Operator Algebras and the Foundations of Quantum Mechanics

By

Abdullah Naeem Malik

Department of Mathematics
Quaid-i-Azam University
Islamabad, Pakistan
2016

Operator Algebras and the Foundations of Quantum Mechanics

By

Abdullah Naeem Malik

Supervised By

Dr. Tayyab Kamran

Department of Mathematics
Quaid-i-Azam University
Islamabad, Pakistan
2016

Operator Algebras and the Foundation of Quantum Mechanics

QUAID-I-AZAM UNIVERSITY
ISLAMABAD

By

Abdullah Naeem Malik

A DISSERTATION SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF PHILOSOPHY
IN
MATHEMATICS

Supervised By
Dr. Tayyab Kamran

Department of Mathematics
Quaid-i-Azam University
Islamabad, Pakistan
2016

Acknowledgments

One can always resort to being an ingrate and ego-centric but convention dictates that we be hackneyed. For most, this page is usually written with an overwhelmed heart as one completes a creatively frustrating, year-long journey. Such journeys are usually isolated ones, which is where one derives comfort by using the word “we” through-out an academic manuscript, indicating that such painstakingly secluded expeditions are not to selfishly leave a mark in the world but to contribute to one unified system of (useful or useless) knowledge.

Such motivations work best when one is given the liberty to paint whatever picture one wants to. Indeed, my supervisor, Dr. Tayyab Kamran, deserves merit for allowing me to toy with whatever idea I wanted to the extent of being ludic, which was complemented by dropping hints all along of believing in me. I earnestly hope I have not been a source of disappointment. I would also like to mention my Bachelor’s supervisor, Dr. Tasawar Abbas, who introduced me to the beautiful world of Quantum Mechanics, for, without him, I would not have been keen on knowing how the world works.

My endeavours in the academia were always fuelled by my father’s kvell. Unfortunately, he passed away midway during the preparation of this report. My family members and my class mates were present to push my spirits whenever they dwindled. Without them, I would not have written this report in the first place.

Abdullah Naeem Malik

To my late father

Operator Algebras and the Foundations of Quantum Mechanics

by

Abdullah Naeem Malik

Submitted to the Department of Mathematics
on 25 January 2016, in partial fulfillment of the
requirements for the degree of
Master of Philosophy

Abstract

The purpose of this thesis is to analyse the Hilbert Space requirement for Quantum Mechanics. In particular, we justify sharp observables but question the requirement of completeness of the inner product space and the underlying field. We view our mathematical framework as a dynamical theory but with a mysterious probabilistic interpretation instead of the otherway round. Whenever we speak of "Quantum Mechanics", we mean Non-relativistic Quantum Mechanics. To make things less messy, we assume associativity through-out. No attempt has been made to refer to QFT and statistical quantum mechanics and we use conventional mathematical symbols instead of Dirac's formalism.

Thesis Supervisor: Dr. Tayyab Kamran
Title: Associate Professor

Contents

Chapter 1

Preface to Hilbert Spaces

Natura non facit saltus - Liebniz

Any working mathematical physicist's indispensable tool is the area of functional analysis in which there exists a natural structure to be able to model physics. The idea of a complete space corresponds to being able to find a solution in an iterated sequence of steps or experiments, provided that the results of the sequence become closer as they progress, enough to perform calculus. The implications are not limited to physics: indeed, functional analysis has a major role in the field of numerical analysis, geometry and even number theory.

The dot product finds major use in geometry. The work of David Hilbert on integral equations in a series of papers between 1904 and 1912 [14] found such a striking similarity between this dot product of vectors and a particular product of functions – the ordinary integration. This initial motivation from work on Integral Equations prompted a formal axiomatic framework of these ideas, which was furnished by John von Neumann [33][1], a by-product of which was the equivalence of the Schrödinger approach and the Hiesenberg approach in Quantum Mechanics [25].

The formulation of a Hilbert Space $\mathcal{H}$, now modern analysis, was a quantum leap from classical analysis. The latter was confined to Real and Complex Analysis and the solution to different equations. After Cantor's paradise and Cauchy's $\epsilon - \delta$ approach, analysis took a different turn. The focus was shifted to foundational aspects. Hints of topology and algebra

[1] Contrary to popular belief, the paper was published in 1930, not 1929, as seen on Springer's official website.

began to reveal themselves in classical analysis. This revelation was not an isolated experience: starting with the work of the giants viz. Gauss, Riemann, Grassman, Hadamard, Fréchet, Hausdroff [28] in the budding for Non-Euclidean Geometry, Vector Spaces, (Analytic) Function Spaces, Metric spaces and finally Topological Spaces, respectively, thereby the new field of modern analysis was born. Functions (and just analytic ones) were now considered as points on a space, much like real numbers on a line.

To see the analogy, consider the popular dot product in finite dimensions

$$\langle a, b\rangle = \sum_{i=1}^{n} a_i \overline{b_i}$$

For the infinite case, one has

$$\langle f, g\rangle = \sum_{k=1}^{\infty} f(x_k)\, \overline{g(x_k)}$$

where f,g are functions $f : \mathbb{N} \longrightarrow \mathbb{R}$ i.e. sequences. In most cases, this is infinite unless one considers functions f such that

$$\sum_{k=1}^{\infty} |f(x_k)|^2 < \infty$$

This is, by current standards, a point in l^2 space. Using point-wise addition and scalar multiplication, one can easily see that this as a vector, an element of a special space once reserved only for the Euclidean space. These analogies were strengthened by the Banach-Mazur theorem, which states that every Banach space is isometrically isomorphic to a subspace of a continuous scalar functions on a compact Hausdroff Space. We adopt the approach of Hilbert himself for a justification. A general form of an integral equation, as conceived by David Hilbert [14], is

$$f(x) = \varphi(x) - \lambda \int_a^b K(s, t)\, \varphi(t)\, dt$$

where φ is unknown, K is kernel of the equation. In current literature, this is called Voltera's Integral Equation of second kind[2].

In terms of countable basis, the integral equation then took the form

[2] Hilbert called this equation Integral Equation of Second Kind [14]

$$f_k + \sum K_{kl} f_l = g_k$$

In terms of looking at K_{kl} as Fourier Coefficients, from Parvesal's relation[3], we can readily obtain

$$\sum_{k\in\mathbb{N}} |f_k|^2 = \int_a^b |f(x)|^2 \, dx$$

We explain this leap: the trouble with direct substitution itself is that then uncountably many basis are considered whereas the sum over an uncountable set is always infinite.

Proof. We prove the contrapositive: let $(x_\alpha)_{\alpha\in A}$ for some uncountable $A \subseteq \mathbb{R}$ be a sequence such that

$$\sum_{\alpha\in A} x_\alpha = M < \infty$$

Consider $S_n = \left\{\alpha \in A : x_\alpha > \frac{1}{n}\right\}$ Then, $M \geq \Sigma_{\alpha\in S_n} x_\alpha > \Sigma_{\alpha\in S_n} 1/n = \frac{N}{n}$ where $N \in \mathbb{N}$ or $N = \infty$ is the number of elements in S_n. Thus, $|S_n| < Mn$. Hence $\{\alpha : \alpha \in A \text{ and } x_\alpha > 0\} = \cup S_n$ is countable. That is, $x_\alpha = 0$ for all but at most countably many α ■

Of course one can, in principle, use the supremum to account for the uncountable case but that would have to be a generalisation in terms of integrals via measure theory and that, too, using the counting measure, thus assuring the exclusion of a large class of functions. Hilbert's genius lay in recognising infinite orthonormal (countable) basis for functions, thus superseding the limitation. Such basis could be, say, in terms of $e_i = x^i$ with Maclaurin's series expansion of a function for arbitrary domains and $e^{2\pi ikx}$ in terms of Fourier Series expansion for the unit interval as a domain. Under this approach, the inner product then took a natural form

$$\langle f, g\rangle = \int_a^b f(x)\,\overline{g(x)}dx$$

implying that $f \in l^2(\mathbb{N}) = L^2(\mathbb{N}, m)$ where m is the counting measure. The justification later came from what is now known as Riesz-Fisher Theorem: for any sequence (c_k) and any

[3] A variant of the 1799 Parvesal's relation says that if two square-integrable complex-valued functions $f(z)$ and $g(z)$ have Fourier series $f(z) = \sum_{n=0}^{\infty} f_n e^{inz}$ and $g(z) = \sum_{n=0}^{\infty} g_n e^{inz}$, then $\sum_{n=0}^{\infty} f_n \overline{g_n} = \int_a^b f(z)\,\overline{g(z)}dz$

orthonormal system $(e_k) \in L^2(a,b)$, there exists a function $f \in L^2(a,b)$ such that $\langle e_k, f \rangle = c_k \iff (c_k) \in l^2$ [26], later furnishing an isomorphism between the two spaces, thus concluding that the "sequential approach" is the same as the "integral approach". Note that, however, not all continuous functions fall under the space $L^2\left(a,b\right)$ – only smooth functions with a compact support do with an added completion [16] without resorting to the definition of a measure. Technically, this implies the required

$$\int_a^b |f\left(x\right)|^2 \, dx < \infty$$

Finally, in order to show that the idea of an abstract Hilbert space was accurately captured in the work on integral equations[14], we mention the following important theorem: every separable Hilbert space is isomorphic to a subspace of L^2. Thus, the importance of square integrable functions, one instance of which is Schrodinger's wave equation, is not without its importance, to which we now turn.

Chapter 2

Formulation and Limitation of Quantum Mechanics

> All of science is uncertain and subject to revision. The glory of science is to imagine more than we can prove. Freeman Dyson

Mathematics is a rigorous enterprise, whereby axioms are formulated and theorems are explored using water-tight reasoning. In a similar vein, physical theories are formulated using the language of mathematics and predictions are made. For those who insist that the job of a physical theory is to explore nature, they are reminded that such theories have to be tested, thanks to the ideas of Karl Popper [38]. They are thus reminded that a mathematically elegant which does not explain facts is useless from an empirical point of view. Competing theories are weiged using Occam's Razor. Any mathematical theory or physical theory is, therefore, accepted if it has best explanatory power (with best compression) using the most primitive axioms. For us, the Hilbert Space requirement of Quantum Mechanics, though mathematically solid (as we shall see), is too much to ask for.

In any physical theory, one needs to be able to precisely mention what physical states are in a phase space, define the evolution of such states and define the measurement of observables. Measurement is usually not a part of a physical theory and the headache is transferred to the experimentalists. In particular, in classical physics, measurement poses no problem since it does not influence the state of the system. In Quantum Mechanics, the situation is quite different and, therefore, such a need has to be taken into consideration, as well[1].

A mathematical language for a physical theory is very important not only because ordinary language is full of ambiguities but because ordinary languages instill in their use an interpretation. Any physical theory based on a mathematical formalism, and therefore of very abstract nature, must be free of any interpretations or at least provide as little interpretations as possible. In the latter case, the situation is of a great advantage if such interpretations are universally agreed upon. Thus, the accommodation of a clear and well-defined notion of a state and an observable with compatible mathematics describing the dynamics of the state ought to be the aim.

For example, in classical mechanics, Newton's laws come to the fore, with an equivalent form in Hamiltonian mechanics. States are described by a position and a momentum coordinate (q, p) with appropriate degrees of freedom in a compact phase space M (ensuring that the finiteness of the human mind is not boggled), at least locally homeomorphic to the Euclidean space (ensuring the applicability of calculus) and a symplectic 2-form, giving rise to the name symplectic manifold M. More accurately, such phase state corresponds to a pure state and with the tools to calculus available, a state of the system characterises questions (relating to energy, momentum etc.) which can be asked and answered conclusively. This is because of the associated differential or integral equations of the system with initial or boundary conditions. Thus, it turns out that these tools suffice for everyday classical mechanics. Observables may be described by continuous functions on M (thus depend upon the state of the system) so that measurements can be made arbitrarily accurate: for any ϵ error of observation we wish to tolerate, we can be assured of finding (q, p) at a $\delta-$distance of the "real" value (q_0, p_0). If the observable turns out to be smooth (e.g. Hamiltonian), it makes life much easier as one can avoid pathologies. At any rate, continuity assures that the function itself is nice enough for analysis to take its stronghold and that, in principle, one can measure to an arbitrary accuracy. In practice, the scale of classical mechanics is large enough to make uncertainities optimistically trivial. It is when we go "down below" that things start to get fuzzy.

Any experimental apparatus simply obtains a value within some uncertainty. Consider a

[1] We shall forego the discussion of such a postulate because of its controversy. For an excellent coverage of the measurement problem from a vantage point, see [27].

simple length measurement in which, besides using a ruler, one might use lasers but this, too, depends upon the wavelength which is determined by the frequency and velocity of the wave, each of which has to be measured or calculated by making other measurements. In a stubborn insistence, an accuracy to 20 digits is still not guarantee enough to allow for observables to be described by continuous functions. Empirical considerations aside, such functions need a mathematical structure to stand upon. For a simple three-dimensional system comprising of n particles, one needs a $3n$-dimensional Euclidean space. Our grandiose assumptions land us right into trouble: gas dynamics and our current computation theories teach us that many problems require resources greater than the universe can offer[1].

It is, therefore, wise to invoke averages of a measurement. This assumption, though clearly needed for quantum mechanics, justifies its use in classical mechanics since if repeated measurements (assuming no changes to the system!) give (nearly) similar observations, then the average stands close to what we observe each time, reduces random errors and provides us with a very low standard deviation.

This approach rests on the supposition that an observable ought to yield a definite value of a system each time it is observed and that the value is close to its "real value". This might make Quantum Physicists subscribing to the Copenhagen interpretation cringe. One cannot be agnostic to an interpretation of a physical theory. For Quantum Mechanics, there is an array of interpretations of Standard Quantum Mechanics. For instance, in order to restore determinism, Bohm's wave mechanics comes to the fore. On the other hand as in [31], one can give due consideration to density matrices altogether and not pure states to conform to the tastes of the experimentalists. Quantum Mechanics can also be formulated as a branch of statistics and information theory. It is not unnatural to expect a now popular opinion: Quantum Mechanics is only seen as a statistical tool used to calculate probabilities, not to make predictions [31]. Obeying a formalist school of thought is, in a mathematical sense, the last resort when all other philosophies appear to be thick forests with roads going nowhere. Worse is the case when there are different formalisms of Quantum Mechanics which stem from different philosophies themselves.

In what follows, we shall lean halfway: for us, Quantum Mechanics is still a statistical theory but with an added problem of interpreting our achieved results. Following Occam's

Razor, we shall start off with the most general mathematical considerations stemming from physical motivations and see where they lead us to.

Usually, the focus is on trying to define a simple system and later on upping the ante by forming a multi-dimensional space using the simple space via tensor products with the following requirements [39]:

1. The tensor product should be linear in each factor so that change of basis in one factor changes the basis in the resulting tensor product
2. The product of operators acting on each factor should have similar behaviour
3. An inner product which assures completeness (or can be completed) by corresponding to the natural structure must be produced.

We begin by considering probabilities: in a spin-half particle, if x represents spin-up and if y represents spin down and if the probabilities associated with them are, respectively, α and β, then experimental work suggests that such a particle will exist in a superposition, if not measured and left to evolve on their own, for otherwise, the experiment does not make sense. This principle of superposition has stood the test of many experiments [4] and we shall adopt it. Our starting point is, therefore, $S = \alpha x + \beta y$ with $\alpha + \beta = 1$, an element of a vector space. These "weights" may be negative and even imaginary, making the correspondence with probabilities less intuitive[2]. The Manhattan Norm corrects this and satisfies $\|S\| = 1$, provided that x and y are taken as our basis. Such a state is pure in the sense that maximal information can be extracted from a (compatible) set of observables from such a vector. A mixed state, which is a statistical mixture of pure states, is what we end up with if we take the outer product of the vector with itself.

The eigenstates themselves are very important for a change of basis does not necessarily guarantee a preservation of the norm[3].

The Euclidean norm is usually selected for its experimental support of Born's rule under the lens of the Copenhagen interpretation. The geometric correspondence, comfortable association

[2] In consideration of the length of the thesis afforded, we shall choose to stay silent on whether or not the individual probabilities are *a priori* or not.

[3] See Example 3.1 in [24] for a counter example

with intensity of a wave and use of unitary operators has customarily set this norm to Euclid's. As of now, we shall be liberal and, therefore, only require that $\|x\| = 1$. This requires a norm to be defined on the vector space, or at least a semi-norm. The insistence not to have a non-zero state with zero norm takes away the freedom of having a state perpendicular to itself. In particular, it loses the correspondence to the Minkowski's product and, therefore, starts off badly without accord to General Relativity[4]. Our concern at the moment is with elements of semi-norm one, only.

And now, for the natural question: under what conditions are x and y the basis? To this, we start with the observation that eigenvalues have played a pivotal role in nearly all of physics, inspired by a famous question of Mark Kac, "Can you hear the shape of a drum?" [21], which gave rise to the differential equation $\Delta u = \lambda u$. The approach to a possible solution is deep: if one can know the eigenvalues, then the eigenstate is completely determined. Thus, in some sense, an eigenvalue "pins down" the value of an eigenstate, implying a tremendous computational ease if such (quantum) eigenstates are our basis for then the (semi-)norm of a pure state will correspond to the total probability.

In vector spaces, operators are the natural candidates for defining the evolution of a state and the observables. As far as the evolution is concerned, we need operators that do not change the unit probability of the state as it evolves in time. We can make do with an isometric operator but these are not necessarily compatible with the inner-product structure (more on this later). Unitary operators serve this gap. We would also want operators that generalise the real and complex numbers to correspond closer to the formulation of Classical Mechanics[5]. Normal operators do the job for the latter and are even more general than unitary operators but for the former, we have symmetric, self-adjoint and Hermitian operators. Which operator to choose? The Uncertainty Principle, which is what makes Quantum Mechanics fundamentally different, is a natural consequence of the fact that observables do not necessarily commute[6], and, therefore, not of primary concern. Our focus should be on the eigenvalues the operators produce. For some vague reason, physicists are more comfortable with real eigenvalues. Hermitian operators

[4] See §5

[5] Technically, quantisation, which is a map Q from the algebra of continuous functions $\mathcal{A}$ to the space of linear operators $L(\mathcal{H})$

[6] See §8.1

are usually resorted to but this, in itself, incorporates linearity. In particular, Schrödinger's correspondence with the Hamilton-Jacobi equation has given away the impression that the Quantum world is fundamentally linear, which has made (linear) differential equations and matrices the favourites of physicists. This is a subtle issue that is in direct contradiction with measurement processes. Why must the evolution of a state be linear and deterministic and its measurement be completely opposite? We remark that non-linear unitary, normal and adjoints have been defined for Quantum Mechanics [13].

As a recollection, we add the oft confounded definitions:

T is symmetric if $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in \mathcal{D}(T)$.

T is Hermitian if it is symmetric and bounded and $\overline{\mathcal{D}(T)} = \mathcal{H}$.

T is said to be self-adjoint if it is symmetric and $\mathcal{D}(T) = \mathcal{D}(T^*)$.

T is said to be isometric if $\|Tx\| = \|x\|$.

T is said to be unitary if it is a surjective isometry.

Here, $\mathcal{D}(T^*)$ is the set of all $y \in \mathcal{H}$ such that $|\langle Tx, y\rangle| \leq k_y \|x\|$ for all $x \in \mathcal{D}(T)$. If T is merely symmetric, it follows that $\mathcal{D}(T) \subset \mathcal{D}(T^*)$. Regardless of boundedness or even completeness, one can still define an adjoint using a dense domain, without using Riesz's Representation Theorem[7].

As far as observables are concerned, Quantum Mechnanics works neatly with eigenvectors as a basis for the underlying space. For any operator A, not all vector spaces will furnish the operator with an eigenvector basis. On $\mathbb{R}^n$ ($\mathbb{C}^n$, respectively) an operator is symmetric (normal, respectively) if and only if it admits an orthonormal basis of $\mathbb{R}^n$ ($\mathbb{C}^n$, respectively) comprising of eigenstates with eigenvalues from the underlying field [26]. The complex case is inevitably involved because in the real case, an operator may have no spectral values [26]. In fact, the spectrum of any (bounded!) operator on a complex space is non-empty and compact [26]. Thus, a generalised move to the complex case is inevitable, which was provided by von Neumann in what is now called the Spectral Decomposition Theorem: if the compact, linear operator A is self-adjoint, then there is such a complete orthonormal system comprising of non-zero eigenvectors of A. As a corollary, we can even decompose self-adjoint operators using projection operators [26] which is where we get projection valued measures from and hence a

[7] See §5

host of conflicting interpretations.

It is easy to define the adjoint of any linear operator T even for Banach spaces: simply define $g\left(T\left(x\right)\right)=\left(T'g\right)\left(x\right)$ for any linear operator T for any linear functional g. This obeys $\|T'\|=\|T\|$ by the Hahn-Banach theorem, for $T\in B\left(X,Y\right)$ where X,Y are Banach spaces. If $T\in L\left(X,Y\right)$, then $T'\in L\left(Y',X'\right)$ by construction where $X'=B\left(X,\mathbb{K}\right)$ is the topological dual. In such a case, we cannot compose these two operators, so we take Hilbert spaces instead of Banach spaces where the composition of an operator with its adjoint can be defined.

Notice how swiftly we've shifted to a requirement of completeness! One reason why this is needed is as follows: in a space where one can have a reasonable norm, completeness is equivalent to the condition that every absolutely convergent series converges. In fact, the exponentiation of an operator, which is a definite requirement for Schrödinger's equation, may not be defined if completeness were taken away.

Not only does one need a complete inner product space but a separable Hilbert space in order to accommodate for a countable basis.

As of now, we are giving the impression that a Hilbert space is a phase space for quantum mechanics. Nothing could be further from the truth because a phase space $(x,p)\subseteq\mathbb{C}^2$ of position and momentum treats both on equal footing. In the words of Max Born, "at every instant a grain of sand has a definite position and velocity. This is not the case with a [quantum particle]" [10].

For a moment, let us assume that the observables ought to be represented by linear, self-adjoint operators. In light of the success of Hamilton-Jacobi equation in classical mechanics, Schrödinger suggested a now famous equation $E\psi=H\psi$ which successfully explained electron orbits for the Hydrogen atom [10]. In the stationary case, E is a real eigenvalue. A similar equation holds for $\psi=\psi\left(t\right)$ and $E=i\frac{h}{2\pi}\frac{\partial}{\partial t}$. The Hamiltonian operator H is constructed, depending on the physical situation, as a sum of the total energy of the system. This will involve kinetic energy and potential energy so that momentum and position are inevitably put on the same footing. These observables are self-adjoint and since the sum of two self-adjoint operators is self-adjoint, the Hamiltonian operator is self-adjoint.

By Stone's theorem, there is a one-to-one correspondence between any (not necessarily bounded) self-adjoint operator A and continuous (in the strong topology) one-parameter unitary operator $U(t)$ such that $U(t+s) = U(t)\,U(s)$ by defining $U(t) = e^{\frac{2\pi i}{h}tA}$ where

$$e^{\frac{2\pi i}{h}tA} = \sum_{n=0}^{\infty} \frac{\left(\frac{2\pi i}{h}tA\right)^n}{n!}$$

In effect, a state $\psi(t)$ evolves with time such that $\psi(t) = U(t)\,\psi(t_0)$ where $\psi(t_0)$ is the initial condition of the state. It is pointed out that this unitary operator is determined by the Hamiltonian of the system. Applying the derivative (possible because of Stone's theorem) on both sides of this equation gives us the Schrödinger equation $E\psi(t) = H\psi(t)$. Thus, $\psi(t) = U(t)\,\psi(t_0)$ if and only if $E\psi(t) = H\psi(t)$

Schrödinger's picture (Wave Mechanics) requires the underlying Hilbert space to be the space of square integrable functions. States are (wave) functions of unit norm and operators do not depend on t. Hiesenberg's idea (Matrix Mechanics) was that the state of a system does not depend on t but the observables do. The only difference is that the observables are not required to be self-adjoint so that we are not guaranteed (eigen)basis of our choice. This corresponds to an active transformation and a passive transformation, with equivalence established via Stone's theorem[8].

Since non-determinism is fundamental to our approach, we must have a way of determining the expectation value for any observable A. In case the vectors and the scalars are of the same nature (i.e. the vector space is a field over itself), then the answer is straight-forward: for any quantum state x, we multiply all the probabilities with the eigenvalues λ of A associated with x and add them up. What are these probabilities? So far, we have not spoken of probabilities but only of weights of any unit vector. These weights become probabilities once we take the square of their absolute values, thanks to Gleason's theorem, with the only requirement being that the underlying space be at least three dimensional and that $\|x\| = 1$. Where do we get this from? To answer this question, we move to the case where the vector space is a not a field over itself. The expectation value has to be a linear functional. By Hahn-Banach theorem, we are guaranteed a linear functional g such that $\|g\| = 1$ and $g(Ax) = \|Ax\| \leq \|A\|$, which is (or its additive inverse is) an eigenvalue of x. What if we do nothing to the state (i.e., $A = I$)? In such

[8] Dirac went a step ahead and assumed that both states and observables vary with t (interaction picture), which formed the basis of Quantum Field Theory.

a case, we must expect $\|Ax\| = 1 = \langle x, x \rangle$. By Riesz Representation theorem, this functional can be an inner product such that $g(Ax) = \langle y, Ax \rangle$ where $\|g\| = \|y\|$. What, indeed, is y? If $y \neq x$, then for $A = I$, we arrive at a contradiction to the required definition of expectation value in general and Gleason's theorem in particular. Thus, $y = x$.

As of now, we have settled the fact that a pure state is an element x of a Hilbert space $\mathcal{H}$ such that $\|x\| = 1$, inflicting a pinch to the liberty of using all of that which is not forbidden but rather by reasoning that a Hilbert space and its beauties place us at an advantage. However, the argument to require a norm and not semi-norm is rather not convincing at the primitive level of the axioms. We also have a sloppy rationale for defining observables. Whatever the definition of observables, we know that not all of them are compact: from $[Q, P] = i\frac{h}{2\pi}I$, it follows that $tr(QP) - tr(PQ) = i\frac{h}{2\pi} = 0$, a contradiction (trace is only defined for compact operators). Thus, some observables will then not map compact sets to relatively compact sets and, therefore, will be unbounded. This seems to be like less of a trouble for physicists, even though unbounded operators may have an empty spectrum, even in the complex plane [26]. At any rate, the set of states according to our definition, is a closed unit ball, which is not precompact in the infinite dimensional case by Riesz's lemma under the norm topology. With this, we can not guarantee the existence of an orthonormal sequence (e_n) such that for any state x, a series of $\langle x, e_n \rangle$ converges, provided $\mathcal{H}$ is separable[9]. In fact, for any orthonormal sequence, using the fact that $\Sigma |\alpha_n|^2$ converges if and only if $\Sigma \alpha_n e_n$ converges (in which case Parvesal's relation and hence the norm by Gleason's theorem may fail, unless the orthonormal sequence is complete), we cannot even guarantee the Fourier expansion of any element! Worse yet is the nightmare when physicists usually replace self-adjoint operators with Hermitian ones. It seems as thought physicists are not concerned with the behaviour of operators on vectors other than those belonging to the unit ball. To dash further folklore regarding Hermitian operators, recall that such operators are densely defined. It is possible for two operators to have a dense but disjoint domain and corresponding range, in which case the composition of the operators becomes rather meaningless. The arguments involving trace of a compact operator is put to

[9] Let E be a precompact subset (in the norm topology) of an infinite dimensional, separable Hilbert space. Then there exists (e_n) such that $\sum_{n=1}^{\infty} \langle x, e_n \rangle < \infty$ for all $x \in E$

test as well: the trace of the identity map is infinity in an infinite dimensional Hilbert space.

There are some interesting consequences of introducing time as a parameter, not as an observable. Assume there was one T. Given the standard formalism of Quantum Mechanics, we should expect that $T\psi(t) = t\psi(t)$ but then this multiplication operator has an empty spectrum on $L^2[0,1]$. If this operator were symmetric, then we are still not guaranteed a real spectrum. Thus, time is taken as an intrinsic and non-relative parameter and not a dynamic variable, standing in direct contradiction to the basics of General Relativity in particular and against a common philosophical point of view, in general: since measurement irreversibly changes the state of the system, therefore upon a measurement of "time", the state of the system after the registration of the first arrival is no longer causally related to the initial state of the system or even just before the measurement was made. In essence, measurement rips apart the philosophy of causality.

According to Derek Lawden [29], it suffices to assume the eigenvalue/eigenvector approach where eigenvalues represent measurement outcomes. We associate with a quantity A the eigenvalues a_1, a_2, ... and eigenvectors ψ_1, ψ_2 , ... and with a compatible quantity B the eigenvalues b_1, b_2, ... and eigenvectors ϕ_1, ϕ_2 , ... Any mutually compatible set of quantities are assumed to completely describe a quantum system if (a) the results of none of the procedures can be inferred from the results of the others and (b) there exists no other procedure compatible with every member of the set, whose outcome is not derivable from the results of observations belonging to the set. This sense of completeness is attributed to Niels Bohr. It is to be remarked that completeness for Einstein meant that the state provides all the necessary information of a quantum system (the idea of a pure state). In order to study dynamics, the problem is to then compute the probabilities that the system moves from an initially observed state ψ_i to ϕ_j. Symbolise these as $P_{ij} = P\left(\psi_i \longrightarrow \phi_j\right)$ with the following assumptions:

1. $\sum P_{ij} = 1$
2. $P\left(\psi_i \longrightarrow \psi_j\right) = \delta_{ij}$
3. $P\left(\psi_i \longrightarrow \phi_j\right) = P\left(\phi_j \longrightarrow \psi_i\right)$

Any proposed dynamic law, apart from being able to determine probabilities $P\left(\psi_i \longrightarrow \phi_i\right)$ and $P\left(\phi_i \longrightarrow \chi_i\right)$ must also apply to a third, compatible quantity C with eigenvalues c_1, c_2,

... and eigenvectors χ_1, χ_2, ... It turns out that these suffice under the machinery of a Hilbert space with unitary and self-adjoint operators being a natural consequence. The first two assumptions are rather natural, given that probabilities are defined as in (1) and that every observable has an orthonormal decomposition of eigenvectors. The third is perhaps the most troubling, with a direct contradiction to the second law of thermodynamics in the global level.

Furthermore, since our idea of a pure state is a vector x of seminorm one, it follows that we cannot distinguish between x and λx where λ is a scalar such that $|\lambda| = 1$. For instance, consider adding a phase global $e^{i\theta}$.

Merely three years after setting a rigorous foundation of Quantum Mechanics, von Neumann himself is known to have had doubts regarding the use of Hilbert spaces in Quantum Mechanics. We quote a famous line, which summarises our arguments thus far:

> I would like to make a confession which may seem immoral: I do not believe absolutely in Hilbert space any more. After all, Hilbert space was obtained by generalising Euclidean space, footing on the principle of 'conserving the validity of all formal rules'. Now we begin to believe that it is not the vectors which matter, but the lattice of all linear (closed) subspaces. Because: 1) The vectors ought to represent the physical states, but they do it redundantly, up to a complex factor, only 2) and besides, the states are merely a derived notion, the primitive (phenomenologically given) notion being the qualities which correspond to the linear closed subspaces [40].

There are three notable successors to the Hilbert space formalism, initiated by von Neumann himself: the algebraic approach to quantum mechanics, quantum logics and rings of operators. The first approach is a particular algebra derived without the associative law – an approach which has begun to find experimental support [9]. The second has perhaps exceeded in producing most literature as it is a shift from sharp (eigen)values to closed subspaces themselves. In particular, the lack of correspondence (with the distributive law) identified by von Neumann was startling and suggested that Quantum Mechanics fundamentally obeys a different set of logic[10]. Of substantially greater success is his last notable contribution to the foundations of

[10] See §6.1

quantum theory, now von Neumann algebra or $W*$-algebra, a weakly closed subalgebra of $C*$-algebras (C stands for compact, a reference to the Banach-Algalou theorem), also known as an exceptional Segal algebra. These have the added advantage of easily accommodating an infinite quantum system. The key difference between Segal algebra and von Neumann algebra is that the former is uniformly closed whereas the latter is weakly closed.

From an empirical point of view, an algebra is necessary so that we can perform repeated measurements of observables. That is, we can form the sum and product of two observables to get another observable. One ought to be able to scale an observable as well: we would definitely want an experiment performed at CERN to be replicated in a small lab. Thus, our observables must be closed under scaler multiplication as well.

Instead of focusing on elements of a Hilbert space as states, the $C*$-algebraic approach targets observables by defining a canonical algebra of bounded operators $B(\mathcal{H})$. Of course this has the danger of excluding the momentum operator P for its unboundedness but there is a way: exponentiation. Then, e^{iP} is not only bounded but also unitary, thanks to Stone and von Neumann's theorem. This approach itself has the advantage of being coordinate free. States in such a setting are positive, unit-preserving functionals ω defined on $B(\mathcal{H})$ of norm one instead of vectors. Thus, states assign expectation values $\omega(A) = tr(\rho A)$ where ρ a density operator obtained by taking the outer product of a vector with itself. This approach, too, has its limitations. In particular, one cannot take the expectation value of two observables combined[11]. There are ways to achieve a correspondence with the Hilbert space approach via the GNS construction[12]. With this, we are guaranteed the existence of a vector corresponding to the vacuum state in particular and of a Hilbert space $\mathcal{H}$ in general. Thus, $\mathcal{H}$ corresponding to a quantum system can be recovered from the observables of the system. Mathematically, therefore, Einstein's and Bohr's approach are same sides of one coin.

There is a particularly striking observation of $B(\mathcal{H})$ which goes thus: such an algebra is non-commutative if and only if $\dim \mathcal{H} > 1$ i.e. if the underlying Hilbert space is not a subspace of the real line, then observables are bound to follow the Hiesenberg uncertainty principle so the only way one can "recover" Newton's physics from Quantum Mechanics is by building

[11] See §5
[12] See §8.2

a correspondence of the machinery of Quantum Mechanics with the real line – this is where the expectation values come in but only for operators with real eigenvalues. This piece of mathematics suggests that the decoherence approach is perhaps best suited as an explanation for a relationship with classical mechanics.

Even with all that standing, we have still remained silent on the choice of the underlying field and, therefore, the admitted topology. Quantum Mechanics is usually formulated on a complex field, which are weights (probability amplitudes in the Copenhagen interpretation), technically scalars. For classical mechanics, observables are real-valued functions so that even their expectation values are real. In Quantum Mechanics, this is described by self-adjoint operators, whose expectation values are real. A modified version of Quantum Mechanics can also work with $\mathbb{H}$ and $\mathbb{O}$. The latter, being non-associative, needs specialised algebra. Indeed, the associative algebras ($\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$) were determined to be on equal footing with a classification theorem:

Theorem 1 *Every normed division $*$-algebra is isomorphic to either $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ or $\mathbb{O}$[46]*

However, there are some subtle issues. For example, the tensor product of two Hilbert spaces over $\mathbb{H}$ is not a Hilbert space [6] and, therefore, such a quantum mechanical theory will only be valid for only the simplest of systems, though there are non-standard ways around this [39].

All in all, we see that self-adjoint operators pose technical problems (let alone Hermitian operators) and that the popular axioms of Quantum Mechanics are a result of a negation of the second law of thermodynamics. The idea of linearity in Quantum Mechanics is also illusory: the Schrödinger equation tells us that the observable A evolves as e^{iAt} but then the observables $A + B$ do not evolve with the phase $e^{i(A+B)t}$ unless A and B commute under multiplication. We can get as far as the following: the "mean" value of A and B combined is equal to their individual values.

Given the resources afforded for this thesis, we shall only attempt at answering two objections: one, why must sharp eigenvalues be considered? Two, why must the underlying field be the field of complex numbers? Thus, our approach will be as follows: we shall consider not necessarily linear operators over generalised fields and see how far the above reservations disappear.

Chapter 3

Preliminaries

The following sets basic terminology and definitions necessary for the main aim of this thesis. Most of the definitions are fairly elementary but are only included for consistency. Furthermore, for the sake of not converting this report into a set of notes, we exclude from this section various results. With this section, however, we hope to draw a route for the machinery of Quantum Mechanics with as little assumptions as possible. Knowledge of field theory is assumed but cited where necessary. Definitions can be found in [26] and [15].

Definition 2 *Let $n \in \mathbb{N}$. An n-ary operator (or a finitary operator) on a non-empty set S is a function $* : S^n \longrightarrow S$*

In such a case, S is said to be closed under $*$. For the special case of $n = 0$, called the nullary operation, $S^0 := \{()\}$, i.e. only one "n-tuple". A function, by definition, maps each element of the domain to exactly one element of the codomain. The nullary operation (an empty function) will map all elements of the empty set – which are non-existent – to S. Thus, the empty function will vacuously assign a value of the complete set to one single value (by definition of function) of S. This implies that the empty function is unique – a function that accepts no arguments and returns only one single value. We, therefore, assign a unique value $*(\varnothing)$. Thus, S is also said to be closed under the nullary operator by vacuous reasoning.

Definition 3 *Let A be a non-empty set. An algebraic system or an algebra is a tuple (A, O) with an operator set O elements of which act on A*

We assume that operators are not multivalued and are $n-$ary. This assumption will be justified in §5. O is always non-empty since it will contain at least the empty function. An example of such an algebraic system is a group G with $O = \{\hat{e}, *, \hat{\imath}\}$ where $\hat{\imath}$ is an 1-ary operator such that $\hat{\imath}(x) = x^{-1}$. The operator e is nullary with range G and, finally, $*$ is our familiar friend, the binary operator. A field does not form an algebra since we must have a restriction on the multiplicative inverse operator.

A non-empty subset B of A is called a subalgebra (B, O_B) if it is closed under the same n-ary operations it inherits. Technically, $O_B \approx O_A$ with the restriction that the domain of the operators in O_B is restricted to that of B.

The condition for A being non-empty is not technical and can be omitted but is only written out of laziness to avoid entering into details of empty functions applying on empty sets. This, however, does not technically stop us from asserting that an arbitrary intersection of subalgebras is an algebra since we've defined subalgebras to be non-empty sets themselves.

3.1 Order

The machinery of Quantum Mechanics (e.g. expectation values) requires an ordered vector space, which in turn needs an ordered underlying field, whether Archimedean or not (see **Definition 8**). Since we are assuming generalised fields, we shall need the following basic notions.

Definition 4 *A* ***prepositive*** *skew field* $(\mathbb{K}, +, \cdot)$ *is a skew field which contains a subset* $C \subset \mathbb{K}$ *such that*

C1 $x, y \in C$ *implies* $x + y, x \cdot y \in C$

C2 $\forall x \in \mathbb{K}$, $x^2 \in C$

C3 $-1 \notin C$

Clearly, $1 \in C$. Some authors explicitly exclude 0, referring to the latter as a blunt cone. In our case, $0 \in C$, which gives a sharp cone.

Also, $x \in C \Longrightarrow x \in \mathbb{K} \Longrightarrow x^{-1} \in \mathbb{K}$. By C1, $x^{-1} \in C$. Hence $C\backslash\{0\}$ is a multiplicative subgroup.

If $\mathbb{K} = C \cup -C$ and $C \cap -C = \{0\}$ where $-C = \{-x : x \in C\}$, then C is called a **positive cone** of $\mathbb{K}$. $x \in C \backslash \{0\}$ is called **positive**. At first sight, it might not be obvious but this definition leads to the usual:

Definition 5 *A skew field* $(\mathbb{K}, +, \cdot)$ *together with a total order* $\leq$ *on* $\mathbb{K}$ *is an* ***ordered skew field*** *if the order satisfies*

O1 if $a \leq b$ *then* $a + c \leq b + c$

O2 if $0 \leq a$ *and* $0 \leq b$ *then* $0 \leq a \cdot b$

To show that this approach is equivalent to the former, we first derive some necessary results: for every $a \in \mathbb{K}$, either $-a \leq 0 \leq a$ or $a \leq 0 \leq -a$.

Proof. If $a = 0$, then there is nothing to prove. Assume $a \neq 0$. Since $\leq$ is a total order, hence for any $a \in \mathbb{K}$, either $a \leq 0$ or $0 \leq a$. We assume both do not hold simultaneously since otherwise by antisymmetricity, $a = 0$, which we're not considering. Focusing on $a \leq 0$, then by O1, $a - a \leq 0 - a$ which implies $0 \leq -a$. Together, $a \leq 0 \leq -a$. The second conclusion follows similarly. ■

Also, for any $a, b, c, d \in \mathbb{K}$, we are allowed to "add inequalities": if $a \leq b$ and $c \leq d$, then $a + c \leq b + d$

Proof. $a \leq b$ and $c \leq d$ imply $0 \leq b - a$ and $\leq d - c$ by O1. By O1 again, $0 \leq b - a + d - c$ so that $a + c \leq b + d$ by O1 again. ■

We are also allowed to "multiply inequalities with positive elements": if $a \leq b$ and $0 \leq c$, then $ac \leq bc$.

Proof. By hypothesis, it follows that $0 \leq b - a$ and $0 \leq c$. By O2, $0 \leq (b - a)\, c = bc - ac$. The result follows by O1 after adding ac on both sides. ■

Lemma 6 *An ordered skew field has characteristic* 0.

Proof. Since $0 \leq 1$, then $0 \leq 1 + 1$ *ad infinitum.* If the skew field had characteristic p^k for some prime p and natural number k, then $-1 = p - 1 \geq 0$, a contradiction. ■

We now show that the positive cone approach and the order approach are equivalent.

Lemma 7 *There exists a bijection between the positive cones of* $\mathbb{K}$ *and total orderings on* $\mathbb{K}$

Proof. We first show that each positive cone gives rise to a total order on $\mathbb{K}$. Let $x, y \in C$. Define $x - y \in C \iff y \leq_C x$

Since $x - y$ is well-defined based on the axioms of a skew field, therefore the equivalence is well-defined. Next, $0 \in C$ hence $0 = x - x \in C$ so that $\leq_C$ is reflexive. Let $x \leq_C y$ and $y \leq_C x$. Then, $x - y, -(x - y) \in C$. That is, $a, -a \in C$. This is only possible if $a = 0$ since $C \cap -C = \{0\}$. Hence $x = y$. Furthermore, let $x \leq_C y$ and $y \leq_C z$. Then, $y - x, z - y \in C$. By C1, $y - x + z - y = z - x \in C$, which satisfies transitivity. Finally, let $x, y \in C$. Then, $x - y \in \mathbb{K}$. Since for every $a \in \mathbb{K}$, $-a \leq 0 \leq a$ or $a \leq 0 \leq -a$ holds, it follows that either $x - y \in C$ or $y - x \in C$. In either case, x is related to y.

Next, we show that each total order gives rise to a cone. This is possible when we define a subset C of $\mathbb{K}$ such that $x \geq 0 \iff x \in C$

It follows that $b - a \geq 0 \iff b - a \in C$. Let $a, b \in C$. Then, either $a \leq b$ or $b \leq a$ since $\mathbb{K}$ is a totally ordered skew field. We focus on only the first case. The proof for the second is similar. $a \leq b \implies 0 \leq b - a \in C$. Since $a \geq 0$, then $0 \leq b - a \leq b + a \in C$

Also, by O2, $ab \in C$. Hence C1 holds. Next, for any $a \in \mathbb{K}$, either $a \leq 0$ or $0 \leq a$. In the latter, by O2, $a^2 \geq 0$ applied on a twice. In the former, $a \leq 0 \implies 0 \leq -a$ by O1 and hence $0 \leq (-a)^2$ by O2 $\implies a^2 \in C$. Now, for $-1 \in \mathbb{K}$, either $-1 \leq 0$ or $0 \leq -1$. In the former, C3 is satisfied trivially. Assume the latter holds. Then, by O1 $1 \leq 0 \implies 1 \notin C$ but then if $0 \leq -1$, then $(-1)(-1) = 1 \in C$, contradiction. Hence $0 \not\leq -1$. ■

Definition 8 *An **Archimedean ordered skew field** is an ordered skew field $\mathbb{K}$ which obeys the Archimedean Property*: $\forall x, y \in \mathbb{K}$, $\exists n \in \mathbb{Z} \setminus \{0\}$ *such that* $nx \geq y$

Compounded by the well-ordering principle, if $\mathbb{K}$ is an Archimedean ordered skew field, we can define a bracket function $[x]$ to be $n - 1$ where n is the least $n \in \mathbb{Z}$ such that $x \leq n$. An ordered skew field that does not satisfy the Archimedean property is said to be **non-Archimedean ordered skew field**: if there does not exist a natural number n for positive x,y such that $nx < y$. In other words, if $0 < yx^{-1} < n^{-1}$ for all n ($\frac{y}{x}$, provided we have commutativity), then we have an infinitesimal, which is an element ϵ such that $0 < \epsilon$ and $\epsilon < r$ for every positive $r \in \mathbb{K}$. By definition, an infinitesimal number is not a real number but belongs to an extension of $\mathbb{R}$, $^*\mathbb{R}$, called the field of hyperreal numbers. Thus, negating the Archimedean Property leads us to numbers which are smaller than any positive number one can ever imagine. We shall not deprive ourselves of this excitement at all costs even if such "evanescent" quantities are not first order and, therefore, not a part of ZFC [32]. For us, our fields are ordered and we shall stay silent on the existence of infinitesimals but are only considered in §6.2.

Any ordering compatible with a skew field does not need to obey the Archimedean property, an example of which is the field of rational functions with real coefficients, apart from the example of the hyperreals.

3.2 Vector Algebra

For definition of vector space, see [26], which we shall call $\mathbb{K}$-vector space with $\mathbb{K}$ acting on the left and right, which we will always assume coincide. Fortunately, when it comes to Hilbert spaces, such actions are not limited to one side, even for skew fields (see Proposition 3.4b of [35]). If $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, then we shall call such a space **linear**. A vector space is an algebra in the sense of **Definition 3** with $O = \left\{\hat{1}, \hat{0}, \cdot_\lambda, +\right\}$ where $\cdot_\lambda$ is a family of unary scalar multiplication operators for each λ.

We can consider an order on a vector space as follows.

Definition 9 *Let V be a vector space over an ordered skew field $\mathbb{K}$. A* ***cone*** *is a subset of a vector space V such that $\mathbf{x}, \mathbf{y} \in M$ implies $\alpha\mathbf{x}+\beta\mathbf{y} \in M$ for any $\mathbf{x} \in M$ and $\forall\alpha, \beta \in \mathbb{K}$ such that $0 \leq \alpha, \beta$*

If $M \cap (-M) = \{0\}$, then the cone is said to be positive. In order to see that this induces an order on the vector space, we must define an order on a vector space first which is compatible with the vector space structure: a vector space V over $\mathbb{K}$ is said to be **ordered** under $\preceq$ if for all $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in V and $0 \leq \alpha \in \mathbb{K}$, we have that $\mathbf{x} \prec \mathbf{y}$ implies $\mathbf{x} + \mathbf{z} \prec \mathbf{y} + \mathbf{z}$ and if $\mathbf{y} \prec \mathbf{x}$ implies $\alpha\mathbf{y} \prec \alpha\mathbf{x}$. A positive cone defines a pre-order on V by defining $\mathbf{x} \prec \mathbf{y} \iff \mathbf{y} - \mathbf{x} \in M$. In such a case, we say that $\mathbf{x} \succeq \mathbf{0} \iff \mathbf{x} \in M$, following a similar series of steps for fields as above. That is, $\prec$ is binary relation that is symmetric, reflexive and transitive i.e. a partial order. We use Zorn's lemma to extend partial orders of a vector space to a total order.

For a vector space V over $\mathbb{K}$ equipped with an additional binary operation $\cdot : V \times V \longrightarrow A$ is a **vector $\mathbb{K}$-algebra** if the following identities hold for any three elements $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ of V, and all scalars α and β of $\mathbb{K}$

1. $(\mathbf{x}+\mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$ (Right Distributivity)

2. $\mathbf{x} \cdot (\mathbf{y}+\mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$ (Left Distributivity)

3. $(\alpha\mathbf{x}) \cdot (\beta\mathbf{y}) = (\alpha\beta)(\mathbf{x} \cdot \mathbf{y})$ (Compatibility with scalars)

These three axioms are another way of saying that the binary operation is bilinear. This binary operation is often referred to as **multiplication** in V. Some vector spaces are naturally algebras. For example, for the vector space $C[a,b]$ of continuous functions, the ordinary (point-wise) multiplication of functions satisfies the above. A more popular example is the space $l^2(\mathbb{R})$ by defining point-wise multiplication of sequences. Unfortunately, this space does not posses the identity $(1, 1, ...)$ where an identity is a vector $\mathbf{e}$ such that $\mathbf{xe} = \mathbf{x} = \mathbf{ex}$ for all $\mathbf{x} \in V$. An algebra is **commutative** if $\mathbf{ab} = \mathbf{ba}$ for all $\mathbf{a}, \mathbf{b}$. If vector multiplication is associative and if there are inverses such that $\mathbf{x}^{-1}\mathbf{x} = \mathbf{e} = \mathbf{x}\mathbf{x}^{-1}$, then V is an **associative division algebra.** Fortunately, von Neumann has had a way to make an algebra without the need of associativity [20].

We can also order an algebra as follows:

Definition 10 *A $\mathbb{K}$-algebra V together with a total order $\preceq$ on V is an **ordered $\mathbb{K}$-algebra** if the order satisfies*

O1' if $\mathbf{x} \preceq \mathbf{y}$ then $\mathbf{x}+\mathbf{z} \preceq \mathbf{y}+\mathbf{z}$

O2' if $0 \preceq \mathbf{x}$ and $0 \preceq \mathbf{y}$ then $0 \preceq \mathbf{xy}$

O3' if $0 \preceq \mathbf{y}$ and $\beta \geq 0$, then $0 \preceq \beta\mathbf{y}$ where $\geq$ is an order on the skew field.

Again, this gives rise to a cone M such that $M \cap -M = \{0\}$ and $M \cup -M = V$

Definition 11 *If X is a vector space over $\mathbb{K}$, then, a **functional** or **linear form** or **1-form** is a mapping $g : X \longrightarrow \mathbb{K}$ such that $g(\mathbf{x}+\mathbf{y}) = g(\mathbf{x}) + g(\mathbf{y})$ and $g(\alpha\mathbf{x}) = \alpha g(\mathbf{x})$. A functional is said to be **positive** if for every positive element $\mathbf{v} \in X$, $g(\mathbf{v}) \geq 0$. A linear functional is said to be **multiplicative** if $g(\mathbf{x})\, g(\mathbf{y}) = g(\mathbf{xy})$, provided V is a $\mathbb{K}$-algebra.*

Example 12 *For* $A = \mathbb{C}^n$*,* $f_i(z_1, z_2, ..., z_n) = z_i$ *and for* $X = C(\Omega)$ *where* Ω *is compact Hausdroff,* $F_x(g) = g(x)$ *is the default functional on* $C(X)$ *[28]*

There are two requirements for the second definition: an order should exist on the vector space and that the underlying field must be ordered as well. This, therefore, excludes finite fields and even the complex numbers and "beyond" since they do not admit a compatible ordering[1].

There is an interesting characterisation of linear functionals being multiplicative: in the commutative case for $\mathbb{K} = \mathbb{C}$, $g \in \Delta(X) \iff g(\mathbf{x}) \in \sigma(x) = \{\lambda : \mathbf{x} - \lambda\mathbf{e} \notin G(X) \text{ and } \lambda \in \mathbb{C}\}$ where $G(X) = \left\{\mathbf{x} \in X : \mathbf{x}^{-1} \in X\right\}$ and $\Delta(X)$ is the collection of multiplicative linear functionals on X [22].

3.3 Semi Norms

Definition 13 *Let* N *be a vector space over* $\mathbb{F}$*. A* ***norm*** *on* N *is a real-valued function* $\|.\| : N \longrightarrow \mathbb{R}$ *such that*

N1: $\|\mathbf{x}\| = 0 \implies \mathbf{x} = 0$ *(non-degneracy)*

N2: $\|\alpha\mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ *for all* $\alpha \in \mathbb{K}$*,* $\forall\mathbf{x} \in N$ *(homogeneity)*

N3: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ *for arbitrary* $\mathbf{x}, \mathbf{y} \in N$ *(triangle inequality) if* $\mathbb{K}$ *is Archimedean*

N3':$\|\mathbf{x} + \mathbf{y}\| \leq \max(\|\mathbf{x}\|, \|\mathbf{y}\|)$ *(strong triangle inequality) if* $\mathbb{K}$ *is non-Archimedean*

A vector space together with a norm defined on it is called a **normed space**. If the space is complete under the given norm, then the norm algebra is referred to as a **Banach Space**. If we take away N1, we are left with what is called a **seminorm**. Thus, all norms are seminorms. It is safe to say that every functional g gives rise to a seminorm $\|.\|$ such that $\|\mathbf{x}\| = |g(\mathbf{x})|$ where $|.|$ is the absolute value of an element $g(\mathbf{x})$ of the underlying field, provided the underlying field has an absolute value defined on it.

Absolute values on Archimedean fields are functions $|.| : \mathbb{K} \longrightarrow \mathbb{R}$ satisfying $|x| = 0 \iff x = 0$, $|x| \geq 0$, $|xy| = |x||y|$ and $|x + y| \leq |x| + |y|$. Non-Archimedean fields instead satisfy the strong triangle inequality $|x + y| \leq \max(|x|, |y|)$ which is an equality if $|x| \neq |y|$. Thus, an absolute value is trivial if $|x| = 1$ for all non-zero x. In such a case, the trivial seminorm is then

[1] As a set, however, any field, finite or infinite, will admit a (well-)ordering without a particular construction, thanks to the Axiom of Choice.

$\|\mathbf{x}\| = 1$ for all $\mathbf{x}$ whereas the trivial norm further requires that $\|\mathbf{x}\| = 0$ for $\mathbf{x} = 0$. Non-trivial absolute values give rise to a metric by defining $d(x, y) = |x - y|$ and we can meaningfully talk about (weak) convergence in a vector space.

A seminorm space N is expanded to a norm space N/W by taking a collection W of all vectors $\mathbf{v}$ such that $\|\mathbf{v}\| = 0$. A new norm has to be defined: $\|\mathbf{x} + W\|_{N/W} = \|\mathbf{x}\|_N$ for all $x + W \in N/W$. Thus every vector space over a field admitting an absolute value can be converted into a norm space. Using the axiom of choice, every field can have a valuation and, therefore, every vector space admits a seminorm. Such arguments, however, only force a norm on a vector space even if there are technical conflicts with the natural structure of the vector space such as the induced norm may not coincide with the topology on the topological vector space. In particular, the metric may not be translation invariant and, furthermore, may be unbounded. Even if the topologies coincide, it may be that the two norms may not be isometric: for instance, the supremum norm and the usual norm on $\mathbb{R}^2$ because in one topology, the "unit ball" is a circle but in the other is a square (though topologically homeomorphic). Apart from topological considerations, completeness is compromised as well: the forced norm on the space of smooth functions from $\mathbb{R}$ to $\mathbb{R}$ is never complete. Therefore, one must be careful in choosing a norm which respects the algebraic, analytic and topological structure.

Lemma 14 *For any seminorm space N and $\mathbf{x} \in N$, the following hold*

1. $\|\mathbf{0}\| = 0$
2. $\|\mathbf{x}\| = \|-\mathbf{x}\|$
3. $\|\mathbf{x}\| \geq 0$

Proof. 1. $\|0\| = \|0\mathbf{x}\| = |0| \|\mathbf{x}\| = 0$
2. $\|-\mathbf{x}\| = |-1| \|\mathbf{x}\| = \|\mathbf{x}\|$
3. $0 = \|\mathbf{x} - \mathbf{x}\| \leq \max\{\|\mathbf{x}\|, \|-\mathbf{x}\|\} = \|\mathbf{x}\|$
Alternatively, $\|\mathbf{x} - \mathbf{x}\| \leq \|\mathbf{x}\| + \|\mathbf{x}\| = 2\|\mathbf{x}\|$ and hence $0 \leq \|\mathbf{x}\|$ ■

These axioms are usually included in the part of the definition of a norm space for brevity but as we can see, these are actually derivable for both Archimedean and non-Archimedean fields, even for seminorms. In the remainder of this report, we shall use $\|.\|$ to stand for a seminorm.

A seminorm on an algebra which respects the algebraic, analytic and topological structure only if

$$\|\mathbf{xy}\| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$$

holds for all $\mathbf{x}, \mathbf{y}$ apart from the other stated axioms. In particular, one has this submultiplicative property even for $\mathbb{K}$-algebras equipped with a seminorm provided that $\left\|\mathbf{x}^2\right\| = \|\mathbf{x}\|^2$ for all $\mathbf{x}$[7].

It also follows that $\|\mathbf{e}\| \geq 1$, provided that the $\mathbb{K}$-algebra is unital.

Proof. $\|\mathbf{x}\| = \|\mathbf{xe}\| \leq \|\mathbf{x}\| \, \|\mathbf{e}\|$

$\|\mathbf{e}\| \geq 1$ ■

Such an algebra is called a **semi-normed algebra**. If the underlying space is complete, then the space will be called a **Weak Banach Algebra**.

A frequently occurring example is that of the space $B(X, Y)$. In order to glue this concept to that of a semi-norm space, we must wait until we justify the use of such operators in the context of Hermitian spaces.

Chapter 4

Algebraic decomposition of vectors

A linear combination of non-zero vectors $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ is **linearly independent** in a vector space V if $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_n\mathbf{x}_n = 0 \iff \alpha_i = 0 \ \forall i$ [26]. A vector space V is said to have **dimension** n, written $\dim V = n$, if every vector $\mathbf{x} \in V$ admits a decomposition in the form of a linear combination of a linearly independent set of exactly n vectors. This n has a great deal of theory fixed behind it but before we say a few words, we show that this definition does not just apply to ordinary physical vectors we're used to: consider $\mathbf{x}_i(t) = t^i \in C[a,b]$ for $i \in \mathbb{N}$ and $t \neq 0$ such that

$$\sum_{i=1}^{n} \alpha_i t^i = 0$$

If a finite polynomial is equated to zero, when are all the scalars zero? High school training has primed us to impulsively find roots of this polynomial. One root, clearly, is $t = 0$ – not the kind of thing we've defined. For $\alpha, t \in \mathbb{R}$ or $\mathbb{Q}$, we cannot always factor this polynomial. Thus $\alpha_i = 0$ is a possible solution and we're done with fields which are not algebraically closed (we won't have to consider the case of finite fields, which are not algebraically closed). In case we have an algebraically closed field, we'll have non-zero scalars with $t \neq 0$ such that we have a linearly dependent combination. The trick is to see that this cannot be repeated forever for otherwise we arrive at the contradiction that the field is not algebraically closed.

For our purposes, one importance lies in the following justification of Quantum Mechanics being a linear theory: consider the (homogenous) Schrödinger equation which has two linearly independent solutions. These are known to form basis for a family of solutions since the equation itself is linear.

We also take this occasion to suggest a similar treatment for a $\mathbb{K}$-algebra as follows: vectors $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ of the form $\sum \alpha_{ij}\mathbf{x}_i\mathbf{x}_j$ will be called a **multiplicative linear combination** and will be **multiplicatively linearly independent** if $\sum \alpha_{ij}\mathbf{x}_i\mathbf{x}_j = 0$ implies $\alpha_{ij} = 0$.

Let X be a (not necessarily finite) subset of a vector space V. Consider $[X]$, the collection of all subalgebras of V containing X. This is non-empty since $V \in [X]$. Closure property of algebra implies that linear combination of elements of X are contained in elements of $[X]$. If V is an algebra, then $[X]$ also consists of a multiplicative linear combination, as well. Now,

$$X \subseteq \bigcap_{A\in[X]} A := \langle X\rangle$$

This subalgebra is called the algebra spanned by or generated by X and X is called its spanning set. The algebra $\langle X\rangle$ is the smallest algebra containing X for if there were another algebra Y containing X, then $Y \in [X]$ and by construction, $Y \supseteq \langle X\rangle$. If $\langle X\rangle = V$, whether multiplication is defined on V or not, then X is called a basis. The tensor product of two vector spaces V, W with $V = \langle X\rangle$ and $W = \langle Y\rangle$ is $V \otimes W = \langle X \times Y\rangle$, whether or not the underlying fields are the same since the tuples are treated distinctively. This, though, is a topological catastrophe, which can be by-passed by requiring the underlying respective fields to have a valuation preserving morphism.

If X is a basis and if every if every finite linear combination of X is linearly independent, then X is called a **Hamel basis**. X will be called an **m-Hamel** basis if every finite multiplicative linear combination of X is multiplicative linearly independent. In both cases, every vector in V can be written as a finite linear/multiplicative-linear combination of elements in X. The rationale for considering only finite linear combinations is as follows: if X is countable and if all we need is a linear/multiplicative-linear combination of vectors in X, then we're left with a series which raises serious questions of convergence and, therefore, requires a metric. Contrarily, if X is uncountable, then the sum always diverges and is, therefore, not defined in the first place.

Example 15 *Instead of* $1, i, j, k$ *as Hamel basis for the quaternion algebra* $\mathbb{H}$ *over* $\mathbb{R}$*, we can have* i, j *because then* $ij = k$ *and* $i^2 = j^2 = -1$ *generates the quaternions.* i, j *are multiplicative-linearly independent.*

Example 16 *Instead of* $e_1 = (1, 0, 0)$*,* $e_2 = (0, 1, 0)$ *and* $e_3 = (0, 0, 1)$*, we can have* $x = (1, 0, 1)$ *and* $y = (0, 1, 1)$ *and using point-wise multiplication, we get* $(1, 0, 1)(0, 1, 1) = (0, 0, 1) = z$*. The tuples* x, y, z *are linearly independent whereas the tuples* x *and* y *are multiplicatively linearly independent.*

X is a basis if and only if X is minimal. That is, deletion of any element from X, except $\mathbf{0}$ if $\mathbf{0} \in X$, does not form a basis

Proof. Let $Z = X \backslash \{\mathbf{v}\}$. For the sake of contradiction, assume $\langle Z \rangle = V$. Since $Z \subset X$, we must have $\langle Z \rangle \subset \langle X \rangle$ but then $V \subset V$

Conversely, suppose X is minimal but not a basis. Then, there exists $\mathbf{v} \in V$ such that $\mathbf{v}$ cannot be formed out of finite algebraic operations of elements in X. This implies $\langle X \rangle = \langle X \rangle \setminus \{\mathbf{v}\} = \langle X \backslash \{\mathbf{v}\} \rangle$. If $\mathbf{v} \in X$, then we contradict the minimality of X. If $\mathbf{v} \notin X$, then for any binary operation T on V, $T(\mathbf{x}, \mathbf{v}) \in V$ for $\mathbf{x} \in X \implies \mathbf{v} \in \langle X \rangle \implies V \subseteq \langle X \rangle$ so that $V = \langle X \rangle$. That is, X is a basis and not a basis, another contradiction. Thus the supposition that X is minimal and not a basis leads to contradictions, implying that the negation "either X is not minimal or X is a basis" is true. That is, X is minimal implies X is a basis. ∎

Lemma 17 *For any two basis* X*,* Y *of* V*,* $|X| = |Y|$

Proof. Without loss of generality, assume $|Y| + \mathfrak{m} = |X|$ for $\mathfrak{m} > 0$. Using a mapping between X and Y, we can observe that $\mathfrak{m}$ elements can be deleted from X to form a basis. This contradicts the minimality of X. ∎

If X and Y are both countable or uncountable, then the above proof fails. The countable case does not bother us since a Hamel basis, if infinite, is always uncountable[1]. The trouble is always with the uncountable case. Now, assuming the Continuum hypothesis, if $|X| > |Y|$, then $|Y|$ is finitely countable (we're excluding the infinitely countable case) but then $\langle Y \rangle = V$ implies dimension of V is finite, a contradiction.

It is easy to see that every subset of a multiplicatively linearly independent set is multiplicatively linearly independent. Since that sounds like a mouthful, we will now shorten multiplicatively linearly independent to m-linearly independent.

[1] For instance, consider the vector space of real sequences.

Every m-Hamel basis gives rise to a Hamel basis

Proof. Let H be an m-Hamel basis. Then, for every $A = \{\mathbf{x}_1, ..., \mathbf{x}_k\} \subseteq H$, $\sum \alpha_{ij}\mathbf{x}_i\mathbf{x}_j = 0$ implies $\alpha_{ij} = 0$ for $1 \leq i, j \leq k$. Now, let $\mathbf{x}_i\mathbf{x}_j = \mathbf{x}_l$. Then, $\sum \alpha_l \mathbf{x}_l = 0$ implies $\alpha_l = 0$ for $1 \leq l \leq k$ ■

Can we be assured of the existence of a smallest set X satisfying $\langle X \rangle = V$? For this, it suffices to prove that every algebra possesses an m-Hamel basis because we have already seen that if $\langle X \rangle = V = \langle Y \rangle$, then $|X| = |Y|$

Proof. Let $\mathcal{L}$ be a set of L m-linearly independent subsets of an algebra V over $\mathbb{K}$. This is non-empty since every vector itself is trivially m-linearly independent. It is easy to see that under set inclusion, $\mathcal{L}$ is partially ordered. For every totally ordered set $C \subseteq \mathcal{L}$, define $\hat{C} = \bigcup\limits_{C \subseteq \mathcal{L}} C$. Clearly, $\hat{C}$ is an upper bound of every C. To verify that $\hat{C}$ is an element of $\mathcal{L}$, Let $X = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\} \subseteq \hat{C}$ be a finite collection of vectors. Then there exist sets $L_1, ..., L_n \in C$ such that $x_i \in L_i$ for all $1 \leq i \leq n$. Since C is a chain, every L_i is related. We can define a join-lattice $(C, \cup)$ by letting $L_i \cup L_j = L_i \iff L_j \subseteq L_i$ for any $1 \leq i, j \leq n$. Thus, there is a number k with $1 \leq k \leq n$ such that $C \ni L_k = \bigcup\limits_{i=1}^{n} L_i$ and thus $X \subseteq L_k$, that is $X \in C$ and, therefore, is m-linearly independent. Since X was arbitrary, therefore $\hat{C} \in \mathcal{L}$.

According to Zorn's lemma $\mathcal{L}$, has a maximal element M which is m-linearly independent. We show now that M is a m-Hamel basis. Assume there exists an $\mathbf{v} \in V \backslash \langle M \rangle$. Let $\{\mathbf{v}_1, ..., \mathbf{v}_n\} \subseteq M$ be a finite collection of vectors such that $\sum \alpha_{ij}\mathbf{v}_i\mathbf{v}_j = 0$ where $\alpha_{ij} \in \mathbb{K}$. Since $\mathbf{v} \in V$, we must have $\alpha\mathbf{v} = \sum \beta_{ij}\mathbf{v}_i\mathbf{v}_j$ for some $\beta_{ij} \in \mathbb{K}$

If $\alpha = 0$, then $\alpha_{ij} = 0$ for $1 \leq i, j \leq n$, making $\{\mathbf{v}\} \cup M$ m-linearly independent in contradiction to the maximality of M. If $\alpha \neq 0$, we would have

$$\mathbf{v} = \alpha^{-1} \sum \beta_{ij}\mathbf{v}_i\mathbf{v}_j$$

But this implies the contradiction $\mathbf{v} \in \langle M \rangle$. Thus such an $\mathbf{v}$ does not exist and $V = \langle M \rangle$, so M is a generating set and hence a m-Hamel basis. ■

Thus, it makes sense to define $\dim V = |X|$ if V is a vector space. If V is an algebra, then we will call m-dim $V := \dim(V) = |X|$ (corresponding to m-Hamel basis). From the above discussion, it follows that $\dim(V) \leq \dim V$

Problem 18 *Find k such that* $\dim(V) + k = \dim V$

Leaving this problem for future generations for lack of skills in combinatorial arguments, we show that these basis have connections to analysis

Theorem 19 *A totally ordered division $\mathbb{K}$-algebra over a skew field is $\mathbb{K}$.*

Proof. Consider a positive non-trivial multiplicative linear functional $f : V \longrightarrow \mathbb{F}$. This is surjective since for any $\alpha \in \mathbb{F}$, $\alpha\mathbf{e} \in V$. Now, let $f(\mathbf{x}) = f(\mathbf{y})$. Then, since for any positive $\mathbf{x} \in V$, we have $\mathbf{x} = \sum \alpha_{ij}\mathbf{v}_i\mathbf{v}_j$, hence we have $\sum \left(\alpha_{ij} - \beta_{ij}\right) f(\mathbf{v}_i) f(\mathbf{v}_j) = 0$ where $\mathbf{x} = \sum \alpha_{ij}\mathbf{v}_i\mathbf{v}_j$ and $\mathbf{y} = \sum \beta_{ij}\mathbf{v}_i\mathbf{v}_j$ are positive elements. Let $a_{ij} = \left(\alpha_{ij} - \beta_{ij}\right) f(\mathbf{v}_i) f(\mathbf{v}_j)$. Then, $\sum a_{ij} = 0 \implies a_{ij} = 0$. Thus, $\left(\alpha_{ij} - \beta_{ij}\right) = 0$, $f(\mathbf{v}_i) = 0$ or $f(\mathbf{v}_j) = 0$. The latter two possibilities imply the contradiction that f is trivial. Hence $\alpha_{ij} = \beta_{ij}$ so that $\mathbf{x} = \mathbf{y}$. For negative vectors, we use the same argument with $g : V \longrightarrow \mathbb{F}$ such that $g = -f$. ■

Using these, we can define known vector multiplication by focusing on the scalars of the basis. That is, $\left(\mathbf{x} = \sum \alpha_{ij}\mathbf{v}_i\mathbf{v}_j\right) . \left(\mathbf{y} = \sum \beta_{ij}\mathbf{v}_i\mathbf{v}_j\right)$
$= \sum \alpha_{ij}\beta_{ij}\mathbf{v}_i\mathbf{v}_j$, which obeys all the familiar laws of a $\mathbb{K}$-algebra.

Definition 20 *Let X be a vector space over $\mathbb{F}$ and Y be vector space over $\mathbb{K}$ and let $\phi : \mathbb{F} \longrightarrow \mathbb{K}$ be a homomorphism. Then, an operator $T : X \longrightarrow Y$ is a ϕ-**vector space homomorphism** between X and Y if for all $\mathbf{x}, \mathbf{y} \in X$ and scalars $\alpha \in \mathbb{F}$, $T(\alpha\mathbf{x} + \beta\mathbf{y}) = \phi(\alpha) T(\mathbf{x}) + \phi(\beta) T(\mathbf{y})$. T is an **isomorphism** if T and ϕ are bijective. A ϕ-**algebra homomorphism** is of the form $T((\alpha\mathbf{x})(\beta\mathbf{y})) = T(\alpha\beta\mathbf{x}\mathbf{y}) = \phi(\alpha\beta) T(\mathbf{x}) T(\mathbf{y})$, which we shall call an **isomorphism** if ϕ and T are bijective.*

Combined, these give us the generalised superposition principle $T(\alpha\mathbf{x} + \beta\mathbf{y}) = \phi(\alpha) T(\mathbf{x}) + \phi(\beta) T(\mathbf{y})$, which makes the application of linear operators a little uneasy. We, therefore, do not call our operator linear but rather a homomorphism of vector spaces. This has the added advantage of being able to meaningfully ascribe "superposition" where a configuration is changed from one system to another and not within. To reify our definition, we remark that there are extensive examples of involutive anti-automorphisms over non-Archimedean (skew) fields [3].

This definition is based for an operator which is single valued. We can define a linear operator for multi-valued mappings, as well, as follows: if $T = \{(x, z) : x \in V, z \in W\}$ is a relation, then $(\alpha\mathbf{x} + \beta\mathbf{y})Tz = \phi(\alpha)\,\mathbf{x}Tz + \phi(\beta)\,\mathbf{y}Tz$.

A straight-forward derivation of this definition is that $T(\mathbf{0}_X) = \mathbf{0}_Y$ and the preservation of (multiplicative) linear dependence, provided that we are guaranteed the injectivity of T only (and not necessarily of ϕ).

Proof. $T(\mathbf{0}_X) = T(\mathbf{x} + (-1\mathbf{x}))$

$= T(\mathbf{x}) + T(-1\mathbf{x}) = T(\mathbf{x}) + \phi(-1)\,T(\mathbf{x})$

$= T(\mathbf{x}) - 1T(\mathbf{x}) = T(\mathbf{x}) - T(\mathbf{x}) = \mathbf{0}_Y$.

Let $\sum \alpha_{nm}\mathbf{x}_n\mathbf{y}_m = \mathbf{0}$ implies $\alpha_{nm} = 0$. Then, $T\left(\sum \alpha_{nm}\mathbf{x}_n\mathbf{y}_m\right) = T(\mathbf{0}_X) = \mathbf{0}_Y$

$\Longrightarrow \sum \alpha_{nm}\mathbf{x}_n\mathbf{y}_m = \mathbf{0}$

$\Longrightarrow \alpha_{nm} = 0$ for all n, m

$\Longrightarrow \phi(\alpha_{nm}) = 0$ for all n, m $\blacksquare$

In a similar vein, one can prove that $T(\mathbf{e}_X) = \mathbf{e}_Y$, provided T is surjective

Proof. $\mathbf{e}_Y \in Y \Longrightarrow \exists\mathbf{x}$ such that $T(\mathbf{x}) = \mathbf{e}$

Now, $\mathbf{e} = T(\mathbf{x}) = T(\mathbf{xe}) = T(\mathbf{x})\,T(\mathbf{e}) = \mathbf{e}T(\mathbf{e}) = T(\mathbf{e})$ $\blacksquare$

It is not necessary that a homomorphism between two algebras will preserve identities. For instance, given any skew field $\mathbb{K}$ (which is an algebra over itself), the $\mathbb{K}$-algebra of $n \times n$ matrices M and $T : \mathbb{K} \longrightarrow M$, a ϕ-algebra homomorphism where ϕ is an automorphism with $T(\mathbf{x}) = (\alpha_{ij})$ such that $\alpha_{11} = \mathbf{x}$ and zero otherwise. Then, $T(\mathbf{xy}) = T(\mathbf{x})\,T(\mathbf{y})$, $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ and $T(\mathbf{0}) = \mathbf{0}$ but $T(\mathbf{e}) \neq I$. Many nice properties for rings in general (and $\mathbb{K}$-algebras in particular) fail if identities are not preserved. We, therefore, force this requirement to be a part of our definition for the ϕ-algebra homomorphism.

Chapter 5

Algebra of Multivalued Operators

Sharp eigenvalues, by definition, are those in which there is no variance of the associated observable[1] whereas fuzzy values are otherwise. In order to be able to pin-down the value of an observable but to move away from the eigenvalue approach, a generic way is to consider multivalued mappings. In this section, we show that such a philosophy stands futile.

Definition 21 *Let X be a vector space over $\mathbb{K}$. A f-**sesquilinear 2-form** is a function $\varphi : X \times X \longrightarrow \mathbb{K}$ such that $\forall \alpha \in \mathbb{K}$ and $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in X$,*

S1 $\varphi(\mathbf{x}+\mathbf{y},\mathbf{z}) = \varphi(\mathbf{x},\mathbf{z}) + \varphi(\mathbf{y},\mathbf{z})$

S2 $\varphi(\mathbf{x},\mathbf{y}+\mathbf{z}) = \varphi(\mathbf{x},\mathbf{y}) + \varphi(\mathbf{x},\mathbf{z})$

S3 $\varphi(\alpha\mathbf{x},\mathbf{y}) = f(\alpha)\,\varphi(\mathbf{x},\mathbf{y})$

S4 $\varphi(\mathbf{x},\alpha\mathbf{y}) = \varphi(\mathbf{x},\mathbf{y})\,\alpha$

where $f : \mathbb{K} \longrightarrow \mathbb{K}$ is an involutive anti-automorphism. Alternatively, we can call φ linear in both arguments provided we defined $\varphi : \overline{X} \times X \longrightarrow \mathbb{K}$ where $\overline{X}$ is a vector space on base set X and conjugated scalars $f(\mathbb{K})$. Through-out this section, f will stand for the involutive anti-automorphism.

Lemma 22 $\varphi(\mathbf{0},\mathbf{y}) = \varphi(\mathbf{x},\mathbf{0}) = 0$

[1] See §8.1

Proof. $\varphi(\mathbf{0},\mathbf{y}) = \varphi(\mathbf{x}-\mathbf{x},\mathbf{y}) = \varphi(\mathbf{x},\mathbf{y}) - \varphi(\mathbf{x},\mathbf{y}) = 0$. Similarly the second part. $\blacksquare$

Thus, every vector is orthogonal to the zero vector.

Note that an anti-automorphism has to be defined over $\mathbb{K}$ in order for $S3$ to make sense. The third argument also says that the function is conjugate linear in the first argument, following a mathematician's rather than a physicist's convention. It turns out that many results relating to orthogonal vectors can be deduced using the above generalised form: two elements $\mathbf{v}$ and $\mathbf{w}$ of vector space space V are said to be $\varphi-$**orthogonal** if $\varphi(\mathbf{v},\mathbf{w}) = 0$. This is written as $\mathbf{v}\perp\mathbf{w}$. We require that this be equivalent to $\mathbf{w}\perp\mathbf{v}$. Subsets A and B of V are orthogonal to each other if every element A is orthogonal to every other element of B. This is written as $A^{\perp} = B$ and B is said to be the **orthogonal complement** of A. B is then a subspace since for $\mathbf{x},\mathbf{y} \in B$, we can have $\alpha\mathbf{x} - \beta\mathbf{y} \in B$. The second orthogonal complement of A is defined as the orthogonal complement of B. From Functional Analysis, we know that $A^{\perp\perp}$ is a closed subspace. It makes sense to define a closure operator $A \longmapsto A^{\perp\perp}$ in the sense of [36] which converts an ordinary set into a subspace by respecting containment of the set (**Lemma 23**), monotonicity (**Corollary 25**) and idempotency (**Corollary 27**):

Lemma 23 *$A \subseteq A^{\perp\perp}$.*

Proof. $\mathbf{x} \in A \implies \varphi(\mathbf{x},\mathbf{y}) = 0$ for $\mathbf{y} \in A^{\perp} \implies \mathbf{x} \in A^{\perp\perp}$ $\blacksquare$

Lemma 24 *$A \subseteq B$, then $B^{\perp} \subseteq A^{\perp}$*

Proof. If $\mathbf{x} \in B^{\perp}$, then $\varphi(\mathbf{x},\mathbf{y}) = 0$ for $\mathbf{y} \in B$
$\implies \varphi(\mathbf{x},\mathbf{y}) = 0$ for $\mathbf{y} \in A$
$\implies \mathbf{x} \in A^{\perp}$ $\blacksquare$

Corollary 25 *$A \subseteq B \implies A^{\perp\perp} \subseteq B^{\perp\perp}$*

Proof. $A \subseteq B \implies B^{\perp} \subseteq A^{\perp}$ by **Lemma 24** $\implies A^{\perp\perp} \subseteq B^{\perp\perp}$ $\blacksquare$

Lemma 26 *$A^{\perp\perp\perp} = A$*

Proof. $A^{\perp} \subseteq \left(A^{\perp}\right)^{\perp\perp} = A^{\perp\perp\perp}$ by **Lemma 23** and $A \subseteq A^{\perp\perp} \implies \left(A^{\perp\perp}\right)^{\perp} = A^{\perp\perp\perp} \subseteq A^{\perp}$ by **Lemma 24** $\blacksquare$

Corollary 27 $A^{\perp\perp\perp\perp} = A^{\perp\perp}$

Proof. $A^{\perp\perp} \subseteq \left(A^{\perp\perp}\right)^{\perp\perp} = A^{\perp\perp\perp\perp}$ by **Lemma 23**. $\mathbf{x} \in A^{\perp\perp\perp\perp}$, then $\varphi(\mathbf{x}, \mathbf{y}) = 0$ for $\mathbf{y} \in A^{\perp\perp\perp} = A$

$\implies \mathbf{x} \in A^{\perp\perp}$ by **Lemma 23** again ■

Theorem 28 $A^{\perp\perp}$ *is the smallest subset containing* A

Proof. Assume there exists a closed B such that $A \subset B \subseteq A^{\perp\perp}$. Then, $B = B^{\perp\perp}$ and $A \subset B^{\perp\perp} \subseteq A^{\perp\perp}$ so that $B^{\perp} \subset A^{\perp}$ and $A^{\perp\perp\perp} = A^{\perp} \subseteq B^{\perp}$ and hence $B^{\perp\perp} = A^{\perp\perp}$. ■

If $S3$ is replaced by $\varphi(\alpha\mathbf{x}, \mathbf{y}) = \alpha\varphi(\mathbf{x}, \mathbf{y})$, then the 2-form is said to be **bilinear**. A sesquilinear form φ such that $\varphi(\mathbf{x}, \mathbf{y}) = -\varphi(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y}$ is said to be **antisymmetric**. The Lie Bracket is such an example. Every 2-form is said to be **alternating** if $\varphi(\mathbf{x}, \mathbf{x}) = 0\ \forall \mathbf{x} \in X$. In such a case, we write $\mathbf{x}\perp\mathbf{x}$ and we say that $\mathbf{x}$ is isotropic. This is important in physical theories. For example, in the Minkowski space $\mathbb{R}^4$ over $\mathbb{R}$, $\varphi(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4$ where $\mathbf{x} = (x_1, x_2, x_3, x_4)$ and $\mathbf{y} = (y_1, y_2, y_3, y_4)$ in which we can have non-zero isotropic vectors. This gives rise to a semi-norm space. It is, therefore, no surprise that the underlying space for relativistic mechanics is fundamentally different from Quantum Mechanics.

Every alternating 2-form is antisymmetric.

Proof. $\varphi(\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y}) = 0$

$\implies \varphi(\mathbf{x}, \mathbf{x} - \mathbf{y}) - \varphi(\mathbf{y}, \mathbf{x} - \mathbf{y}) = 0$

$\implies \varphi(\mathbf{x}, \mathbf{x}) - \varphi(\mathbf{y}, \mathbf{y}) - \varphi(\mathbf{y}, \mathbf{x}) + \varphi(\mathbf{y}, \mathbf{y}) = 0$

$\implies -\varphi(\mathbf{x}, \mathbf{y}) - \varphi(\mathbf{y}, \mathbf{x}) = 0$

$\implies \varphi(\mathbf{x}, \mathbf{y}) = -\varphi(\mathbf{y}, \mathbf{x})$ ■

The converse holds if the characteristic of the underlying field is equal to 2.

Lemma 29 *If char*$\mathbb{K}$ $=2$*, then* $\varphi(\mathbf{v}, \mathbf{v}) = 0 \iff \varphi(\mathbf{v}, \mathbf{w}) = -\varphi(\mathbf{w}, \mathbf{v})$ *for all* $\mathbf{v}, \mathbf{w}$

Proof. Neccesity has already been proven. To show sufficiency, let $\mathbf{x} \in X$. Write $\mathbf{x} = \mathbf{v} + \mathbf{w}$

Then, $\varphi(\mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}) = \varphi(\mathbf{v}, \mathbf{v}) + \varphi(\mathbf{w}, \mathbf{v}) + \varphi(\mathbf{v}, \mathbf{w}) + \varphi(\mathbf{w}, \mathbf{w}) = 0$ ■

Thus, no norm can be defined on a vector space from the sesquilinear form φ if char$\mathbb{K}$ $=2$.

A 2-form is said to be **symmetric** if $\varphi(\mathbf{v}, \mathbf{w}) = \varphi(\mathbf{w}, \mathbf{v})$ for all vectors $\mathbf{v}, \mathbf{w}$. A 2-fom is **non-degenerate** if $\varphi(\mathbf{u}, \mathbf{v}) = 0$ for all $\mathbf{v}$ implies $\mathbf{u} = 0$ and is degenerate if there exists $\mathbf{u} \neq \mathbf{0}$

such that $\varphi(\mathbf{u},\mathbf{v})=0$ for all $\mathbf{v}$. A 2-form is **symplectic** if it is bilinear, alternating and non-degenerate. A form is **positive** if $\varphi(\mathbf{x},\mathbf{x})\geq 0$ for all $\mathbf{x}$. A positive Hermitian form is said to be **separating** or **positive definite** if $\mathbf{x}\neq 0$ implies $\varphi(\mathbf{x},\mathbf{x})>0$. The triple $(X,\mathbb{K},\varphi)$ is said to be a **Hermitian space** if the sesquilinear 2-form φ is Hermitian i.e. $\varphi(\mathbf{x},\mathbf{y})=f(\varphi(\mathbf{y},\mathbf{x}))$ for all $\mathbf{x},\mathbf{y}$[2]. The triple $(X,\mathbb{K},\varphi)$ is said to be a **skew-Hermitian space** if the sesquilinear 2-form φ is skew-Hermitian i.e. if $-\varphi(\mathbf{x},\mathbf{y})=f(\varphi(\mathbf{y},\mathbf{x}))$

From $\varphi(\mathbf{x},\mathbf{x})=f(\varphi(\mathbf{x},\mathbf{x}))$, we see that $\varphi(\mathbf{x},\mathbf{x})$ has to be a real number. In fact, more can be said:

Lemma 30 *A sesquilinear form φ is Hermitian if and only if $\varphi(\mathbf{x},\mathbf{x})$ is real*

Proof. ($\Longleftarrow$) Applying f on both sides of

$\varphi(\mathbf{x},\mathbf{x})-\varphi(\mathbf{x},\mathbf{y})-\varphi(\mathbf{y},\mathbf{x})+\varphi(\mathbf{y},\mathbf{y})=\varphi(\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y})$

We get $\varphi(\mathbf{x},\mathbf{x})-f(\varphi(\mathbf{x},\mathbf{y}))-f(\varphi(\mathbf{y},\mathbf{x}))+\varphi(\mathbf{y},\mathbf{y})=\varphi(\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y})$

Subtract both to get

$$f(\varphi(\mathbf{x},\mathbf{y}))+f(\varphi(\mathbf{y},\mathbf{x}))=\varphi(\mathbf{x},\mathbf{y})+\varphi(\mathbf{y},\mathbf{x}) \tag{5.1}$$

Now, assume $\varphi(\mathbf{x},\mathbf{y})\neq f(\varphi(\mathbf{y},\mathbf{x}))$

Then, $\varphi(\mathbf{y},\mathbf{x})\neq f(\varphi(\mathbf{x},\mathbf{y}))$

So that when we add these two, we get $f(\varphi(\mathbf{x},\mathbf{y}))+f(\varphi(\mathbf{y},\mathbf{x}))\neq\varphi(\mathbf{x},\mathbf{y})+\varphi(\mathbf{y},\mathbf{x})$, which contradicts (5.1). ■

We now prove a generalised version of the Cauchy-Schwarz inequality

Proof. Let φ be a positive form and let $\mathbf{x},\mathbf{y}$ be non-zero vectors. Then, $0\leq\varphi(\mathbf{x}-\alpha\mathbf{y},\mathbf{x}-\alpha\mathbf{y})$

$=\varphi(\mathbf{x},\mathbf{x}-\alpha\mathbf{y})-f(\alpha)\,\varphi(\mathbf{y},\mathbf{x}-\alpha\mathbf{y})$

$=\varphi(\mathbf{x},\mathbf{x})-\varphi(\mathbf{x},\mathbf{y})\alpha-f(\alpha)\,[\varphi(\mathbf{y},\mathbf{x})-\varphi(\mathbf{y},\mathbf{y})\alpha]$

Let $\alpha=[\varphi(\mathbf{y},\mathbf{y})]^{-1}\,\varphi(\mathbf{y},\mathbf{x})$

Since $f(\varphi(\mathbf{y},\mathbf{x}))=\varphi(\mathbf{x},\mathbf{y})$, $f(\alpha\beta)=f(\beta)\,f(\alpha)$ and $f(\varphi(\mathbf{x},\mathbf{x}))=\varphi(\mathbf{x},\mathbf{x})$, then we have $\varphi(\mathbf{x},\mathbf{x})-\varphi(\mathbf{x},\mathbf{y})\,[\varphi(\mathbf{y},\mathbf{y})]^{-1}\,\varphi(\mathbf{y},\mathbf{x})-\varphi(\mathbf{x},\mathbf{y})\,[\varphi(\mathbf{y},\mathbf{y})]^{-1}\left[\varphi(\mathbf{y},\mathbf{x})-\varphi(\mathbf{y},\mathbf{y})\,[\varphi(\mathbf{y},\mathbf{y})]^{-1}\,\varphi(\mathbf{y},\mathbf{x})\right]$

[2] **Caution: Hermitian spaces for us do not have a seperating, positive 2-form**

$\varphi(\mathbf{x},\mathbf{x}) - \varphi(\mathbf{x},\mathbf{y})\left[\varphi(\mathbf{y},\mathbf{y})^{-1}\right]\varphi(\mathbf{y},\mathbf{x}) = \varphi(\mathbf{x}-\alpha\mathbf{y},\mathbf{x}-\alpha\mathbf{y})$

$\implies \varphi(\mathbf{x},\mathbf{x}) \geq \varphi(\mathbf{x},\mathbf{y})\varphi(\mathbf{y},\mathbf{x})\left[\varphi(\mathbf{y},\mathbf{y})^{-1}\right]$ (because $\varphi(\mathbf{y},\mathbf{y})$ is real)

$\implies \varphi(\mathbf{x},\mathbf{x})\varphi(\mathbf{y},\mathbf{y}) \geq \varphi(\mathbf{x},\mathbf{y})\varphi(\mathbf{y},\mathbf{x})$ ■

With every sesquilinear, positive Hermitian form φ, we can associate a semi-norm $\varphi(\mathbf{x},\mathbf{x}) := \|\mathbf{x}\|^2$, provided that f is valuation preserving.

Proof. $\|\alpha\mathbf{x}\|^2 = \varphi(\alpha\mathbf{x},\alpha\mathbf{x})$

$= f(\alpha)\,\alpha\varphi(\mathbf{x},\mathbf{x})$ since $\varphi(\mathbf{x},\mathbf{x})$ is real

$= |\alpha|^2\,\|\mathbf{x}\|^2$

Also, $\|\mathbf{x}+\mathbf{y}\|^2 = \varphi(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})$

$= \varphi(\mathbf{x},\mathbf{x}) + \varphi(\mathbf{x},\mathbf{y}) + \varphi(\mathbf{y},\mathbf{x}) + \varphi(\mathbf{y},\mathbf{y})$

$\leq |\varphi(\mathbf{x},\mathbf{x})| + |\varphi(\mathbf{x},\mathbf{y})| + |\varphi(\mathbf{y},\mathbf{x})| + |\varphi(\mathbf{y},\mathbf{y})|$

$= \|\mathbf{x}\|^2 + 2\,|\varphi(\mathbf{x},\mathbf{y})| + \|\mathbf{y}\|^2$

$\leq \|\mathbf{x}\|^2 + 2\,\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2$

$= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$

Here, we have used the triangle inequality valid only for spaces over Archimedean fields to establish N3. For N3'

$\|\mathbf{x}+\mathbf{y}\|^2 = \varphi(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})$

$= \varphi(\mathbf{x},\mathbf{x}) + \varphi(\mathbf{x},\mathbf{y}) + \varphi(\mathbf{y},\mathbf{x}) + \varphi(\mathbf{y},\mathbf{y})$

$\leq \max(|\varphi(\mathbf{x},\mathbf{x})|, |\varphi(\mathbf{x},\mathbf{y})|, |\varphi(\mathbf{y},\mathbf{x})|, |\varphi(\mathbf{y},\mathbf{y})|)$

Thus, $\|\mathbf{x}+\mathbf{y}\| \leq \sqrt{\max(|\varphi(\mathbf{x},\mathbf{x})|, |\varphi(\mathbf{y},\mathbf{x})|, |\varphi(\mathbf{y},\mathbf{y})|)}$

Now, if $\varphi(\mathbf{x},\mathbf{y}) = a + b$ for $a, b \in \mathbb{K}$ such that $f(a) = a$ and $f(b) \neq b$, then $|a|$ and $|b| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|$ [3] so that $|\varphi(\mathbf{x},\mathbf{y})| = \max\{|a|, |b|\} \leq \|\mathbf{x}\|\,\|\mathbf{y}\|$. It follows that $\max(|\varphi(\mathbf{x},\mathbf{x})|, |\varphi(\mathbf{y},\mathbf{x})|, |\varphi(\mathbf{y},\mathbf{y})|) = \max\{\|\mathbf{x}\|, \|\mathbf{y}\|\}$ ■

This is a norm if φ is separating: let $\|\mathbf{x}\| = 0$ for $\mathbf{x} \neq 0$ but then $\varphi(\mathbf{x},\mathbf{x}) = 0 > 0$. Now, being positive definite (or negative definite) is an equivalent condition for the non-existence of isotropic vectors relative to the 2-form. Without resorting to the axioms of norm spaces, we can safely say that we still cannot have isotropic vectors if we require our 2-form to be positive definite (or negative definite). Thus, if one needs isotropic vectors, one must not require the 2-form to be definite, a need we shall follow.

A 2-form $\varphi(\mathbf{x},\mathbf{y})$ is said to be bounded if there exists an $m \in \mathbb{R}^{+}$ such that $|\varphi(\mathbf{x},\mathbf{y})| \leq m\|\mathbf{x}\|\|\mathbf{y}\|$ for all $\mathbf{x},\mathbf{y}$. We have seen that positive Hermitian forms are bounded, by the Cauchy-Schwarz inequality. From this boundedness, we have continuity.

Proof. If $|\varphi(\mathbf{x},\mathbf{y})| \leq m\|\mathbf{x}\|\|\mathbf{y}\|$, then letting $\mathbf{x}_n \longrightarrow \mathbf{x}$ and $\mathbf{y}_n \longrightarrow \mathbf{y}$, we have $|\varphi(\mathbf{x},\mathbf{y}) - \varphi(\mathbf{x}_n,\mathbf{y}_n)|$

$= |\varphi(\mathbf{x},\mathbf{y}) - \varphi(\mathbf{x}_n,\mathbf{y}_n) + \varphi(\mathbf{x}_n,\mathbf{y}) - \varphi(\mathbf{x}_n,\mathbf{y})|$

$= |\varphi(\mathbf{x}-\mathbf{x}_n,\mathbf{y}) + \varphi(\mathbf{x}_n,\mathbf{y}-\mathbf{y}_n)| \leq |\varphi(\mathbf{x}-\mathbf{x}_n,\mathbf{y})| + |\varphi(\mathbf{x}_n,\mathbf{y}-\mathbf{y}_n)|$

$\leq M(\|\mathbf{x}-\mathbf{x}_n\| + \|\mathbf{y}-\mathbf{y}_n\|) \longrightarrow 0$ ■

Thus, it is meaningful to talk about topologies even in Hermitian spaces. We are now in a position to state that for any $A \subseteq X$, $A^{\perp}$ is closed, the proof of which can be found in any routine Hilbert Space book.

In what follows, bold notation for vectors will be dropped.

Let $(X,\mathbb{K},\varphi)$ be a sesquilinear space. We require φ to be antisymmetric. Let T be a relation on X. We say that T is closed (subspace) if $T = T^{\perp\perp}$ where $\perp$ is defined relative to $X \oplus X$. Through-out this section, only closed relations will be considered. This has the following advantage: a closed relation T is linear

Proof. Let $(x,u),(y,v) \in T$. Then, $(x+y,u+v) \in T$ because T is a subspace of $X \oplus X$. Thus, if $T(x) = u$ and $T(y) = v$, we have $T(x+y) = u+v = T(x)+T(y)$. Next, we know that $(x,y) \in T$ implies that $\alpha(x,y) = (\alpha x,\alpha y) \in T$. That is, $T(\alpha x) = \alpha T(x)$. ■

In order to require $T(\alpha x) = f(\alpha)T(x)$, we define scalar multiplication as $\alpha(x,y) := (\alpha x, f(\alpha)y)$. It can be shown that this does not violate any of the axioms of scalar multiplication for vectors and vector spaces constructed from old ones.

It follows that if T is closed, then $\ker(T)$ is a closed subspace of $X \oplus X$.

Equipped with these tools, we now modify the approach of [37].

Lemma 31 *The sum and composition of two closed relations is closed.*

Proof. We define $T+S = \{(x,y) : y = s+t \text{ for } s \in \mathcal{R}(S), t \in \mathcal{R}(T)\}$ and $T \circ S = TS = \{(x,z) : (x,y) \in S \text{ and } (y,z) \in T \text{ for some } y\}$. To show closure, we note that $(T+S) \subseteq (T+S)^{\perp\perp}$ and $TS \subseteq (TS)^{\perp\perp}$ always holds and the reverse inclusion follows from the fact that $T^{\perp\perp} = T$ and $S^{\perp\perp} = S$. ■

If T is not closed, then the closure $\overline{T} = T^{\perp\perp}$ is the smallest extension of T, which exists regardless of Zorn's lemma. Furthermore, a single-valued adjoint of densely defined T will always exist, be single-valued and linear, even if T is not single valued.

Proof. Part A

We first define the adjoint of a relation on a Hermitian space using a function U on $X \times X$. Define $U : X \times X \longrightarrow X \times X$ by $U(x, y) = (-y, x)$. This is well defined and injective since $(x_1, y_1) = (x_2, y_2) \iff x_1 = x_2$ and $y_1 = y_2 \iff x_1 = x_2$ and $-y_1 = -y_2 \iff (-y_1, x_1) = (-y_2, x_2)$. Also, since X is a vector space, every element has an additive inverse. Therefore, U is surjective. Finally, U is defined everywhere so that $U^{-1}(x, y) = (y, -x)$ exists. For $z, w \in X \times X$ with $z = (z_1, z_2)$ and $w = (w_1, w_2)$, we define $\Phi(z, w) := (\varphi \oplus \varphi)(z, w)_{X \times X} = \varphi(z_1, w_1) + \varphi(z_2, w_2)$. In this case, we note that

$\Phi(U(z), w) = \Phi\left(z, U^{-1}(w)\right)$

To show this, $\Phi(U(z), w) = (\varphi \oplus \varphi)((-z_2, z_1), (w_1, w_2)) = -\varphi(z_2, w_1) + \varphi(z_1, w_2)$ on the LHS whereas on the RHS we have $(\varphi \oplus \varphi)((z_1, z_2), (w_2, -w_1)) = \varphi(z_1, w_2) - \varphi(z_2, w_1)$

Furthermore, for $M \subseteq X \times X$, we have $U\left(M^{\perp}\right) = U(M)^{\perp}$: let $x \in U\left(M^{\perp}\right)$. Then, $x = (x_1, x_2) = U(y_1, y_2)$ for $y = (y_1, y_2) \in M^{\perp} \iff \Phi(y, z) = 0$ for $z \in M$

$\iff \Phi\left(U^{-1}(x), z\right) = 0$

$\iff \Phi(x, U(z)) = 0$

$\iff x \in U(M)^{\perp}$

Now, for any relation T on X, define a relation T^* on X by $T^* = U(T)^{\perp} = U\left(T^{\perp}\right)$. Note that T^* is automatically closed and hence linear. With this, for $\alpha = (x, z) \in T$ and $\beta = (y, w) \in T^* = U(T)^{\perp} \implies \Phi(\beta, \gamma) = 0$ for $\gamma \in U(T)$. Note that $\gamma = U(\alpha) = (-z, x)$. Hence $-\varphi(y, z) + \varphi(w, x) = 0$ so that $\varphi(y, z) = \varphi(w, x) \implies \varphi(z, y) = \varphi(x, w)$. Identifying $z = Tx$ and $w = T^*y$, we have $\varphi(Tx, y) = \varphi(x, T^*w)$. The converse holds as well.

Part B

Now we prove certain relations for the above construction

1. $\ker T^* = \mathcal{R}(T)^{\perp}$

   Let $x \in \ker T^*$. Then, $(x, 0) \in T^*$. By the end of **Part A**, we have $\varphi(b, x) = \varphi(a, 0) = 0$ for $(a, b) \in T$. Hence $b \in \mathcal{R}(T)$ and $x \in \mathcal{R}(T)^{\perp}$. The converse follows similarly.

2. $(\lambda T)^* = f(\lambda) T^*$

   Let $\alpha = (x, y) \in U(\lambda T)^{\perp}$. Then, $\Phi(\alpha, \beta) = 0$ for $\beta \in U(\lambda T)$ where $\beta = (-\lambda v, u)$ for $(u, \lambda v) \in \lambda T$. Hence $\varphi(\lambda v, x) = f(\lambda)\varphi(v, x) = \varphi(u, y)$ so that $(u, f(\lambda) v) \in f(\lambda) T$ implying $(x, y) \in f(\lambda) T^*$. The converse follows similarly.

3. $\left(T^{-1}\right)^* = (T^*)^{-1}$

   For $\alpha = (x, z) \in T^{-1} \Longrightarrow (z, x) \in T$ $\beta = (y, w) \in \left(T^{-1}\right)^* = U\left(T^{-1}\right)^{\perp} \implies \Phi(\beta, \gamma) = 0$ for $(-x, z) = \gamma \in U\left(T^{-1}\right)$. Hence $\varphi(z, w) = \varphi(x, y)$ so that $(w, y) \in T^*$ and hence $(y, w) \in (T^*)^{-1}$. The converse follows similarly.

4. $T^* = \left(-T^{-1}\right)^{\perp}$, $T = T^{**}$, $(RT)^* = T^* R^*$, $\mathcal{D}(T) = \mathcal{R}(T^*)$

   Established similarly by definitions.

We cannot have $I^* = I$ because I is not closed in a general Hermitian space. If it were, then by Bullet 1, $\ker I^* = X^{\perp} = \{0\}$, implying that X has no isotropic vectors, which we have not established. However, I is closed on a collection of anisotropic vectors since such vectors give rise to a Hausdroff space[3]

Part C

By densely defined T, we mean $\mathcal{D}(T)^{\perp\perp} = X$. This condition is, in fact, equivalent to T^* being single valued. Of course this does not mean that X is separable since $\mathcal{D}(T)$ is not necessarily a subspace nor necessarily countable.

Assume that $\beta_1 = (p, q)$ and $\beta_2 = (p, r) \in T^* = \left(-T^{-1}\right)^{\perp}$. If $p \in \ker T^*$, then $q = r$ and there is nothing to prove. If $p \notin \ker T^*$. Then, $p \notin \mathcal{R}(T)^{\perp}$ and $\varphi(z, p) \neq 0$ for $z \in \mathcal{R}(T)$. For $(x, z) \in T$ (that is, $x \in E^{\perp}$), we have $\varphi(z, p) = \varphi(x, q) = \varphi(x, r)$. Using, $x \in X^{\perp}$ and $q, r \in X$, we get the contradiction that $\varphi(z, p) = 0$

Conversely, assume that T^* is single-valued. Recall that a relation R is single valued if and only if $RR^{-1} = I_{\mathcal{R}(R)} = \{(x, x) : x \in \mathcal{R}(R)\}$. Thus, we have that $T^* T^{*-1} = I_{\mathcal{R}(T^*)}$. Now, assume that there exists $x \in X \backslash \mathcal{D}(T)^{\perp\perp}$. Then, $x \notin \mathcal{D}(T) = \mathcal{R}(T^*)$ so that $T^* T^{*-1} \neq I_{\mathcal{R}(T^*)}$.

■

[3] $T = \{(x, y) : x \in X, y \in Y\}$ is closed iff T is continuous provided that Y is Hausdroff

This takes care of the important question of having a canonical star operation on a collection of multivalued operators. This, the additive identity $O = \{(x,0)\}$, multiplicative identity $I = \{(x,x)\}$ and the identification $\lambda T = \{(x,\lambda y) : (x,y) \in T\}$ for $\lambda \in \mathbb{K}$ still do not tell us that the collection of such relations forms an algebra since we are short of the distributive law

Lemma 32 *Let R, S, T be closed relations. Then, $RS + RT \subseteq R(S+T)$ where reverse inclusion holds if $\mathcal{D}(R) = X$ and $(S+T)R \subseteq SR + TR$ where reverse inclusion holds if R is single valued.*

Proof. Let $(x,y) \in RS+RT$ where $y = s+t$ for $s \in \mathcal{R}(RS), t \in \mathcal{R}(RT)$ or that $(x,s) \in RS$ and

$(x,t) \in RT \implies \exists y_1, y_2$ such that $(x,y_1) \in S$, $(y_1,s) \in R$ and $(x,y_2) \in T$, $(y_2,t) \in R$ so that $(x, y_1+y_2) \in S+T$. Note that $(y_1+y_2, s+t) \in R$ because R is a subspace of $X \oplus X$ so that $(x,y) \in R(S+T)$. Conversely, let $(x,y) \in R(S+T)$. Then, $\exists z$ such that $(x,z) \in S+T$ and $(z,y) \in R$. Let $z = s+t$ for $s \in \mathcal{R}(S), t \in \mathcal{R}(T)$. Thus, $(x,s) \in S$ and $(x,t) \in T$. Now, since $\mathcal{D}(R) = X$, then $s,t \in \mathcal{D}(R)$ and hence $(s,y),(t,y) \in R$. Hence we have $(x,y) \in RS$ and RT so that $(x,y) \in RS+RT$ because RS and RT are closed.

$(x,s_1) \in RS$ and $(x,t_1) \in RT$ so that $(x,y) \in RS+RT$

For the second inclusion, let $(x,y) \in (S+T)R$. Then $\exists z$ such that $(x,z) \in R$ and $(z,y) \in S+T$. Let $y = s+t$. Then, $(z,s) \in S$ and $(z,t) \in T$. Hence $(x,s) \in SR$ and $(x,t) \in ST$ which implies $(x,y) \in SR+TR$. Conversely, let $(x,y) \in SR+TR$. Then, $y = a+b$ for $a \in \mathcal{R}(SR)$ and $b \in \mathcal{R}(TR)$ so that $(x,a) \in SR$ and $(x,b) \in TR$. Thus, there exists y_1, y_2 such that $(x,y_1) \in R$, $(y_1,a) \in S$, $(x,y_2) \in R$ and $(y_2,b) \in T$. The single valuedness of R implies $y_2 = y_1 = z$ (say) so that we are left with $(x,z) \in R$, $(z,a) \in S$ and $(z,b) \in T$. Thus, $(z,y) \in S+T$ and $(x,y) \in (S+T)R$ ■

Worse, even if $SR = RS$ and $TR = RT$, we are still not guaranteed $(S+T)R = R(S+T)$. For example, take $S = I$, $T = -I$ and $\mathcal{D}(R) \neq X$. There is a slight way around these: if S and T are such that $\ker T = \ker S$ and $\mathcal{R}(S) = \mathcal{R}(T)$, then $S \subset T$ implies $S = T$. However, for our purposes, this is quite useless since it is rare for two observables to have the same kernel and range, even if some of them are compatible with others.

In order to form a $*$-algebra, one is forced to consider only single-valued operators defined with domains in X (not dense). In order to have a restricted algebra, we can either take densely defined operators or single-valued operators and not both. In either case, spectral theory one has to take up is one which combines both linear and non-linear operators. No such formluation is currently known to the author.

In general, a way to impose a star operation on an algebra is as follows: one considers what is called the **conjugate linear involution** $* : X \longrightarrow X$ such that $\forall \mathbf{a}, \mathbf{b} \in X$ and $\alpha \in \mathbb{K}$, we have

1. $*(\mathbf{a}+\mathbf{b}) = *\,(\mathbf{a}) + *\,(\mathbf{b})$
2. $*\,(\alpha\mathbf{a}) = f\,(\alpha) *\,(\mathbf{a})$
3. $*\,(*\,(\mathbf{a})) = \mathbf{a}$
4. $*\,(\mathbf{ab}) = *\,(\mathbf{b}) *\,(\mathbf{a})$.

where $f : \mathbb{K} \longrightarrow \mathbb{K}$ is an involutive, valuation preserving anti-automorphism. In such a case, $*\,(\mathbf{a})$ is called the adjoint of a. Needless to state, if $*\,(\mathbf{a}) = \mathbf{a}$, then $\mathbf{a}$ is **self-adjoint**. A $\mathbb{K}$-algebra X equipped with such a conjugate linear involution is called a $*$-algebra. For brevity, we shall write $*\,(\mathbf{a}) = \mathbf{a}^*$. If $|f\,(\alpha)| = |\alpha|$ [4], we can use the definition of operator seminorm for $*$. Now, from our definition, it follows that $\|*\,(\mathbf{a}^*\mathbf{a})\| = \|\mathbf{a}^*\mathbf{a}\|$. The admission of fixed points and the definition $\|T\| = \inf\{k : \|T\mathbf{x}\| \leq k\,\|\mathbf{x}\|, \forall \mathbf{x}\}$ for a linear operator implies $\|*\| \geq 1$. Conversely, $\|\mathbf{a}\| = 1$ implies $\|\mathbf{a}^*\| = \|*\,(\mathbf{a})\| \leq \|*\| = \sup_{\|\mathbf{a}\|=1} \|\mathbf{a}^*\|$. Taking supremum on both sides over $\mathbf{a}$, we end up $\|*\| \leq 1$. Thus, $\|*\| = 1$. It follows that $\|\mathbf{a}\| = \|\mathbf{a}^*\|$. Thus, for any Banach Algebra with an isometric involution, we have $\|\mathbf{a}^*\mathbf{a}\| \leq \|\mathbf{a}\|^2$. The converse does not necessarily follow [43]. If it does, then the norm is called a $*$-norm. If, with this norm, X is complete, then X will be called a $\mathbb{K}*$-algebra.

If X is a vector space, we can always define a $*$ operation (without condition 4, obviously) as follows: for any $\mathbf{x} \in V$, define $*\,(\mathbf{x}) = *\left(\sum \alpha_i \mathbf{v}_i\right) := \sum f\,(\alpha_i)\,\mathbf{v}_i$ where f is any involutive, anti-automorphism. If, however, X is an algebra, we may try to define $*\,(\mathbf{x}) = *\left(\sum \alpha_{ij}\mathbf{v}_i\mathbf{v}_j\right) =$

[4] Some anti-involutive automorphisms do not preserve order and hence are not isometric. See §8.3

$\sum f(\alpha_{ij})\,\mathbf{v}_i\mathbf{v}_j$. Provided that we are willing to accept the Axiom of Choice, this definition will even cater for matrix algebra of square matrices with $*$ given by the conjugate transpose but then we shall be losing the preservation of valuation i.e. $\|*\| = 1$ may not hold. Thus, a similar treatment for $\mathbb{K}$-algebras seems impossible at the moment, unless one specifies how the basis ought to be permuted, as well. Our choice of the word "permuted" stems from the idea that an automorphism happens to be a permutation of the algebraic structure.

Sadly, there are no multiplicative functionals on $B_\phi(X)$

Proof. Not all multiplicative linear functionals preserve identities – only onto ones do, as was proved. For any $\lambda \in \mathbb{K}$, we have $\lambda I \in B_\phi(X)$. Thus, for any multiplicative linear functional g, it must be onto and that $g(I) = e$. Let us consider orthogonal projection operators P and $Q \in B_\phi(X)$ such that $\dim P(X) = \dim Q(X)$. Let $T : P(X) \longrightarrow Q(X)$ be a partial isometry. If $P \neq T^*T$, $Q \neq TT^*$, then we arrive at the contradiction that $P(x) \notin P(X)$ and $Q(x) \notin Q(X)$. Thus, $P = T^*T$, $Q = TT^*$. Then, as $PQ = 0$, we have $g(P)g(Q) = 0$, so at least one of them is zero. By construction, we get $g(P) = g(T^*T) = g(T)\,g(T^*) = g(Q)$ so both are zero. By construction, we must also have $P + Q = I$. Then $e = g(I) = g(P+Q) = g(P) + g(Q) = 0$, a contradiction. ■

We can try to collect all operators (linear or not) and turn them into an algebra. Unfortunately, one part of the distributive law ($R(S+T) = RS + RT$) does not necessarily hold, as we have already seen. There are ways around this by, say, defining a slash-product instead of the usual composition [41]. Not all operators can be defined with a complete domain but this isn't a problem since one is usually interested in the unit ball only. The added problem of considering a canonical algebra is to have a compatible seminorm that turns the algebra into a seminormed algebra with a compatible topology. The norm

$$\|T\| = \limsup_{\|\mathbf{x}\|\to\infty} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|}$$

yields $\|RT\| \geq \|R\|\,\|T\|$[47].

We keep things simple for now: consider bounded single-valued operators $T : X \longrightarrow Y$ where Y is a Banach space over $\mathbb{F}$ and X is a normed space over $\mathbb{K}$ with $\phi : \mathbb{K} \longrightarrow \mathbb{F}$ being a valuation preserving homomorphism. We turn this into a vector space by defining vector addition as point-wise addition of the linear operators. Scalar multiplication is also defined as point-wise scaling.

Recall the definition of supremum. It is an upper bound and the lowest of the upper bounds of a set. Thus, we can collect vectors $\mathbf{x}$ such that $\frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} \leq c$ with $\|\mathbf{x}\| \neq 0$ and define a supremum out of it. If we can find a smallest such c, then we have the seminorm of a bounded linear operator T, denoted by $\|T\|$, defined as $\|T\| = \sup\limits_{\|\mathbf{x}\|\neq 0} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|}$.

Since $\|T\| = \sup\limits_{\|\mathbf{x}\|\neq 0} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|}$, we can safely say that $\|T\| \geq \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|}$ for all $\mathbf{x} \in X$ so that we have for ourselves the inequality

$$\|T\mathbf{x}\| \leq \|T\| \, \|\mathbf{x}\|$$

The following definitions are equivalent:

$$\begin{aligned} \|T\| \quad &: \quad = \sup_{\|\mathbf{x}\|\neq 0} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} = \sup_{\|\mathbf{x}\|=1} \|T\mathbf{x}\| = \sup_{0<\|\mathbf{x}\|\leq 1} \|T\mathbf{x}\| \\ &= \sup_{0<\|\mathbf{x}\|<1} \|T\mathbf{x}\| = \inf\{k : \|T\mathbf{x}\| \leq k\,\|\mathbf{x}\|, \forall \mathbf{x}\} \end{aligned}$$

Let $A = \left\{ \frac{\|T(x)\|}{\|x\|} : x \in X \text{ and } \|x\| \neq 0 \right\}$

$B = \{\|T(x)\| : x \in X \text{ and } \|x\| = 1\}$

$C = \{\|T(x)\| : x \in X \text{ and } \|x\| \leq 1\}$

$D = \{\|T(x)\| : x \in X \text{ and } \|x\| < 1\}$

Since equal sets have the same supremum, we will show that $A = B = C = D$

Clearly, A contains B, C and D.

Let $a \in A$

$\iff a = \frac{\|T(x)\|}{\|x\|}$ for some $\mathbf{x} \in X$ such that $\|\mathbf{x}\| \neq 0$

Since X is a seminorm space and closed under scalar multiplication, we can let $\mathbf{y}\,\|\mathbf{x}\| = \mathbf{x}$

$\iff \|\mathbf{y}\| = 1$ so that $a = \|T(\mathbf{y})\|$ for some $\mathbf{y} \in X$

$\iff a \in B$

$\iff A = B$

It is clear that $D \subseteq C$ and that $B \cup D = C$ so that $B \subseteq C$ as well.

Further, $B = A \subseteq C$ so that we have $B = A = C$

To show that $B \subseteq D$

$a \in B$

$\implies a = \|T(\mathbf{x})\|$ for some $\mathbf{x} \in X$ such that $\|\mathbf{x}\| = 1$

Assume that $\exists \mathbf{x}_n$ such that $\mathbf{x}_n \longrightarrow \mathbf{x}$.

Let $\mathbf{y}_n = \frac{n-1}{n}\mathbf{x}_n$. Then, $\mathbf{y}_n \longrightarrow \mathbf{y}$. Furthermore, $\|\mathbf{y}_n\| < \|\mathbf{x}_n\|$ for all n so that $\|\mathbf{x}\| = 1$ implies $\|\mathbf{y}\| < 1$

Then, $a_n = \|T\mathbf{y}_n\| = \left|\frac{n-1}{n}\right| \|T\mathbf{x}_n\| \longrightarrow \|T\mathbf{y}\| = a$

Finally, we show that $\|T\| = \inf\{k : \|T\mathbf{x}\| \leq k\|\mathbf{x}\|, \forall \mathbf{x}\}$

Assume that $\sup\limits_{\|\mathbf{x}\| \neq 0} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} = \alpha$

Then, $\|T\mathbf{x}\| \leq \alpha \|\mathbf{x}\|$

$\implies \sup\limits_{\|\mathbf{x}\| \neq 0} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} = \alpha \geq \inf\{k : \|T\mathbf{x}\| \leq k\|\mathbf{x}\|, \forall \mathbf{x}\}$

Next, $\inf\{k : \|T\mathbf{x}\| \leq k\|\mathbf{x}\|, \forall \mathbf{x}\} \geq \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} \geq \alpha - \frac{1}{n}$ for all n

So that $\inf\{k : \|T\mathbf{x}\| \leq k\|\mathbf{x}\|, \forall \mathbf{x}\} = \alpha = \sup\limits_{\|\mathbf{x}\| \neq 0} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|}$

This is a seminorm:

For N1, $\|T\| \geq \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} \geq 0$. $\|T\| = 0$ if and only if

$$\sup \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} = 0$$

which implies $\sup\|T\mathbf{x}\| = 0$. Since we have a supremum of non-negative numbers and this is equal to zero, therefore $\|T\| = 0$ if and only if $\|T\mathbf{x}\| = 0$ for all $\mathbf{x}$. Considering that $T\mathbf{x}$ is a vector whose seminorm is equal to zero, we cannot guarantee that $T\mathbf{x}$ is equal to zero unless we're considering a norm space N instead of a seminorm space.

For N2, $\|\alpha T\| = \sup\limits_{\|\mathbf{x}\| \neq 0} \frac{\|\alpha T\mathbf{x}\|}{\|\mathbf{x}\|} = \sup\limits_{\|\mathbf{x}\| \neq 0} \frac{|\alpha|\|T\mathbf{x}\|}{\|\mathbf{x}\|} = |\alpha| \sup\limits_{\|\mathbf{x}\| \neq 0} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|} = |\alpha|\|T\|$. In the second step, the homogenity property is applied because of the norm of $\mathcal{R}(T)$. In the third step, the scalar can be factored out because it has no role in the supremum since it does not depend on $\mathbf{x}$.

For N3, $\sup\|(T_1 + T_2)\mathbf{x}\| = \sup\|T_1\mathbf{x} + T_2\mathbf{x}\|$. Since

$$\|T_1\mathbf{x} + T_2\mathbf{x}\| \leq \|T_1\mathbf{x}\| + \|T_2\mathbf{x}\|$$

and

$$\|T_1\mathbf{x} + T_2\mathbf{x}\| \leq \max\{\|T_1\mathbf{x}\|, \|T_2\mathbf{x}\|\}$$

and so also their supremum, thus $\sup \|T_1\mathbf{x}+T_2\mathbf{x}\| \leq \sup \|T_1\mathbf{x}\| + \sup \|T_2\mathbf{x}\|$ and $\sup \|T_1\mathbf{x}+T_2\mathbf{x}\| \leq \sup \max \{\|T_1\mathbf{x}\|, \|T_2\mathbf{x}\|\}$

$= \max \{\sup \|T_1\mathbf{x}\|, \sup \|T_2\mathbf{x}\|\}$

This space is complete, regardless of the completion of X, provided that Y is complete. Vector multiplication is defined as the composition of operators, which do not necessarily commute. The distributive law holds only if the domain of such single-valued operators is X. These definitions obey our axioms, making this space a bona fide Weak Banach Algebra.

There is a technical detail that needs to be dealt with in order to satisfy the homogeneity requirement i.e. $\|\alpha T\| = |\phi(\alpha)| \|T\|$: we must have a valuation preserving homomorphism ϕ between the skew field. This is because even if ϕ is an automorphism, assuming the AC, we can get many "wild" automorphisms which do not preserve order so that the basic assumption of being able to scale appropriately is destroyed[5]. If the field is non-Archimedean, then the situation may change [42].

It now makes sense to introduce a notation – for operators on a space X, we have $B_\phi(X,X) = B_\phi(X)$. In the spirit of Banach-Mazur theorem, we have the following characterisation:

Proposition 33 *Let $(X, \|.\|)$ be a unital Weak Banach algebra. Then X is a closed subalgebra of $B_\phi(X)$[43]*

Proof. We have already showed that $B_\phi(X)$ is an algebra. We show that X is homomorphic to $B_\phi(X)$: let $L_\mathbf{x}(\mathbf{y}) = \mathbf{xy}$. Then, $L_\mathbf{x}(\mathbf{y}_1+\mathbf{y}_2) = x(\mathbf{y}_1+\mathbf{y}_2) = \mathbf{xy}_1 + \mathbf{xy}_2 = L_\mathbf{x}(\mathbf{y}_1) + L_\mathbf{x}(\mathbf{y}_2)$ and $L_\mathbf{x}(\alpha\mathbf{y}) = x(\alpha\mathbf{y}) = \alpha\mathbf{xy} = \alpha L_\mathbf{x}(\mathbf{y})$ so that this operator is linear. Hence $B_\phi(X) \neq \varnothing$

Next, let $L : X \longrightarrow B_\phi(X)$ be defined as $L(\mathbf{x}) = L_\mathbf{x}$

Then, $L(\mathbf{x}+\mathbf{y}) = L_{\mathbf{x}+\mathbf{y}}$. Now, $L_{\mathbf{x}+\mathbf{y}}(\mathbf{z}) = (\mathbf{x}+\mathbf{y})z = \mathbf{xz}+\mathbf{xz} = L_\mathbf{x}(\mathbf{z}) + L_\mathbf{y}(\mathbf{z}) = (L_\mathbf{x}+L_\mathbf{y})(\mathbf{z})$ for any $\mathbf{z}$ so that $L_{\mathbf{x}+\mathbf{y}} = L_\mathbf{x} + L_\mathbf{y}$

and hence $L(\mathbf{x}+\mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y})$

Next, $L(\alpha\mathbf{y}) = L_{\alpha\mathbf{y}}$

Thus, $L_{\alpha\mathbf{y}}(\mathbf{z}) = (\alpha\mathbf{y})(\mathbf{z}) = \alpha(\mathbf{yz}) = \alpha L_\mathbf{y}(\mathbf{z})$ for any $\mathbf{z}$ so that $L_{\alpha\mathbf{y}} = \alpha L_\mathbf{y}$ and, therefore, $L(\alpha\mathbf{y}) = \alpha L(\mathbf{y})$.

Hence L is a homomorphism and thus $L(X)$ is a subspace of $B_\phi(X)$.

[5] See §8.3

We define a seminorm $\|.\|_o$ on X to be the restriction of the operator norm of $B_\phi(X)$ to $L(X)$, that is $\|\mathbf{x}\|_o := \|L_\mathbf{x}\| = \sup_{\|\mathbf{y}\|\leq 1} \|\mathbf{xy}\|$

Since $L_\mathbf{x}$ is linear operator without deception or fraud, $\|.\|_o$ obeys the axioms of a seminorm. We now show that $\|.\|$ and $\|.\|_o$ are equivalent.

For $\|\mathbf{y}\| \leq 1$, we have $\|\mathbf{xy}\| \leq \|\mathbf{x}\| \|\mathbf{y}\| \leq \|\mathbf{x}\|$. This shows that $\|\mathbf{x}\|_o \leq \|\mathbf{x}\|$. On the other hand, we have

$$\|\mathbf{x}\| = \|\mathbf{xe}\| \leq \sup_{\|\mathbf{y}\|\neq 0} \|\mathbf{xy}\| = \|\mathbf{x}\|_o$$

This shows that $\|\mathbf{x}\| \leq \|\mathbf{x}\|_o$ for all $\mathbf{x} \in X$ and shows that the seminorms are equivalent. In addition, this guarantees us the completeness of the new seminorm and X retains its Banach algebra character but with the help of operators. Since the norm is similar, we must have $(X, \|.\|_o) \simeq L(X)$ is a subspace of $B_\phi(X)$

Since X is a complete subspace of a complete space $B_\phi(X)$, X is closed. ■

This piece of machinery allows us to confidently suggest that $\|\mathbf{e}\|_o = 1$ by definition so that we do not have to worry about having a unit with a seminorm not equal to 1. Furthermore, it also justifies studying $\mathbb{K}$-algebras from the point of view of operators. However, it is worth remembering that this is not a complete characterisation since many properties do not "carry over". For instance, any projection operator is linear and bounded (thus are members of $B_\phi(X)$) and for orthogonal subspaces[6] M, N, two projection operators P_1, P_2 such that $P_1(X) = M$ and $P_2(X) = N$ together imply that $P_1 P_2 = 0$. Thus, $B_\phi(X)$ is not a division algebra because it has zero divisors, even if X is.

What about ordinary linear functionals? They have an interesting story to tell.

5.1 Riesz Representation Theorem on Hermitian Spaces

Theorem 34 *Let X be defined as in the previous chapter. Then there exists an injective linear functional g between X and X^*, the algebraic dual of X such that $\mathcal{R}(g) = X'$, the collection of all continuous functionals on X. Moreover, if I is a collection of anisotropic vectors, then the*

[6] The condition of orthogonality is important: for $\mathbf{n} \in N$ and $\mathbf{m} \in M$, there exists $\mathbf{x}, \mathbf{y} \in X$ such that $P_1(\mathbf{x}) = \mathbf{n}$ and $P_2(\mathbf{x}) = \mathbf{m}$ so that $\varphi(P_2 P_1(\mathbf{x}), \mathbf{y}) = \varphi(P_1(\mathbf{x}), P_2(\mathbf{y})) = \varphi(\mathbf{n}, \mathbf{m}) = 0$. Similarly, $\varphi(\mathbf{x}, P_1 P_2(\mathbf{y})) = 0$. Hence $P_1 P_2 = 0$.

kernel of each element of $g(I)$ is splitting.

Before we embark on the proof, we remark that for us, a linear functional will be continuous if it has a closed kernel. That is, $\ker g = \ker g^{\perp\perp}$. Furthermore, we call a subspace F splitting[7] if $X = F \oplus F^{\perp}$

Proof. Let $y \in X$ be arbitrary. Define $g_y : X \longrightarrow X^*$ such that $g_y(x) = \varphi(y, x)$. This is well defined and injective since for $x_1 = x_2 \iff x_1 - x_2 = 0 \iff \varphi(y, 0) = 0 = \varphi(y, x_1 - x_2) = \varphi(y, x_1) - \varphi(y, x_2)$. By definition, $\ker g_y$ consists of vectors (and their span) orthogonal to y. Thus, $\ker g_y = \{ky : k \in \mathbb{K}\}^{\perp}$, which is closed subspace of X. This implies that $\mathcal{R}(g) \subseteq X'$. To show the converse, let $h \in X'$. If $h = 0$, then $g_0 = h$ so that $h \in \mathcal{R}(g)$. If $h \neq 0$, then $\dim h = 1$ so that $\ker h$ has codimension 1 which assures us of the existence of different vectors kv such that $X = \ker h \oplus \{kv : k \in \mathbb{K}\}$. Choose $0 \neq z \in \ker h^{\perp}$ and $z \notin \{kv : k \in \mathbb{K}\}^{\perp}$ such that $\varphi(z, v) \neq 0$. Letting $w = f^{-1}\left(\varphi(v, z)^{-1} h(v)\right) z$ gives us $h(v) = \varphi(v, w)$. Now let $x \in X$. Then, there exists $x_1 \in \ker h$ and $\alpha \in \mathbb{K}$ such that $x = x_1 + \alpha v$. Applying h on both sides gives us $h(x) = \alpha h(v)$ and $\varphi(x, w) = \alpha\varphi(v, w)$ so that $h = g_w$ and hence $h \in \mathcal{R}(g)$.

If y is anisotropic, then $y \notin \{ky : k \in \mathbb{K}\}^{\perp}$ hence there exists g_y such that $\ker g_y = \{ky : k \in \mathbb{K}\}^{\perp}$ is closed and $X = \{ky : k \in \mathbb{K}\} \oplus \{ky : k \in \mathbb{K}\}^{\perp}$ so that $\ker g_y$ is splitting. ■

Initially, this theorem was only included to show that the completeness requirement for Quantum Mechanics can be weakened. It turns out, however, that this is of no use:

Corollary 35 *φ admits nonzero isotropic vectors, then there are closed subspaces of X that are not splitting.*

Proof. If $0 \neq y \in X$ such that $\varphi(y, y) = 0$, then $\{ky : k \in \mathbb{K}\} \oplus \{ky : k \in \mathbb{K}\}^{\perp} \subset X$ ■

In other words, if there are no closed subspaces of X that are not splitting (i.e. X is orthomodular), then φ does not admit isotropic vectors. More can be said: if every closed subspace of X is splitting (i.e. X is orthomodular) with an orthogonal sequence, then X is a Hilbert Space (see next chapter) with the underlying field being either $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$. It might come as a surprise that two innocent assumptions can fully describe the space. Furthermore, as is evident in the proof of Solèr's theorem, the existence of an orthogonal sequence is rather made heavy use of, instead of the orthomodularity assumption.

[7] Not to be confused with a split space in theory of Quadratic Forms, a completely opposite concept

Thus, in essence, the fundamental difference between Relativity and Quantum Mechanics turns out to stem from the existence of isotropic vectors.

If X is orthomodular, then for any subset F of X, it can be proved that $\langle F\rangle = F^{\perp\perp}$

Proof. Since $F \subseteq F^{\perp\perp}$, it follows that $F^{\perp\perp} \in [F]$ and hence $\langle F\rangle \subseteq F^{\perp\perp}$. The converse follows from the fact that φ is continuous and so are vector addition (and multiplication) and scalar multiplication in a Hermitian space. ■

This satisfies Definition 6[8] of [11]. In particular, the definition follows by Definition 1 of [11] by assuming that $\varphi(x, y) = 0$ implies $x \neq y$ or, equivalently, $\varphi(x, x) \neq 0$ for any x. There is no mention of zero vectors since for the cited source, X is just an ordinary set, not a Hermitian space and, instead, $x \perp y$ is taken as a binary relation on X.

A collection of all sets of the form $F^{\perp\perp}$forms an atomic ortholattice. That is, for all subsets F and $x \in X \backslash F^{\perp\perp}$, $(F \cup \{x\})^{\perp\perp}$ covers $F^{\perp\perp}$. Hence by Piron's theorem, such a lattice is isomorphic to the lattice of closed subspaces of a Hilbert space over an arbitrary Archimedean skew field.

Thus, orthomodular Hermitian Spaces will admit only Archimedean skew fields. In particular, the infinitesimals have no room. This is explicitly shown in the proof of Solèr's theorem, to which we now turn.

[8] For lack of time and considering the length of this report, we shall not enter into a detailed discussion of the cited paper.

Chapter 6

Underlying field

We have seen that the choice of the underlying field influences the structure of the vector space itself in general. Hence, the choice of the underlying field is not trivial. So, which field should one choose?

One answer to this question is rather non-conventional and involves formalising our way of making sense of Quantum Mechanics at the level of logic. In order to have a reasonably well-developed form of logic, one must have a model of syntax and semantics. Lattice Theory plays the part for the provision of syntax (language) and the truth values offer an interpretation (semantics) via a valuation. We will not be concerned with the latter. The former algebra happens to be converted to propositions via a mapping. Thus, x' becomes $\neg x$ whereas $x \& y$ is converted to $x \wedge y$. This requires us to develop a rule for formulating strings, which is taken care of in Lattice Theory.

The idea now is straight-forward: our classical world, the dynamics of which are modelled by a Poisson Manifold, is understandable from the point of view of classical logic – that is, Boolean Lattice, which is an orthocomplemented, distributive lattice. By orthocomplementation, we mean the following: let L be a lattice. For each $x \in$L, let $M_x \subseteq M$ be the set of complements of x. L is said to be **orthocomplemented** if there is a function f :L$\longrightarrow M$, called an **orthocomplementation**, such that

1. $f(x) \in M_x$

2. $f(f(x)) = x$

3. $x \leq y \implies f(y) \leq f(x)$ for all $x, y \in$L

The two element $\{0,1\}$ Boolean algebra can model electrical circuits. Another type of Boolean algebra is that on the set of positive divisors of n for any natural number n with $a \wedge b := \gcd(a,b)$ and $a \vee b := \operatorname{lcm}(a,b)$. This gives rise to partial order with $a \leq b \iff a \mid b$. The complement of a is given by $\frac{n}{a}$. Another Boolean Algebra known as the field of sets is an algebra $(\mathcal{P}(X), \cup, \cap)$ where X is any non-empty set. In fact, by Stone's representation theorem[1], every Boolean Algebra is isomorphic to a field of sets [19]. Boolean lattice model *classical logic* because such lattices capture the notions of classical logic.

6.1 Orthomodularity

For Boolean lattices, distributivity guarantees the uniqueness of complements. Distributivity has a physical correspondence in the classical case. However, from the point of the view of Quantum Mechanics, the distributive law fails. Consider the Hilbert space $\mathbb{R}$ and a particle moving on on the real line and then have the following propositions:

$a =$ "The particle has momentum in the interval [-1/12, 1/12]"

$b =$ "the particle is in the interval [-1, 1]"

$c =$"the particle is in the interval [1, 3]"

These intervals will correspond to some converted units with Planck's constant equal to unity. Thus, $\Delta p \Delta x \geq \frac{1}{2}$ must always hold. In the case of a, $\Delta p = \frac{1}{6}$. In the case of b and c, $\Delta x = 2$. It can be the case that the particle's momentum is between 0 and +1/6, and its position is between -1 and +3 because this does not violate the uncertainty relation. Thus, a and (b or c) is true. On the other hand, the propositions "a and b" and "a and c" are both false since they do not satisfy the uncertainty principle. So, (a and b) or (a and c) is false. Thus the distributive law fails. In summary, the most notable difference that clearly distinguishes Quantum Logic from classical logic is the failure of the distributive law. We are then forced to consider weaker laws.

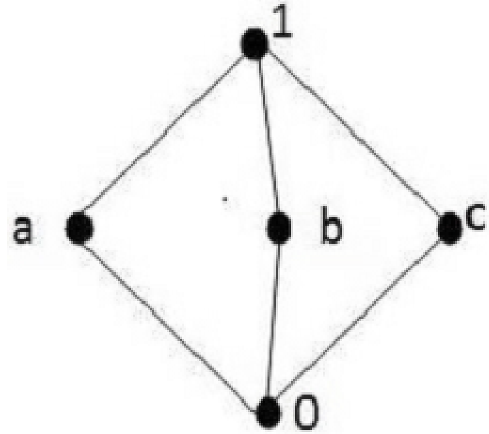
Definition 36 *A lattice L is* ***modular*** *if $x \leq z$ implies $x \vee (y \wedge z) = (x \vee y) \wedge z$ for all $x, y, z \in L$*

[1] Every Boolean algebra is isomorphic to $(\mathcal{P}(X), \cup \cap, ^c)$

Lemma 37 *Every distributive lattice is modular but the converse is not true*

Proof. Let $x \leq z$. Then, $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) = (x \vee y) \wedge z$ ■

Example 38 *The diamond lattice M_5 is an example for a modular lattice which is not distributive where M_5 is the lattice*



$a \vee (b \wedge c) = a \vee e = a$ *but* $(a \vee b) \wedge (a \vee c) = I \wedge I = I$. *Therefore,* M_5 *is not distributive.*

Definition 39 *A lattice L is* ***orthomodular*** *if $x \leq z$ implies $x \vee (x' \wedge z) = z$ for all $x, z \in$L[45]*

Lemma 40 *Every modular lattice is orthomodular*

Proof. Let $x \leq z$. Then, $x \vee (y \wedge z) = (x \vee y) \wedge z$ for all $x, y, z \in$L. Take $y = x'$. Then, $(x \vee y) \wedge z = (x \vee x') \wedge z = 1 \wedge z = z$ ■

The collection of closed subspaces of (countable dimensional) Hilbert Space H over field $\mathbb{K}$ form an orthomodular lattice that is not modular.

Now, which lattice should be a natural candidate for Quantum Logic? Orthomodular lattices or modular ones?

Based on the idea that every eigenvector is an eigenstate, the corresponding subspace generated by each eigenstate therefore corresponds to experimental propositions. In particular, we consider the lattice of closed subspaces of a Hilbert Space. Since the intersection of two subspaces is a subspace, a natural candidate for the meet operation is the set-theoretic intersection. However, the union of two subspaces is not necessarily a subspace. Hence we discard the set-theoretic union as a natural candidate for our join operation and instead, call the closed-span of the union of two subspaces as our join operation. This collection of subspaces is bounded below by the trivial subspace $\{0\}$ and the bounded above the trivial subspace H. Our complement operation has to ensure that the complement of any closed subspace ought to be a closed subspace. This is the case if we define the complement of M as $M^{\perp}$. The closed subspaces of (countable dimensional) Hilbert Space H over field $\mathbb{K}$ form an orthomodular lattice that is not modular. To relate this lattice theoretic definition with that of the Hilbert space, we have the following theorem:

Theorem 41 *The following are equivalent:*

1. *H is orthomodular: that is, if for all subspaces* $X \subset \mathcal{H} : X = X^{\perp\perp} \Rightarrow \mathcal{H} = X \oplus X^{\perp}$.
2. *Lattice of closed subspaces* $C(\mathcal{H})$ *is orthomodular*

Proof. ($2 \implies 1$)Notice that $A \vee B = \overline{span\,(A \cup B)} = A \oplus B$ where A, B are closed subspaces of H. The direct sum of two closed subspaces is closed. Now, for $\forall A, B \in C(\mathcal{H})$, $A \leq B \implies A \vee \left(A^{\perp} \wedge B\right) = B$. In particular, for B =H, we have $A \oplus \left(A^{\perp} \cap \mathcal{H}\right) = A \oplus A^{\perp}$ =H.

($1 \implies 2$)Let $A, B \in C(\mathcal{H})$ such that $A \subseteq B$ ($\implies B^{\perp} \subseteq A^{\perp}$). We have to show that $A \oplus \left(A^{\perp} \cap B\right) = B$. Let $0 \neq z \in A \oplus \left(A^{\perp} \cap B\right)$. Then, there exists $x \in A$ and $y \in \left(A^{\perp} \cap B\right)$ such that $x + y = z$. In particular, $y \in B$. Since $A \subseteq B$, then $x \in B \implies x + y \in B$. Hence $A \oplus \left(A^{\perp} \cap B\right) \subseteq B$. To show that B is a subset of $A \oplus \left(A^{\perp} \cap B\right)$, let $z \in B$ =H$\cap B$ = $\left(A \oplus A^{\perp}\right) \cap B$, then $z = x + y$ where $z \in B$, $x \in A$ and $y \in A^{\perp}$. From $A \subseteq B$, it follows that $x \in B$ and that $y = z - x \in B$. That is, $y \in A^{\perp} \cap B$. Therefore, $z = x + y \in A \oplus \left(A^{\perp} \cap B\right)$. ■

In the countably infinite case, every such (separable) Hilbert space is isomorphic to $l^2(\mathbb{C})$ which is again orthomodular. Separability is important for otherwise, with uncountable basis, we end up with all sorts of trouble, including summing elements of the interval $[0, 1]$.

For modularity, finite dimensional spaces are exempted: lattice of all orthogonally closed subspaces of $\mathcal{H}$ is modular iff $\mathcal{H}$ if finite dimensional [8]. What if weaken the structure? That is, what about the lattice of subspaces of a Hermitian space? In 1980, Hans Keller proved that the lattice of all orthogonally closed subspaces of a Hermitian space $(E, \varphi, \mathbb{K})$ is modular if and only if E is finite dimensional, where φ is non-degenerate, Hermitian form [23]. Even if finite dimensional spaces are more well-behaved than infinite dimensional ones yet the historical significance of infinite dimensional spaces in general and L^2 in particular cannot be discarded. Therefore, one must drop modularity, as the following counter-example attests to:

Example 42 *Consider* $H = L^2[0,1]$ *and form the lattice of closed subspaces* $C(\mathcal{H})$*. Recall that* $C[0,1], L^\infty[0,1], L^1[0,1] \subseteq L^2[0,1]$ *are subspaces which are not closed. If the closed subsets are not orthogonal to each other, then it is possible to find subspaces such that* $cl\left(L^1[0,1]\right) \subset cl(C[0,1])$ *but* $(cl(C[0,1]) \vee cl(L^\infty[0,1])) \cap cl\left(L^1[0,1]\right) \supset cl(C[0,1]) \vee (cl(L^\infty[0,1]) \cap cl\left(L^1[0,1]\right))$*.*

$C(\mathcal{H})$ is not distributive.

Example 43 *Let* $H = \mathbb{R}^2$*. Consider subspaces* $X = (x,0)$*,* $Y = (0,y)$ *and* $Z = (mx, y)$ *where* $y = mx$*. Then,* $(X \oplus Y) \cap Z = Z$ *but* $(X \cap Z) \oplus (Y \cap Z) = \{0\}$*.*

This approach focuses on states. However, there is an equivalent approach to measurement of such states: there is a one-to-one correspondence between a projection operator and a closed subspace of H.

Proof. If P is a projection operator, then the space $\mathrm{H}_p = P(\mathcal{H})$ is closed. We show that $\ker P^\perp = P(\mathcal{H})$ where $\ker P = \{y : y \in \mathcal{H} \text{and } P(y) = 0\}$. If $x \in P(\mathcal{H})$, then $P(x) = x$ so that $P(x) \neq 0$ and $x \notin \ker P \implies x \in \ker P^\perp$. Thus, $P(\mathcal{H}) \subseteq \ker P^\perp$. It follows that $\ker P \subseteq P(\mathcal{H})^\perp$.

To show the converse, let $x \in \ker P^\perp$. Then, for all $y \in \ker P$, $\langle x, y\rangle = 0$ and $y \in P(\mathcal{H})^\perp$ so that $x \in P(\mathcal{H})$ and hence $\ker P^\perp \subseteq P(\mathcal{H})$.

Vice versa, since every vector $z \in$H can be uniquely decomposed as $z = x + y$ where $x \in \mathcal{V}$ and $y \in \mathcal{V}^\perp$. The linear map defined via $P_\mathcal{V}(z) = x$ is then a projection.

We shall write $P_\mathcal{V} = P$ for simplicity. Let $z \in$H. Then, $P^2(z) = P(x) = x$ since $x = x + 0$ and every subspace is orthogonal to the trivial subspace $\{0\}$. Also, $P(z) = x$ by definition. That is $P(z) = P^2(z)$ for all z. Hence $P = P^2$. Second, let $z_1, z_2 \in$H with $z_1 = x_1 + y_1$ and $z_2 = x_2 + y_2$ with $x_1, x_2 \in \mathcal{V}$ and $y_1, y_2 \in \mathcal{V}^\perp$. Then, $\langle P(z_1), z_2\rangle = \langle x_1, z_2\rangle = \langle x_1, x_2 + y_2\rangle = \langle x_1, x_2\rangle + \langle x_1, y_2\rangle$ whereas $\langle z_1, P(z_2)\rangle = \langle z_1, x_2\rangle = \langle x_1 + y_1, x_2\rangle = \langle x_1, x_2\rangle + \langle y_1, x_2\rangle$. It remains to show that $\langle x_1, y_2\rangle = \langle y_1, x_2\rangle$. This holds because $\langle x_1, y_2\rangle = \langle y_1, x_2\rangle = 0$. ■

Around the introduction of Quantum Logic in 1936 [8] by von Neumann and Birkhoff, an abstract definition of orthomodular lattices had not been formulated but various instances were known, including the canonical Hilbert lattice C($\mathcal{H}$). This natural existence implied the selection of orthomodular lattices to model the logic of quantum mechanics yet this was not the choice of the founding fathers of Quantum Logic for technical reasons. Instead, von Neumann wished to preserve modularity. Von Neumann's letter of January 19,1935 writes:

> Using the operator-description, $a \vee b, a \wedge b$ can be formed, if the physically significant operators form a ring [of operators]. This, I think should be assumed anyhow, even if one does not require that all operators are physically significant. But we need probably not insist on this point too much [40].

By November 1935, P($\mathcal{H}$) was then given consideration again:

> I am somewhat scared to consider all physical quantities = bounded self-adjoint operators as a lattice [40].

This was because for unbounded, densely defined operators, we are not guaranteed a non-empty intersection of domains and, therefore, von Neumann discards $\mathcal{P}(\mathcal{H})$ as a natural candidate. Von Neumann emphasizes:

> Examples could be constructed which make no use of operator theory, but I think that this example shows more clearly "what it's all about": it is the existence of "pathological" operators – like X,Y above – in Hilbert space, which destroys the [modular] lattice characters [40]

This pathology, for von Neumann is a serious problem because it prohibits adding and composing these operators in general so that the essential requirement of closure for an algebra is not satisfied, which appeared to him rather un-physical [30]. This character against the hope to achieve an algebra of observables was one of the main reasons why, in his famous talk on "Unsolved Problems in Mathematics" in 1954 suggested that a restricted set of operators would be a more suitable mathematical framework for quantum mechanics than Hilbert space theory [34].

In this situation von Neumann saw two options:

(I) Either we define the "join" by $\cup$ (as a simple linear sum)... but we must admit all (not-necessarily-closed) linear subspaces,

(II) or we define the "join" by $\vee$ (closure of the linear sum), then [modularity] is lost [40].

A collection of closed subspaces is rather natural since all probabilities in the state K are equal to those in the state $K^{\perp\perp}$ yet $K \cap\ L \neq K^{\perp\perp} \cap\ L^{\perp\perp}$ does not hold in general. Given this and the inclination of von Neumann, one would expect that (I) would have been chosen: on the contrary, the orthogonal complement K still has the property $K^{\perp} + L^{\perp} = (K \cap\ L)^{\perp}$ but $K^{\perp} \cap L^{\perp} = (K \oplus\ L)^{\perp}$ and $K^{\perp\perp} = K$ are not necessarily true; for instance, in L^2 spaces. This is because, assuming anisotripicity, we have $K \cap K^{\perp} = \{0\}$, while $K \oplus K^{\perp}$ is everywhere dense but not necessarily equal to H for Hilbert Spaces over general fields $\mathbb{K}$.

We are, therefore, left with the canonical orthomodular lattice of closed subspaces $C(\mathcal{H})$ of a Hilbert space to model Quantum Logic, as envisioned in the pioneering paper by von Neumann and Birkhoff in 1936 [8]. Here, the "and" operation is defined as the ordinary set intersection and the "or" operation of two subspaces is their direct sum. Complements are defined by orthogonal complements. To qualify as a logic, $C(\mathcal{H})$ must be a language consisting of propositions and connectives (operations) axioms and a rule of inference.

The easiest way to understand propositions of Quantum Logic is in terms of logic of experimental propositions. That is one can define explicitly some connectives for a certain special class of well-suited propositions relating to idealised quantum mechanical tests with an aim to obtain a logical system satisfying certain formal requirements. If this is accomplished, one says that one has introduced non-classical or non-Boolean logic. Provided that the formal requirements are rigorously met, this accomplishment should be non-controversial. The controversies, however, surround the underlying assumptions and the pragmatism of a non-classical logic.

A proposition a in Quantum Logic is represented by means of the closed subspace M_a of Hilbert space H used to describe the quantum entity under consideration. An alternate way is by means of the orthogonal projection operator P_a on this closed subspace, which we have seen is in one-to-one correspondence, and, therefore, $\mathcal{P}(\mathcal{H})$, the collection of projection operators on a Hilbert space, forms our required lattice.

Following the approach of [2], let P be the set of propositions of a quantum system. A quantum state q is then represented by a vector $u_q \in$H such that $\|u_q\| = 1$.

Taking analogy with the Boolean Lattice case, for a proposition $a \in P$, we can assign a closed subspace M_a. This vector belongs to a subspace, which is associated with a proposition in Quantum Logic (i.e., $u_q \in M_a$). If M_b is another closed subspace associated with the proposition b, we define Quantum Logical operations as follows:

$$a \Longrightarrow b \underline{\longleftrightarrow} M_a \subseteq M_b$$

$$a \wedge b \underline{\longleftrightarrow} M_a \wedge M_b = M_{a \wedge b} = M_a \cap M_b \tag{6.1}$$

$$a \vee b \underline{\longleftrightarrow} M_a \vee M_b = M_{a \vee b} = cl(M_a \cup M_b) \tag{6.2}$$

$$N\left(a\right) \underline{\longleftrightarrow} M_{\neg a} = M_a^{\perp} \tag{6.3}$$

$cl(M_a \cup M_b)$ is the topological closure of the linear space generated by the set $M_a \cup M_b$. It can be shown that $M_a \oplus M_b = cl(M_a \cup M_b)$. Here, the closure operation depends on the topology generated by the inner product. In either definition, $M_a \vee M_b$ is the smallest closed subspace of H that contains both M_a and M_b.

Instead of assigning a valuation for semantics, we can take advantage of the Hilbert structure of Quantum Mechanics. To this end, we say that a proposition $a \in P$ is true if the associated quantum state q upon "inquiring a" has unit probability. From the axioms of quantum mechanics, we can see that a unit probability corresponds to the fact that the vector $u_q \in M_a$. This is equivalent to $Pu_q = u_q$ where P is a projection operator such that $P_a\left(\mathcal{H}\right) = M_a$. According to these criterion, a positive result will correspond to P_a whereas a negative result will correspond to $I - P_a$, where I is the unit operator. This makes sense since $\left(I - P_a\right)\left(P_a\right) = P_a - P_a^2 = 0$ and so $a \wedge \sim a$ is always false. On the other hand, $\left(I - P_a\right)\left(\mathcal{H}\right) =H- M_a = M_a^{\perp}$

Equipped with these and the equivalences above, it is then no surprise that Quantum Logical conjunction and implication behave in a manner similar to that of classical logic. Disjunction is different; since $M_b \subset cl(M_a \cup M_b)$ and $M_a \subset cl(M_a \cup M_b)$, it follows that if a is true, then $a \vee b$ is true. Similarly, if b is true, then $a \vee b$ is true as well. However, we cannot admit a converse. This is because $cl(M_a \cup M_b)$ will contain vectors other than those contained in either M_a or

M_b. We conclude that the truth table of classical and quantum conjunction and implication are the same but that for disjunction is different from classical and Quantum Logic.

The Quantum Logic negation is also not same as the classical counterpart: consider a proposition $a \in P$ and suppose $\sim a$ is true. Associate the quantum proposition a represented by a vector u_a belonging to the closed subspace M_a. This vector belongs to orthogonal complement of M_a. That is, $u_a \in M_a^{\perp}$. but at the same time, this vector does not belong to M_a that is $u_a \notin M_a$ unless it is zero since $M_a^{\perp} \cap M_a = \{0\}$ .Hence a is not true. Thus, if a proposition is false, then it's negation is true. In summary, if the quantum negation of a proposition is true, then the classical negation of this proposition is true. The converse does not hold: assume that $\sim a$ is false. Then, $u_a \notin M_a^{\perp}$. Even if $\mathcal{H} = M_a^{\perp} \oplus M_a$, then we yet cannot conclude that $u_a \in M_a$ (it could belong to the span of both) so that we cannot say that a is true.

Our idea of saying that a proposition is true might give away the impression that quantum logic is fundamentally multivalued. If Quantum Logic is to make sense of the real world, valuations must correspond to reality. According to the Copenhagen interpretation, any particle assumes a superposition of infinite values. In particular, a spin-half particle can encode entire texts of Shakespeare. Jauch and Piron, however, beg to differ: they argued that if Quantum Logic is not infinite valued, then for any two proposition p and q there must exist a conditional proposition $p \implies q$. In fact, they proved that for a specific valuation, the interval $[0, 1]$ is reduced to $\{0, 1\}$ provided that we introduce a conditional proposition and that monotonic valuations are closed under the formation of mid-points [18]. These two assumptions may seem mild but are rather disturbing to the Kochen-Specker theorem: any orthomodular lattice admits total homomorphisms onto $\{0, 1\}$ if it is distributive. In other words, there are no homomorphisms from $\mathcal{C}(\mathcal{H})$ to $\mathbb{Z}_2$ – this is the famous Kochen-Specker theorem which, in one swift theorem, discards Einstein's hypothesis of the EPR paradox as mutually contradictory. In conclusion, Quantum Logic is neither fuzzy nor Boolen and thus demands a radical shift in thinking.

6.2 Solèr's theorem

With this in mind, we see that the requirement of orthomodularity of a lattice is rather forced for technical reasons. Thus, orthomodularity in the sense of a Hilbert space is an essential consideration. It turns out that orthomodularity is key factor one must take into account in order to determine that the Hermitian Space is Hilbert, thanks to Soler's theorem, which we now set the stage for. As expected, Hermitian space $(X, \mathbb{K}, \varphi)$ is called **orthomodular** if for all subspaces $F \subset X : F = F^{\perp\perp} \Rightarrow X = F \oplus F^{\perp}$. Not all Hermitian spaces are orthomodular, as was established in **Corollary 35**.

Note that in such a space, we are not even assuming positivity of the 2-form!

Let I be any indexing set. $(f_i)_{i \in I}$ is called λ-**orthogonal system** if $\varphi(f_i, f_k) = 0$, $(i \neq k)$ and if $\varphi(f_i, f_i) = \lambda$ for a $\lambda \in \mathbb{K}$. For $\lambda = 1$, $(f_i)_{i \in I}$ is called **orthonormal system**.

Assume that the orthomodular space $(X, \mathbb{K}, \varphi)$ has a γ-orthogonal system. That is, i.e. $\varphi(e_i, e_i) = \gamma$ for all $i \in \mathbb{N}$ and for $\gamma \in \mathbb{K}$. We can convert this into an orthonormal system as follows: we define a new involution $\alpha \longmapsto \tilde{\alpha}$ and a new form

$$\tilde{\varphi} : X \times X \longrightarrow \mathbb{K} \text{ such that } (x, y) \longmapsto \tilde{\varphi}(x, y) := \varphi(x, y)\, \gamma^{-1}$$

Then $(X, \mathbb{K}, \tilde{\varphi})$ is a Hermitian space relative to the involution "$\tilde{\cdot}$"and it is orthomodular because $\varphi(x, y) = 0 \iff \tilde{\varphi}(x, y) = 0$ for all $x, y \in X$. From hereon, $(e_i)_{i \in \mathbb{N}}$ will be treated as an orthonormal system in $(X, \mathbb{K}, \varphi)$.

Definition 44 *A positive definite, non-degenerate, Hermitian 2-from is called an* ***inner product*** *and* $(X, \mathbb{K}, \varphi)$ *is called an* ***inner product space****.*

$\mathbb{K}$ must have a compatible ordering, even if it is not Archimedean. Thus, we cannot have an inner product space over a finite field, implying an impossibility of "ordinary" Quantum Mechanics over a finite field, agreeing with a second approach in [12]. In particular, this means that we cannot define a norm over a finite field. We, therefore, choose to exclude them from our discussion from hereon.

Lemma 45 *If* φ *is an inner product, then* $\varphi(x, x) = 0 \iff x = 0$

Proof. ($\Longleftarrow$) Holds for any sesquilinear 2-form

($\Longrightarrow$) Since φ is non-degenerate, $\varphi(x, y) = 0$ for all x implies $y = 0$. In particular, it is valid for $x = y$. Hence $\varphi(x, x) = 0 \implies x = 0$ ■

Not all normed spaces are inner product spaces. The simplest example is that of $B(X)$ with X a real normed space. It is easy to see that $\|T - S\|^2 + \|T + S\|^2 \neq 2\left(\|T\|^2 + \|S\|^2\right)$: let $T, S : l^2 \longrightarrow l^2$ be projection operators. That is, $T(x_n) = x_n$ and $S(x_n) = x_m$. Then, $\|S\|, \|T\| \neq 0$ whereas $\|T - S\| = \|T + S\| = 0$ for $n \neq m$. That is, $B(X)$ does not obey the parallelogram law. On the other hand, for non-Archimedean Hermitian space, $\|T - S\|^2 + \|T + S\|^2 \leq 2\max\left(\|T\|^2, \|S\|^2\right)$, which is rather routine to verify. Therefore, an inner product cannot be made from the normed space of operators, if the underlying field is Archimedean but our hands are not tied when it comes to spaces over a non-Archimedean field.

A space X is called **half-normal** if it is orthomodular and if there exists an orthogonal system $(e_i)_{i\in\mathbb{N}} \subset X$. X is **normal** if it is half-normal and $X = \left((e_i)_{i\in\mathbb{N}}\right)^{\perp\perp}$. Topologically, X is a **normal space** if and only if, given any disjoint closed sets A and B, there are neighbourhoods U of A and V of B that are also disjoint. In fancier terms, this condition says that A and B can be separated by neighbourhoods, something akin to the Hausdroff topology but for sets instead of points. Thus, our first step will be to approach normality.

Orthomodularity guarantees the existence of closed sets whereas the orthogonal system and its span guarantee their separation. The topology is generated by the norm on X.

For our purposes, we shall shorten $\varphi(x, x) = \langle x, x\rangle = \langle x\rangle$, which we call the **length** of a vector, following Solèr. There is no danger of confusion with the expectation value as x is not necessarily an operator. We will also replace $\varphi(x, y)$ with $\langle x, y\rangle$ to make the proof look less messy than it already is.

Theorem 46 (M. P. Solèr's Theorem) *Let $(X, \mathbb{K}, \varphi)$ be an infinite dimensional orthomodular space over a skew field $\mathbb{K}$ which contains an orthonormal system $(e_i)_{i\in\mathbb{N}}$. Then $\mathbb{K}$ is either $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ and $(X, \mathbb{K}, \varphi)$ is a Hilbert space [44]*

Proof. Preparations

In the proof that follows, we shall assume that X is normal and then replace this condition without $X = \left((e_i)_{i\in\mathbb{N}}\right)^{\perp\perp}$ and forgo completeness until then.

Let x be any element of $\mathbb{K}$. Then, $nx = \langle\sum_{i=0}^{n} e_i\rangle\, x = 0 \iff \langle\sum_{i=0}^{n} e_i\rangle = 0 \iff n = 0$

Thus, the characteristic of the underlying field of the half-normal space X is zero. Thus, $\mathbb{K}$ is infinite and therefore, $\mathbb{Z} \subset \mathbb{K}$ but since $\mathbb{K}$ is a skew field, $\mathbb{Q} \subset \mathbb{K}$.

Part A

Let $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$. First we show that for every sequence $(\alpha_i)_{i\in\mathbb{N}^*} \in \mathbb{Q}^{\mathbb{N}^*}$ with $\alpha := \sum_{i=0}^{\infty} \alpha_i^2 \in \mathbb{Q}$ there is a vector $x = \sum_{i\in\mathbb{N}^*} \alpha_i e_i \in X$, with $\langle x\rangle = \alpha$. For the rationals, we also have $\mathbb{Q}^+ \subset \mathbb{Q}^2$. Without loss of generality, we assume $\alpha_i \neq 0$ for all $i \in \mathbb{N}$. We define

$$\lambda_0 := \alpha$$

$$\lambda_n := \frac{\alpha^2}{\alpha - \sum_{i=1}^{n} \alpha_i^2} - \sum_{i=0}^{n-1} \lambda_i \ (n \geq 1)$$

so that

$$\sum_{i=0}^{n} \lambda_i = \frac{\alpha^2}{\alpha - \sum_{i=1}^{n} \alpha_i^2}$$

Clearly, $\lambda_i \in \mathbb{Q}$ and so does $\sum\limits_{i=0}^{n} \lambda_i$. However, the sum $\sum_{i=0}^{\infty} \lambda_i$ diverges in $\mathbb{R}$ and therefore in $\mathbb{Q}$ because the denominator $\alpha - \sum_{i=1}^{n} \alpha_i^2 \longrightarrow 0$. Clearly, we are currently considering only one such sequence and not every. This will be accommodated in **Part B** of this proof.

Since X contains orthonormal systems, one finds an orthogonal system $x_0, x_1, x_2, ...$ in X with $\langle x_i\rangle = \lambda_i$ for all $i \geq 0$. Let $F := ((x_i)_{i\in\mathbb{N}})^{\perp\perp}$ be a normal space and

$$y_n := x_0 + ... + x_{n-1} - \frac{\lambda_0 + ... + \lambda_{n-1}}{\lambda_n} x_n \ (n \geq 1)$$

Since $\langle x_i, x_j\rangle = 0$ for for $i \neq j$, we have $y_i \perp y_j$ for $i \neq j$. Thus,

$$\begin{aligned} \langle x_0, y_n\rangle &= \left\langle x_0, \sum_{j=0}^{n-1} x_j - \frac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \right\rangle \\ &= \left\langle x_0, \sum_{j=0}^{n-1} x_j \right\rangle - \left\langle x_0, \frac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \right\rangle \\ &= \langle x_0\rangle = \lambda_0 = a \end{aligned}$$

for $n \geq 1$ and

$$
\begin{aligned}
\langle y_n \rangle &= \left\langle \sum_{j=0}^{n-1} x_j - \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ \right\rangle \\
&= \left\langle \sum_{j=0}^{n-1} x_j - \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ , \sum_{j=0}^{n-1} x_j - \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ \right\rangle \\
&= \left\langle \sum_{j=0}^{n-1} x_j \ , \sum_{j=0}^{n-1} x_j - \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ \right\rangle - \left\langle \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ , \sum_{j=0}^{n-1} x_j - \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ \right\rangle \\
&= \left\langle \sum_{j=0}^{n-1} x_j \ , \sum_{j=0}^{n-1} x_j \right\rangle - \left\langle \sum_{j=0}^{n-1} x_j \ , \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ \right\rangle + \left\langle \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ , \sum_{j=0}^{n-1} x_j \right\rangle + \left\langle \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ , \tfrac{x_n}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \ \right\rangle \\
&= \sum_{j=0}^{n-1} \langle x_j \rangle + \langle x_n \rangle \left(\tfrac{1}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \right)^2 \\
&= \sum_{j=0}^{n-1} \lambda_j + \lambda_n \left(\tfrac{1}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \right)^2 \\
&= \sum_{j=0}^{n-1} \lambda_j + \tfrac{1}{\lambda_n} \left(\sum_{j=0}^{n-1} \lambda_j \right)^2 \\
&= \left(\sum_{j=0}^{n-1} \lambda_j \right) \left(1 + \tfrac{1}{\lambda_n} \sum_{j=0}^{n-1} \lambda_j \right) \\
&= \left(\tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \right) \left(1 + \tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \left(\tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n} \alpha_i^2} - \tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \right)^{-1} \right) \\
&= \left(\tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \right) \left(1 + \tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \left(\tfrac{\alpha^2 \left(\alpha - \sum_{i=1}^{n-1} \alpha_i^2\right) - \alpha^2 \left(\alpha - \sum_{i=1}^{n} \alpha_i^2\right)}{\left(\alpha - \sum_{i=1}^{n} \alpha_i^2\right)\left(\alpha - \sum_{i=1}^{n-1} \alpha_i^2\right)} \right)^{-1} \right) \\
&= \left(\tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \right) \left(1 + \tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \tfrac{\left(\alpha - \sum_{i=1}^{n} \alpha_i^2\right)\left(\alpha - \sum_{i=1}^{n-1} \alpha_i^2\right)}{\alpha^2 \left(\alpha - \sum_{i=1}^{n-1} \alpha_i^2\right) - \alpha^2 \left(\alpha - \sum_{i=1}^{n} \alpha_i^2\right)} \right) \\
&= \left(\tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \right) \left(1 + \alpha^2 \tfrac{\left(\alpha - \sum_{i=1}^{n} \alpha_i^2\right)}{-\alpha^2 \sum_{i=1}^{n-1} \alpha_i^2 + \alpha^2 \sum_{i=1}^{n} \alpha_i^2} \right) \\
&= \left(\tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \right) \left(1 + \alpha^2 \tfrac{\left(\alpha - \sum_{i=1}^{n} \alpha_i^2\right)}{\alpha^2 \alpha_n^2} \right) \\
&= \left(\tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \right) \left(\tfrac{\alpha_i^2 + \alpha - \sum_{i=1}^{n} \alpha_i^2}{\alpha_n^2} \right) \\
&= \left(\tfrac{\alpha^2}{\alpha - \sum_{i=1}^{n-1} \alpha_i^2} \right) \left(\tfrac{\alpha - \sum_{i=1}^{n-1} \alpha_i^2}{\alpha_n^2} \right) = \alpha^2 \alpha_n^{-2} \text{ for all } n \geq 1.
\end{aligned}
$$

Since $(y_i)_{i \in \mathbb{N}^*}$ is orthogonal, we set $f_i := \alpha^{-1} \alpha_i y_i$ for $i \geq 1$ and obtain an orthonormal system $(f_i)_{i \in \mathbb{N}^*}$. In this case,

$$
\begin{aligned}
\langle x_0, f_i \rangle &= \left\langle x_0, \alpha^{-1} \alpha_i y_i \right\rangle \\
&= \langle x_0, y_i \rangle \, \alpha^{-1} \alpha_i \\
&= \alpha \alpha^{-1} \alpha_i = \alpha_i
\end{aligned}
$$

In other words, there is a vector

$$f=\sum_{i=1}^{\infty}\alpha_i f_i \in ((f_i)_{i\in\mathbb{N}^*})^{\perp\perp} \subseteq F=((x_i)_{i\in\mathbb{N}^*})^{\perp\perp}\subset X$$

The crucial point is to show now that $\langle f\rangle = a$. A direct calculation can show this but it misses the important condition that $f=x_0$. Thus we have to show that $(f_i)_{i\in\mathbb{N}^*}$ is maximal in F. That is, $((f_i)_{i\in\mathbb{N}^*})^{\perp\perp}=F$. Assume by way of contradiction that the reverse containment does not hold. Then, there is $0\neq z\in F$ with $z\perp y_i$ for $i>1$ so that $z\perp f_i$ and hence $z\perp F^{\perp}$. We cannot have $x_0\perp z$. If this were the case, then $\langle z,y_1\rangle=\langle z,x_0\rangle+\frac{\lambda_0}{\lambda_1}\langle z,x_1\rangle=0$ implies $\langle z,x_1\rangle=0$ and $x_i\perp z$ for all i by induction. Now, $f_n=\alpha^{-1}\alpha_n y_n=\alpha^{-1}\alpha_n\sum_{j=0}^{n-1}x_j-\alpha^{-1}\alpha_n\frac{x_n}{\lambda_n}\sum_{j=0}^{n-1}\lambda_j$ implies $z\perp F=((x_i)_{i\in\mathbb{N}})^{\perp\perp}$ so that $z\in F^{\perp\perp}\cap F^{\perp}=\{0\}$. Thus, $z=0$, which is a contradiction. By scaling z, we may assume that $\langle x_0,z\rangle=\alpha$. Definition of the y_i entails

$$\begin{aligned}
0 &= \langle y_i,z\rangle\\
&= \left\langle \alpha^{-1}\alpha_n\sum_{j=0}^{n-1}x_j-\alpha^{-1}\alpha_n\frac{x_n}{\lambda_n}\sum_{j=0}^{n-1}\lambda_j, z\right\rangle\\
&= \alpha^{-1}\alpha_n\sum_{j=0}^{n-1}\langle x_j,z\rangle-\frac{\alpha^{-1}\alpha_n}{\lambda_n}\sum_{j=0}^{n-1}\lambda_j\langle x_n,z\rangle
\end{aligned}$$

i.e.

$$\alpha^{-1}\alpha_n\sum_{j=0}^{n-1}\langle x_j,z\rangle=\frac{\alpha^{-1}\alpha_n}{\lambda_n}\sum_{j=0}^{n-1}\lambda_j\langle x_n,z\rangle$$

or

$$\lambda_n\sum_{j=0}^{n-1}\langle x_j,z\rangle=\sum_{j=0}^{n-1}\lambda_j\langle x_n,z\rangle$$

$\implies \langle x_i,z\rangle=\lambda_i=\langle x_i\rangle$ for $i\geq 0$, so that

$$z=\sum_{i\in\mathbb{N}}x_i$$

Such a vector then does not exist. To this end, we show that a vector representation in terms of its basis is immaterial from a topological point of view: assume that $(a_i)_{i\in I}$ and $(b_i)_{i\in I}$ are orthogonal systems in X with $(a_i)_{i\in I}\perp(b_i)_{i\in I}$ and $\langle a_i\rangle=\lambda\langle b_i\rangle$ for all $i\in I$ and for a $\lambda\in\mathbb{K}$.

Let further be $x = \sum_{i\in I} \zeta_i a_i \in X$ with $\zeta_i \lambda = \lambda \zeta_i$ for all $i \in I$. This assumption might be unwarranted but later on, we shall select $\lambda = 1$ to show that this assumption does not matter. However, we cannot have $\lambda = -1$ for otherwise then $\langle a_i \rangle = -\langle b_i \rangle$ or $\langle b_i \rangle + \langle a_i \rangle = 0$ which implies

$$\begin{aligned}
\langle a_i + b_i \rangle &= \langle a_i + b_i, a_i + b_i \rangle \\
&= \langle a_i \rangle + \langle b_i \rangle + \langle b_i, a_i \rangle + \langle a_i, b_i \rangle \\
&= \langle b_i \rangle + \langle a_i \rangle = 0 \\
&\Longrightarrow a_i + b_i = 0 \\
&\Longrightarrow a_i = -b_i \\
&\Longrightarrow a_i \not\perp b_i, \text{ a contradiction.}
\end{aligned}$$

Hence $\lambda \neq -1$. We can therefore have $(1+\lambda)^{-1}$. Other than that, we have

$$\begin{aligned}
\langle a_i + b_i \rangle &= \langle a_i + b_i, a_i + b_i \rangle \\
&= \langle b_i \rangle + \langle a_i \rangle + \langle b_i, a_i \rangle + \langle a_i, b_i \rangle \\
&= \langle b_i \rangle + \langle a_i \rangle \\
&= (1+\lambda) \langle b_i \rangle
\end{aligned}$$

In order to construct the vector $y = \sum_{i\in I} \zeta_i b_i$, we first consider some required equalities. For all $i \in I$, $\lambda \langle b_i \rangle = \langle a_i \rangle$

$\Longrightarrow \overline{\lambda \langle b_i, b_i \rangle} = \overline{\langle a_i, a_i \rangle}$

$\Longrightarrow \overline{\langle b_i, b_i \rangle \lambda} = \overline{\langle a_i, a_i \rangle}$

$\Longrightarrow \langle b_i, b_i \rangle \overline{\lambda} = \langle a_i, a_i \rangle$

$\Longrightarrow \langle b_i \rangle \overline{\lambda} = \langle a_i \rangle$

Thus, $\lambda \langle b_i \rangle = \langle b_i \rangle \overline{\lambda}$

From this, we have $\overline{\lambda}^{-1} \langle b_i \rangle^{-1} = \langle b_i \rangle^{-1} \lambda^{-1}$

Thus, by adding $\langle b_i \rangle$ on both sides of the above two, we can have $\langle b_i \rangle \left(1 + \overline{\lambda}\right) = (1+\lambda) \langle b_i \rangle$ and $\left(1 + \overline{\lambda}\right)^{-1} \langle b_i \rangle^{-1} = \langle b_i \rangle^{-1} (1+\lambda)^{-1}$.

Note that $\langle a_i + b_i, a_i - \lambda b_i \rangle$

$= \langle a_i, a_i - \lambda b_i\rangle + \langle b_i, a_i - \lambda b_i\rangle$

$= \langle a_i, a_i\rangle - \langle a_i, b_i\rangle \overline{\lambda} + \langle b_i, a_i\rangle - \langle b_i, b_i\rangle \overline{\lambda}$

Since $\langle b_i\rangle \overline{\lambda} = \langle a_i\rangle$, therefore $\langle a_i + b_i, a_i - \lambda b_i\rangle = 0$

Using these, we can define the orthogonal system

$$H := (a_i + b_i, a_i - \lambda b_i)_{i\in I}$$

Next, a_i can be written as $(1+\lambda)^{-1}(1+\lambda)a_i$

$= (1+\lambda)^{-1}\left[(1+\lambda)a_i + \lambda b_i - \lambda b_i\right]$

$= (1+\lambda)^{-1}\left[a + \lambda a_i + \lambda b_i - \lambda b_i\right]$

$= (1+\lambda)^{-1}\left[\lambda(a_i + b_i) + (a_i - \lambda b_i)\right]$ whence $((a_i)_{i\in I})^{\perp\perp} \subset H^{\perp\perp}$.

Now, we let $x = x_1 + x_2 \in X$ with $x_1 \in ((a_i + b_i)_{i\in I})^{\perp\perp}$ and $x_2 \in ((a_i - \lambda b_i)_{i\in I})^{\perp\perp}$

Since $\langle x_1, a_i + b_i\rangle \langle a_i + b_i\rangle^{-1}$

$= \langle x - x_2, a_i\rangle \langle a_i + b_i\rangle^{-1} = \langle x, a_i + b_i\rangle \langle a_i + b_i\rangle^{-1} - \langle x_2, a_i + b_i\rangle \langle a_i + b_i\rangle^{-1}$

$= \langle x, a_i\rangle \langle a_i + b_i\rangle^{-1} + \langle x, b_i\rangle \langle a_i + b_i\rangle^{-1}$

$= \langle x, a_i\rangle \langle a_i + b_i\rangle^{-1}$

$= \zeta_i \langle a_i\rangle \langle a_i + b_i\rangle^{-1}$

$= \zeta_i \lambda \langle b_i\rangle \left((\lambda+1)\langle b_i\rangle\right)^{-1}$

$= \zeta_i \lambda \langle b_i\rangle \langle b_i\rangle^{-1} (\lambda+1)^{-1}$

$= \zeta_i \lambda (\lambda+1)^{-1}$

We therefore have

$$x_1 = \sum_{i\in I} \zeta_i \lambda (\lambda+1)^{-1}(a_i + b_i)$$

Further we decompose $x_1 = a + b$, where $a \in ((f_i)_{i\in I})^{\perp\perp}$ and $b \in ((g_i)_{i\in I})^{\perp\perp}$.

We can write

$$\begin{aligned} b_i &= (1+\lambda)^{-1}(1+\lambda)b_i \\ &= (1+\lambda)^{-1}(b_i + \lambda b_i) \\ &= (1+\lambda)^{-1}(a_i - a_i + b_i + \lambda b_i) \\ &= (1+\lambda)^{-1}((a_i + b_i) - (a_i - \lambda b_i)) \end{aligned}$$

Thus, $\langle b, b_i\rangle \langle b_i\rangle^{-1} = \langle x_1 - a, b_i\rangle \langle b_i\rangle^{-1}$

$= \langle x_1, b_i\rangle \langle b_i\rangle^{-1} - \langle a, b_i\rangle \langle b_i\rangle^{-1}$

$= \langle x_1, b_i\rangle \langle b_i\rangle^{-1} = \left\langle x_1, (1+\lambda)^{-1}\left((a_i+b_i) - (a_i - \lambda b_i)\right)\right\rangle \langle b_i\rangle^{-1}$

$= \left\langle x_1, (1+\lambda)^{-1}(a_i+b_i)\right\rangle \langle b_i\rangle^{-1} - \left\langle x_1, (1+\lambda)^{-1}(a_i-\lambda b_i)\right\rangle \langle b_i\rangle^{-1}$

$= \left\langle x_1, (1+\lambda)^{-1}(a_i+b_i)\right\rangle \langle b_i\rangle^{-1}$

Since $\langle x_1, a_i+b_i\rangle \langle a_i+b_i\rangle^{-1} = \zeta_i\lambda(\lambda+1)^{-1}$, $\left(1+\overline{\lambda}\right)^{-1}\langle b_i\rangle^{-1} = \langle b_i\rangle^{-1}(1+\lambda)^{-1}$ and $\langle b_i + b_i\rangle = (1+\lambda)\langle b_i\rangle$, we therefore have

$\zeta_i\lambda(\lambda+1)^{-1}\langle a_i+b_i\rangle(1+\overline{\lambda})^{-1}\langle b_i\rangle^{-1} = \zeta_i\lambda(\lambda+1)^{-1}(1+\lambda)\langle b_i\rangle\langle b_i\rangle^{-1}(1+\lambda)^{-1}$

$= \zeta_i\lambda(\lambda+1)^{-1}(1+\lambda)(1+\lambda)^{-1}$

$= \zeta_i\lambda(\lambda+1)^{-1}$

We can therefore have

$$b = \sum_{i\in I}\zeta_i\lambda(\lambda+1)^{-1}b_i$$

Now we construct the vector y by defining $y := (\lambda+1)\lambda^{-1}b = \sum_{i\in I}\zeta_i b_i$. From here, we have that $\langle y\rangle = (\lambda+1)\lambda^{-1}\langle b\rangle = (1+\lambda^{-1})\langle b\rangle$. We now calculate the length of y by focusing on $\langle b\rangle$ and therefore on $\langle a\rangle$. We have

$$\begin{aligned}
\langle a, a_i\rangle\langle a_i\rangle^{-1} &= \langle x_1, a_i\rangle\langle a_i\rangle^{-1} - \langle b, a_i\rangle\langle a_i\rangle^{-1}\\
&= \langle x_1, a_i\rangle\langle a_i\rangle^{-1}\\
&= \left\langle x_1, (1+\lambda)^{-1}\left[\lambda(a_i+b_i) + (a_i-\lambda b_i)\right]\right\rangle\langle a_i\rangle^{-1}\\
&= \left\langle x_1, (1+\lambda)^{-1}\lambda(a_i+b_i)\right\rangle\langle a_i\rangle^{-1} + \left\langle x_1, (1+\lambda)^{-1}(a_i-\lambda b_i)\right\rangle\langle a_i\rangle^{-1}\\
&= \left\langle x_1, (1+\lambda)^{-1}\lambda(a_i+b_i)\right\rangle\langle a_i\rangle^{-1}\\
&= \langle x_1, (a_i+b_i)\rangle(1+\overline{\lambda})^{-1}\overline{\lambda}\langle a_i\rangle^{-1}\\
&= \zeta_i\lambda(\lambda+1)^{-1}\langle a_i+b_i\rangle(1+\overline{\lambda})^{-1}\overline{\lambda}\langle a_i\rangle^{-1}\\
&= \zeta_i\lambda(\lambda+1)^{-1}\langle b_i+b_i\rangle(1+\overline{\lambda})^{-1}\langle b_i\rangle^{-1}\\
&= \zeta_i\lambda(\lambda+1)^{-1}(1+\lambda)\langle b_i\rangle(1+\overline{\lambda})^{-1}\langle b_i\rangle^{-1}\\
&= \zeta_i\lambda(\lambda+1)^{-1}(1+\lambda)\langle b_i\rangle\langle b_i\rangle^{-1}(1+\lambda)^{-1}\\
&= \zeta_i\lambda(\lambda+1)^{-1}(1+\lambda)(1+\lambda)^{-1}\\
&= \zeta_i\lambda(\lambda+1)^{-1}
\end{aligned}$$

Hence we can have

$$\begin{aligned} a &= \sum_{i\in I} \zeta_i \lambda(\lambda+1)^{-1} a_i \\ &= \lambda(\lambda+1)^{-1} \sum_{i\in I} \zeta_i a_i \\ &= \lambda(\lambda+1)^{-1} x \end{aligned}$$

For the length of x we get

$$\begin{aligned} \langle x\rangle &= \langle x, x\rangle \\ &= \langle x_1 + x_2, x_1 + x_2\rangle \\ &= \langle x_1, x_1 + x_2\rangle + \langle x_2, x_1 + x_2\rangle \\ &= \langle x_1, x_1\rangle + \langle x_1, x_2\rangle + \langle x_2, x_1\rangle + \langle x_2, x_2\rangle \\ &= \langle x_1\rangle + \langle x_2\rangle \\ &= \langle a+b, a+b\rangle + \langle x_2\rangle \\ &= \langle a, a\rangle + \langle a, b\rangle + \langle b, a\rangle + \langle b, b\rangle + \langle x_2\rangle \\ &= \langle a\rangle + \langle b\rangle + \langle x_2\rangle \end{aligned}$$

Since $x_2 = x - x_1 = (\lambda+1)\lambda^{-1}a - a - b = \lambda^{-1}a - b$, we can then compute $\langle x_2\rangle$ as follows:

$$\begin{aligned} \langle x_2\rangle &= \left\langle \lambda^{-1}a - b, \lambda^{-1}a - b\right\rangle \\ &= \left\langle \lambda^{-1}f, \lambda^{-1}a - b\right\rangle - \left\langle b, \lambda^{-1}a - b\right\rangle \\ &= \lambda^{-1}\langle a, a-b\rangle \overline{\lambda}^{-1} - \langle b, a-b\rangle \\ &= \lambda^{-1}\langle a, a\rangle \overline{\lambda}^{-1} - \lambda^{-1}\langle a, b\rangle \overline{\lambda}^{-1} - \langle a, b\rangle \overline{\lambda}^{-1} + \langle b, b\rangle \\ &= \lambda^{-1}\langle a\rangle \overline{\lambda}^{-1} + \langle b\rangle \end{aligned}$$

Thus, the $\langle x\rangle$ becomes $\langle a\rangle + 2\langle b\rangle + \lambda^{-1}\langle a\rangle \overline{\lambda}^{-1}$. This can be re-written as

$$\lambda\langle x\rangle \overline{\lambda} = \lambda\langle a\rangle \overline{\lambda} + 2\lambda\langle b\rangle \overline{\lambda} + \langle a\rangle$$

We now use $\langle a\rangle = \lambda \langle b\rangle$, its derivands and $\langle y\rangle = (1+\lambda^{-1}) \langle b\rangle$ to get

$$(1+\lambda) \langle x\rangle (1+\overline{\lambda}) = \lambda \langle x\rangle \overline{\lambda} + 2\lambda \langle y\rangle \overline{\lambda} + \langle x\rangle$$

Solving for $\langle y\rangle$, we get

$$\langle y\rangle = \frac{1}{2}\left(\langle x\rangle \overline{\lambda}^{-1} + \lambda^{-1} \langle x\rangle\right) \tag{6.4}$$

Now choose $\lambda = 1$ to get $\langle y\rangle = \langle x\rangle$

Thus, we are assured that the vector representation does not depend on the choice of the orthogonal system. Using this, we will show that any infinite dimensional orthomodular space can be written as a direct sum of orthogonal copies of it, even if it is not normal. We define $F_1 := ((f_{2i})_{i\in\mathbb{N}})^{\perp\perp}$ and $F_2 := ((f_{2i+1})_{i\in\mathbb{N}})^{\perp\perp}$. We have $F = F_1 \overset{\perp}{\oplus} F_2$ and according to what we have just proved, we must have $F_1 \cong F_2$. This is because if we define a function $\varphi : F_1 \longrightarrow F_2$ such that

$$x = \sum_{i\in\mathbb{N}} \alpha_i f_{2i} \longmapsto \sum_{i\in\mathbb{N}} \alpha_i f_{2i+1}$$

then, φ is isometric: let $a, b \in F_1$ and $a' = \varphi(a)$. Then,

$$\langle a\rangle = \sum_{i\in\mathbb{N}} \alpha_i \langle f_{2i}\rangle = \sum_{i\in\mathbb{N}} \alpha_i \langle f_{2i+1}\rangle = \langle a'\rangle$$

Furthermore, if we have $b' = \varphi(b) \in E_2$, then

$$\begin{aligned}
\langle a+b\rangle &= \sum_{i\in\mathbb{N}} (\alpha_i + \beta_i) \langle f_{2i}\rangle \\
&= \sum_{i\in\mathbb{N}} (\alpha_i + \beta_i) \langle f_{2i+1}\rangle \\
&= \langle a' + b'\rangle
\end{aligned}$$

We therefore have

$$\begin{aligned}
\langle a+b\rangle-\langle a\rangle-\langle b\rangle &= \langle a+b,a+b\rangle-\langle a\rangle-\langle b\rangle\\
&= \langle a,a+b\rangle+\langle b,a+b\rangle-\langle a\rangle-\langle b\rangle\\
&= \langle a,a\rangle+\langle a,b\rangle+\langle b,a\rangle+\langle b,b\rangle-\langle a\rangle-\langle b\rangle\\
&= \langle a,b\rangle+\langle b,a\rangle
\end{aligned}$$

That is,

$$\langle a+b\rangle-\langle a\rangle-\langle b\rangle=\left\langle a',b'\right\rangle+\left\langle b',a'\right\rangle.$$

Likewise we get for any $\lambda\in\mathbb{K}$,

$$\langle \lambda a,b\rangle+\langle b,\lambda a\rangle=\left\langle \lambda a',b'\right\rangle+\left\langle b',\lambda a'\right\rangle$$

$\Longrightarrow$

$$\langle \lambda a,b\rangle-\left\langle \lambda a',b'\right\rangle=\left\langle b',\lambda a'\right\rangle-\langle b,\lambda a\rangle$$

$\Longrightarrow$

$$\begin{aligned}
\lambda\left(\langle a,b\rangle-\left\langle a',b'\right\rangle\right) &= \left(\left\langle b',a'\right\rangle-\langle b,a\rangle\right)\overline{\lambda}\\
&= -(\langle b,a\rangle-\left\langle b',a'\right\rangle)\overline{\lambda}\\
&= -(\overline{\langle a,b\rangle-\langle a',b'\rangle})\overline{\lambda}
\end{aligned}$$

If $\langle a,b\rangle\neq\langle a',b'\rangle$, then for $\lambda=(\langle a,b\rangle-\langle a',b'\rangle)^{-1}$ we would get $1=-1$.

F_1 and F_2 are further decomposed as follows: $F_1=F_{11}\overset{\perp}{\oplus}F_{12}$ and $F_2=F_{21}\overset{\perp}{\oplus}F_{22}$, where $F_{11}:=((e_{4i})_{i\in\mathbb{N}})^{\perp\perp}, F_{12}:=((f_{4i+2})_{i\in\mathbb{N}})^{\perp\perp}, F_{21}:=((f_{4i+1})_{i\in\mathbb{N}})^{\perp\perp}$ and $F_{22}:=((f_{4i+3})_{i\in\mathbb{N}})^{\perp\perp}$. Again, we have $F_2\cong F_{11}$ and thus $F_1\cong F_{11}$.Further $F_2\cong F_{12}$. Altogether we have $F=F_1\overset{\perp}{\oplus}F_2\cong F_{11}\overset{\perp}{\oplus}F_{12}=F_1$, thus $F\cong F\overset{\perp}{\oplus}F$.

Finally, for this part, we now show that there is no vector $z=\sum_{n\in\mathbb{N}}x_n$ in F for $\langle x_i\rangle=1$.

Assume for the sake of contradiction that there is a vector $z = \sum_{n\in\mathbb{N}} x_i$ in F. Then we have

$$\langle z \rangle = \left\langle \sum_{i\in\mathbb{N}} x_i \right\rangle = \left\langle \sum_{i\in\mathbb{N}} x_{2i} \right\rangle + \left\langle \sum_{i\in\mathbb{N}} x_{2i+1} \right\rangle$$

These three lengths are equal since $X \cong X \overset{\perp}{\oplus} X$. Since they are not zero, we have arrived at a contradiction. Hence the orthonormal system $(f_i)_{i\in\mathbb{N}^*}$ is maximal in F so that we have a vector $f = \sum_{i\in\mathbb{N}} \alpha_i f_i$ in X such that $\langle f \rangle = \alpha$.

Part B

Thus, for every sequence $(\alpha_i)_{i\in\mathbb{N}} \in \mathbb{Q}^{\mathbb{N}^*}$ with $\alpha := \sum_{i=1}^{\infty} \alpha_i^2 \in \mathbb{R}$ we find a vector $x = \sum_{i\in\mathbb{N}^*} \alpha_i e_i \in X$. It can be shown that the choice of the selection of the sequence is immaterial.

To show that $\mathbb{R} \subset \mathbb{K}$, we first define

$$\psi : \mathbb{R}^+ \longrightarrow \mathbb{K}$$

such that $\alpha = \sum_{i=0}^{\infty} \alpha_i^2 \longmapsto \left\langle \sum_{i\in\mathbb{N}} \alpha_i e_i \right\rangle$ where $(\alpha_i)_{i\in\mathbb{N}}$ is any sequence in $\mathbb{Q}^{\mathbb{N}}$ with $\alpha = \sum_{i=0}^{\infty} \alpha_i^2 \in \mathbb{R}$. ψ is well defined: we know that the field element $\left\langle \sum_{i\in\mathbb{N}} \alpha_i e_i \right\rangle$ does not depend on the choice of the orthonormal system $(e_i)_{i\in\mathbb{N}}$. Let $(\alpha_i')_{i\in\mathbb{N}} \in \mathbb{Q}^{\mathbb{N}}$ have $\sum_{i=0}^{\infty} \alpha_i'^2 = \alpha$. Pick $\rho > \alpha$ in $\mathbb{Q}$ and $(\beta_i)_{i\in\mathbb{N}} \in \mathbb{Q}^{\mathbb{N}}$ with $\sum_{i=0}^{\infty} \beta_i^2 = \rho - \alpha$.

If $x := \sum_{i\in\mathbb{N}} \alpha_i e_{2i} + \sum_{i\in\mathbb{N}} \beta_i e_{2i+1}$, and $x' := \sum_{i\in\mathbb{N}} \alpha_i' e_{2i} + \sum_{i\in\mathbb{N}} \beta_i e_{2i+1}$

then $\langle x \rangle = \rho = \langle x' \rangle$ by **Part A**, we obtain the asserted independence of $\left\langle \sum_{i\in\mathbb{N}} \alpha_i e_i \right\rangle$ from the choice of the α_i.

We turn now to the algebraic properties of ψ. Let $\alpha,\, \beta \in \mathbb{R}^+$, $\alpha,\, \beta \neq 0$. We have

1. $\psi(\alpha + \beta) = \psi(\alpha) + \psi(\beta)$

   This is routine to verify, considering that for $x = \sum_{i\in\mathbb{N}} \alpha_i e_i$ and $y = \sum_{i\in\mathbb{N}} \beta_i e_i$, we have $x + y = \sum_{i\in\mathbb{N}} (\alpha_i + \beta_i)\, e_i$

2. $\psi(1) = 1$

   This is also fairly easy to see: if $1 = \sum_{i=0}^{\infty} \alpha_i^2$, then $\left\langle \sum_{i\in\mathbb{N}} \alpha_i e_i \right\rangle = \sum_{i\in\mathbb{N}} \alpha_i^2 \langle e_i \rangle = 1 = \psi(1)$

3. $2\psi(\alpha\beta) = \psi(\alpha)\psi(\beta) + \psi(\beta)\psi(\alpha)$

In order to show multiplicativity, let $(\alpha_i)_{i\in\mathbb{N}}, (\beta_i)_{i\in\mathbb{N}}$ be sequences with convergent series such that square sums α and β are in $\mathbb{R}^+$, respectively. For $y = \sum_{i\in\mathbb{N}} \beta_i e_i$ and $x = \sum_{i\in\mathbb{N}} \alpha_i e_i \in X$. Fix some orthonormal system $(e_i^{(j)})_{i,j\in\mathbb{N}}$. We have already seen that, for each $i \in \mathbb{N}$, there exists the vector y_i such that $\langle y_i\rangle = \langle y\rangle$ with $y_i = \sum_{j\in\mathbb{N}} \beta_j e_i^{(j)}$. In this case, $\langle e_i\rangle = \langle y\rangle^{-1} \langle y_i\rangle$. Note that $\alpha_i \langle y\rangle^{-1} = \langle y\rangle^{-1} \alpha_i$, for all $i \in \mathbb{N}$ so that we can apply $\lambda = \langle x\rangle^{-1} = \overline{\lambda}$ in **Eq** (6.4) to get a vector $\sum_{i\in\mathbb{N}} \alpha_i y_i = \sum_{i,j\in\mathbb{N}} \alpha_i \beta_j e_i^{(j)}$ in X with length $\frac{1}{2}(\langle x\rangle \langle y\rangle + \langle y\rangle \langle x\rangle)$. Thus, we are allowed to assume the existence of a vector $z = \sum_{i,j\in\mathbb{N}} \alpha_i \beta_j e_i^{(j)}$ with $\sum_{i,j=0}^{\infty} (\alpha_i \beta_j)^2 = \alpha\beta$ so that $\langle z\rangle = \psi(\alpha\beta)$ and

$$\langle z\rangle = \frac{1}{2}\left[\psi(\alpha)\psi(\beta) + \psi(\beta)\psi(\alpha)\right] = \psi(\alpha)\psi(\beta)$$

For $\alpha = \beta$ we have $2\psi(\alpha^2) = 2\psi(\alpha)^2$ so that we get:

4. $\psi(\alpha^2) = \psi(\alpha)^2$

   and for $\beta = \alpha^{-1}$ we get

5. $2\psi(\alpha\alpha^{-1}) = 2\psi(1) = 2 = \psi(\alpha)\psi(\alpha^{-1}) + \psi(\alpha^{-1})\psi(\alpha)$

   Multiplying on the left and right by $\psi(\alpha)$, we get

$$\psi(\alpha)2 = \psi(\alpha)^2\psi(\alpha^{-1}) + \psi(\alpha)\psi(\alpha^{-1})\psi(\alpha)$$

   and $2\psi(\alpha) = \psi(\alpha)\psi(\alpha^{-1})\psi(\alpha) + \psi(\alpha^{-1})\psi(\alpha)^2$. Since addition is commutative, we must have $\psi(\alpha)2 = 2\psi(\alpha)$. This means that we can equate both sides, then cancel the common $\psi(\alpha)\psi(\alpha^{-1})\psi(\alpha)$ to get

$$\psi(\alpha)^2\psi(\alpha^{-1}) = \psi(\alpha^{-1})\psi(\alpha)^2$$

   We can replace $\psi(\alpha)^2$ with $\psi(\alpha)$.

$$\psi(\alpha)\psi(\alpha^{-1}) = \psi(\alpha^{-1})\psi(\alpha)$$

   By Bullet 4 and Bullet 5 we get $1 = \psi(\alpha)\psi(\alpha^{-1}) = \psi(\alpha^{-1})\psi(\alpha)$. Thus

6. $\psi(\alpha)^{-1} = \psi(\alpha^{-1})$

Now by Bullet 3 and Bullet 4

$$1 = \psi(\alpha\beta)\psi\left(\beta^{-1}\alpha^{-1}\right) = \psi\left(\alpha\beta\right)\psi\left(\alpha^{-1}\beta^{-1}\right)$$

$$= \frac{1}{2}\left(\psi\left(\alpha\right)\psi\left(\beta\right) + \psi\left(\beta\right)\psi\left(\alpha\right)\right)\frac{1}{2}\left(\psi\left(\alpha^{-1}\right)\psi\left(\beta^{-1}\right) + \psi\left(\beta^{-1}\right)\psi\left(\alpha^{-1}\right)\right)$$

$$= \frac{1}{4}\left(\begin{array}{c} \psi\left(\alpha\right)\psi\left(\beta\right)\psi\left(\alpha^{-1}\right)\psi\left(\beta^{-1}\right) + \psi\left(\alpha\right)\psi\left(\beta\right)\psi\left(\beta^{-1}\right)\psi\left(\alpha^{-1}\right) \\ +\psi\left(\beta\right)\psi\left(\alpha\right)\psi\left(\alpha^{-1}\right)\psi\left(\beta^{-1}\right) + \psi\left(\beta\right)\psi\left(\alpha\right)\psi\left(\beta^{-1}\right)\psi\left(\alpha^{-1}\right) \end{array}\right)$$

$$= \frac{1}{4}(\rho + 1 + 1 + \rho^{-1})$$

where $\rho = \psi\left(\alpha\right)\psi(\beta)\psi\left(\alpha^{-1}\right)\psi(\beta^{-1})$. Thus, $4 = \rho + 2 + \rho^{-1}$ from which we have $\rho + \rho^{-1} = 2$. Multiply both sides by ρ to get $\rho^2 - 2\rho + 1 = (\rho - 1)^2 = 0$. This simplification is obtained using the commutativity of addition, not multiplication.

It follows $\rho = 1$ and $\psi\left(\alpha\right)\psi(\beta) = \psi(\beta)\psi\left(\alpha\right)$, thus by Bullet 3, $\psi(\alpha\beta) = \psi\left(\alpha\right)\psi(\beta)$

Now for $\alpha < 0$, we send α to $-\psi(-\alpha)$, and further $0 \longmapsto 0$. Hence ψ is extended to all of $\mathbb{R}$; this extension is an embedding of fields. We identify $\psi(\mathbb{R}) \subset \mathbb{K}$ now as an embedding of ψ to $\mathbb{K}$ from the proof above with $\mathbb{R}$ and we consider now $\mathbb{R}$ as a fixed subfield of $\mathbb{K}$. In particular, this means that $\mathbb{R} \subset S \subset \mathbb{K}$ where $S = \{x \mid x = \overline{x}\}$

In the proof of **Part A**, we have taken $(\alpha_i)_{i\in\mathbb{N}} \in l_2\left(\mathbb{Q}\right)$. By now, we have that $\mathbb{R} \subset \mathbb{K}$. Therefore, the construction in the proof of **Part A** may now be run as $\mathbb{R} \subset \mathbb{K}$ with $\mathbb{R}$ in the role of $\mathbb{Q}$. This will prove that for every sequence $(\alpha_i)_{i\in\mathbb{N}} \in l_2\left(\mathbb{R}\right)$ with $\alpha := \sum_{i=0}^{\infty}\alpha_i^2$, there exists a vector $a = \sum_{i\in\mathbb{N}}\alpha_i e_i \in X$ such that $\langle a \rangle = \alpha$

Part C

We now show that $\mathbb{R} \subset Z = \{x \mid xy = yx, \forall y \in \mathbb{K}\}$. To this end, we show that for $\alpha \in S \subset \mathbb{K}$ with $\alpha^2 \neq \pm 1$, $\alpha \neq 0$, there exists either the vector

$$\sum_{i\in\mathbb{N}}\alpha^i e_i \in X \text{ with the length } (1 - \alpha^2)^{-1}$$

or

$$\sum_{i\in\mathbb{N}}(\alpha^{-1})^i e_i \in X \text{ with the length } (1 - (\alpha^{-1})^2)^{-1}$$

We define for all $n \in \mathbb{N}$

$$x_n := \sum_{i=0}^{n} \alpha^i e_i$$

$$y_n := x_n - (1-\alpha^2)^{-1} e_0$$

$$a_n := x_{2n+1} - \alpha^2 x_{2n}$$

$$b_n = y_{2n+2} - \alpha^2 y_{2n+1}$$

From this, we have

$$y_{n-1} = x_{n-1} - (1-\alpha^2)^{-1} e_0$$

$$\begin{aligned}
\langle x_n, y_n \rangle &= \left\langle \sum_{i=0}^{n} \alpha^i e_i, x_n - (1-\alpha^2)^{-1} e_0 \right\rangle \\
&= \left\langle \sum_{i=0}^{n} \alpha^i e_i, \sum_{i=0}^{n} \alpha^i e_i - (1-\alpha^2)^{-1} e_0 \right\rangle \\
&= \left\langle \sum_{i=0}^{n-1} \alpha^i e_i + \alpha^n e, \sum_{i=0}^{n-1} \alpha^i e_i + \alpha^n e + (1-\alpha^2)^{-1} e_0 \right\rangle \\
&= \langle x_{n-1} + \alpha^n e_n, y_{n-1} + \alpha^n e_n \rangle
\end{aligned}$$

Thus,

$$\begin{aligned}
\langle x_n, y_n \rangle &= \langle x_{n-1} + \alpha^n e_n, y_{n-1} + \alpha^n e_n \rangle \\
&= \langle x_{n-1}, y_{n-1} + \alpha^n e_n \rangle + \langle \alpha^n e_n, y_{n-1} + \alpha^n e_n \rangle \\
&= \langle x_{n-1}, y_{n-1} \rangle + \langle x_{n-1}, \alpha^n e_n \rangle + \langle \alpha^n e_n, y_{n-1} \rangle + \langle \alpha^n e_n, \alpha^n e_n \rangle \\
&= \begin{array}{c} \left\langle \sum\limits_{i=0}^{n-1} \alpha^i e_i, \sum\limits_{i=0}^{n-1} \alpha^i e_i - (1-\alpha^2)^{-1} e_0 \right\rangle + \left\langle \sum\limits_{i=0}^{n-1} \alpha^i e_i, \alpha^n e_n \right\rangle \\ + \left\langle \alpha^n e_n, \sum\limits_{i=0}^{n-1} \alpha^i e_i - (1-\alpha^2)^{-1} e_0 \right\rangle + \langle \alpha^n e_n, \alpha^n e_n \rangle \end{array}
\end{aligned}$$

$$= \begin{array}{c} \left\langle \sum\limits_{i=0}^{n-1} \alpha^i e_i, \sum\limits_{i=0}^{n-1} \alpha^i e_i \right\rangle - \left\langle \sum\limits_{i=0}^{n-1} \alpha^i e_i, (1-\alpha^2)^{-1} e_0 \right\rangle + 0 \\ + \left\langle \alpha^n e_n, \sum\limits_{i=0}^{n-1} \alpha^i e_i \right\rangle - \langle \alpha^n e_n, (1-\alpha^2)^{-1} e_0 \rangle + \alpha^n \langle e_n, e_n \rangle \overline{\alpha^n} \end{array}$$

$$
\begin{aligned}
&= \sum_{i=0}^{n-1}\alpha^i \left\langle e_i, \sum_{i=0}^{n-1}\alpha^i e_i \right\rangle - \left\langle \sum_{i=0}^{n-1}\alpha^i e_i, e_0 \right\rangle (1-\alpha^2)^{-1} + \alpha^n\overline{\alpha^n} \\
&= \sum_{i=0}^{n-1}\alpha^i \langle e_i, e_j\rangle \sum_{j=0}^{n-1}\overline{\alpha^j} - (1-\alpha^2)^{-1} + \alpha^n\alpha^n \\
&= \sum_{i=0}^{n-1}\alpha^{2i} - (1-\alpha^2)^{-1} + \alpha^{2n} \\
&= \sum_{i=0}^{n}\alpha^{2i} - (1-\alpha^2)^{-1}
\end{aligned}
$$

It follows that $\langle x_{n-1}, y_{n-1}\rangle = \sum_{i=0}^{n-1}\alpha^{2i} - (1-\alpha^2)^{-1}$ so that

$$
\begin{aligned}
\langle x_n, y_n\rangle \langle x_{n-1}, y_{n-1}\rangle^{-1} &= \left(\sum_{i=0}^{n}\alpha^{2i} - (1-\alpha^2)^{-1}\right)\left(\sum_{i=0}^{n-1}\alpha^{2i} - (1-\alpha^2)^{-1}\right)^{-1} \\
&= \left(\alpha^{2n} + \sum_{i=0}^{n-1}\alpha^{2i} - (1-\alpha^2)^{-1}\right)\left(\sum_{i=0}^{n-1}\alpha^{2i} - (1-\alpha^2)^{-1}\right)^{-1} \\
&= \alpha^{2n}\left(\sum_{i=0}^{n-1}\alpha^{2i} - (1-\alpha^2)^{-1}\right)^{-1} + 1
\end{aligned}
$$

$$= 1 + \alpha^{2n}\left(1 - (1-\alpha^2)^{-1} + \alpha^2 + ... + \alpha^{2(n-1)}\right)^{-1}$$

$$= 1 + \alpha^{2n}\left(1-\alpha^2\right)^{-1}\left(1 - \alpha^2 - 1 + \left(1-\alpha^2\right)\alpha^2 + ... + \left(1-\alpha^2\right)\alpha^{2(n-1)}\right)^{-1}$$

$$= 1 + \alpha^{2n}.\left(1 - (1-\alpha^2)^{-1}.\left(1 - \alpha^2 - 1 + (1-\alpha^2)\alpha^2 + ... + (1-\alpha^2)\alpha^{2(n-1)}\right)\right)^{-1}$$

$$= 1 + \alpha^{2n}(-\alpha^{2n})(1-\alpha^2) = \alpha^2$$

In the same way we get $\langle y_n, x_n\rangle \langle y_{n-1}, x_{n-1}\rangle^{-1} = \alpha^2$. Using the same routine verification, we get

$$\langle a_n, b_m\rangle = 0$$

for all $n, m \in \mathbb{N}$. We build

$$A := \{x \in X \mid \langle x, b_n\rangle = 0, \forall n \in \mathbb{N}\}$$

This set is clearly non-empty. Also, because of the continuity of the Hermitian product, we have that A is closed. We can therefore have $X = A \oplus A^{\perp}$.

Since $(1-\alpha^2)^{-1}e_0 \in X$, we can decompose it as $a - b$, with $a \in A$ and $b \in A^{\perp}$. Since we can write any vector in X as an infinite sum for basis e_n, we decompose $a \in A \subset X$ as

$$a = \sum_{i \in \mathbb{N}} (\alpha^n + \varepsilon_n) e_n$$

This choice is justified as follows: a uses a_n as a basis which in turn uses $x_n = \sum_{k=0}^{n} \alpha^k e_k$ as a basis. We therefore introduce an "error term" $\varepsilon_n \in \mathbb{K}$. If we let $\langle b \rangle = (1-\alpha^2)^{-1}\varepsilon$ for some $\varepsilon \in \mathbb{K}$, we get from $(1-\alpha^2)^{-1}e_0 = a - b \implies$

$$\begin{aligned}
\left\langle (1-\alpha^2)^{-1} e_0 \right\rangle &= (1-\alpha^2)^{-1} \\
&= \langle a - b \rangle \\
&= \langle a \rangle - \langle b \rangle \\
&= \sum_{i \in \mathbb{N}} (\alpha^n + \varepsilon_n) \langle e_n \rangle - \langle b \rangle
\end{aligned}$$

$\implies$

$$\begin{aligned}
\langle a \rangle &= (1-\alpha^2)^{-1} + \langle b \rangle \\
&= (1-\alpha^2)^{-1} + (1-\alpha^2)^{-1}\varepsilon \\
&= (1-\alpha^2)^{-1}(1+\varepsilon)
\end{aligned}$$

It follows that $(1-\alpha^2)^{-1} \langle e_0, a \rangle = \langle a \rangle = (1-\alpha^2)^{-1}(1+\varepsilon)$. Thus, $\langle e_0, a \rangle = (1+\varepsilon)$.

Since $b = a - (1-\alpha^2)^{-1}e_0$, for every $n \in \mathbb{N}$ we have

$$0 = \langle b, a_n \rangle = \left\langle b, x_{2n+1} - \alpha^2 x_{2n} \right\rangle$$

$$= \left\langle a - (1-\alpha^2)^{-1} e_0, x_{2n+1} - \alpha^2 x_{2n} \right\rangle$$

$$= \left\langle a, x_{2n+1} - \alpha^2 x_{2n} \right\rangle - (1-\alpha^2)^{-1} \left\langle e_0, x_{2n+1} - \alpha^2 x_{2n} \right\rangle$$

After some simplification, we get

$$= -\varepsilon_{2n+2} + 2\alpha^{2n+2} + \varepsilon_{2n+1} + \varepsilon_0$$

thus

$$\varepsilon_{2n+2} = \varepsilon_{2n+1}\alpha + \varepsilon_0\alpha^{-(2n+2)}$$

Likewise, from

$$0 = \left\langle \alpha, y_{2n+2} - \alpha^2 y_{2n+1} \right\rangle$$

we get

$$\varepsilon_{2n+3} = \varepsilon_{2n+2} - \varepsilon_0\alpha^{-(2n+3)}$$

We can therefore arrive at

$$\varepsilon_n = \varepsilon_{n-1}\alpha + (-1)^n\varepsilon_0\alpha^{-n} \text{ for } n \geq 2$$

If $\varepsilon_0 = 0$, then $\varepsilon, \varepsilon_n = 0$ for all $n \in \mathbb{N}$. Thus $a = \sum_{i\in\mathbb{N}} \alpha^n e_n$ and $\langle a \rangle = (1-\alpha^2)^{-1}$

If $\varepsilon_0 \neq 0$, then let $\zeta := \alpha\varepsilon_0\alpha^{-1}\varepsilon_0^{-1}$. Then $\zeta\varepsilon_0\alpha = \alpha\varepsilon_0$ and thus $\zeta\varepsilon_n\alpha = \alpha\varepsilon_n$ for all $n \geq 0$.

It can be shown that $\zeta = 1$ and this leads to a contradiction. Thus, we are guaranteed the existence of both vectors.

Now, let $r \in \mathbb{R}$, $r > 0$ and $0 \neq \lambda \in \mathbb{K}$. By **Part B**, r can be represented as the length of the vector $a = \sum_{i\in\mathbb{N}} \alpha_i e_i$, where $(e_i)_{i\in\mathbb{N}}$ is an orthonormal system and $\alpha_i \in \mathbb{Q}$ for all $i \in \mathbb{N}$. That is, $\langle a \rangle = r$. On the one hand we have $\langle \lambda a \rangle = \lambda r \overline{\lambda}$. On the other, $\langle \lambda a \rangle = \left\langle \sum_{i\in\mathbb{N}} \alpha_i \lambda e_i \right\rangle$. Now, $\langle e_i \rangle = (\lambda\overline{\lambda})^{-1} \langle \lambda e_i \rangle$ and $(\lambda\overline{\lambda})^{-1}\alpha_i = \alpha_i(\lambda\overline{\lambda})^{-1}$ for all $i \in \mathbb{N}$. Thus using **Eq** (6.4) we get,

$$\langle \lambda a \rangle = \frac{1}{2}(r\lambda\overline{\lambda} + \lambda\overline{\lambda}r)$$

Hence

$$\lambda r \overline{\lambda} = \frac{1}{2}(r\lambda\overline{\lambda} + \lambda\overline{\lambda}r) \tag{6.5}$$

We do the same for $1/r$ and $\overline{\lambda}^{-1}$ to get

$$\overline{\lambda}^{-1}\frac{1}{r}\lambda^{-1} = \frac{1}{2}(\frac{1}{r}\overline{\lambda}^{-1}\lambda^{-1} + \overline{\lambda}^{-1}\lambda^{-1}\frac{1}{r}) \tag{6.6}$$

Clearly, $\overline{\lambda}^{-1}\frac{1}{r}\lambda^{-1} = \left(\lambda r\overline{\lambda}\right)^{-1}$. We multiply sides of the equations **Eq** (6.5) and **Eq** (6.6) to get

$$1 = \frac{1}{4}(r\lambda\overline{\lambda}\frac{1}{r}\overline{\lambda}^{-1}\lambda^{-1} + \overline{\lambda}^{-1}\lambda^{-1}\frac{1}{r})$$

As in the proof of **Part B** we get $r\lambda\overline{\lambda} = \lambda\overline{\lambda}r$. That is, $rk = kr$ for $k \in \mathbb{K}$.

Part D

We now show that $\mathbb{R} = S$. Let $\gamma \in S$. We consider the (commutative) field $\mathbb{R}(\gamma) \subset \mathbb{K}$ of the real numbers extended with $\gamma \in \mathbb{K}$. $\mathbb{R}(\gamma)$ is formally real:

If $\sum_{i=0}^{n} \xi_i^2 = 0$ for some $\xi_i \in \mathbb{R}$, $(i = 0, 1, ..., n)$, we have $\xi_i = 0$ for all $i = 0, 1, ..., n$. In this case, $\langle \sum_{i=0}^{n} \xi_i e_i \rangle = 0$ by **Part B.**

Let now

$S \subseteq P := \left\{ \langle x \rangle \mid 0 \neq x = \sum\limits_{i \in \mathbb{N}} \xi_i e_i,\ \xi_i \in \mathbb{R}(\gamma) \forall i \in \mathbb{N} \text{ and } \langle x \rangle \in \mathbb{R}(\gamma) \right\} \subseteq \mathbb{R}(\gamma)$

We have

1. $0 \notin P$
2. $1 \in P$
3. $P + P \subseteq P$
4. $P^2 \subseteq P$
5. $(\mathbb{R}(\gamma))^2 . P \subseteq P$

Bullets 1, 2 are obvious. Bullets 3 and 5 follow from the fact that an orthomodular, infinite dimensional Hermitian space contains copies of itself, the latter requiring an additional Bullet 4. We prove only Bullet 4: let $x = \sum_{i \in \mathbb{N}} \xi_i e_i$, $y = \sum_{i \in \mathbb{N}} \mu_i e_i$ and $\langle x \rangle$, $\langle y \rangle$, ξ_i, $\mu_i \in \mathbb{R}$ for all $i \in \mathbb{N}$. Let further $(e_i^{(j)})_{i,j \in \mathbb{N}}$ be an orthonormal system. In **Part B**, we have already shown that there exists $z := \sum_{i \in \mathbb{N}} \xi_i \mu_i e_i^{(j)} \in X$ with the length $\langle x \rangle \langle y \rangle$, thus $\langle x \rangle \langle y \rangle = \langle z \rangle \in P$.

Using Zorn's Lemma we can extend P to a positive cone and we get a total ordering " $\leq$ "on $\mathbb{R}(\gamma)$. Now we choose a $\delta \in \mathbb{R}(\gamma)$ with $0 < \delta < 1$. If δ is infinitesimal relative to " $\leq$ ", then for each $n \in \mathbb{N}\backslash\{0\}$ we have $(1-(n\delta)^{-2})^{-1} < 0$ and thus by the vectors we have shown to exist in **Part C**, we have $\sum_{i\in\mathbb{N}}(n\delta)^i e_i \in X$ with the length $(1-(n\delta)^{-2})^{-1}$. We can use this to construct an orthogonal system $(h_n)_{n\in\mathbb{N}}$ with $\langle h_n\rangle = (1-n^2\delta^2)$ $\forall n \in \mathbb{N}$, which is orthogonal to an orthonormal system $(e\prime_n)_{n\in\mathbb{N}}$. For each $n \in \mathbb{N}$ we define $f_n := h_n + n\delta e\prime_n$ and take, $g_n \in \text{span}_K(h_n, e\prime_n)$ with $g_n \perp f_n . (f_n)_{n\in\mathbb{N}}$ is an orthonormal system.

By **Part B**, there exists the vector $x = \sum_{n=1}^{\infty}\frac{1}{n}e\prime_n \in X$, which we can decompose to $x = f+g$ with $f \in ((f_n)_{n\in\mathbb{N}})^{\perp\perp}$, $g \in ((g_n)_{n\in\mathbb{N}})^{\perp\perp}$, and we consider f

$$(f, f_k) = \left\langle \sum_{n=1}^{\infty}\frac{1}{n}e\prime_n, f_k \right\rangle = \left\langle \frac{1}{k}e\prime_k, h_k + k\delta e\prime_k \right\rangle = \delta$$

for all $k \geq 1$. Thus $f = \sum_{n=1}^{\infty}\delta\langle f_n\rangle^{-1} = \sum_{n=1}^{\infty}\delta f_n$. This leads us to a contradiction in **Part A** of the non-existence of such vectors.

We have shown δ is not infinitesimal, hence "$\leq$" is Archimedean and we have $\mathbb{R}(\gamma) = \mathbb{R}$.

Part E

Finally, we show that $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ and $X \cong l_2(\mathbb{R}), l_2(\mathbb{C})$ or $l_2(\mathbb{H})$. We know that $\lambda\overline{\lambda}$ and $\lambda+\overline{\lambda}$ are elements of $S = \mathbb{R}$, each $\lambda \in \mathbb{K}$ is zero of a polynomial of degree two over $\mathbb{R}$. If $\mathbb{R} \subsetneqq \mathbb{K}$:

Let $\lambda \in \mathbb{K}\setminus\mathbb{R}$ Then $\mathbb{R}(\lambda)$ is a (commutative) field and $\mathbb{R}(\lambda) \cong \mathbb{C}$, without loss of generality we can say $\mathbb{C} \subset \mathbb{K}$.

If $\mathbb{C} \subsetneqq \mathbb{K}$

Let $\lambda \in \mathbb{K}\setminus\mathbb{C}$. We consider the $\mathbb{C}$-left-vector space $\mathbb{C}+\mathbb{C}\lambda$. Let $\alpha, \beta, \alpha', \beta' \in \mathbb{C}$. Then,

$$\lambda\alpha = (\alpha^{-1}\lambda^{-1})^{-1} = r(\alpha^{-1}\lambda^{-1}) = r(\alpha^{-1}(r'\lambda + s'))$$

$$= r\alpha^{-1}r'\lambda + r\alpha^{-1}s' \in \mathbb{C}\lambda + \mathbb{C}$$

For certain, $r, r\prime, s' \in \mathbb{R}$ note that $[\mathbb{R}(\alpha^{-1}\lambda^{-1}) : \mathbb{R}]$, $[\mathbb{R}(\lambda^{-1}) : R] \leq 2$. Thus

$$(\alpha+\beta\lambda)(\alpha'+\beta'\lambda) \in \mathbb{C}+\mathbb{C}\lambda$$

and

$$1 = (\alpha+\beta\lambda)\left(\overline{\alpha+\beta\lambda}\right)t^{-1} = \left(\overline{\alpha+\beta\lambda}\right)t^{-1}(\alpha+\beta\lambda)$$

with $\overline{\alpha+\beta\lambda}t^{-1} \in \mathbb{C}+\mathbb{C}\lambda$ and $t{=}(\alpha+\beta\lambda)\left(\overline{\alpha+\beta\lambda}\right) \in \mathbb{R}$. $\mathbb{C}+\mathbb{C}\lambda$ is therefore a skew field, thus an associative, finite dimensional real division algebra and by Frobenius' Theorem[2] isomorphic to $\mathbb{H}$, thus, without loss of generality $\mathbb{H} \subset \mathbb{K}$.

If $\mathbb{H} \subsetneqq \mathbb{K}$

Let $\lambda \in \mathbb{K} \setminus \mathbb{H}$. We build $\mathbb{H}+\mathbb{H}\lambda$ and we get as above $\mathbb{H}+\mathbb{H}\lambda \cong \mathbb{H}$, a contradiction. Hence we have $\mathbb{K} = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$.

In the case $\mathbb{K} = \mathbb{R}$ we get $X \cong l_2(\mathbb{R})$ with **Part A**: assume by way of contradiction that there is a vector $x = \sum_{i\in\mathbb{N}} \alpha_i e_i$ in X with $\sum_{i=0}^{\infty} \alpha_i^2 = \infty$, then $\langle x\rangle = \langle \sum_{i=0}^{n} \alpha_i e_i\rangle + \left\langle \sum_{i>n} \alpha_i e_i\right\rangle > \sum_{i=0}^{n} \alpha_i^2$ for all $n \in \mathbb{N}$ thus $\langle x\rangle = \infty$.

In the case $\mathbb{K} = \mathbb{C}$ we have $\bar{\cdot} : \mathbb{C} \longrightarrow \mathbb{C}$. That is, $r+is \longmapsto r-is$, $(i = \sqrt{-1})$. Let $(\alpha_n)_{n\in\mathbb{N}} \in l_2(\mathbb{C})$ with $\alpha_n = r_n + is_n$ $(n \in \mathbb{N})$ and $\sum_{n=0}^{\infty} \alpha_n\overline{\alpha_n} = \sum_{n=0}^{\infty}(r_n^2+s_n^2)$. We build

$$x := \sum r_n e_n + i\sum s_n e_n = \sum \alpha_n e_n$$

x has the length

$$\left\langle \sum_{n\in\mathbb{N}} r_n e_n \right\rangle + \left\langle \sum_{n\in\mathbb{N}} r_n e_n, \sum_{n\in\mathbb{N}} s_n e_n \right\rangle \bar{i} + i\left\langle \sum_{n\in\mathbb{N}} s_n e_n, \sum_{n\in\mathbb{N}} r_n e_n \right\rangle + i\left\langle \sum_{n\in\mathbb{N}} s_n e_n \right\rangle \bar{i}$$

By the Cauchy-Schwarz inequality, which we can apply because we've established parts of the underlying field which gives us positiviy of the 2-form, we have

$$\left|\left\langle \sum_{i>\mathbb{N}} r_i e_i, \sum_{i>\mathbb{N}} s_i e_i \right\rangle\right|^2 \leq \left\langle \sum_{i>\mathbb{N}} r_i e_i \right\rangle \left\langle \sum_{i>\mathbb{N}} s_i e_i \right\rangle$$

and

$$\left\langle \sum_{n\in\mathbb{N}} r_n e_n, \sum_{n\in\mathbb{N}} s_n e_n \right\rangle = \left\langle \sum_{i=0}^{n} r_i e_i, \sum_{i=0}^{n} s_i e_i \right\rangle + \left\langle \sum_{i>n} r_i e_i, \sum_{i>n} s_i e_i \right\rangle$$

for all $n \in \mathbb{N}$. Hence

$$\lim_{n\to\infty} \left\langle \sum_{i>n} r_i e_i, \sum_{i>n} s_i e_i \right\rangle < \infty$$

[2] Every finite-dimensional associative division algebra over the real numbers is isomorphic to either $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$

and

$$\left\langle \sum_{n\in\mathbb{N}} r_n e_n, \sum_{n\in\mathbb{N}} s_n e_n \right\rangle = \lim_{n\to\infty} \left\langle \sum_{i=0}^{n} r_i e_i, \sum_{i=0}^{n} s_i e_i \right\rangle = \sum_{i=0}^{\infty} r_i s_i \in \mathbb{R}$$

Thus

$$\langle x \rangle = \sum_{n=0}^{\infty} r_n^2 + \sum_{n=0}^{\infty} s_n^2 = \sum_{n=0}^{\infty} \alpha_n \overline{\alpha_n}$$

Hence as above we have $X \cong l_2(\mathbb{C})$.

For the case $\mathbb{K} = \mathbb{H}$, we consider analogous to the previous one. With the basis consisting of

$$I_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$I_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

$$I_2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

$$I_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

which are the well-known Pauli matrices.

$\mathbb{H}$ is an $\mathbb{R}$-vector space

We have $\bar{.} : \mathbb{H} \longrightarrow \mathbb{H}$ such that

$$\sum_{i=0}^{3} r_i I_i \longmapsto r_0 I_0 - \sum_{i=0}^{3} r_i I_i$$

with $r_i \in \mathbb{R}$. Let $(\alpha_n)_{n\in\mathbb{N}} \in l_2(\mathbb{H})$ with $\alpha_n = \sum_{i=0}^{3} r_i^{(n)} I_i$. We build

$$x = \sum_{i=0}^{3} I_i . \sum_{n\in\mathbb{N}} r_i^{(n)} e_n = \sum_{n\in\mathbb{N}} \alpha_n e_n .$$

Again we have $\langle x \rangle = \sum_{n=0}^{\infty} \alpha_n \overline{\alpha_n}$, and we get $X \cong l_2(\mathbb{H})$. ■

We can now apply this result to half-normal spaces

Theorem 47 *Let $(X, \mathbb{K}, \varphi)$ be a half-normal space over the skew field $\mathbb{K}$. Then $\mathbb{K} = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ and X is a classical Hilbert space.*

Proof. Let $(e_n)_{n\in\mathbb{N}} \subset X$ be an orthonormal system. Let $C := ((e_n)_{n\in\mathbb{N}})^{\perp\perp}$.

Then $(C, \mathbb{K}, \varphi_{C\times C})$ is a normal space. Thus $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ and $C \cong l_2(\mathbb{R}), l_2(\mathbb{C})$ or $l_2(\mathbb{H})$.

Let $(x_i)_{i\in\mathbb{N}}$ be a Cauchy sequence in X. We consider $F := ((e_n)_{n\in\mathbb{N}})^{\perp\perp}$. If $\dim F < \infty$ the sequence converges in F. If $\dim F = \infty$. Then again we have a maximum countable orthogonal basis for F. Thus $(F, \langle.,.\rangle_{F\times F}) \cong l_2(\mathbb{R}), l_2(\mathbb{C})$ or $l_2(\mathbb{H})$ and the sequence converges in F. ■

6.3 Concluding Remarks

The fundamental reason we started off with orthomodular lattices was because of the Spectral Decomposition Theorem for Hilbert Spaces. Since the span of each eigenbasis generates a closed subspace, therefore, such an approach corresponds to a set of yes-no experimental propositions of Quantum Mechanics. Of course this stems from the standard formulation of Quantum Mechanics.

From **Corollary 35**, we see that even the existence of one non-zero isotropic vector gives spaces which are not splitting. Thus, our focus should stem on the fact that the geometry of Quantum Mechanics and Relativity cannot be unified even if we "loosen up" some basic notions.

These notions are important because the state of a quantum entity is represented by an eigenbasis which are orthogonal. From what we see so far, no eigenbasis (eigenstates) will be orthogonal to itself if it is the same. This might offer computational and physical clarity but this basic assumption is what makes quantum mechanics fundamentally different.

In conclusion, the choice of the underlying field, if Archimedean, is rather immaterial from a geometric point of view and that one cannot hope to formulate a Quantum Mechanics over a Hermitian space, even with all its generality, given the condition of orthomodularity, which has physical motivations. In particular, the fundamental geometry of General Relativity is different than the fundamental geometry of Quantum Mechanics. Thus, as of yet, mathematics tells us that the manner in which Quantum Mechanics and Relativity are currently formulated are essentially incompatible and that a unification will require a major revision and not a tweaking of either. The hope to formulate Quantum Mechanics on non-Archimedean fields must justify not taking orthomodularity into account on physical and logical grounds. Some pieces of mathematics do point us in the same direction: if $\mathbb{K}$ is the field p-adic numbers and $\dim X \geq 5$, then we are guaranteed the existence of isotropic vectors.

Chapter 7

Conclusion

In summary, we see that Quantum Mechanics is fundamentally different and that the hope to achieve a unified theory of physics will require a completely revamped system based on the grounds of geometry (solely on the non-existence of isotropic vectors), which implies the current logic of Quantum Mechanics. In particular, the basic postulate of quantum mechanics of the formulation of states (complete in either sense of Einstein or Bohr) are a drastic departure from unification. This was achieved by justifying the use of sharp eigenvalues.

Chapter 8

Appendix

8.1 Uncertainty Principle

Definition 48 *Let $\hat{A}$,$\hat{B}$ be two self-adjoint operators on H. The* ***commutator*** $\left[\hat{A},\hat{B}\right]$ *of* $\hat{A}$,$\hat{B}$ *is defined as* $\hat{A}\hat{B}-\hat{A}\hat{B}$ *and the anti-commutator* $\left\{\hat{A},\hat{B}\right\}$ *is defined as* $\hat{A}\hat{B}+\hat{A}\hat{B}$

An interesting consequence of these properties are the uncertainty relations, from which stems Hiesenberg's uncertainty relation

Definition 49 ***Deviation*** $\triangle\hat{A}$ *of an operator* $\hat{A}$ *is defined as*

$$\triangle\hat{A}=\hat{A}-\langle\psi,\hat{A}\psi\rangle\hat{I}$$

where $\left\langle\hat{A}\right\rangle=\langle\psi,\hat{A}\psi\rangle$ *denotes the expectation value of* $\hat{A}$ *where* $\psi\in H$ *is a state (that is,* $\|\psi\|=1$*). The*

Lemma 50 *If* $\hat{A}$ *is self-adjoint, then so is* $\triangle\hat{A}$

Proof. $\hat{A}^{*}=\hat{A}$ implies $\triangle\hat{A}^{*}=\hat{A}^{*}-\overline{\langle\psi,\hat{A}\psi\rangle}\hat{I}^{*}=\hat{A}-\langle\hat{A}\psi,\psi\rangle\hat{I}=\hat{A}-\langle\psi,\hat{A}\psi\rangle\hat{I}$ ■

Lemma 51 $\left\langle\left(\triangle\hat{A}\right)^{2}\right\rangle=\langle\hat{A}^{2}\rangle-\langle\hat{A}\rangle^{2}$

Proof. From the definition of $\triangle\hat{A}$, it follows that

$$\begin{aligned}\left(\triangle\hat{A}\right)^2 &= \left(\hat{A}-\left\langle\hat{A}\right\rangle\hat{I}\right)^2\\ &= \left(\hat{A}-\left\langle\hat{A}\right\rangle\hat{I}\right)\left(\hat{A}-\left\langle\hat{A}\right\rangle\hat{I}\right)\\ &= \hat{A}^2-\left\langle\hat{A}\right\rangle\hat{A}-\hat{A}\left\langle\hat{A}\right\rangle+\left\langle\hat{A}\right\rangle^2\hat{I}\\ &= \hat{A}^2-2\left\langle\hat{A}\right\rangle\hat{A}+\left\langle\hat{A}\right\rangle^2\hat{I}\end{aligned}$$

Now, $\left\langle\left(\triangle\hat{A}\right)^2\right\rangle = \langle\psi,\left(\triangle\hat{A}\right)^2\psi\rangle = \langle\psi,\left(\hat{A}^2-2\hat{A}\left\langle\hat{A}\right\rangle+\left\langle\hat{A}\right\rangle^2\right)\psi\rangle$

$= \langle\psi,\hat{A}^2\psi\rangle - 2\left\langle\hat{A}\right\rangle\langle\psi,\hat{A}\psi\rangle+\left\langle\hat{A}\right\rangle^2\langle\psi,\psi\rangle$

$= \langle\psi,\hat{A}^2\psi\rangle - 2\left\langle\hat{A}\right\rangle^2+\left\langle\hat{A}\right\rangle^2$

$= \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2$ ■

Using this, we can define the **uncertainty** ΔA of $\hat{A}$: $\Delta A = \sqrt{\left\langle\left(\triangle\hat{A}\right)^2\right\rangle} = \sqrt{\langle\hat{A}^2\rangle-\langle\hat{A}\rangle^2}$.
And now, for the celebrated Uncertainty Relation

Theorem 52 *Let $\hat{A},\hat{B}$ be any two Hermitian operators. Then,* $\Delta A\triangle B \geq \frac{1}{2}\left|\left\langle[\hat{A},\hat{B}]\right\rangle\right|$

Proof. If $\left(\triangle\hat{A}\right)\psi = \left(\hat{A}-\left\langle\hat{A}\right\rangle\hat{I}\right)\psi=\chi$ and $\left(\triangle\hat{B}\right)\psi = \left(\hat{B}-\left\langle\hat{B}\right\rangle\hat{I}\right)\psi=\phi$, then
$\langle\left(\triangle\hat{A}\right)^2\rangle\langle\left(\triangle\hat{B}\right)^2\rangle = \langle\psi,\left(\triangle\hat{A}\right)^2\psi\rangle\langle\psi,\left(\triangle\hat{B}\right)^2\psi\rangle$
$= \langle\chi,\chi\rangle\langle\phi,\phi\rangle$
$\geq |\langle\chi,\phi\rangle|^2$ by the Cauchy-Schwarz inequality
$= \left|\left\langle\left(\triangle\hat{A}\right)\psi,\left(\triangle\hat{B}\right)\psi\right\rangle\right|^2$
$= \left|\left\langle\psi,\left(\triangle\hat{A}\triangle\hat{B}\right)\psi\right\rangle\right|^2 = \left|\langle\triangle\hat{A}\triangle\hat{B}\rangle\right|^2$
From a direct calculation, it can be inferred $\left|\left\langle\triangle\hat{A}\triangle\hat{B}\right\rangle\right|^2 = \frac{1}{4}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|^2+\frac{1}{4}\left|\left\langle\left\{\hat{A},\hat{B}\right\}\right\rangle\right|^2$
$\Rightarrow \left|\left\langle\triangle\hat{A}\triangle\hat{B}\right\rangle\right|^2 \geq \frac{1}{4}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|^2$
Since $\left\langle\left(\triangle\hat{A}\right)^2\right\rangle\left\langle\left(\triangle\hat{B}\right)^2\right\rangle \geq \left|\left\langle\triangle\hat{A}\triangle\hat{B}\right\rangle\right|^2$ and $\left|\left\langle\triangle\hat{A}\triangle\hat{B}\right\rangle\right|^2 \geq \frac{1}{4}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|^2$
we have $\Delta A\Delta B \geq \frac{1}{4}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|$ ■
In particular, for position P and momentum Q, $[P,Q] = i\frac{\mathcal{H}}{2\pi}$

8.2 GNS Construction

For a coverage of the GNS construction according to our requirements and assumptions, see [17]

8.3 Non-isometric involutions

One of the best known bits of mathematical folklore is that there are infinitely many automorphisms of complex numbers i.e. the complex numbers can be permuted in many ways (besides the familiar conjugation) that preserve addition and multiplication. It might hit as a surprise that these other automorphisms, which we will call "wild" in line with [48], rely on the use of the AC. In particular, in [5], it is claimed without proof that the automorphisms of $\mathbb{C}$ are $2^{2^{\aleph_0}}$. Note that this is the same as the set of all complex-valued mappings, which even includes constant functions! We use essentially the same arguments to show that the same is valid for involutive anti-automorphisms. Herein lies the paradox.

Since the latter claim relies on a non-constructive axiom (AC), the anti-automorphisms which will be constructed are going to be non-constructive, as well, making it difficult to imagine almost all of them (pun intended).

Clearly the identity map which reverses order of multiplication on a subfield of an infinite skew field $\mathbb{K}$, $I_{\mathbb{K}}$ is an involutive anti-automorphism of $\mathbb{K}$, the trivial anti-automorphism of $\mathbb{K}$. All other involutive anti-automorphisms of $\mathbb{K}$ are called non-trivial.

We shall first prove that there are only two automorphisms by using the fact that for any $\mathbb{K}$, if $AS(\mathbb{K}) = \{\alpha : \alpha^* = -\alpha\}$ and $S(\mathbb{K}) = \{\alpha : \alpha^* = -\alpha\}$, then $\mathbb{K} = S(\mathbb{K}) \oplus AS(\mathbb{K})$ so that $\alpha = a + b$ uniquely for $a \in S(\mathbb{K})$ and $b \in AS(\mathbb{K})$ for any $\alpha \in \mathbb{K}$ so that if $AS(\mathbb{K}) \neq \varnothing$, then for $i \in AS(\mathbb{K})$, we have the unique decomposition $\alpha = a_1 + ia_2$

Theorem 53 *Let $\varphi : \mathbb{K} \longrightarrow \mathbb{K}$ be an involutive anti-automorphism. Then φ is either equal to the identity or to conjugation*

Proof. Every automorphism sends 0 and 1 to themselves and from this it follows that every automorphism sends the rational numbers $\mathbb{Q} \subset \mathbb{K}$ to itself. Furthermore, if $a \in \mathbb{Q}$ is non-zero and $\alpha \in \mathbb{K}$ satisfies $\alpha^2 = a$, then we also have $\varphi(\alpha)^2 = \varphi(a) = a$, and since $\pm\alpha$ are the only two numbers such that $\alpha^2 = a$ we must have $\varphi(\alpha) = \pm\alpha$. Now, $\varphi(\alpha) = \varphi(a_1 + ia_2) = a_i + \varphi(i) a_2 = \pm(a_i + ia_2)$. It follows that either $\varphi(i) = i$ or $\varphi(i) = -i$ ■

Theorem 54 *Any involutive anti-automorphism between subfields of $\mathbb{K}$ extends $I_{\mathbb{Q}}$, the identity map on $\mathbb{Q}$.*

Proof. Let ϕ be an involutive anti-automorphism and let $\mathbb{F} = \{a : \phi(a) = a\}$. It is easy to show that $\mathbb{F}$ is a subfield of $\mathbb{K}$. Since $\mathbb{Q}$ is contained in any subfield, ϕ must extend $I_{\mathbb{Q}}$ [48]. ■

Theorem 55 *If ϕ is an involutive anti-automorphism with domain $\mathbb{K}$, then ϕ can be extended to $\mathbb{K}^a$.*

Proof. Let $\mathcal{F} = \{\theta : \theta$ is an involutive anti-automorphism extending ϕ to a subfield of $\mathbb{K}^a\}$.We shall show that $\mathcal{F}$ satisfies the three hypotheses of Zorn's Lemma. $\mathcal{F}$ is nonempty since ϕ itself extends to $\mathbb{K}$. Clearly, $\mathcal{F} \subseteq \mathbb{K} \times \mathbb{K}$. Let $\mathcal{S}$ be a chain in $\mathcal{F}$ and let σ be the union of all θ in $\mathcal{S}$. $\mathcal{S}$ as a chain, is nonempty; hence it contains atleast one involutive anti-automorphism and thus $\langle 0, 0\rangle$ and $\langle 1, 1\rangle$ are in σ. Let $\langle a, b\rangle$ and $\langle x, y\rangle$ be in σ. Then $\langle a, b\rangle \in \theta_1$ and $\langle x, y\rangle \in \theta_2$ for some $\theta_1, \theta_2 \in \mathcal{S}$. Since $\mathcal{S}$ is a chain, either $\theta_1 \supseteq \theta_2$ or $\theta_1 \subseteq \theta_2$ and thus the two ordered pairs are both in the larger one of θ_1 and θ_2. From this, it follows easily that σ is a one-to-one function which preserves algebric operations. The involutive anti-automorphism σ is in the family $\mathcal{F}$ since it clearly extends ϕ and its domain, the union of subfields of $\mathbb{K}^a$, is contained in $\mathbb{K}^a$. We apply Zorn's Lemma and let ψ be a maximal member of $\mathcal{F}$. We must show that the domain and range of ψ are $\mathbb{K}^a$.

If the domain of ψ is not all of $\mathbb{K}^a$, then there is atleast one element α in $\mathbb{K}^a$ but not in the domain of ψ. Since α is algebriac over $\mathbb{K}$ and $\mathbb{K}^a$ is algebraically closed there is at least one β in $\mathbb{K}^a$ which is the root of the ψ transform of the minimal polynomial of α over $\mathbb{K}$. Thus there is atleast one way of extending ψ to a larger involutive anti-automorphism still in $\mathcal{F}$. This is a contradiction to the maximality of ψ and thus $\mathbb{K}^a$ is the domain of ψ.

Since $\mathbb{K}^a$ is algebraically closed and ψ is an involutive anti-automorphism, the range of ψ is an algebraically closed subfield of $\mathbb{K}^a$ contains $\mathbb{K}$. But the only such subfield of $\mathbb{K}^a$ is $\mathbb{K}^a$ itself; hence $\mathbb{K}^a$ is the range of ψ and the proof is complete. ■

Theorem 56 *Wild, involutive anti-automorphisms do not preserve order*

Proof. Let ϕ be an involutive anti-automorphism between the subfields of $\mathbb{K}$. We first show that ϕ preserves order in $S(\mathbb{K})$. If $x < y$, then there is a number w such that $w \neq 0$ and $y - x = w^2$ but when $\phi(y) - \phi(x) = [\phi(w)]^2$ so that $\phi(w) \in S(\mathbb{K})$ and $\phi(w) \neq 0$. Hence $\phi(y) - \phi(x)$ is positive i.e $\phi(x) < \phi(y)$. Now extend ϕ to $\mathbb{K}$ and assume $a \in \mathbb{K}$ but that $\phi(a) \neq a$.Choose a symmetric number q between a and $\phi(a)$ such that $a < q < \phi(a)$ and apply ϕ: the ordering between a and q is reversed. ■

Corollary 57 *$|\phi(a)| \neq |a|$ for some a.*

Proof. Take $\mathbb{K} = \mathbb{R}$ and $S(\mathbb{K}) = \mathbb{Q}$ with ϕ extended to $\mathbb{R}$ ■

Bibliography

[1] S. J. Aaronson, *Limits on Efficient Computation in the Physical World*, PhD Thesis, University of California, Berkeley (2004)

[2] D Aerts, E. D'Hondt, L. Gabora, *Why the Disjunction in Quantum Logic is Not Classical*, Found. Phys. **30** 9 (2000) pp. 1473–1480

[3] J. A. Alvarez, *$C*$-algebras of operators in non-archimedean Hilbert spaces*, Comment. Math. Univ. Carolin. **33** 4 (1992) pp. 573–580

[4] M. Ardnt, K. Hornberger, *Testing the limits of quantum mechanical superpositions*, Nature Phys. **10** (2014) pp. 271–277

[5] R. Baer *Linear Algebra And Projective Geometry.* Academic Press, New York (1952) pp. 63

[6] J. Baez, *Divison Algebras and Quantum Theory*, Found. Phys. **42** 7 (2011)

[7] S. J. Bhatt, *A Seminorm with square property on a Banach Algebra is submultiplicative*, Proc. Am. Math. Soc. **117** 2 (1993) pp. 435–438

[8] G. D. Birkhoff, J. von Neumann, *The logic of quantum mechanics*, Ann. of Math. **37** (1936) pp. 823–843.

[9] M. Bojowald, S. Brahma, U. Büyükçam, *Testing Nonassociative Quantum Mechanics*, Phys. Rev. Lett. **115** 22 (2015) pp. 22–27

[10] M. Born, *The Statistical Interpretation of Quantum Mechanics*, Nobel Lecture, December 11, 1954 http://www.nobelprize.org/nobel_prizes/physics/laureates/1954/born-lecture.pdf

[11] O. Brunet, *Orthogonality and Dimensionality*, Axioms, **2** (2013) pp. 477–489

[12] J. P. Eckmann, P. C. Zabey, *Impossibility of Quantum Mechanics in a Hilbert Space over a Finite Field*, Hel. Ph. Acta. **42** (1969) pp. 420–424

[13] H. J. Efinger, *A Nonlinear Unitary Framework for Quantum State Reduction*, Department of Scientific Computing Technical Report Series (2005) https://www.cosy.sbg.ac.at/research/tr/2005-03_Efinger.pdf

[14] D. Hilbert, *Foundations of a General Theory of Linear Integral Equations* in Fredholm, Hilbert, Schmidt: Three Fundamental Papers on Integral Equations, Translat. G. W. Stewart (2011) http://www.cs.umd.edu/~stewart/FHS.pdf

[15] T. W. Hungerford, *Algebra*, Springer (2003) ISBN: 978-1-4612-6103-2

[16] Y. Ilyashenko, *Topics in Analysis Math 6130*, Lecture Notes, Cornell University, Lecture 7, retrieved from http://www.math.cornell.edu/~web6130/King'sway.pdf

[17] A. Jaffe, *The GNS Construction*, http://isites.harvard.edu/fs/docs/icb.topic1466080.files/GNS.pdf

[18] J. M. Jauch, C. Piron, *On the Structure of Quantal Proposition Systems* pp. 427–436 in The Logico-Algebraic Approach to Quantum Mechanics by C. A. Hooker (Ed.), Springer Netherlands (1975)

[19] T. Jech. *Set theory*, Springer, 3rd Ed. (2006) ISBN: 978-3-540-44761-0 pp. 149–150

[20] P. Jordan, J. v. Neumann and E. Wigner, *On an algebraic generalization of the quantum mechanical formalism*, Ann. Math. **35** 1 (1934) pp. 29–64

[21] M. Kac, *Can one hear the shape of a drum?*, Amer. Math. Monthly. **73** (1966) pp. 1–23.

[22] J. P. Kahane and W. Zelazko, *A characterization of maximal ideals in commutative Banach algebras*, Studia Math. **29** (1968) pp. 339–343

[23] H. Keller, *On the lattice of all closed subspaces of a Hermitian space*, Pac. J. of Math. **89** 1 (1980) pp. 105–110

[24] A. Y. Khrennikov, *Non-Archimedean Analysis: Quantum Paradoxes, Dynamical Systems and Biological Models*, Springer Netherlands (1997) ISBN 978-94-009-1483-4

[25] M. Kline, *Mathematical Thought from Ancient to Modern Times*, 3, 3rd Ed., Oxford Univ. Press (1990) ISBN: 978-0-19-506137-6.

[26] E. Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons (1989) ISBN: 978-0-47-150459-7

[27] H. Krips, *Measurement in Quantum Theory*, Stanford Encyclopedia of Philosophy (Fall 2013 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2013/entries/qt-measurement/

[28] N. P. Landsman, *Lecture Notes on Hilbert Spaces and Quantum Mechanics*, CreateSpace Independent Publishing Platform (2014) ISBN: 978-1-50-561348-3

[29] D. F. Lawden, The Mathematical Principles of Quantum Mechanics, Dover Publications (2005) ISBN: 978-0-48-644223-5

[30] F. J. Murray, J. v. Neumann, *On rings of operators*, Annals of Math. **37** (1936) pp. 116–229.

[31] W. M. d Muynck, *Foundations of quantum mechanics: an empiricist approach*, Kluwer Academic Publishers, Dordrecht, Boston, London (2002) ISBN 1-4020-0932-1

[32] E. Nelson, *Internal Set Theory: A New Approach To Nonstandard Analysis*, Am. Math. Soc. **83** 6 (1977) pp. 1165–1198

[33] J. v. Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton Landmarks in Mathematics and Physics, Princeton Univ. Press (1996) ISBN: 978-0-69-102893-4

[34] J. v. Neumann, *Unsolved problems in mathematics* in John von Neumann and the Foundations of Quantum Physics, M. Rédei and M. Stöltzner (Eds.), Institute Vienna Circle

Yearbook, pp. 231–245. Kluwer Academic Publishers, Dordrecht (2001) ISBN: 978-0-7923-6812-0

[35] C-K. Ng, *On quaternionic functional analysis*, Math. Proc. Cambridge Philos. Soc. **143** (2007) pp. 391–406

[36] R. Piziak, *Mackey Closure Operator*, J. Lon. Math. Soc. **2** 4:1 (1971) pp. 33–38

[37] R. Piziak, *Sesquilinear forms in infinite dimensions*, Pac. J. Math. **43** 2 (1972) pp. 475–481

[38] K. Popper, *The Logic of Scientific Discovery.* New York, NY: Basic Books (1959) ISBN 978-0-0911-1721-4.

[39] A. Razon, L. P. Horwitz, *Tensor product of quaternion Hilbert Modules*, Acta Appl. Math. **24** 2 (1991) pp. 141–178

[40] M. Rédei, (Editor) *John von Neumann: Selected Letters*, **27:** History of Mathematics, Rhode Island, Am. Math. Soc. and Lon. Math. Soc. (2005)

[41] C. Schwartz, *Nonlinear operators II*, J. of Math. Phys. **38** pp. 3841–3862 (1997)

[42] K. Shamseddine, *Nontrivial order preserving automorphisms of non-Archimedean fields*, Cont. Math. **547** (2011) pp. 217–225

[43] V. Shirbisheh, *Lectures on $C*$-algebras*, arXiv preprint arXiv:1211.3404 (2012)

[44] M. P. Solèr, *Characterisation Of Hilbert Spaces by Orthomodular Spaces*, Comm. In Alg. **23**:1 (1995) pp. 219–243

[45] K. Svozil, *Quantum Logic*, Springer-Verlag Singapore (1998) ISBN: 978-981-4021-07-4

[46] K. Urbanik, F. B. Wright, *Absolute-valued Algebras*, Bull. Acad. Polon. Sci. Śer. Sci. Math. Astronom. Phys. **8** (1960) pp. 285–6

[47] A. Widder, *Spectral Theory for Nonlinear Operators*, Master's Thesis, Vienna Institute of Technology

[48] P. B. Yale, *Automorphisms of Complex Numbers*, Math. Assoc. of Am. Math. Mag. **39** (1966) pp. 135–141