; \forall s, t \in \mathbb{R}.$$

Similarly, we must also have $f^+_{\nu,\eta}(\frac{i}{2}) = f^-_{\nu,\eta}\big(-\frac{i}{2}\big)$, and using $J_{\mathcal{A}}$ and $J_{\mathcal{B}}$ to flip the order in the inner product, we obtain

$$J_{\mathcal{A}} U(s) J_{\mathcal{B}} = U(s+\tfrac{i}{2}).$$

□

Wiesbrock in [49] constructs a strongly continuous one-parameter group of unitaries that satisfy the conditions of Borchers' theorem starting from a given 'good' von Neumann subalgebra $\mathcal{B} \subset \mathcal{A}$. The pair $(\mathcal{B} \subset \mathcal{A}, \Omega)$ is called a $\pm$half-sided modular inclusion if

$$\Delta_{\mathcal{A}}^{-it}\mathcal{B}\Delta_{\mathcal{A}}^{it} \subset \mathcal{B}, \;\; \forall t \in \mathbb{R}_{\mp}.$$

Suppose $U(s)$ is a group of unitaries for $s \in \mathbb{R}$ satisfying the conditions of Borchers theorem, that is there exists $P \geq 0$ such that $U(s) = e^{isP}$ and $U(s)\Omega = \Omega$ such that

$$U(s)\mathcal{A}U(s)^\dagger \subset \mathcal{A}, \;\; \forall s \in \mathbb{R}^+.$$

Then by Borchers theorem, we have

$$\Delta_{\mathcal{A}}^{it} U(s) \Delta_{\mathcal{A}}^{-it} = U(e^{-2\pi t} s),\ \forall s, t \in \mathbb{R}.$$

By Tomita's theorem we have $\Delta_{\mathcal{A}}^{it} \mathcal{A} \Delta_{\mathcal{A}}^{-it} = \mathcal{A}$. Combining the two we have

$$\begin{aligned}
\Delta_{\mathcal{A}}^{-it}\big(U(s)\mathcal{A}U(s)^\dagger\big)\Delta_{\mathcal{A}}^{it} &= (\Delta_{\mathcal{A}}^{-it} U(s) \Delta_{\mathcal{A}}^{it})\big(\Delta_{\mathcal{A}}^{-it} \mathcal{A} \Delta_{\mathcal{A}}^{it}\big)(\Delta_{\mathcal{A}}^{-it} U(s)^\dagger \Delta_{\mathcal{M}}^{it}) \\
&= U(e^{2\pi t} s)\big(\Delta_{\mathcal{A}}^{-it} \mathcal{A} \Delta_{\mathcal{A}}^{it}\big) U(-e^{2\pi t} s) \\
&= U(s)\big(U((e^{2\pi t} - 1)s)\mathcal{A}U((e^{2\pi t} - 1)s)^\dagger\big)U(s)^\dagger.
\end{aligned}$$

Since $(e^{2\pi t} - 1)s \geq 0\ \forall s, t \in \mathbb{R}^+$ and by assumption $U(s)\mathcal{A}U(s)^\dagger \subset \mathcal{A}$, $\forall s \in \mathbb{R}^+$, we have

$$\Delta_{\mathcal{A}}^{it}\Big[U(s)\mathcal{A}U(s)^\dagger\Big]\Delta_{\mathcal{A}}^{-it} \subset \Big[U(s)\mathcal{A}U(s)^\dagger\Big],\ \forall s, t \in \mathbb{R}^+.$$

So, $(U(s)\mathcal{A}U(s)^\dagger \subset \mathcal{A}, \Omega)$ is a $-$half-sided modular inclusion for each $s \in \mathbb{R}^+$. So unitary groups satisfying conditions of Borchers theorem give rise to half-sided modular inclusions. That is, whenever $(\mathcal{A}, U, \Omega)$ is a Borchers triple then $\forall s \in \mathbb{R}^+, (U(s)\mathcal{A}U(s)^\dagger \subset \mathcal{A}, \Omega)$ is a half-sided modular inclusion.

Appendices for the Chapter

2.A | The Haar Measure

A left Haar measure on $\mathcal{G}$ is a non-zero Radon measure λ, such that for all $x \in \mathcal{G}$,

$$\lambda(xH) = \lambda(H).$$

where H is a Borel subset of $\mathcal{G}$ and $x \in \mathcal{G}$. Radon measures are measures that respect the topology of the underlying space. In case of locally compact spaces the Riesz representation theorem tells us that Radon measures are in bijection with linear functionals on the space of compactly supported continuous functions, given by $f \mapsto \int f d\lambda$, so the invariance translated to this setting means,

$$\int L_x f d\lambda = \int f d\lambda.$$

where $(L_x f)(y) = f(x^{-1}y)$. Haar measures are very useful, and they allow much of the analysis that was possible with Lebesgue measures on $\mathbb{R}^n$.

Theorem 2.A.1. *Every locally compact group has a unique left Haar measure.*

Proof

Existence of Haar measures on locally compact groups is a very important tool for analysis of locally compact groups. We need to construct a functional on the space of compactly supported continuous functions that is translation invariant. All that we have to work with is the set of continuous functions on the locally commpact group. Compactly supported functions are well behaved, so, we can try to exploit their nice properties.

Let f, g be two compactly supported positive functions. Note that the support of each of these functions can be covered by finitely many open sets. We can relate the two functions using this. We can dominate the function f by a finite linear combination of scaled translations of g. Note that this is possible because f attains a maximum, and g is a non-zero positive function. So, there exists some $x \in \mathcal{G}$ such that $g(x) > 0$, and we can find a number λ such that $\lambda g(x) > \|f\|_{\sup}$. By continuity of translation we can find a neighborhood around x where this holds. We can now transfer via translation L_y all the problematic parts to this neighborhood. Since the support of f is compact, we can cover it with finitely many translations of this neighborhood, hence we can make sense of a sum. So, there exists a finite collection of real numbers λ_i and translations L_{x_i} such that,

$$f \leqq \textstyle\sum_i \lambda_i (L_{x_i}(g)).$$

Consider the set of all finite collections of numbers $\{\lambda_i\}$ such that $f \leqq \sum_{i=1}^{n} \lambda_i(L_{x_i}(g))$. This is nonempty due to the above reason. Denote this set by $\mathcal{D}(f,g)$. The intuitive meaning here is that $\sum_i \lambda_i(L_{x_i}g)$ is an approximation of f in terms of translations of the function g. Using this we can define the 'best' approximation based on how little scaling we have to do with the translations of g. This defines some sort of index of the function f with respect to g.

Since each element of this set is a finite collection of real numbers, we can add them up. So, we have a 'sum' function on $\mathcal{D}(f,g)$,

$$\begin{aligned} \textstyle\sum : \mathcal{D}(f,g) &\to \mathbb{R} \\ \{\lambda_i\} &\mapsto \textstyle\sum_i \lambda_i. \end{aligned}$$

Using this function, we can define a new object,

$$(f,g) \coloneqq \inf_{\{\lambda_i\}\in\mathcal{D}(f,g)} \big(\textstyle\sum_i \lambda_i\big).$$

We will call this the 'approximation index' of the function f with respect to g. This approximation index allows us to reduce the problem of constructing a special linear functional on the space of compactly supported continuous functions to doing some stuff with some 'reference function'. This will be made precise below. Note that this is where the 'locally compact' part has been important. The Riesz representation tells us that the construction of Radon measures is equivalent to constructing some special linear functionals on compactly supported continuous functions. Then the compactness of the supports of the functions ensure that the sets $\mathcal{D}(f,g)$ are non-empty for each compactly supported functions f and g.

Note that the translation L_y is a linear map on the space of continuous functions, so we have,

$$L_{y^{-1}}\big(\textstyle\sum_i \lambda_i(L_{x_i}g)\big) = \sum_i \lambda_i(L_{y^{-1}x_i}g).$$

Since the approximation index is an infimum, the approximation index of any translation L_y of the function f with respect to g will be the same. So, we have,

$$(f,g) = (L_y f, g). \tag{invariance}$$

Scaling the functions f by λ just scales the sets $\mathcal{D}(f,g)$, so we have,

$$(\lambda f, g) = \lambda(f,g). \tag{scaling}$$

If $\{\lambda_i\} \in \mathcal{D}(f,g)$ and $\{\mu_j\} \in \mathcal{D}(h,g)$, then the disjoint union, $\{\lambda_i\} \coprod \{\mu_j\} \in \mathcal{D}(f+h,g)$. So, we have,

$$(f+h,g) \leqq (f,g) + (h,g). \tag{subadditivity}$$

Note that if this was additive, we would be done, we would have constructed a linear functional that is invariant under left translation. It is however possible choose g such that the inequality gets closer to an equality. The key lies in how well we can approximate the given function f, in terms of linear combinations of translations of g. So, the problem is now choosing an appropriate g. We will need some further properties of the functionals (f,g) to do this.

If $f \leq h$ then any upper approximation of h by translations of g is also an upper approximation of f by translations of g. Hence we have,

$$f \leq h \Rightarrow (f,g) \leqq (h,g). \tag{monotonicity}$$

Let f attain a maximum at x, that is $f(x) = \|f\|_{\sup}$ and g attain a maximum at y, that is, $g(y) = \|g\|_{\sup}$. Then the translation $L_{yx^{-1}}(f)$ attains a maximum at y, by $(f,g) = (L_{yx^{-1}}f, g)$, we obtain,

$$(f,g) \geqq \|f\|_{\sup}/\|g\|_{\sup}. \qquad \text{(boundedness)}$$

Suppose $f \leqq \sum_i \lambda_i (L_{x_i} g)$ and $g \leqq \sum_j \mu_j (L_{y_j} h)$, then by linearity of translations, we have, $f \leqq \sum_{i,j} \lambda_i \mu_j (L_{x_i y_j} h)$. That is to say, $\{\lambda_i \mu_j\} \in \mathcal{D}(f,h)$ whenever $\{\lambda_i\} \in \mathcal{D}(f,g)$ and $\{\mu_j\} \in \mathcal{D}(g,h)$. So, we have,

$$(f,h) \leqq (f,g)(g,h). \qquad \text{(product)}$$

Now we choose a 'reference function' F, and define a normalized index,

$$I_g(f) \coloneqq (f,g)/(F,g).$$

This gives us a functional $I_g : C_c^+(\mathcal{G}) \to \mathbb{R}$, that is left invariant, subadditive and monotone. By, $(F,g) \leq (F,f)(f,g)$ and $(f,g) \leq (f,F)(F,g)$ we have,

$$f_F \equiv (F,f)^{-1} \leqq I_g(f) \leqq (f,F) \equiv f^F \qquad \text{(rescaling)}$$

Now the problem is to choose an appropriate g such that (subadditivity) becomes an equality. In order to choose such a g, we have to understand what makes the approximations (f,g) not 'close' to the function f. The function is determined by its value at each point. So, given a number, we can scale it appropriately at each point to get f. What we are trying to do is to approximate the function f, by scaling a bunch of functions on some open sets. So, smaller open sets means better approximation. So, to make the approximation closer to the actual value of f, we could use g that has smaller support.

Let f_i, f_j be two compactly supported continuous functions. Let g be a positive compactly supported continuous function such that $g \equiv 1$ on the support of the functions $(f_i + f_j)$. Let $\delta > 0$ be a number (will be specified a bit later). Then we can consider the function,

$$h = f_i + f_j + \delta \cdot g$$

Using this we can construct new functions,

$$h_i = f_i/h, \;\; h_j = f_j/h.$$

This implies that $h_i \in \mathcal{C}_c^+(\mathcal{G})$, and $h_i \equiv 0$ whenever $f_i \equiv 0$. By continuity of h_i, for every $\delta > 0$ we can find a small enough neighborhood V of $1 \in \mathcal{G}$ such that

$$|h_i(x) - h_i(y)| \leqq \delta$$

whenever $x, y \in V$. For any function φ with support in V, consider an approximate of h by finitely many translations of φ, $h \leqq \sum_k \lambda_k L_{x_k} \varphi$, then we have,

$$f_i(x) = h(x)h_i(x) \leqq \sum_j \lambda_k L_{x_k}\varphi(x) h_i(x) \leqq \sum_k \lambda_k \varphi(x_k^{-1}x) h_i(x) \leqq \sum_k \lambda_k \varphi(x_k^{-1}x)(h_i(x_k) + \delta)$$

whenever $x, x_k \in V$. Since $h_i + h_j \leqq 1$, we get,

$$(f_i,\varphi) + (f_j,\varphi) \leqq \sum_k \lambda_k \varphi(x_k^{-1}x)(h_i(x_k) + \delta) + \sum_k \lambda_k \varphi(x_k^{-1}x)(h_j(x_k) + \delta) \leqq (1 + 2\delta) \sum_k \lambda_k.$$

Hence, by taking the infimum of all such $\sum_k \lambda_k$, we get,

$$I_\varphi(f_i) + I_\varphi(f_j) \leqq (1+2\delta) I_\varphi(h) \leqq (1+2\delta)(I_\varphi(f_i) + I_\varphi(f_j) + \delta I_\varphi(g)) \\ = (I_\varphi(f_i + f_j)) + \delta(2I_\varphi(f_i + f_j) + I_\varphi(g)) + 2\delta^2 I_\varphi(g).$$

The last step is by (subadditivity) and (scaling). Then by (rescaling) we can choose δ small enough such that the last part is less than ϵ, $\delta(2I_\varphi(f_i + f_j) + I_\varphi(g)) + 2\delta^2 I_\varphi(g) < \epsilon$. Hence we can choose a function φ with a small support such that,

$$I_\varphi(f_i) + I_\varphi(f_j) \leqq I_\varphi(f_i + f_j) + \epsilon.$$

The choice of δ determines the support of the function used for approximation. This can be chosen such that $I_\varphi(f_i) + I_\varphi(f_j)$ is as close to $I_\varphi(f_i + f_j)$ as we want. However this choice of δ depends on the functions f_i and f_j.

Each function $\varphi \in \mathcal{C}_c^+(\mathcal{G})$ gives us a map $f \mapsto I_\varphi(f) \in \mathbb{R}$. The goal is to find a functional that is linear. For each φ the value of $I_\varphi(f)$ lies in the interval $X_f \equiv [f_F, f^F]$. So, we can embed all such functionals inside the compact Hausdorff space,

$$X = \prod_{f \in \mathcal{C}_c^+(\mathcal{G})} X_f.$$

This is compact Hausdorff space because it is a product of compact Hausdorff spaces. It consists of all real valued functions from the space of compactly supported functions whose value at f lies in the interval X_f.

Consider for each neighborhood V of $1 \in \mathcal{G}$, the compact subsets $K(V)$ of X, consisting of closures of the collection of all functionals of the form I_φ with support of φ inside V.

$$X = \bigcup_{1 \in V} K(V)$$

Clearly, $V \subseteq W$ implies $K(V) \subseteq K(W)$. So, these sets satisfy the 'finite intersection property' $\bigcap_{i=1}^n K(V_i) \supset K(\bigcap_{i=1}^n V_i)$. This is a strict subset because there always exist functions with domain that is larger than $\bigcap_i V_i$. Since X is compact, the intersection $\bigcap_i K(V_i)$ is nonempty.[4] So there exists some point I in X that lies in all of the $K(V)$'s.

Every neighborhood of I contains I_φ's. So, for any collection of functions $\{f_i\}_{i=1}^n \subset \mathcal{C}_c^+(\mathcal{G})$ and $\epsilon > 0$, there exists some $\varphi \in \mathcal{C}_c^+(\mathcal{G})$ such that $|I(f_i) - I_\varphi(f_i)| < \epsilon$, by definition of neighborhoods in product topology.

This implies that I is left translation invariant $L_y I = I$, and commutes with scaling. Since by adjusting the support for all $\epsilon > 0$, we can find a function φ such that,

$$I_\varphi(f_i) + I_\varphi(f_j) \leqq I_\varphi(f_i + f_j) + \epsilon.$$

This implies that, $I(f_i) + I(f_j) \leqq I(f_i + f_j) + \epsilon$ for all ϵ, and since I belongs to every $K(V)$, ϵ can be arbitrarily small, hence,

$$I(f_i) + I(f_j) = I(f_i + f_j).$$

This has a unique extension to $\mathcal{C}_c(\mathcal{G})$, because any function can be written as $f = g - h$ with $g, h \in \mathcal{C}_c^+(\mathcal{G})$, and if $f = g' - h'$, then we have $g + h' = g' + h$ and hence, $I(g) + I(h') = I(g') + I(h)$, or, $I(f) = I(g) - I(h)$ is well defined. This linear functional corresponds to a left Haar measure.

[4] This is an alternate characterization of compactness. The equivalence of the two definitions is basically rewriting the compactness definition of open sets and unions in terms of closed sets and intersections using De Morgan is laws.

Theorem 2.A.2. *Left invariant Haar measure is unique upto scaling.*

Proof

Suppose λ and μ are two left Haar measures on $\mathcal{G}$. To prove uniqueness we have to show that the ratio $\int f d\lambda / \int f d\mu$ is the same for all compactly supported functions f. So, it is sufficient to show for any $\epsilon > 0$,

$$\left| \frac{\int f d\lambda}{\int f d\mu} - \frac{\int g d\lambda}{\int g d\mu} \right| < \epsilon$$

Let f be a compactly supported function, let $K_f = \text{supp}(f)$. Since f is a continuous function, for every $\epsilon > 0$ there exists a neighborhood U_x of x such that,

$$|f(z) - f(x)| < \tfrac{1}{2}\epsilon.$$

for any $z \in U_x$. We can now translate this neighborhood to the identity $L_{x^{-1}} U_x$, and hence we have for every ϵ, a neighborhood of 1 such that $|f(yx) - f(x)| < \frac{1}{2}\epsilon$ for all y in the neighborhood. Choose a symmetric neighborhood V_x of 1 such that $V_x V_x \subset L_{x^{-1}} U_x$. The sets xV_x cover the compact set $K_f = \text{supp}(f)$, and hence there must exist a finite subcover, $K \subseteq \bigcup_{i=1}^{n} x_i V_{x_i}$.

Consider the neighborhood of 1, $V = \bigcap_{i=1}^{n} V_{x_i}$. Let $y \in V$. If $x \in K_f$ then there exists some i such that $x_i^{-1} x \in V_{x_i}$, so that $xy = x_i(x_i^{-1} x) \in U_{x_i}$. Then,

$$|f(xy) - f(x)| \leqq |f(xy) - f(x_i)| + |f(x_i) - f(x)| < \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon.$$

If both x and yx are not in K then $f(x) = f(yx) = 0$. Hence if f is a compact supported function, then it's left uniformly continuous, that is $\|L_y f - f\|_{\sup} \to 0$ uniformly.

Let $f, g \in \mathcal{C}_c^+(\mathcal{G})$, let V_0 be a symmetric, compact neighborhood of 1, and let $V_f = V_0 K_f \cup K_f V_0$, and $V_g = V_0 K_g \cup K_g V_0$. V_f and V_g are compact sets. For any $y \in V_0$, we have continuous functions $x \mapsto f(xy) - f(yx)$ and $x \mapsto g(xy) - g(yx)$ with supports V_f and V_g respectively. Note that their value at 1 is zero.

Hence for any $\epsilon > 0$, there exists a neighborhood $V \subset V_0$ of 1 such that,

$$|f(xy) - f(yx)| < \epsilon, \ |g(xy) - g(yx)| < \epsilon$$

Pick a nice function h such that $h(x) = h(x^{-1})$, with support in V. Then,

$$\int h d\mu \int f d\lambda = \int_{\mathcal{G}} \int_{\mathcal{G}} h(y) f(x) d\lambda(x) d\mu(y) = \int_{\mathcal{G}} \int_{\mathcal{G}} h(y) f(yx) d\lambda(x) d\mu(y).$$

Using $h(x) = h(x^{-1})$ we get,

$$\begin{aligned}
\int h d\lambda \int f d\mu &= \int_{\mathcal{G}} \int_{\mathcal{G}} h(x) f(y) d\lambda(x) d\mu(y) = \int_{\mathcal{G}} \int_{\mathcal{G}} h(y^{-1} x) f(y) d\lambda(x) d\mu(y) \\
&= \int_{\mathcal{G}} \int_{\mathcal{G}} h(x^{-1} y) f(y) d\mu(y) d\lambda(x) = \int_{\mathcal{G}} \int_{\mathcal{G}} h(y) f(yx) d\mu(y) d\lambda(x) \\
&= \int_{\mathcal{G}} \int_{\mathcal{G}} h(y) f(xy) d\lambda(x) d\mu(y).
\end{aligned}$$

Therefore,

$$\left| \int h d\lambda \int f d\mu - \int h d\mu \int f d\lambda \right| = \left| \int_{\mathcal{G}} \int_{\mathcal{G}} h(y) \big[f(xy) - f(yx) \big] d\lambda(x) d\mu(y). \right| \leqq \epsilon \lambda(V_f) \int h d\mu.$$

Similarly,

$$\left| \int h d\lambda \int f d\mu - \int h d\mu \int f d\lambda \right| \leqq \epsilon \lambda(V_f) \int h d\mu.$$

Dividing these inequalities by $\int h d\mu \int f d\mu$ and $\int h d\mu \int g d\mu$ and adding them, we get,

$$\left| \frac{\int f d\lambda}{\int f d\mu} - \frac{\int g d\lambda}{\int g d\mu} \right| < \epsilon \left[\frac{\lambda(V_f)}{\int f d\mu} + \frac{\lambda(V_g)}{\int g d\mu} \right].$$

Since ϵ is arbitrary, the difference must be zero, and hence λ and μ are unique upto a constant multiple. □

If $\mathcal{G}$ is non-commutative, the left Haar measures allows us to quantify the 'amount of non-commutativity'. For each $x \in \mathcal{G}$, define, $\lambda_x(H) = \lambda(Hx)$. This is again a left Haar measure, and by uniqueness of Haar measures, there exists a number $\Delta(x) > 0$ such that,

$$\lambda_x = \Delta(x)\lambda.$$

The number $\Delta(x)$ is independent of the choice of λ by the uniqueness Haar measure. So, $x \mapsto \Delta(x)$ is a well defined function called the modular function. For any $x, y \in \mathcal{G}$, and $H \subset \mathcal{G}$, we have, $\lambda_{xy}(H) = \lambda(Hxy) = \Delta(y)\Delta(x)\lambda(H)$, so,

$$xy \mapsto \Delta(x)\Delta(y),$$

or, Δ is a group homomorphism from the group $\mathcal{G}$ to the multiplicative group of positive numbers denoted by $\mathbb{R}_\times$. If χ_E is the characteristic function on H, then we have, $\chi_H(xy) = \chi_{Hy^{-1}}(x)$. So,

$$\int_{\mathcal{G}} \chi_H(xy) d\lambda(x) = \lambda(Hy^{-1}) = \Delta(y^{-1})\lambda(H) = \Delta(y^{-1}) \int_{\mathcal{G}} \chi_H(x) d\lambda(x).$$

Since a general function $f \in L^1(\mathcal{G})$ can be approximated by simple functions, we have,

$$\int R_y f d\lambda = \Delta(y^{-1}) \int f d\lambda.$$

or equivalently, this can be abbreviated as,

$$d\lambda(xy) = \Delta(y) d\lambda(x).$$

Since the map $y \mapsto R_y f$ is uniformly continuous, the map $y \mapsto \int R_y f d\lambda$ is a continuous map. So, Δ is continuous. So, the homomorphism Δ also respects the topological structure.

To each left Haar measure, we can associate a right Haar measure, $\rho(H) = \lambda(H^{-1})$. The modular modular function relates the two measures. If $f \in \mathcal{C}_c(\mathcal{G})$, then we have,

$$\int_{\mathcal{G}} R_y f(x) \Delta(x^{-1}) d\lambda(x) = \Delta(y) \int_{\mathcal{G}} f(xy) \Delta(xy^{-1}) d\lambda(x) = \int_{\mathcal{G}} f(x) \Delta(x^{-1}) d\lambda(x).$$

In the last step we changed the variable. Thus the functional $f \mapsto \int f(x)\Delta(x^{-1})d\lambda(x)$ is right invariant, and hence its associated Radon measure is a right Haar measure. Since ρ is a right Haar measure it must vary from the Radon measure associated to this functional by a constant, say κ. On symmetric neighborhoods, the left and right Haar measures are equal.

If the constant κ is not 1, then by continuity of Δ, we can choose a small enough symmetric neighborhood of $1 \in \mathcal{G}$ such that $|\Delta(x^{-1}) - 1| \leqq \frac{1}{2}|\kappa - 1|$ on that neighborhood.

$$|\kappa - 1|\lambda(U) = |\kappa\rho(U) - \lambda(U)| = \left| \int_U [\Delta(x^{-1}) - 1] d\lambda(x) \right| \leqq \tfrac{1}{2}|\kappa - 1|\lambda(U).$$

which can only happen if $\kappa = 1$. This gives us,

$$d\rho(x) = \Delta(x^{-1}) d\lambda(x)$$

This can be restated in a more convinient way as,

$$d\lambda(x^{-1}) = \Delta(x^{-1}) d\lambda(x).$$

If $\Delta \equiv 1$, in such a case the left Haar measures do not measure any non-commutativity, such a locally compact group $\mathcal{G}$ is called unimodular. Clearly, abelian groups are unimodular. For discrete groups, every element of the group has same measure, and hence whether we left translate or right translate, it would still have the same measure. So, discrete groups are also unimodular. Since Δ is a continuous homomorphism, it takes compact groups to compact subgroups of $\mathbb{R}_\times$, which can only be $\{1\}$. So, compact groups are unimodular, and if $K \subset \mathcal{G}$ is a compact subgroup, then $\Delta|_K \equiv 1$.

Let $[\mathcal{G}, \mathcal{G}]$ be the smallest closed subgroup of $\mathcal{G}$ containing all elements of the form, $[x, y] = xyx^{-1}y^{-1}$, called the commutator subgroup of $\mathcal{G}$. It is a normal subgroups since, $z[x, y]z^{-1} = zxyx^{-1}y^{-1}z^{-1} = [zxz^{-1}, zyz^{-1}]$. Since Δ is a group homomorphism, we have,

$$\Delta([x, y]) = [\Delta(x), \Delta(y)]$$

since $\mathbb{R}_\times$ is abelian, we have $[\Delta(x), \Delta(y)] = 1$. So, by the isomorphism theorem of groups, the group homomorphism $\Delta : \mathcal{G} \to \mathbb{R}_\times$ must factor through $\mathcal{G}/[\mathcal{G}, \mathcal{G}]$. This implies that if $\mathcal{G}/[\mathcal{G}, \mathcal{G}]$ is compact, then $\mathcal{G}$ is unimodular. This is intuitively obvious, the modular function was built to measure the non-commutativity and this just means that, $[\mathcal{G}, \mathcal{G}]$ does not contain any information about the non-commutativity contained in the group $\mathcal{G}$. Intuitively, the important thing to note about the modular function is that the left Haar measures can only measure non-commutativity of 'large' groups in some sense.

Summaries

Abstract

In the first chapter we reconstruct algebraic quantum theory, and prove Sakai's characterisation of von Neumann algebras following the approach of Takesaki and Tomiyama. In the second chapter we impose constraints of relativity on the objects of quantum theory and study the consequences. In particular we study the constraints imposed by the positivity of energy, and its relation with the notion of locality. In particular we prove a generalised version of the Reeh-Schlieder theorem. We then review the Tomita-Takesaki modular theory and a related theorem due to Borchers.

Zusammenfassung

Im ersten Kapitel rekonstruieren wir die algebraische Quantentheorie und beweisen Sakais Charakterisierung von von-Neumann-Algebren, wobei wir dem Ansatz von Takesaki und Tomiyama folgen. Im zweiten Kapitel betrecten wir zusätzliche Beschränkungen aus der Relativitäts theorie die Objekte der Quantentheorie fest und untersuchen die Konsequenzen. Insbesondere befassen wir uns mit den Beschränkungen, die sich aus der Positivität der Energie ergeben, und deren Beziehung zur Begrifflichkeit der Lokalität. Insbesondere beweisen wir eine verallgemeinerte Version des Reeh-Schlieder-Theorems. Anschließend geken wir einen Überblic über die modulare Theorie von Tomita-Takesaki und einen verwandten Satz von Borchers.

[1] E. Alfsen, Compact Convex Sets and Boundary Integrals. Springer-Verlag, 1971

[2] H. Araki, Einführung in die axiomatische Quantenfeldtheorie, Vorlesungen an der E.T.H. Zürich, I.W.S, 1961

[3] W. B. Arveson, On groups of automorphisms of operator algebras, J. Funct. Anal. 15, 21,7 1974

[4] L. Ballentine, Quantum Mechanics: A Modern Development. World Scientific, 1998

[5] V. Barbu, T. Precupanu, Convexity and Optimization in Banach Spaces. Springer, 2020

[6] H. J. Borchers, Translation Group and Particle Representations in Quantum Field Theory, Springer, 1996

[7] H. J. Borchers, On the Converse of the Reeh-Schlieder Theorem, Commun. math. Phys. 10, 269-273, 1968

[8] H. J. Borchers, The CPT-Theorem in two-dimensional theories of local observables, Commun. Math. Phys. 143, 315, 1992

[9] O. Bratteli, D. Robinson, Operator Algebras and Quantum Statistical Mechanics I, Second Edition, Springer, 1987

[10] P. Busch, M. Grabowski, P. Lahti, Operational Quantum Physics. Springer, 1995

[11] R. Correa, A. Hantoute, M. López, Fundamentals of Convex Analysis and Optimization. Springer, 2023

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

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

[14] M. Florig, Geometric Modular Action, PhD Thesis, 1999

[15] G. Folland, A Course in Abstract Harmonic Analysis, CRC Press, 1995

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

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

[18] R. Haag, D. Kastler, An Algebraic Approach to Quantum Field Theory, Journal of Mathematical Physics, 5 (7): 848–861, 1964

[19] R. Haag, Local Quantum Physics: Fields, Particles, Algebras, 2nd Edition, Springer, 1992

[20] P. R. Halmos, S. Givant, Logic as Algebra. Dolciani Mathematical Expositions. Vol. 21, 1998

[21] L. Hörmander, The Analysis of Linear Partial Differential Operators I, Springer-Verlag, 1980

[22] L. Hörmander, Notions of Convexity. Birkhäuser, 1994

[23] V. F. R. Jones, Von Neumann Algebras, Lecture Notes, 2010

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

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

[26] S. Kontorovitz, A. Viselter, Introduction to Modern Analysis. Oxford University Press, 2022

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

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

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

[30] G. Ludwig, Foundations of Quantum Mechanics 2. Springer-Verlag, 1983

[31] C. L. Merdy, An Operator Space Characterization of Dual Operator Algebras. American Journal of Mathematics, Vol. 121, No. 1 (Feb., 1999), pp. 55-63

[32] G. Murphy, C^*-Algebras and Operator Theory. Academic Press Inc., 1990

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

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

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

[36] S. Sakai, C^*-Algebras and W^*-Algebras. Springer, 1970

[37] S. Sherman, The Second Conjugate of a C^*-Algebra. Proc. Inter. Congress Math. 470. 1950

[38] Z. Takeda, Conjugate Spaces of Operator Algebras. Proc. Japan Acad. 30, 1954

[39] M. Takesaki, On the Conjugate Space of Operator Algebra. Tohoku Math J, 10, 1958

[40] M. Takesaki, Theory of Operator Algebras I, Springer, 2000

[41] M. Takesaki, Theory of Operator Algebras II, Springer, 2000

[42] J. Tomiyama, On the Projection of Norm One in W^*-Algebras. Proc. Japan Acad. 33, 1957

[43] J. Tomiyama, On the Projection of Norm One in W^*-Algebras II. Tohoku Math. J.. 33, 10., 1958

[44] J. Tomiyama, Tensor products and projections of norm one in von Neumann algebras. Lecture notes, University of Copenhagen, 1970.

[45] T. Turumaru, On The Direct Product of Operator Algebras III. Tohoku Math J. (2), 6(2-3), 208-211, 1954.

[46] S. Vickers, Topology via Logic. Cambridge University Press, 1989

[47] J. Voigt, A Course on Topological Vector Spaces. Birkhäuser, 2020

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

[49] H. W. Wiesbrock, Half-Sided Modular Inclusions of von-Neumann-Algebras, Commun. Math. Phys. 157, 83-92, 1993

[50] E. C. Zeeman, Causality Implies the Lorentz Group, Journal of Mathematical Physics, 5, 490, 1964