Construction of a Consistent Quantum Mechanical System with non-Hermitian Set of Operators

by

Seher Özçelik

A Thesis Submitted to the
Graduate School of Sciences & Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Mathematics

Koç University

June 2006

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Seher Özçelik

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Committee Members:

Prof. Dr. Ali Mostafazadeh

Prof. Dr. Tekin Dereli

Prof. Dr. Varga Kalantarov

Date:

To my family

ABSTRACT

Conventionally, observables of quantum mechanical systems are represented by Hermitian linear operators acting in a separable Hilbert space. This is mainly because Hermitian operators have real eigenvalues and eigenvalues of an observable represent experimental results which are also real numbers. Recently, non-Hermtian operators with real eigenvalues have been employed in the description of quantum mechanical systems. The following key observations makes this possible: 1) A diagonalizable linear operator H with a discrete spectrum has real eigenvalues if and only if H is η-pseudo-Hertmitian, i.e.,$H^{\dagger} = \eta H \eta^{-1}$, for a bounded, Hermitian, invertible and positive-definite η; 2) If we define a new inner product, namely $\langle\langle\xi|\rho\rangle\rangle_{\eta} = \langle\xi|\eta|\rho\rangle \ \forall|\xi\rangle, |\rho\rangle$ in Hilbert space, then H is Hermitian with respect to this new inner product. η is not unique, but can be uniquely determined up to an arbitrary multiplicative constant provided that one chooses an irreducible set of operators with real spectra and require them to be η-pseudo-Hertmitian. We give a detailed discussion of the related mathematical concepts and a proof of the uniqueness theorem for η. Furthermore we show, by explicit calculations, how this theorem applies for quantum systems with a two-dimensional Hilbert space. In particular, for the latter systems we obtain the general form of a diagonalizable operator H with a real spectrum, calculate the most general η that makes it η-pseudo-Hertmitian, determine the form of any other diagonalizable operator O with a real spectrum that is η-pseudo-Hertmitian, and finally demonstrate how choosing a particular O that together with H form an irreducible set fix η up to a multiplicative constant.

ÖZET

Geleneksel olarak, bir quantum mekaniği sistemindeki gözlenirler (observable) Hilbert uzayındaki Hermitik operatörler ile gösterilir. Bunun temel sebebi Hermitik operatörlerin özdeğerlerinin gerçel olması ve bu özdeğerlerin gerçel sayılar olan deney verilerini temsil ediyor olmasıdır. Son zamanlarda, Hermitik olmayan ancak özdeğerleri gerçel olan operatörler quantum mekaniksel sistemleri ifade etmekte kullanılıyor. Bunu mümkün kılan sebepler şunlardır: 1) Ayrık spektruma sahip diagonalleştirilebilir lineer operatör H'nin özdeğerleri gerçel olur ancak ve ancak H sahte-Hertmitik (pseudo-Hermitian) ise yani $H^\dagger = \eta H \eta^{-1}$, burada η sınırlı (bounded), Hermitik, tersi sınırlı ve tüm uzayda tanımlı (invertible) olan ve tamamen-pozitif (positive-definite) bir lineer operatördür; 2) Eğer yeni bir iç çarpım tanımlarsak, şöyle ki; Hilbert uzayındaki tüm $|\xi\rangle, |\rho\rangle$ için $\langle\langle\xi|\rho\rangle\rangle_\eta = \langle\xi|\eta|\rho\rangle$, H bu yeni iç çarpıma göre Hermitik olur. η tekil değildir, ancak gerçel özdeğerlere sahip operatörler seti seçilip η-sahte-Hermitik (η-pseudo-Hermitian) olması sağlanırsa, η sabit çarpımına kadar tekil olarak belirlenebilir. Burada gerekli matematik temeli ayrıntılı olarak oluşturulduktan sonra η için tekillik teoremi verildi. Ayrıca bu teorem iki boyutlu kompleks Hilbert uzayındaki bir quantum sistemi için ayrıntılı hesaplamalarla uygulandı. Bunun için diagonalleştirilebilir ve gerçel özdeğerlere sahip operatör H'nin genel formu elde edildi, H'nin η-sahte-Hermitik olduğu en genel η bulundu, diagonalleştirilebilir, gerçel özdeğerlere sahip ve η-sahte-Hermitik diğer bir operatör O'nun formu belirlendi ve son olarak O ve H'nin indirgenmez (irreducible) bir set olarak seçiminin η'yı sabit çarpımına kadar tekil olarak belirlediği gösterildi.

ACKNOWLEDGMENTS

First I would like to thank my supervisor Prof. Ali Mostafazadeh for sharing his valuable knowledge and time, for his understanding and for helping me in all stage of the work.

I thank to members of my thesis committee for critical reading of this thesis and for their valuable comments.

Finally I am grateful to my family and friends for their encouragement and support all the time.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

Recently, there has been some interest in a new vision of quantum mechanics in which one can describe a quantum mechanical system using certain non-Hermitian operators. In conventional approach to quantum mechanics, one starts with an inner product and a vector space which is complete with respect to the metric defined by this inner product, i.e., it is a Hilbert space. Then operators which are Hermitian with respect to this inner product are chosen to construct a quantum mechanical system. In [20] Scholtz *et.al.* state that this is a restrictive point of view. They claim that, instead, one can first choose a set of not necessarily Hermitian operators with real eigenvalues and then find (if exists) an inner product or Hilbert space where this set of operators are Hermitian. Naturally this new inner product should be uniquely determined so that one can calculate physical quantities such as an expectation value. This uniqueness problem is our main interest in this thesis. In [12] Mostafazadeh shows that the eigenvalues of a diagonalizable[1] linear operator A are real if and only if there is a bounded, Hermitian, invertible and positive-definite linear operator η such that A is η-pseudo-Hermitian; that is[2] $A^\dagger = \eta A \eta^{-1}$. If we define a new inner product in terms of our usual or reference inner product and η according to $\langle\langle \xi|\rho\rangle\rangle_\eta = \langle \xi|\eta|\rho\rangle$ for all $|\xi\rangle, |\rho\rangle$ in the Hilbert $\mathcal{H}$, then A is Hermitian with respect to this new inner product. However, in [14] Mostafazadeh shows that this η is not unique. Scholtz *et.al.* state that if a set of η-pseudo-Hermitian operators are irreducible on the Hilbert space $\mathcal{H}$, then this η can be uniquely determined up to a constant multiple. We give detailed discussion of the proof of this statement in this thesis.

---

[1] If A is a linear operator on a separable Hilbert space $\mathcal{H}$ with a discrete spectrum and admits a bounded complete biorthonormal system of eigenvectors then A is said to be diagonalizable.

[2] $\dagger$ denotes Hilbert adjoint of the operator.

In Chapter 2, we give some important definitions and theorems to constitute a satisfactory mathematical background for the following chapters. In this chapter we generally follow Kreyszig [9], Hislop *et.al.* [8] and Schmeidler [19].

Chapter 3 begins with some basic notions of quantum mechanics and in Section 3.2 one of the main concepts of this thesis, namely pseudo-Hermiticity, is discussed. We generally follow Bohm [1] for the basic notions and Mostafazadeh [11] [12] [13] [14] [15] for pseudo-Hermiticity.

Chapter 4 consists of a detailed discussion of the uniquness theorem for η as proposed by Scholtz *et.al.*

In Chapter 5, we recall that η is not unique for a η-pseudo-Hermitian operator A and calculate the most general η for a diagonalizable linear operator A with real eigenvalues in finite-dimensions. Then we take another operator S and assume that $\{S, H\}$ is an irreducible set where S is pseudo-Hermitian with respect to one such η and study the effect of perturbing η on S.

In Chapter 6 we study two examples. We use Example 1 to provide an explicit verification of the uniqueness theorem in two dimensions. Here we calculate the most general η for a diagonalizable linear operator A with real eigenvalues. Then we consider a general operator S and require it to be η-pseudo-Hermitian. We then obtain the form of S such that if we fix one such S then, $\{S, H\}$ will be a set of η-pseudo-Hermitian operators. We observe that for the case $det[S, H] \neq 0$ [3] which is equivalent to the irreducibility of the set $\{S, H\}$ on $\mathbb{C}^2$, then η is determined uniquely up to a constant multiple which verifies the uniqueness theorem. Example 2 is a concrete application of the results of Section 5.2.

[3] $det[S, H] = det(SH - HS) \neq 0$.

Chapter 2

MATHEMATICAL BACKGROUND

In this chapter, we review some basic notions that we will use in subsequent chapters. We will generally follow Ref. [9] in our presentation. The proofs of all the theorems and propositions of this chapter are given in Appendix A.1.

2.1 Preliminaries

Definition 2.1.1. *(Vector Space or Linear Space)*
A vector (linear) space over a field K ($\mathbb{R}$ or $\mathbb{C}$) is a nonempty set X of elements $x, y, \cdots$ (called vectors) together with two algebraic operations. The elements of X are called vectors and the binary operations are called vector addition and multiplication of vectors by scalars, that is, by elements of K. They are postulated to have the following properties: Vector addition associates with every ordered pair (x, y) of vectors a vector $x + y$, called the sum of x and y, in such a way that the following properties hold. Vector addition is commutative and associative, that is, for all vectors x, y, z we have $x+y = y+x$ and $x+(y+z) = (x+y)+z$; furthermore there exist a vector 0, called the zero vector, such that for all vectors x we have $x + 0 = x$ and for every vector x there exist a vector $-x$, such that $x + (-x) = 0$. Multiplication by scalars associates with every vector x and scalar α a vector αx (also written $x\alpha$), called the product of α and x, in such a way that for all vectors x, y and scalars α, β we have $\alpha(\beta x) = (\alpha\beta)x$ and $1x = x$ and the distributive laws $\alpha(x + y) = \alpha x + \alpha y$ and $(\alpha + \beta)x = \alpha x + \beta x$ hold.

If X is defined on $\mathbb{R}$ then X is called a real vector space and if X is defined on $\mathbb{C}$ then it is called a complex vector space.

Definition 2.1.2. *[6](Span, Linear Combination)*

Let X be a vector space over K. The vectors in a subset $S = \{v_i | i \in I : \textit{index set}\}$ of X span (or generate) X if for every $x \in X$, we have

$$x = \alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n} \tag{2.1}$$

for some $\alpha_j \in K$ and $v_{i_j} \in S$, $j = 1, 2, \cdots n$. $\sum_{j=1}^{n} \alpha_j v_{i_j}$ is a linear combination of the v_{i_j}.

Definition 2.1.3. *[6](Independence)*
The vectors in a subset $S = \{v_i | i \in I\}$ of a vector space X over a field K are linearly independent over K if the only way to express the zero vector as a linear combination of the vectors v_i is to have all scalar coefficients equal to 0, that is,

$$\alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n} = 0 \quad \Rightarrow \quad \alpha_1 = \alpha_2 = ... = \alpha_n = 0. \tag{2.2}$$

If the vectors are not linearly independent over K, they are linearly dependent over K.

Definition 2.1.4. *[6](Basis)*
If X is a vector space over a field K, the vectors in a subset $B = \{b_i | i \in I\}$ of X form a basis for X over K if they span X and are linearly independent.

Definition 2.1.5. *[6](Dimension)*
Every vector space has a basis. Any two basis have the same cardinality which is called the dimension.

Definition 2.1.6. *(Normed Space[1])*
A normed space X is a vector space with norm defined on it. Here a norm on a (real or complex) vector space X is a real-valued function on X whose value at an $x \in X$ is denoted by $\|x\|$ (read "norm of x") and which has the following properties.

$$\|x\| \geq 0, \tag{2.3}$$

$$\|x\| = 0 \Leftrightarrow x = 0, \tag{2.4}$$

[1]It is also called "normed vector space" or "normed linear space".

$$\|\alpha x\| = |\alpha|\|x\|, \tag{2.5}$$

$$\|x + y\| \leq \|x\| + \|y\| \quad (Triangle\ inequality), \tag{2.6}$$

where x and y are arbitrary vectors in X and α is any scalar.

Proposition 2.1.7. *Let X be a normed vector space. Then for any $x, y \in X$ we have*

$$|\ \|x\| - \|y\|\ | \leq \|x - y\|. \tag{2.7}$$

Definition 2.1.8. *(Metric or Distance Function Defined by the Norm)*
The norm of a normed vector space X defines a metric or distance function d on X which is given by

$$d(x, y) = \|x - y\| \quad (x, y \in X) \tag{2.8}$$

and is called the metric induced by the norm. Recall that a metric or distance function is defined as a real-valued function of ordered pairs (x, y) of elements of X that satisfies the conditions

$$\begin{gathered} d \text{ is real-valued and nonnegative;} \\ d(x, y) = 0 \text{ if and only if } x = y; \\ d(x, y) = d(y, x) \quad \text{(Symmetry);} \\ d(x, y) \leq d(x, z) + d(z, y) \quad \text{(Triangle inequality).} \end{gathered} \tag{2.9}$$

Definition 2.1.9. *(Open Ball, Open Set, Interior point)*
Given a point x_0 in a normed vector space X and a real number $r > 0$, then

$$B(x_0; r) = \{x \in X \mid d(x, x_0) = \|x - x_0\| < r\} \tag{2.10}$$

is called an open ball.
A subset M of a normed vector space X is said to be open if it contains an open ball about each of its points.
An open ball $B(x_0; \varepsilon)$ of radius ε is often called an ε-neighborhood of x_0. By a neighborhood x_0 we mean any subset of X which contains an ε-neighborhood of x_0. Obviously every neighborhood of x_0 contains x_0, and if N is a neighborhood of x_0 and $N \subseteq M$, then M is also a neighborhood of x_0.

We call x_0 an interior point of a set $M \subseteq X$ if M is a neighborhood of x_0. i.e., M contains an ε-neighborhood of x_0 (i.e., contains $B(x_0; \varepsilon)$). The interior of M is the set of all interior points of M.

Definition 2.1.10. *(Limit or Accumulation Point, Closure)*
Let M be any subset of a normed vector space X. A limit or accumulation point x_0 of M is a point in X that every neighborhood of x_0 contains at least one point of M distinct from x_0. Elements of M and all limit points of M, together constitute the closure of M which is denoted by $\overline{M}$. Note that the set $\overline{M} \setminus M$ does not contain any interior point.

Definition 2.1.11. *(Convergence of a Sequence, Limit)*
A sequence $\{x_n\}$ in a normed vector space X is said to converge or to be convergent if there is an $x \in X$ such that

$$\lim_{n \to \infty} d(x_n, x) = \lim_{n \to \infty} \|x_n - x\| = 0, \tag{2.11}$$

or equivalently,

$$\forall\ \varepsilon > 0, \quad \exists\ N \in \mathbb{N} \quad \textit{such that, for all } n \geq N, \quad d(x_n, x) = \|x_n - x\| < \varepsilon. \tag{2.12}$$

x is called the limit of $\{x_n\}$ and we write

$$\lim_{n \to \infty} x_n = x \tag{2.13}$$

or simply

$$x_n \to x \tag{2.14}$$

We say that $\{x_n\}$ converges to x or has the limit x.

Note that by (2.7) we have $|\ \|x_n\| - \|x\|\ | \to 0$ and hence

$$\|x_n\| \to \|x\|. \tag{2.15}$$

Theorem 2.1.12. *Let M be a nonempty subset of a normed vector space X, then $x \in \overline{M}$ if and only if there is a sequence $\{x_n\}$ in M such that $x_n \to x$.*

Definition 2.1.13. *(Closed Subset)*
A subset M of a normed vector space X is said to be closed if $M = \overline{M}$. Moreover by above theorem we see that M is closed if and only if for any convergent sequence $\{x_n\} \subset M$ with $x_n \to x$, the limit point $x \in M$.

Definition 2.1.14. *(Dense Subset)*
A subset M is said to be dense in the normed vector space X if $\overline{M} = X$.

Definition 2.1.15. *(Cauchy Sequence, Completeness)*
A sequence $\{x_n\}$ in a normed vector space X is said to be Cauchy if

$$\forall\ \varepsilon > 0,\ \exists N \in \mathbb{N} \quad \textit{such that}, \quad \forall\ m, n \geq N \quad d(x_m - x_n) = \|x_m - x_n\| < \varepsilon, \tag{2.16}$$

(i.e., $\|x_m - x_n\| \to 0$ as $m, n \to \infty$).

A normed vector space is said to be complete, if every Cauchy sequence converges in this space, i.e., has a limit which is an element of this space.

Proposition 2.1.16. *Every convergent sequence is Cauchy, but converse is not generally true.*

Definition 2.1.17. *(Inner Product or pre-Hilbert Space)*
An inner product or pre-Hilbert space is a vector space X with an inner product defined on it. Here, an inner product on X is a mapping of $X \times X$ [2] into the field K of X; that is, with every pair of vectors x and y there is associated an element of K which is written as $\langle x, y \rangle$ and is called the inner product of x and y, such that for all vectors x, y, z and scalars α we have

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \tag{2.17}$$

$$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle, \tag{2.18}$$

$$\langle x, y \rangle = \langle y, x \rangle^*, \tag{2.19}$$

$$\begin{gathered} \langle x, x \rangle \ \textit{is real and non-negative}, \\ \langle x, x \rangle = 0 \ \Leftrightarrow \ x = 0. \end{gathered} \tag{2.20}$$

[2] Set of all ordered pairs (x_1, x_2), where $x_1, x_2 \in X$.

Here * denotes the complex conjugate of the scalar. Note that

$$\langle x, \alpha y_1 + y_2\rangle = \langle x, \alpha y_1\rangle + \langle x, y_2\rangle = \alpha^*\langle x, y_1\rangle + \langle x, y_2\rangle$$

where $y_1, y_2 \in X$ *because,* $\langle x, \alpha y_1 + y_2\rangle = \langle \alpha y_1 + y_2, x\rangle^* = (\langle \alpha y_1, x\rangle + \langle y_2, x\rangle)^* = \langle \alpha y_1, x\rangle^* + \langle y_2, x\rangle^* = (\alpha\langle y_1, x\rangle)^* + \langle y_2, x\rangle^* = \alpha^*\langle y_1, x\rangle^* + \langle y_2, x\rangle^* = \alpha^*\langle x, y_1\rangle + \langle x, y_2\rangle$.

An inner product on X defines a norm on X given by

$$\|x\| = \sqrt{\langle x, x\rangle} \tag{2.21}$$

and a metric on X given by

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y\rangle}. \tag{2.22}$$

Definition 2.1.18. *(Hilbert Space)*

A Hilbert space $\mathcal{H}$ is an inner product space which is complete in the metric defined by its inner product.

Notation: In what follows we will use $\mathcal{H}$ to denote a Hilbert space.

Theorem 2.1.19. *Every closed subspace[3] M of a Hilbert Space $\mathcal{H}$ is again a Hilbert Space.*

Definition 2.1.20. *(Orthogonality)*

An element x of an inner product space X is said to be orthogonal to an element $y \in X$ if

$$\langle x, y\rangle = 0 \tag{2.23}$$

We also say that x and y are orthogonal, and we write $x \perp y$. Similarly, for subsets $A, B \subseteq X$ we write $x \perp A$ if $x \perp a$ for all $a \in A$, and $A \perp B$ if $a \perp b$ for all $a \in A$ and all $b \in B$.

A set is said to be orthogonal if every two distinct elements are orthogonal to each other.

[3] Subset that satisfies the conditions of a vector space.

Definition 2.1.21. *[8] (Orthogonal complement)*
Let M be a closed subspace of a Hilbert Space $\mathcal{H}$. The orthogonal complement of M in $\mathcal{H}$, denoted as $M^\perp$, is the set of vectors in $\mathcal{H}$ which are orthogonal to M.

$$M^\perp = \{x \in \mathcal{H} | \langle x, m \rangle = 0,\ \forall m \in M\} \tag{2.24}$$

Proposition 2.1.22. *[8] Let M be a closed subspace of $\mathcal{H}$. Then the subset $M^\perp$ is a closed subspace of $\mathcal{H}$ and is therefore a Hilbert Space. Furthermore $M \cap M^\perp = \{0\}$.*

Proposition 2.1.23. *[8] If M is a subspace of $\mathcal{H}$ which is not closed, then we can still define the orthogonal complement of M as in Definition (2.1.21) and $M^\perp$ is still a closed subspace of $\mathcal{H}$. Moreover $(M^\perp)^\perp = \overline{M}$ and $M^\perp = \overline{M}^\perp$.*

Theorem 2.1.24. *[8] (The Projection Theorem)*
Let M be a closed subspace of $\mathcal{H}$. Then any $x \in \mathcal{H}$ can be written uniquely as $x = y+z$ with $y \in M$ and $z \in M^\perp$, i.e., $\mathcal{H}$ is the direct sum of M and $M^\perp$ (which are themselves Hilbert spaces), and we write $\mathcal{H} = M \oplus M^\perp$.

Proposition 2.1.25. $\mathcal{H}^\perp = \{0\}$.

Proposition 2.1.26. *For any subset $W \neq \emptyset$ of a Hilbert space $\mathcal{H}$, the space M of all (finite) linear combinations of elements of W, that is $M = spanW$, is dense in $\mathcal{H}$ if and only if $W^\perp = \{0\}$.*

Definition 2.1.27. *(Fundamental Set)*
A fundamental set in a Hilbert space $\mathcal{H}$ is a subset $W \subseteq \mathcal{H}$ whose span is dense in $\mathcal{H}$, that is $W^\perp = \{0\}$.

Definition 2.1.28. *[8](Orthonormal Set)*
A collection $\{x_\alpha\}$ of vectors in a Hilbert space $\mathcal{H}$ is called an orthonormal set if

$$\langle x_\alpha, x_\beta \rangle = \delta_{\alpha\beta} = \begin{cases} 1 & \text{for } \alpha = \beta; \\ 0 & \text{for } \alpha \neq \beta, \end{cases} \tag{2.25}$$

that is, it is an orthogonal set and all its elements have norm 1.

An orthogonal set $\{y_\alpha\}$ in a Hilbert space $\mathcal{H}$ can be made into an orthonormal set by dividing all the elements by their norms.

Definition 2.1.29. *[8] (Orthonormal Basis)*

An orthonormal set B in a Hilbert space $\mathcal{H}$ is called an orthonormal basis if there exist no other orthonormal set B' in $\mathcal{H}$ such that B is a proper subset of B'. Hence an orthonormal basis B in $\mathcal{H}$ is a maximal orthonormal set and clearly, $B^{\perp} = \{0\}$. That is, B is a fundamental set in $\mathcal{H}$. It is important to note that this is not a basis, in the sense of algebra unless $\mathcal{H}$ is finite dimensional[9].

Let $\mathcal{H}$ be a finite-dimensional Hilbert space and $\{x_\alpha\}_{\alpha=1}^{n}$ be an orthonormal basis for $\mathcal{H}$. Then any $x \in \mathcal{H}$ can be written uniquely as

$$x = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n \tag{2.26}$$

in terms of scalars $c_1, c_2, \cdots, c_n$. Then, for any x_α from the basis,

$$\langle x, x_\alpha \rangle = \Big\langle \sum_{\beta=1}^{n} c_\beta x_\beta, x_\alpha \Big\rangle = \sum_{\beta=1}^{n} \langle c_\beta x_\beta, x_\alpha \rangle = \sum_{\beta=1}^{n} c_\beta \langle x_\beta, x_\alpha \rangle = \sum_{\beta=1}^{n} c_\beta \delta_{\alpha\beta} = c_\alpha. \tag{2.27}$$

Hence

$$x = \langle x, x_1 \rangle x_1 + \langle x, x_2 \rangle x_2 + \cdots + \langle x, x_n \rangle x_n = \sum_{\alpha=1}^{n} \langle x, x_\alpha \rangle x_\alpha. \tag{2.28}$$

The norm of x is,

$$\begin{aligned} \|x\| = \sqrt{\langle x, x \rangle} &= \sqrt{\Big\langle \textstyle\sum_{\alpha=1}^{n} c_\alpha x_\alpha, \sum_{\beta=1}^{n} c_\beta x_\beta \Big\rangle} \\ &= \sqrt{\textstyle\sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} c_\alpha c_\beta^{*} \langle x_\alpha, x_\beta \rangle} = \sqrt{\textstyle\sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} c_\alpha c_\beta^{*} \delta_{\alpha\beta}} \\ &= \sqrt{\textstyle\sum_{\alpha=1}^{n} c_\alpha c_\alpha^{*}} = \sqrt{\textstyle\sum_{\alpha=1}^{n} |c_\alpha|^2} = \sqrt{\textstyle\sum_{\alpha=1}^{n} |\langle x, x_n \rangle|^2}, \end{aligned} \tag{2.29}$$

briefly,

$$\|x\|^2 = \sum_{\alpha=1}^{n} |c_\alpha|^2 = \sum_{\alpha=1}^{n} |\langle x, x_\alpha \rangle|^2. \tag{2.30}$$

Let $y = d_1 x_1 + d_2 x_2 + \cdots + d_n x_n$ be any vector in the same space, then

$$\langle x, y \rangle = \Big\langle \sum_{\alpha=1}^{n} c_\alpha x_\alpha, \sum_{\beta=1}^{n} d_\alpha x_\alpha \Big\rangle = \sum_{\alpha=1}^{n} c_\alpha d_\alpha^{*}. \tag{2.31}$$

Note that we can represent the vector x with the sequence $(c_1, c_2, \cdots, c_n)$ of scalars in this basis.

Definition 2.1.30. *[8](Separable Hilbert Space)*

A Hilbert space is said to be separable if it has a countable orthonormal basis.

Theorem 2.1.31. *Every finite dimensional inner product space is a separable Hilbert space.*

This is based on the theorem that; every finite dimensional subspace Y of a normed vector space X is complete. In particular, every finite dimensional normed vector space is complete.

In a separable Hilbert space we see that an orthonormal basis is a fundamental set in this space. Representation of vectors in terms of an orthonormal basis in an infinite dimensional separable Hilbert space is given by the following theorem.

Theorem 2.1.32. *[8] Let $B = \{x_n\}$ be an orthonormal basis for an infinite-dimensional separable Hilbert space $\mathcal{H}$. Then for any $x \in \mathcal{H}$*

$$x = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n \tag{2.32}$$

where the sum converges absolutely and

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2. \tag{2.33}$$

Conversely, if $\{c_n\}$ is a collection of complex numbers satisfying $\sum_{n=1}^{\infty} |c_n| < \infty$ then $\sum_{n=1}^{\infty} c_n x_n$ converges absolutely to a vector in $\mathcal{H}$.

Definition 2.1.33. *(Linear Operator)*
Let X and Y be two vector spaces. A linear operator T is a mapping $T : X \to Y$ such that

(i) the domain $\mathcal{D}(T)$ of T is a vector space and the range $R(T) := T(\mathcal{D}(T))$ lies in a vector space over the same field,

(ii) for all $x, y \in \mathcal{D}(T)$ and scalars α,

$$T(x + y) = Tx + Ty \tag{2.34}$$

$$T(\alpha x) = \alpha Tx. \tag{2.35}$$

By definition, the null space or kernel $ker(T)$ of T is the set of all $x \in \mathcal{D}(T)$ such that $Tx = 0$, i.e., $ker(T) := \{x \in \mathcal{D}(T) | Tx = 0\}$.

Definition 2.1.34. *(Inverse of a Linear Operator)*
Let X and Y be two vector spaces and $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, be a linear operator. T is said to be injective or one-to-one if different points in the domain have different images, that is, if for any $x_1, x_2 \in \mathcal{D}(T)$,

$$x_1 \neq x_2 \quad \Rightarrow \quad Tx_1 \neq Tx_2; \tag{2.36}$$

or equivalently,

$$Tx_1 = Tx_2 \quad \Rightarrow \quad x_1 = x_2. \tag{2.37}$$

In this case there exist the mapping

$$T^{-1} : R(T) \longrightarrow \mathcal{D}(T) \tag{2.38}$$

which maps every $y_0 \in R(T)$ onto the $x_0 \in \mathcal{D}(T)$ for which $Tx_0 = y_0$. The mapping T^{-1} is called the inverse of T.

Proposition 2.1.35. *Let X and Y be two vector spaces and $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, be a linear operator. Then if T^{-1} exists, it is a linear operator.*

Theorem 2.1.36. *Let X and Y be two vector spaces and $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, be a linear operator. Then $T^{-1} : R(T) \to \mathcal{D}(T)$ exists if and only if*

$$Tx = 0 \quad \Rightarrow \quad x = 0, \tag{2.39}$$

i.e., ker(T)= $\{0\}$.

Definition 2.1.37. *(Bounded Linear Operator, Operator Norm)*
Let X and Y be two normed vector spaces and T be a linear operator $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$. The operator T is said to be bounded, if there is a positive real number c such that for all $x \in \mathcal{D}(T)$,

$$\|Tx\| \leq c\|x\|. \tag{2.40}$$

We can define the operator norm for a bounded linear operator according to

$$\|T\| := \sup_{x \in \mathcal{D}(T),\ x \neq 0} \frac{\|Tx\|}{\|x\|}. \tag{2.41}$$

Hence

$$\|Tx\| \leq \|T\|\|x\|. \tag{2.42}$$

Equivalently, we have $\|T\| = \sup_{x\in\mathcal{D}(T),\ \|x\|=1} \|Tx\|$.

Definition 2.1.38. *(Continuous Linear Operator)*
Let X and Y be two normed vector spaces. A linear operator $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, is continuous at a point $x_0 \in \mathcal{D}(T)$ if for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\|Tx - Tx_0\| < \varepsilon \quad \textit{for all } x \textit{ satisfying} \quad \|x - x_0\| < \delta. \tag{2.43}$$

T is said to be continuous if it is continuous at every point of $\mathcal{D}(T)$.

Theorem 2.1.39. *Let X and Y be two normed vector spaces. A linear operator $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, is continuous at a point $x_0 \in \mathcal{D}(T)$ if and only if*

$$x_n \to x_0 \quad \textit{implies} \quad Tx_n \to Tx_0. \tag{2.44}$$

That is for a continuous operator T, $\lim_{n\to\infty} Tx_n = T(\lim_{n\to\infty} x_n) = Tx$.

Proposition 2.1.40. *Let X and Y be two normed vector spaces and $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, be a linear operator. If for each sequence $\{x_n\} \subset \mathcal{D}(T)$ with $\lim_{n\to\infty} x_n = 0$, we have $\lim_{n\to\infty} Tx_n = 0$ then T is continuous.*

Theorem 2.1.41. *Let X and Y be two normed vector spaces and $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, be a linear operator. Then the following two statements are equivalent:*

1. *T is continuous;*

2. *T is bounded .*

Definition 2.1.42. *[8](Invertible Linear Operator)*
Let $T : \mathcal{D}(T) \to \mathcal{H}_2$ be a linear operator, with $\mathcal{D}(T) \subseteq \mathcal{H}_1$, where $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces. Then T is said to be invertible if it has a bounded inverse defined on all of $\mathcal{H}_2$.

Definition 2.1.43. *(Discrete Spectrum of a Linear Operator, Eigenvalue, Eigenvector)*
Let X be a normed vector space and $T : \mathcal{D}(T) \to X$ be a linear operator with domain $\mathcal{D}(T) \subseteq X$. Let $\lambda \in \mathbb{C}$ such that $(T - \lambda\mathbb{I})^{-1}$ does not exist, i.e., $ker(T - \lambda\mathbb{I}) \neq \{0\}$, where $\mathbb{I}$ is the identity operator[4]*, then λ is said to belong to the discrete spectrum (or point spectrum) $\sigma_p(T)$ of T. It is said to be an eigenvalue of T.*

If $(T - \lambda\mathbb{I})x = 0$ with $x \neq 0$. Then λ is an eigenvalue of T, and x is said to be an eigenvector of T for the eigenvalue λ.

The subspace of $\mathcal{D}(T)$ consisting of 0 and all eigenvectors of T corresponding to an eigenvalue λ of T is called eigenspace of T corresponding to that eigenvalue λ.

Definition 2.1.44. *(Linear Functional, Bounded Linear Functional)*
A linear functional f is a linear operator with domain in a vector space X and range in the field K of X; thus, $f : \mathcal{D}(f) \to K$, where $K = \mathbb{R}$ if X is real and $K = \mathbb{C}$ if X is complex. A bounded linear functional f is a linear functional that is bounded, i.e., there exist a real number c such that for all $x \in \mathcal{D}(f)$,

$$|f(x)| \leq c\|x\|. \tag{2.45}$$

Furthermore the norm of f is

$$\|f\| = \sup_{x\in\mathcal{D}(f),\ x\neq 0} \frac{|f(x)|}{\|x\|} = \sup_{x\in\mathcal{D}(f),\ \|x\|=1} |f(x)|. \tag{2.46}$$

So we have,

$$|f(x)| \leq \|f\|\|x\|. \tag{2.47}$$

Definition 2.1.45. *(Dual Space X')*
Let X be a normed vector space. Then the set of all bounded linear functionals defined on all of X, i.e., $\mathcal{D}(f) = X$, constitutes a normed space, with the norm defined by Equation (2.46); it is called the dual space (adjoint space or conjugate space) of X and is denoted by X'.

[4]The identity operator $\mathbb{I} : X \to X$ is defined by $\mathbb{I}x = x$ for all $x \in X$.

Definition 2.1.46. *(Adjoint Operator $T^{\times}$ of a Bounded Linear Operator T)*
Let $T : X \to Y$, with $\mathcal{D}(T) = X$, be a bounded linear operator, where X and Y are normed vector spaces. Let g be a bounded linear functional on Y ($g \in Y'$). Then g defines a functional f on X, according to

$$f(x) := g(Tx) \qquad x \in X. \tag{2.48}$$

f is linear since g and T are linear. f is bounded because

$$|f(x)| = |g(Tx)| \leq \|g\|\|Tx\| \leq \|g\| \; \|T\| \; \|x\|$$

$$\sup_{x\in X, \|x\|=1} |f(x)| \leq \sup_{x\in X, \|x\|=1} \|g\| \; \|T\| \; \|x\|$$
$$\|f\| \leq \|g\| \; \|T\|. \tag{2.49}$$

Since f is linear and bounded, $f \in X'$.
Adjoint operator

$$T^{\times} : Y' \to X' \tag{2.50}$$

is defined by,

$$(T^{\times}g)(x) := g(Tx) = f(x) \qquad (g \in Y'). \tag{2.51}$$

Theorem 2.1.47. *[8](Riesz's Theorem)*
Every bounded linear functional $f : \mathcal{H} \to K$ on a Hilbert space $\mathcal{H}$ with $\mathcal{D}(f) = \mathcal{H}$ can be represented in terms of the inner product as

$$f(x) = \langle x, z \rangle \tag{2.52}$$

where z is uniquely determined by f, and satisfies

$$\|z\| = \|f\|. \tag{2.53}$$

Definition 2.1.48. *(Isomorphism)*
First note that, a linear operator $A : \mathcal{D}(A) \to Y$, where $\mathcal{D}(A) \subseteq X$ and X and Y are vector spaces, is called bijective or bijection if it is one-to-one[5] and onto[6].

---

[5]See 2.1.34.

[6]Range $R(A) = Y$.

An isomorphism of a normed vector space X onto a normed vector space $\tilde{X}$ is a bijective linear operator $T : X \to \tilde{X}$, with $\mathcal{D}(T) = X$, which preserves the norm, that is for all $x \in X$,

$$\|Tx\| = \|x\|. \tag{2.54}$$

(Hence T is isometric.) X is then called isomorphic with $\tilde{X}$, and X and $\tilde{X}$ are called isomorphic normed vector spaces.

Proposition 2.1.49. *Dual space $\mathcal{H}'$ of a Hilbert space $\mathcal{H}$ and $\mathcal{H}$ are isomorphic normed vector spaces. One can express this more briefly by saying that dual space $\mathcal{H}'$ is $\mathcal{H}$.*

Definition 2.1.50. *(Hilbert-adjoint Operator of a Bounded Linear Operator)*
Let $T : \mathcal{H}_1 \to \mathcal{H}_2$, with $\mathcal{D}(T) = \mathcal{H}_1$, be a bounded linear operator where $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert Spaces. Then the Hilbert-adjoint operator $T^\dagger$ of T is the operator;

$$T^\dagger : \mathcal{H}_2 \to \mathcal{H}_1, \tag{2.55}$$

with $\mathcal{D}(T^\dagger) = \mathcal{H}_2$, such that for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$,

$$\langle Tx, y\rangle = \langle x, T^\dagger y\rangle. \tag{2.56}$$

Next, we explain the relation between the adjoint of a bounded linear operator T; $T^\times$, and Hilbert-adjoint of a bounded linear operator T; $T^\dagger$. Equation (2.48) in Definition (2.1.46) states that

$$f(x) := g(Tx) \qquad x \in \mathcal{H}_1. \tag{2.57}$$

By Riesz's Theorem (2.1.47) we can represent these functionals uniquely as

$$f(x) = \langle x, z\rangle \quad z \in \mathcal{H}_1 \qquad \textit{and} \qquad g(Tx) = \langle Tx, y\rangle \quad y \in \mathcal{H}_2. \tag{2.58}$$

Hence Equation (2.57) is transformed into

$$\langle x, z\rangle := \langle Tx, y\rangle \qquad z \in \mathcal{H}_1 \quad \textit{and} \quad y \in \mathcal{H}_2. \tag{2.59}$$

Hence by Equation (2.51) and by one-to-one correspondence $f \leftrightarrow z$ between $\mathcal{H}_1' \leftrightarrow \mathcal{H}_1$ and one-to-one correspondence $g \leftrightarrow y$ between $\mathcal{H}_2' \leftrightarrow \mathcal{H}_2$ we write

$$T^{\times} g := f \quad \longleftrightarrow \quad T^{\dagger} y := z. \tag{2.60}$$

Hence

$$\langle x, T^{\dagger} y \rangle = \langle Tx, y \rangle.$$

Theorem 2.1.51. *(Existence)*
The Hilbert-adjoint operator $T^{\dagger}$ of T in Definition (2.1.50) exists, is unique and is a bounded linear operator with norm

$$\|T^{\dagger}\| = \|T\|. \tag{2.61}$$

Theorem 2.1.52. *Let $\mathcal{H}_1$, $\mathcal{H}_2$ be Hilbert spaces, $S : \mathcal{H}_1 \to \mathcal{H}_2$ and $T : \mathcal{H}_1 \to \mathcal{H}_2$, with $\mathcal{D}(S) = \mathcal{D}(T) = \mathcal{H}_1$, bounded linear operators and α any scalar. Then we have*

$$\langle T^{\dagger} y, x \rangle = \langle y, Tx \rangle \qquad (x \in \mathcal{H}_1, y \in \mathcal{H}_2), \tag{2.62}$$

$$(\alpha T)^{\dagger} = \alpha^* T^{\dagger}, \tag{2.63}$$

$$(T^{\dagger})^{\dagger} = T, \tag{2.64}$$

$$\|T^{\dagger} T\| = \|T T^{\dagger}\| = \|T\|^2, \tag{2.65}$$

$$T^{\dagger} T = 0 \quad \Leftrightarrow \quad T = 0, \tag{2.66}$$

$$(ST)^{\dagger} = T^{\dagger} S^{\dagger} \qquad (\textit{assuming } \mathcal{H}_2 = \mathcal{H}_1), \tag{2.67}$$

where * denotes the complex conjugate of the scalar.

Definition 2.1.53. *(Hilbert-adjoint Operator of a Possibly Unbounded Linear Operator on $\mathcal{H}$)*
Let $T : \mathcal{D}(T) \to \mathcal{H}$, with $\mathcal{D}(T) \subseteq \mathcal{H}$, be a (possibly unbounded) densely defined[7] linear operator in a complex Hilbert space $\mathcal{H}$. Then the Hilbert-adjoint operator

[7] $\mathcal{D}(T)$ is dense in $\mathcal{H}$.

$T^\dagger : \mathcal{D}(T^\dagger) \to \mathcal{H}$ *of* T *is defined as follows. The domain* $\mathcal{D}(T^\dagger)$ *of* $T^\dagger$ *consists of all* $y \in \mathcal{H}$ *such that there is a* $\tilde{y} \in \mathcal{H}$ *satisfying*

$$\langle Tx, y \rangle = \langle x, \tilde{y} \rangle \tag{2.68}$$

for all $x \in \mathcal{D}(T)$*. For each such* $y \in \mathcal{D}(T^\dagger)$ *the Hilbert-adjoint operator* $T^\dagger$ *is defined by*

$$\tilde{y} = T^\dagger y. \tag{2.69}$$

It is easy to show that $T^\dagger$ *is a linear operator.*

Proposition 2.1.54. *The adjoint* $T^\dagger$ *of a linear operator* $T : \mathcal{D}(T) \to \mathcal{H}$ *with* $\mathcal{D}(T) \subseteq \mathcal{H}$ *and* $\mathcal{H}$ *a Hilbert space is not well-defined unless* T *is densely defined, i.e.,* $\mathcal{D}(T)$ *is dense in* $\mathcal{H}$*.*

Definition 2.1.55. *(Hermitian or Self-adjoint Bounded/Unbounded Linear Operator on* $\mathcal{H}$*)*
A linear operator $T : \mathcal{H} \to \mathcal{H}$*, with* $\mathcal{D}(T) \subseteq \mathcal{H}$*, on a Hilbert space* $\mathcal{H}$ *is said to be Hermitian or self-adjoint if* [8]

$$T^\dagger = T, \tag{2.70}$$

that is, $\langle Tx, y \rangle = \langle x, Ty \rangle$ *for all* $x, y \in \mathcal{D}(T) = \mathcal{D}(T^\dagger)$*.*

Theorem 2.1.56. *(Hellinger-Toeplitz Theorem)*
If a linear operator $T : \mathcal{H} \to \mathcal{H}$ *is defined on all of a complex Hilbert Space* $\mathcal{H}$*, and satisfies* $\langle Tx, y \rangle = \langle x, Ty \rangle$ *for all* $x, y \in \mathcal{H}$*, then* T *is bounded.*

As a result of this theorem, an unbounded linear operator $T : \mathcal{D}(T) \to \mathcal{H}$ with $\mathcal{D}(T) \subseteq \mathcal{H}$ that satisfies $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in \mathcal{D}(T)$, cannot be defined on all of $\mathcal{H}$. That is, $\mathcal{D}(T) \neq \mathcal{H}$.

Definition 2.1.57. *[8](Positive-definite Operator, Positive Operator)*
A Hermitian linear operator $T : \mathcal{D}(T) \to \mathcal{H}$*, with* $\mathcal{D}(T) \subseteq \mathcal{H}$*, is said to be positive-definite, if*

$$\langle Tx, x \rangle > 0 \quad \textit{for all} \;\; x \in \mathcal{D}(T) - \{0\}. \tag{2.71}$$

[8]Two operators T_1 and T_2 are defined to be equal, and written $T_1 = T_2$, if they have the same domain $\mathcal{D}(T_1) = \mathcal{D}(T_2)$ and if $T_1 x = T_2 x$ for all $x \in \mathcal{D}(T_1) = \mathcal{D}(T_2)$.

T is said to be positive, if

$$\langle Tx, x\rangle \geq 0 \quad \textit{for all} \;\; x \in \mathcal{D}(T). \tag{2.72}$$

Note that, $\langle Tx, x\rangle = \langle x, Tx\rangle$, so $\langle Tx, x\rangle$ is real.

Theorem 2.1.58. *If a normed space X is finite dimensional, then every linear operator on X is bounded and linear operators on finite dimensional vector spaces can be represented in terms of matrices.*

2.2 Biorthonormal Systems

Definition 2.2.1. *[19](Biorthonormal System)*
Finite or countably infinite set of elements u_α and v_β of a Hilbert space $\mathcal{H}$ that satisfy

$$\langle u_\alpha, v_\beta\rangle = \delta_{\alpha\beta} \tag{2.73}$$

are said to form a biorthonormal system $\{u_\alpha, v_\alpha\}$ in $\mathcal{H}$.

Proposition 2.2.2. *[19] The sets $\{u_\alpha\}$ and $\{v_\alpha\}$ are linearly independent.*

Proof. Let

$$c_1 u_{\alpha_1} + c_2 u_{\alpha_2} + \cdots + c_n u_{\alpha_n} = 0, \tag{2.74}$$

then

$$\begin{gathered} \left\langle \textstyle\sum_{i=1}^{n} c_i u_{\alpha_i}, v_{\alpha_j} \right\rangle = \langle 0, v_{\alpha_j}\rangle = 0 \Rightarrow \\ \textstyle\sum_{i=1}^{n} c_i \langle u_{\alpha_i}, v_{\alpha_j}\rangle = \sum_{i=1}^{n} c_i \delta_{ij} = c_j = 0 \quad (j = 1, 2, \cdots, n). \end{gathered} \tag{2.75}$$

Hence the set $\{u_\alpha\}_{\alpha\in\mathbb{N}}$ is linearly independent. Linear independence of the set $\{v_\alpha\}$ can be shown similarly. □

Theorem 2.2.3. *[19] Let $\mathcal{H}$ be an n-dimensional Hilbert space over a field K. For any linearly independent set of elements $u_1, u_2, \cdots, u_n$ in $\mathcal{H}$, we can always find a set of elements $v_1, v_2, \cdots, v_n$ such that these two sets constitute a biorthonormal system $\{u_\alpha, v_\alpha\}_{\alpha=1}^{n}$ in $\mathcal{H}$.*

Proof. First we orthogonalize the $\{u_\alpha\}_{\alpha=1}^{n}$ with the Gram-Schmidt orthogonalization procedure[22], and construct an orthonormal basis $\{\Phi_\alpha\}_{\alpha=1}^{n}$ for $\mathcal{H}$;

$$\begin{aligned}
\Phi_1' &= u_1 && \text{and} \quad \Phi_1 = \Phi_1'/\|\Phi_1'\| \\
\Phi_2' &= u_2 - \langle u_2, \Phi_1\rangle\Phi_1 && \text{and} \quad \Phi_2 = \Phi_2'/\|\Phi_2'\| \\
\Phi_3' &= u_3 - \langle u_3, \Phi_1\rangle\Phi_1 - \langle u_3, \Phi_2\rangle\Phi_2 && \text{and} \quad \Phi_3 = \Phi_3'/\|\Phi_3'\| \\
&\vdots && \qquad\quad \vdots \\
\Phi_n' &= u_n - \textstyle\sum_{\alpha=1}^{n-1}\langle u_n, \Phi_\alpha\rangle\Phi_\alpha && \text{and} \quad \Phi_n = \Phi_n'/\|\Phi_n'\|.
\end{aligned} \tag{2.76}$$

Then, $\langle \Phi_\alpha, \Phi_\beta\rangle = \delta_{\alpha\beta}$ by construction, and we write,

$$\begin{aligned}
u_1 &= \|u_1\|\ \Phi_1 \\
u_2 &= \langle u_2, \Phi_1\rangle\Phi_1 + \|u_2 - \langle u_2, \Phi_1\rangle\Phi_1\|\ \Phi_2 \\
u_3 &= \langle u_3, \Phi_1\rangle\Phi_1 + \langle u_3, \Phi_2\rangle\Phi_2 + \|u_3 - \langle u_3, \Phi_1\rangle\Phi_1 - \langle u_3, \Phi_2\rangle\Phi_2\|\ \Phi_3 \\
&\vdots \\
u_n &= \textstyle\sum_{\alpha=1}^{n-1}\langle u_n, \Phi_\alpha\rangle\Phi_\alpha + \|u_n - \sum_{\alpha=1}^{n-1}\langle u_n, \Phi_\alpha\rangle\Phi_\alpha\|\ \Phi_n.
\end{aligned} \tag{2.77}$$

That is,

$$u_\alpha = \sum_{\beta=1}^{\alpha} a_{\alpha\beta}\Phi_\beta. \tag{2.78}$$

Here, $a_{\alpha\alpha} = \|u_\alpha - \sum_{\beta=1}^{\alpha-1}\langle u_\alpha, \Phi_\beta\rangle\Phi_\beta\| \neq 0$ because $\{u_\alpha\}_{\alpha=1}^{n}$ is a linearly independent set. The determinant of the triangular matrix $(a_{\alpha\beta})$ is nonzero since the determinant is the product of $a_{\alpha\alpha}$'s. Hence $(a_{\alpha\beta})$ is invertible. Solving the equation (2.78) for the Φ_α, we get

$$\Phi_\alpha = \sum_{\gamma=1}^{\alpha} A_{\alpha\gamma}u_\gamma, \tag{2.79}$$

for some $A_{\alpha\gamma} \in K$. Observe that,

$$u_\lambda = \sum_{\alpha=1}^{\lambda} a_{\lambda\alpha}\Phi_\alpha = \sum_{\alpha=1}^{\lambda} a_{\lambda\alpha}\Big(\sum_{\gamma=1}^{\alpha} A_{\alpha\gamma}u_\gamma\Big) = \sum_{\alpha=1}^{\lambda}\sum_{\gamma=1}^{\alpha} a_{\lambda\alpha}A_{\alpha\gamma}u_\gamma. \tag{2.80}$$

Then, since $\{u_\alpha\}$ is linearly independent,

$$\text{if } \lambda \neq \gamma \text{ then } \sum_{\alpha=1}^{\lambda} a_{\lambda\alpha}A_{\alpha\gamma_0} = 0, \text{ and if } \lambda = \gamma \text{ then } \sum_{\alpha=1}^{\lambda} a_{\lambda\alpha}A_{\alpha\lambda} = 1. \tag{2.81}$$

Hence

$$\sum_{\alpha=1}^{\lambda} a_{\lambda\alpha} A_{\alpha\gamma} = \delta_{\lambda\gamma}. \tag{2.82}$$

Similarly,

$$\Phi_\alpha = \sum_{\gamma=1}^{\alpha} A_{\alpha\gamma} u_\gamma = \sum_{\gamma=1}^{\alpha} A_{\alpha\gamma} \Big(\sum_{\beta=1}^{\gamma} a_{\gamma\beta} \Phi_\beta \Big) = \sum_{\gamma=1}^{\alpha} \sum_{\beta=1}^{\gamma} A_{\alpha\gamma} a_{\gamma\beta} \Phi_\beta \tag{2.83}$$

and since $\{\Phi_\alpha\}$ is linearly independent,

$$\sum_{\gamma=1}^{\alpha} A_{\alpha\gamma} a_{\gamma\beta} = \delta_{\alpha\beta}. \tag{2.84}$$

Here $(a_{\alpha\beta})$ and $(A_{\alpha\beta})$ are "triangular" matrices and each is a left (and right) inverse of the other.

Now using (2.26) - (2.28) we write

$$\begin{aligned} v_\beta &= \textstyle\sum_{\alpha=1}^{n} \langle v_\beta, \Phi_\alpha \rangle \Phi_\alpha = \sum_{\alpha=1}^{n} \Big\langle v_\beta, \sum_{\gamma=1}^{\alpha} A_{\alpha\gamma} u_\gamma \Big\rangle \Phi_\alpha \\ &= \textstyle\sum_{\alpha=1}^{n} \sum_{\gamma=1}^{\alpha} A^*_{\alpha\gamma} \langle v_\beta, u_\gamma \rangle \Phi_\alpha = \sum_{\alpha=1}^{n} \sum_{\gamma=1}^{\alpha} A^*_{\alpha\gamma} \delta_{\beta\gamma} \Phi_\alpha \\ &= \textstyle\sum_{\alpha=1}^{n} A^*_{\alpha\beta} \Phi_\alpha. \end{aligned} \tag{2.85}$$

Therefore,

$$\begin{aligned} \langle u_\alpha, v_\beta \rangle &= \Big\langle \textstyle\sum_{\gamma=1}^{\alpha} a_{\alpha\gamma} \Phi_\gamma, \sum_{\lambda=1}^{n} A^*_{\lambda\beta} \Phi_\lambda \Big\rangle = \sum_{\gamma=1}^{\alpha} \sum_{\lambda=1}^{n} a_{\alpha\gamma} A_{\lambda\beta} \langle \Phi_\gamma, \Phi_\lambda \rangle \\ &= \textstyle\sum_{\gamma=1}^{\alpha} \sum_{\lambda=1}^{n} a_{\alpha\gamma} A_{\lambda\beta} \delta_{\gamma\lambda} = \sum_{\gamma=1}^{\alpha} a_{\alpha\gamma} A_{\gamma\beta} = \delta_{\alpha\beta}. \end{aligned} \tag{2.86}$$

for all $\alpha, \beta = 1, 2, \cdots, n$. □

Note that any vector x in $\mathcal{H}$ can be written as,

$$\begin{gathered} x = \textstyle\sum_{\alpha_1}^{n} c_\alpha u_\alpha \qquad c_\alpha : \text{constant} \\ \langle x, v_\beta \rangle = \langle \sum_{\alpha_1}^{n} c_\alpha u_\alpha, v_\beta \rangle = \sum_{\alpha_1}^{n} c_\alpha \langle u_\alpha, v_\beta \rangle = \sum_{\alpha_1}^{n} c_\alpha \delta_{\alpha\beta} = c_\beta, \end{gathered} \tag{2.87}$$

since the inner product is linear in its first entry. As a result,

$$x = \sum_{\alpha=1}^{n} \langle x, v_\alpha \rangle u_\alpha, \tag{2.88}$$

and similarly

$$x = \sum_{\alpha=1}^{n} \langle x, u_\alpha \rangle v_\alpha. \tag{2.89}$$

Theorem 2.2.4. *[19] Let $\{u_\alpha\}_{\alpha=1}^{\infty}$ be a countably infinite set of linearly independent elements in a separable Hilbert space $\mathcal{H}$. Then, we can find a set of elements $\{v_\alpha\}_{\alpha=1}^{\infty}$ such that these two sets constitute a biorthonormal system $\{u_\alpha, v_\alpha\}_{\alpha=1}^{\infty}$ in $\mathcal{H}$, if and only if the coefficients $A_{\alpha\gamma}$, which arise in the orthogonalization of the u_γ when we form $\Phi_\alpha = \sum_{\gamma=1}^{\alpha} A_{\alpha\gamma} u_\gamma$ satisfy*

$$\sum_{\alpha=1}^{\infty} |A_{\alpha\gamma}|^2 < \infty \quad \textit{for all} \quad \gamma = 1, 2, \cdots. \tag{2.90}$$

Proof. We start as the preceding proof in finite-dimension and orthogonalize the set $\{u_\alpha\}$.

$$u_\alpha = \sum_{\beta=1}^{\alpha} a_{\alpha\beta} \Phi_\beta \quad (\alpha = 1, 2, \cdots). \tag{2.91}$$

We construct the orthonormal set $\{\Phi_\alpha\}_{\alpha=1}^{\infty}$. Since $(a_{\alpha\beta})$ is a triangular matrix and $a_{\alpha\alpha}$ are nonzero, columns of the matrix $(a_{\alpha\beta})$ are linearly independent and so the kernel of the matrix as a linear operator is $ker(a_{\alpha\beta}) = \{0\}$. Hence by Theorem (2.1.36), $(a_{\alpha\beta})^{-1} = (A_{\alpha\beta})$ exists.

$$\Phi_\alpha = \sum_{\gamma=1}^{\alpha} A_{\alpha\gamma} u_\gamma \quad (\alpha = 1, 2, \cdots). \tag{2.92}$$

Then in view of Equation (2.85) we define

$$v_\beta := \sum_{\alpha=1}^{\infty} A^*_{\alpha\beta} \Phi_\alpha. \tag{2.93}$$

We claim that (2.90) is necessary and sufficient condition to construct the biorthonormal system $\{u_\alpha, v_\alpha\}$. Now calculate

$$\begin{aligned} \|v_\beta\| &= \left\langle \textstyle\sum_{\alpha=1}^{\infty} A^*_{\alpha\beta}\Phi_\alpha, \sum_{\lambda=1}^{\infty} A^*_{\lambda\beta}\Phi_\lambda \right\rangle = \textstyle\sum_{\alpha=1}^{\infty}\sum_{\lambda=1}^{\infty} A^*_{\alpha\beta} A_{\lambda\beta} \langle \Phi_\alpha, \Phi_\lambda \rangle \\ &= \textstyle\sum_{\alpha=1}^{\infty}\sum_{\lambda=1}^{\infty} A^*_{\alpha\beta} A_{\lambda\beta} \delta_{\alpha\lambda} = \sum_{\alpha=1}^{\infty} A^*_{\alpha\beta} A_{\alpha\beta} = \sum_{\alpha=1}^{\infty} |A_{\alpha\beta}|^2. \end{aligned} \tag{2.94}$$

For necessity; if biorthonormal system $\{u_\alpha, v_\alpha\}$ exists then $\|v_\beta\|$ should be finite so $\|v_\beta\| = \sum_{\alpha=1}^{\infty} |A_{\alpha\beta}|^2 < \infty$. (2.90) is sufficient because if (2.90) satisfied then $\|v_\beta\| < \infty$ and so $v_\beta \in \mathcal{H}$ and by construction $\{u_\alpha, v_\beta\}$ is a biorthonormal system in $\mathcal{H}$.

Hence $\{v_\beta\} \in \mathcal{H}$ and $\{u_\alpha, v_\beta\}$ is a biorthonormal system in $\mathcal{H}$ iff

$$\sum_{\alpha=1}^{\infty} |A_{\alpha\beta}|^2 < \infty \quad \text{for all} \quad \beta = 1, 2, \cdots. \tag{2.95}$$

We can show that

$$\begin{aligned}\langle u_\alpha, v_\beta\rangle &= \langle \textstyle\sum_{\lambda=1}^{\alpha} a_{\alpha\lambda}\Phi_\lambda, \sum_{\gamma=1}^{\infty} A^*_{\gamma\beta}\Phi_\gamma\rangle = \sum_{\lambda=1}^{\alpha}\sum_{\gamma=1}^{\infty} a_{\alpha\lambda}A_{\gamma\beta}\langle\Phi_\lambda,\Phi_\gamma\rangle \\ &= \textstyle\sum_{\lambda=1}^{\alpha}\sum_{\gamma=1}^{\infty} a_{\alpha\lambda}A_{\gamma\beta}\delta_{\lambda\gamma} = \sum_{\lambda=1}^{\alpha} a_{\alpha\lambda}A_{\lambda\beta} = \delta_{\alpha\beta}\end{aligned} \tag{2.96}$$

for all $\alpha, \beta = 1, 2, \cdots$. □

Definition 2.2.5. *(Hilbert Sequence Space $\mathcal{H}_0$)*

Let $\mathcal{H}_0$ be the set of sequences $x = (x_\alpha) = (x_1, x_2, \cdots)$ of scalars $x_\alpha \in K$ such that $\sum_{\alpha=1}^{\infty} |x_\alpha|^2 < \infty$. Then $\mathcal{H}_0$ with inner product defined by

$$\langle x, y\rangle = \sum_{\alpha=1}^{\infty} x_\alpha y_\alpha^* \tag{2.97}$$

as a separable Hilbert space. The norm is defined by

$$\|x\| = \langle x, x\rangle^{1/2} = \Big(\sum_{\alpha=1}^{\infty} |x_\alpha|^2\Big)^{1/2}. \tag{2.98}$$

Note that two algebraic operations in this linear space is given by

$$\begin{aligned}(x_1, x_2, \cdots) + (y_1, y_2, \cdots) &= (x_1 + y_1, x_2 + y_2, \cdots) \\ r(x_1, x_2, \cdots) &= (rx_1, rx_2, \cdots)\end{aligned} \tag{2.99}$$

where $x = (x_\alpha), y = (y_\alpha) \in \mathcal{H}_0$ and $r \in K$.

Definition 2.2.6. *[5](Bounded Infinite Matrix)*

An infinite matrix with elements $A_{\alpha\beta}$ is bounded if the sesquilinear form (see A.1.5) $\sum_{\alpha=1}^{\infty}\sum_{\beta=1}^{\infty} A_{\alpha\beta}x_\beta y_\alpha^*$ is bounded for all sequences of scalars (x_α) and (y_α) in $\mathcal{H}_0$; that is

$$\Big|\sum_{\alpha=1}^{\infty}\sum_{\beta=1}^{\infty} A_{\alpha\beta}x_\beta y_\alpha^*\Big| \leq \mathcal{C}\sqrt{\sum_{\alpha=1}^{\infty}|x_\alpha|^2\sum_{\alpha=1}^{\infty}|y_\alpha|^2} \tag{2.100}$$

for some constant $\mathcal{C}$.

Lemma 2.2.7. *[19] (Landau)*

Let $c = (c_1, c_2, \cdots)$ be any sequence of scalars $c_\alpha \in K$. If

$$\Big|\sum_{\alpha=1}^{\infty} c_\alpha a_\alpha^*\Big| \leq k\|a\| \tag{2.101}$$

for some $k \in \mathbb{R}$ and for all $a = (a_1, a_2, \cdots)$ in $\mathcal{H}_0$ then $c = (c_1, c_2, \cdots)$ is in $\mathcal{H}_0$.

Proposition 2.2.8. *[19] If $(A_{\alpha\beta})$ is a bounded matrix then (2.90) is satisfied.*

Proof. Let $(A_{\alpha\beta})$ be a bounded matrix, then

$$\Big|\sum_{\alpha=1}^{\infty}\sum_{\beta=1}^{\infty} A_{\alpha\beta}x_{\beta}y_{\alpha}^{*}\Big| \leq \mathcal{C}\sqrt{\sum_{\alpha=1}^{\infty}|x_{\alpha}|^2\sum_{\alpha=1}^{\infty}|y_{\alpha}|^2} \tag{2.102}$$

for some constant $\mathcal{C}$ and all $x = (x_\alpha),\ y = (y_\alpha) \in \mathcal{H}_0$. Now we choose any $x^{\beta_0} \in \mathcal{H}_0$ such that $(0, 0, \cdots, 0, 1, 0, 0, \cdots)$; all scalars x_β are zero except ${\beta_0}^{th}$ which is equal to 1. This shows that

$$\Big|\sum_{\alpha=1}^{\infty} A_{\alpha\beta_0}y_{\alpha}^{*}\Big| \leq \mathcal{C}\|y\| \tag{2.103}$$

for all $y \in \mathcal{H}_0$ and any fixed $\beta = \beta_0$. Then, by using Lemma (2.2.7) we have $(A_{\alpha\beta_0}) \in \mathcal{H}_0$, which means $\sum_{\alpha=1}^{\infty}|A_{\alpha\beta_0}|^2 < \infty$ for any fixed $\beta = \beta_0$. Hence

$$\sum_{\alpha=1}^{\infty}|A_{\alpha\beta}|^2 < \infty \quad \text{for all} \quad (\beta = 1, 2, \cdots). \tag{2.104}$$

□

Definition 2.2.9. *[19](Bounded Biorthonormal System)*
If both $(A_{\alpha\beta})$ and $(a_{\alpha\beta})$ are bounded. Then the biorthonormal system is said to be a bounded biorthonormal system.

Definition 2.2.10. *[19](Complete Biorthonormal System)*
A biorthonormal system $\{u_\alpha, v_\alpha\}$ is said to be complete if $\{u_\alpha\}$ and $\{v_\alpha\}$ are fundamental sets [9] in the separable Hilbert space $\mathcal{H}$.

Theorem 2.2.11. *[19](Bounded Complete Biorthonormal System)*
Any element x in a separable Hilbert space $\mathcal{H}$ can be represented by means of a bounded complete biorthonormal system $\{u_\alpha, v_\alpha\}_{\alpha=1}^{\infty}$ as follows:

$$x = \sum_{\alpha=1}^{\infty}\langle x, v_\alpha\rangle u_\alpha = \sum_{\alpha=1}^{\infty}\langle x, u_\alpha\rangle v_\alpha. \tag{2.105}$$

[9] That is, $\overline{span\{u_\alpha\}} = \overline{span\{v_\alpha\}} = \mathcal{H}$. See Definition (2.1.27)

Proof. Let $\{\Phi_\beta\}_{\beta=1}^{\infty}$ be an orthonormal basis of $\mathcal{H}$, such that

$$u_\alpha = \sum_{\lambda=1}^{\alpha} a_{\alpha\lambda}\Phi_\lambda \quad (\alpha = 1, 2, \cdots) \tag{2.106}$$

then we have

$$\Phi_\alpha = \sum_{\gamma=1}^{\alpha} A_{\alpha\gamma}u_\gamma \quad \text{and} \quad v_\beta = \sum_{\gamma=1}^{\infty} A^*_{\gamma\beta}\Phi_\gamma, \tag{2.107}$$

where $(a_{\alpha\beta})$ and $(A_{\alpha\beta})$ are as in the proof of Theorem (2.2.4). Write, (since $(a_{\alpha\beta})$ and $(A_{\alpha\beta})$ are bounded)

$$\sum_{\beta=1}^{\infty} a^*_{\beta\alpha}v_\beta = \sum_{\beta=1}^{\infty} a^*_{\beta\alpha}\Big(\sum_{\gamma=1}^{\infty} A^*_{\gamma\beta}\Phi_\gamma\Big) = \sum_{\gamma=1}^{\infty}\Big(\sum_{\beta=1}^{\infty} A^*_{\gamma\beta}a^*_{\beta\alpha}\Big)\Phi_\gamma = \sum_{\gamma=1}^{\infty}\delta_{\gamma\alpha}\Phi_\gamma = \Phi_\alpha. \tag{2.108}$$

Now using (2.32), for any $x \in \mathcal{H}$, we have

$$x = \sum_{\alpha=1}^{\infty}\langle x, \Phi_\alpha\rangle\Phi_\alpha = \sum_{\alpha=1}^{\infty}\Big\langle x, \sum_{\beta=1}^{\infty} a^*_{\beta\alpha}v_\beta\Big\rangle\Big(\sum_{\gamma=1}^{\infty} A_{\alpha\gamma}u_\gamma\Big) \tag{2.109}$$

where we used (2.107) and (2.108). Because for triangular matrix A we have: $\gamma > \alpha \Rightarrow A_{\alpha\gamma} = 0$, we can write $\sum_{\gamma=1}^{\alpha} A_{\alpha\gamma}u_\gamma = \sum_{\gamma=1}^{\infty}\delta_{1,sign(\alpha-\gamma+1)}A_{\alpha\gamma}u_\gamma$[10]. This together with (2.82) and (2.109) imply:

$$\begin{aligned} x &= \textstyle\sum_{\alpha=1}^{\infty}\Big\langle x, \sum_{\beta=1}^{\infty} a^*_{\beta\alpha}v_\beta\Big\rangle\Big(\sum_{\gamma=1}^{\infty}\delta_{1,sign(\alpha-\gamma+1)}A_{\alpha\gamma}u_\gamma\Big) \\ &= \textstyle\sum_{\beta=1}^{\infty}\sum_{\gamma=1}^{\infty}\sum_{\alpha=1}^{\infty} a_{\beta\alpha}A_{\alpha\gamma}\delta_{1,sign(\alpha-\gamma+1)}\langle x, v_\beta\rangle u_\gamma \\ &= \textstyle\sum_{\beta=1}^{\infty}\sum_{\gamma=1}^{\infty}\delta_{\beta\gamma}\langle x, v_\beta\rangle u_\gamma = \sum_{\beta=1}^{\infty}\langle x, v_\beta\rangle u_\beta. \end{aligned} \tag{2.110}$$

Hence, we can represent any $x \in \mathcal{H}$ as,

$$x = \sum_{\beta=1}^{\infty}\langle x, v_\beta\rangle u_\beta. \tag{2.111}$$

Equality $x = \sum_{\beta=1}^{\infty}\langle x, u_\beta\rangle v_\beta$ can be shown similarly. $\square$

---

[10] sign function is a function of $\mathbb{R}$, such that $sgn(x) = \begin{cases} -1, & \text{for } x < 0; \\ 0, & \text{for } x = 0; \\ 1, & \text{for } x > 0. \end{cases}$, hence here $\delta_{1,sign(\alpha-\gamma+1)} = \begin{cases} 0, & \text{for } \gamma > \alpha; \\ 1 & \text{for } \gamma \le \alpha. \end{cases}$

Chapter 3

QUANTUM MECHANICAL BACKGROUND

3.1 Introduction

Definition 3.1.1. *[3](Bra-ket Notation (Dirac Notation))*
In the Dirac bra-ket notation, a vector ψ *in the Hilbert Space* $\mathcal{H}$ *is written as* $|\psi\rangle$, *and said to be the ket* ψ. $\mathcal{H}$ *is called the ket space. Elements in the dual space* $\mathcal{H}'$ *of the Hilbert space* $\mathcal{H}$ *are written as* $\langle\phi|$ *and said to be the bra* ϕ. $\mathcal{H}'$ *is called the bra space. In the previous chapter we see that the dual space* $\mathcal{H}'$ *is isomorphic to* $\mathcal{H}$ *and this is the reason why we represent the elements in the* $\mathcal{H}'$ *with vectors in* $\mathcal{H}$. *The natural isomorphism from* $\mathcal{H}$ *onto* $\mathcal{H}'$; *or from* $\mathcal{H}'$ *onto* $\mathcal{H}$, *is called "taking the adjoint" by physicist[3]. That is,* $(|\psi\rangle)^\dagger = \langle\psi|$ *and* $(\langle\phi|)^\dagger = |\phi\rangle$. *The expression* $\langle\phi|A|\psi\rangle$ *can be thought as,* A *acting on* $|\psi\rangle$ *on the right as a matrix times a column vector, or* A *acting on* $\langle\phi|$ *on the left as a row vector times a matrix [3]. Here, the vector* ϕ, *is in the domain of the* $A^\dagger$, *since* $\Big(\langle\phi|A\ \Big)^\dagger = A^\dagger|\phi\rangle$. *Note that, the Hilbert-adjoint of a matrix is the complex conjugate transpose of that matrix, that is* $[(a_{ij})]^\dagger = ([(a_{ij})]^T)^* = (a_{ji})^*$*[9], and* $(AB)^\dagger = B^\dagger A^\dagger$, *where,* B *is another matrix or vector. An example can make it clear. Let* $\mathcal{H} = \mathbb{C}^2$, $A : \mathbb{C}^2 \to \mathbb{C}^2$ *be identified with its matrix representation*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{3.1}$$

in the standard basis[1] *of* $\mathbb{C}^2$*, and* $|\phi\rangle$ *be identified with its representation* $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ *in this basis. Then* $\langle\phi| = \begin{pmatrix} c_1^* & c_2^* \end{pmatrix}$ *and*

$$\langle\phi|A = \begin{pmatrix} c_1^* & c_2^* \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} c_1^* a_{11} + c_2^* a_{21} & c_1^* a_{12} + c_2^* a_{22} \end{pmatrix}$$

$$A^\dagger = (A^T)^* = \begin{pmatrix} a_{11}^* & a_{21}^* \\ a_{12}^* & a_{22}^* \end{pmatrix} \tag{3.2}$$

$$A^\dagger|\phi\rangle = \begin{pmatrix} a_{11}^* & a_{21}^* \\ a_{12}^* & a_{22}^* \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} c_1 a_{11}^* + c_2 a_{21}^* \\ c_1 a_{12}^* + c_2 a_{22}^* \end{pmatrix} = \Big(\langle\phi|A\Big)^\dagger$$

where T *denotes transpose, and* * *complex conjugate.*

The inner product in the bra-ket notation is given by

$$\langle\phi|\psi\rangle := \langle\psi, \phi\rangle. \tag{3.3}$$

Note that, $\langle\ \cdot \mid \cdot\ \rangle$ *is linear in the second entry, where* $\langle\ \cdot\ ,\ \cdot\ \rangle$ *is linear in the first entry. The inner product of two vectors in the mathematicians convention is,* $\langle\phi, \psi\rangle = \phi^T\psi^*$*, but in the bra-ket notation* $\langle\psi|\phi\rangle = \psi^\dagger\phi$*. If* $\phi = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ *and* $\psi = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ *in* $\mathbb{C}^2$*, then,* $\langle\phi, \psi\rangle = \begin{pmatrix} c_1 & c_2 \end{pmatrix} \begin{pmatrix} d_1^* \\ d_2^* \end{pmatrix} = c_1 d_1^* + c_2 d_2^*$ *and* $\langle\phi|\psi\rangle \begin{pmatrix} c_1^* & c_2^* \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = c_1^* d_1 + c_2^* d_2$*. Norm of any vector in* $\mathcal{H}$ *is,* $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$*.*

An operator A *and its Hilbert-adjoint* $A^\dagger$ *satisfy*

$$\langle\phi|A\psi\rangle = \langle A^\dagger\phi|\psi\rangle. \tag{3.4}$$

Theorem 3.1.2. *[17] If* A *is a Hermitian linear operator on a Hilbert space* $\mathcal{H}$*, then all its eigenvalues (if exist) are real, and eigenvectors corresponding to different eigenvalues are orthogonal.*

Proof. [17] Let $|x\rangle \neq 0$ be an eigenvector of A and $A|x\rangle = \lambda|x\rangle$. Then

$$\lambda\|x\|^2 = \lambda\langle x|x\rangle = \langle x|\lambda x\rangle = \langle x|Ax\rangle = \langle Ax|x\rangle = \langle\lambda x|x\rangle = \lambda^*\langle x|x\rangle = \lambda^*\|x\|^2. \tag{3.5}$$

[1] $\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$.

Since $|x\rangle \neq 0$, $\lambda = \lambda^*$, so λ is real. Now let $|y\rangle \neq 0$ be another eigenvector of A with a different eigenvalue and $A|y\rangle = \mu|y\rangle$. Then

$$\lambda\langle y|x\rangle = \langle y|\lambda x\rangle = \langle y|Ax\rangle = \langle Ay|x\rangle = \langle \mu y|x\rangle = \mu^*\langle y|x\rangle = \mu\langle y|x\rangle. \tag{3.6}$$

Now since $\lambda \neq \mu$ we have $\langle y|x\rangle = 0$ which shows orthogonality. □

Note that, the eigenvalues of an n-rowed square matrix $A = (\alpha_{jk})$ are given by the solutions of the characteristic equation,

$$\det(A - \lambda\mathbb{I}) = det\begin{pmatrix} \alpha_{11}-\lambda & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22}-\lambda & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn}-\lambda \end{pmatrix} = 0,$$ of A. Hence A has at least one eigenvalue (and at most n numerically different eigenvalues). All matrices representing a given linear operator $T : X \to X$ on a finite dimensional normed space X relative to various bases for X have the same eigenvalues. Hence a linear operator on a finite dimensional complex normed space X has at least one eigenvalue[9].

Moreover, let $T : \mathcal{H} \to \mathcal{H}$ be a Hermitian linear operator on a finite dimensional Hilbert space $\mathcal{H}$. Then there exist an orthonormal basis for $\mathcal{H}$ which is constructed by the eigenvectors of T[4]. If $T : \mathcal{D}(T) \to \mathcal{H}$ be a densely defined Hermitian linear operator with a discrete spectrum, i.e., spectrum consists of eigenvalues of H, on a separable Hilbert space $\mathcal{H}$, then $\mathcal{H}$ has an orthonormal basis consisting of eigenvectors for T.

If there are two or more linearly independent eigenvectors for an eigenvalue, then this eigenvalue is said to be degenerate. In the following, we assume that there is no degeneracy for simplicity. The operators appearing in the examples we shall study are non-degenerate.

Definition 3.1.3. *[1](Eigenvector Expansion, Spectral Resolution of Identity Operator, Spectral Resolution of a Hermitian Operator)*
Let $A : \mathcal{H} \to \mathcal{H}$ *be a Hermitian linear operator on an* n*-dimensional Hilbert space* $\mathcal{H}$*, and so*

$$A|e_i\rangle = e_i|e_i\rangle \quad (i = 1, 2, \cdots, n). \tag{3.7}$$

Let the eigenvectors $\{|e_i\rangle\}$ *are normalized, that is* $\langle e_i|e_j\rangle = \delta_{ij}$. *Hence* $\{|e_i\rangle\}_{i=1}^{n}$ *constitute an orthonormal basis for* $\mathcal{H}$. *We can write any vector* $|x\rangle \in \mathcal{H}$ *such that* $|x\rangle = c_1|e_1\rangle + c_2|e_2\rangle + \cdots + c_n|e_n\rangle$, *then,* $\langle e_i|x\rangle = \langle e_i|c_1e_1\rangle + \langle e_i|c_2e_2\rangle + \cdots + \langle e_i|c_ne_n\rangle = c_1\langle e_i|e_1\rangle + c_2\langle e_i|e_2\rangle + \cdots + c_n\langle e_i|e_n\rangle = c_i$, *since the inner product, in this notation, is linear in the second entry. Then* $|x\rangle = \sum_{i=1}^{n}\langle e_i|x\rangle\ |e_i\rangle$ *or*

$$|x\rangle = \sum_{i=1}^{n} |e_i\rangle\langle e_i|x\rangle. \tag{3.8}$$

Equation (3.8) is called the eigenvector expansion of $|x\rangle$.

Now let $A : \mathcal{D}(A) \to \mathcal{H}$ *be a densely defined Hermitian linear operator on a separable Hilbert space* $\mathcal{H}$ *with a discrete spectrum. Hence the eigenvectors* $\{|e_i\rangle\}$ *of* A *constitute an orthonormal basis for* $\mathcal{H}$. *Then the eigenvector expansion of any* $|x\rangle \in \mathcal{H}$ *is*

$$|x\rangle = \sum_{i=1}^{\infty} |e_i\rangle\langle e_i|x\rangle. \tag{3.9}$$

If we omit the arbitrary vector $|x\rangle$, *in (3.9), then we obtain the identity operator*

$$\mathbb{I} = \sum_{i=1}^{\infty} |e_i\rangle\langle e_i|, \tag{3.10}$$

which is called the spectral resolution of identity.

We can apply A *to any* $|x\rangle \in \mathcal{D}(A)$ *as in (3.9). This yields*

$$A|x\rangle = \sum_{i=1}^{\infty} A|e_i\rangle\langle e_i|x\rangle = \sum_{i=1}^{\infty} e_i|e_i\rangle\langle e_i|x\rangle. \tag{3.11}$$

If the arbitrary $|x\rangle$ *is omitted, then the spectral resolution of the Hermitian operator* A, *having discrete spectrum, is obtained. That is,*

$$A = \sum_{i=1}^{\infty} e_i|e_i\rangle\langle e_i|. \tag{3.12}$$

Definition 3.1.4. *[8](Projection Operator)*
A bounded operator P *is called a projection operator, if* $P^2 = P$. *If, in addition,* P *is Hermitian* $P^\dagger = P$, *then* P *is called an orthogonal projection.*

In general, if $|\psi\rangle$ *is any normalized vector in the Hilbert space* $\mathcal{H}$, *then* $|\psi\rangle\langle\psi|$ *defines an orthogonal projection operator on* $\mathcal{H}$, *since* $|\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi|$, *and*

$(|\psi\rangle\langle\psi|)^\dagger = |\psi\rangle\langle\psi|$. *The operator* $\wedge = |\psi\rangle\langle\psi|$ *projects to the one dimensional subspace spanned by* $|\psi\rangle$.

In particular, take the projection operator $\wedge_j = |e_j\rangle\langle e_j|$, *where* $|e_j\rangle$ *is one of the eigenvectors of A given in (3.7) and apply this operator to (3.8), then*

$$|e_j\rangle\langle e_j|x\rangle = \sum_{i=1}^{n} |e_j\rangle\langle e_j|e_i\rangle\langle e_i|x\rangle = \sum_{i=1}^{n} |e_j\rangle\delta_{ij}\langle e_i|x\rangle = |e_j\rangle\langle e_j|x\rangle = c_j|e_j\rangle. \tag{3.13}$$

The projection of $|x\rangle = c_1|e_1\rangle + c_2|e_2\rangle + \cdots + c_n|e_n\rangle$ *to the one-dimensional subspace spanned by* $|e_j\rangle$ *is the vector* $c_j|e_j\rangle$,

$$|e_j\rangle\langle e_j|x\rangle = c_j|e_j\rangle. \tag{3.14}$$

Definition 3.1.5. *[1](Trace of an Operator)*
Let $\{|e_i\rangle\}$ *be an orthonormal basis for the separable Hilbert space* $\mathcal{H}$, *then the trace of the operator* $A : \mathcal{D}(A) \to \mathcal{H}$, *with* $\mathcal{D}(A) \subseteq \mathcal{H}$, *is defined by*

$$Tr(A) = \sum_{i=1}^{\infty} \langle e_i|A|e_i\rangle. \tag{3.15}$$

$Tr(A)$ *is independent of the orthonormal basis chosen.* $Tr(A)$ *need not be finite. Some properties of trace are*

$$\begin{aligned} &i.\ Tr(AB) = Tr(BA), \quad \text{in general; } Tr(A_1A_2\cdots A_n) = Tr(A_2A_3\cdots A_nA_1) \\ &ii.\ Tr((A+B)C) = Tr(AC) + Tr(BC) \end{aligned} \tag{3.16}$$

Note that the trace of a square matrix $M = (m_{ij})$ *is the sum of its diagonal elements. That is* $Tr(M) = \sum_i m_{ii}$.

Definition 3.1.6. *[1](Commutator)*
Let X be a vector space and $A : X \to X$, *with* $\mathcal{D}(A) = X$, *and* $B : X \to X$, *with* $\mathcal{D}(B) = X$, *are two linear operators on X. A and B are said to commute if*

$$[A, B] := AB - BA = 0. \tag{3.17}$$

$[A, B]$ *is called the commutator of A and B.*

Mathematical Representation of a Quantum Mechanical System[1]: We will start with an explanation of a quantum mechanics experiment[2]. The purpose of this explanation is to show how a quantum mechanical system works. In this experiment an ensemble of CO molecules which are quantum mechanical oscillators is used . First let us think of a classical harmonic oscillator. Assume that we have two masses connected to each other by a spring. If one of the masses is excited by a projectile, then the system would start a vibration. The amplitude of the vibration depends on the energy transferred by the projectile. This classical mechanical system can have any energy value in $[0, \infty]$. In the corresponding quantum mechanical system the picture is different. A quantum mechanical oscillator, such as a CO molecule, can have just discrete energy values. Furthermore, we are not able to measure one single molecule, but only an ensemble of molecules in a microscopic system.

An electron beam in very narrow energy band is sent to the CO gas. After the collisions of electrons with the CO molecules new energies of electrons are measured. If an electron collides with a molecule elastically, it leaves the gas with the same energy. In contrast, an inelastic collision causes the electron to transfer an amount of its energy to the CO molecule, and then the CO molecule starts to oscillate with this energy[3]. It is observed that after the collision the energies of the electrons tend to take discrete values. Hence CO molecules can be in various discrete energy levels; some of them oscillate with energy E_1 while some of the others oscillate with energy E_2, and so on. It is observed that the difference between two consecutive energy levels are the same ($\approx 0.26eV$). In the experiment reported in Ref. [1], eight energy levels were detected. They are $E_0 = 0$ which is the ground state, $E_1 = 0.26eV$, $E_2 = 2E_1$, $E_3 = 3E_1 \cdots$. If all electrons had the same energy after the collision, which means that all CO molecules oscillate with the same energy, then it would be said that the ensemble of CO molecules is in a pure state. In this experiment we have eight energy levels, therefore this ensemble is in a mixture of states. The probability

[2]From G.J. Schulz, Phys. Rev. 135, (1964) as quoted [1].

[3]Multiple scattering of electrons and CO molecules is negligible because the intensity of the electron current and the density of CO molecules are sufficiently low[1].

w_n for any molecule to have the energy E_n is obtained by the measurement of the relative intensity of the electron current which has lost the energy E_n. That is, if it is possible to pick one molecule from the ensemble, the probability of picking the molecule that has energy E_n is w_n.

In quantum mechanics a measurement changes the state of the system. In this experiment, the electron beam prepares and measures the state. The preparation of the state consists of setting up the apparatus and sending the electron beam in very narrow energy band to the CO gas in a definite density. All experiments prepared under the same conditions gives the same result.

In a quantum mechanical system, a pure state is represented by a vector in a separable (infinite or finite-dimensional) Hilbert space $\mathcal{H}$.[4] An observable, like energy, is represented by a Hermitian linear operator A acting in the $\mathcal{H}$. Eigenvalues $\{\lambda_i\}$ of the observable A are values which are obtained in the measurement of A. An eigenvector $|\lambda_i\rangle$, labelled with its eigenvalue, represents the state after the measurement of A is performed and the value λ_i is obtained.

Let a Hermitian linear operator $H : \mathcal{D}(H) \to \mathcal{H}$ with a dense domain $\mathcal{D}(H) \subseteq \mathcal{H}$ be the energy observable of a quantum mechanical oscillator such that H has a discrete spectrum. Then there is a set of eigenvectors of H that constitute an orthonormal basis in the Hilbert space $\mathcal{H}$. Eigenvalues of the Hermitian operator H are real and these values are obtained in the measurement of the energy. Let $E_0, E_1, \cdots$ be the eigenvalues of H which is assumed to be non-degenerate and $|E_0\rangle, |E_1\rangle \cdots$ are a set of the corresponding eigenvectors. Let $w_0, w_1, \cdots$ be the probability for any oscillator to be in the states $|E_0\rangle, |E_1\rangle, \cdots$, respectively. Clearly,

$$\sum_i w_i = 1. \tag{3.18}$$

The state of the system is represented by a positive Hermitian operator

$$W = \sum_i w_i |E_i\rangle\langle E_i| = \sum_i w_i \wedge_i, \tag{3.19}$$

[4]A pure state is a 1-dimensional subspace defined by a nonzero vector $|\psi\rangle \in \mathcal{H}$. $|\psi\rangle$ is called a state vector. Often one abuses the terminology and calls $|\psi\rangle$ a state.

which is called a statistical operator. Note that trace of the statistical operator is given by

$$\begin{aligned} Tr(W) &= \sum_j \left\langle E_j \middle| \left(\sum_i w_i |E_i\rangle\langle E_i| \right) \middle| E_j \right\rangle = \sum_j \sum_i w_i \langle E_j|E_i\rangle\langle E_i|E_j\rangle \\ &= \sum_j \sum_i w_i \langle E_j|E_i\rangle \delta_{ij} = \sum_j w_j \langle E_j|E_j\rangle = \sum_j w_j = 1. \end{aligned} \tag{3.20}$$

If the system is in a pure state, the statistical operator of the system is a projection operator onto a one-dimensional subspace. Assume that, the system is in the pure state $|E_k\rangle$. Then the statistical operator is

$$W = |E_k\rangle\langle E_k| = \wedge_k. \tag{3.21}$$

We write H using its spectral resolution,

$$H = \sum_i E_i |E_i\rangle\langle E_i| = \sum_i E_i \wedge_i . \tag{3.22}$$

The expectation value of H, that is the average energy value of the ensemble of oscillators, in the state W is

$$\langle H \rangle := Tr(HW) = \sum_i \langle E_i|HW|E_i\rangle = \sum_i w_i E_i. \tag{3.23}$$

The last equality follows from

$$\begin{aligned} Tr(HW) &= \sum_i \langle E_i|HW|E_i\rangle \\ &= \sum_i \left\langle E_i \middle| \left(\sum_j E_j |E_j\rangle\langle E_j| \right) \left(\sum_k w_k |E_k\rangle\langle E_k| \right) \middle| E_i \right\rangle \\ &= \sum_i \left\langle E_i \middle| \left(\sum_j \sum_k E_j w_k |E_j\rangle\langle E_j|E_k\rangle\langle E_k| \right) \middle| E_i \right\rangle \\ &= \sum_i \left\langle E_i \middle| \left(\sum_j \sum_k E_j w_k |E_j\rangle \delta_{jk} \langle E_k| \right) \middle| E_i \right\rangle \\ &= \sum_i \left\langle E_i \middle| \left(\sum_j E_j w_j |E_j\rangle\langle E_j| \right) \middle| E_i \right\rangle \\ &= \sum_i \sum_j E_j w_j \langle E_i|E_j\rangle\langle E_j|E_i\rangle \\ &= \sum_i \sum_j E_j w_j \langle E_i|E_j\rangle \delta_{ij} \\ &= \sum_i E_i w_i \langle E_i|E_i\rangle \\ &= \sum_i E_i w_i, \end{aligned} \tag{3.24}$$

where we used the fact that the inner product in this notation is linear in the second entry and the eigenvectors $\{|E_i\rangle\}$ of the Hermitian operator H form an orthonormal basis of $\mathcal{H}$.

Assume that a quantum mechanical system is in the pure state $W = \wedge_k = |E_k\rangle\langle E_k|$, then the expectation value of H is given by

$$
\begin{aligned}
\langle H\rangle &= Tr(HW) = Tr(H\wedge_k) = \textstyle\sum_i \langle E_i|H \wedge_k |E_i\rangle \\
&= \textstyle\sum_i \Big\langle E_i \Big| \Big(\sum_j E_j |E_j\rangle\langle E_j|E_k\rangle\langle E_k| \Big) \Big| E_i \Big\rangle \\
&= \textstyle\sum_i \Big\langle E_i \Big| \Big(\sum_j E_j |E_j\rangle \delta_{jk} \langle E_k| \Big) \Big| E_i \Big\rangle = \sum_i \langle E_i|(E_k|E_k\rangle\langle E_k|)|E_i\rangle \\
&= \textstyle\sum_i E_k \langle E_i|E_k\rangle\langle E_k|E_i\rangle = \sum_i E_k \langle E_i|E_k\rangle \delta_{ik} = E_k\langle E_k|E_k\rangle = E_k.
\end{aligned}
\tag{3.25}
$$

The projection operator $\wedge_k = |E_k\rangle\langle E_k|$ is also an observable, and the expectation value of it, $\langle \wedge_k\rangle = Tr(\wedge_k W)$, gives the probability for an oscillator in the ensemble to be in the state $|E_k\rangle$. That is,

$$
\langle \wedge_k\rangle = Tr(\wedge_k W) = \sum_i \Big\langle E_i \Big| \Big(|E_k\rangle\langle E_k| \Big) \Big(\sum_j w_j |E_j\rangle\langle E_j| \Big) \Big| E_i \Big\rangle = w_k. \tag{3.26}
$$

As mentioned above, in a quantum mechanical system measurement changes the state. The measurement process both prepares and measures the state. Let the statistical operator W describes the state of the system after a measurement is performed. Experimental results of an immediate repetition of the measurement before the state W has a chance to develop by time or to change by someway will agree with the calculated value by W. State W can change by measurement of another observable. We can not measure all the observables simultaneously. Hence by measurement of another observable, W may change and turn into another state W'.

Let a system be in the state represented by the statistical operator W. If another observable B is measured, where $|b_1\rangle, |b_2\rangle, \ldots$ are eigenvectors of this observable labelled by the eigenvalues of B, then the new state of the system is

$$
W' = \sum_i |b_i\rangle\langle b_i|W|b_i\rangle\langle b_i| = \sum_i \wedge_{b_i} W \wedge_{b_i} . \tag{3.27}
$$

If the measurement yield (eigenvalue) b for the value of B, then the system would be in the state,

$$
W'' = \frac{1}{Tr(\wedge_b W \wedge_b)} \wedge_b W \wedge_b . \tag{3.28}
$$

As we mentioned above, a pure state is represented by a nonzero vector or one-dimensional subspace of the Hilbert space $\mathcal{H}$. A system in a pure state $W = |\psi\rangle\langle\psi| = \wedge_\psi$ would yield a set of different numbers in the measurement of the observable B if this pure state is not an eigenvector of B.

If two observables $A = \sum_i a_i |a_i\rangle\langle a_i| = \sum_i a_i \wedge_{a_i}$ and $B = \sum_i b_i |b_i\rangle\langle b_i| = \sum_i b_i \wedge_{b_i}$ commute, then the projection operators $\wedge_{a_i}$ and $\wedge_{b_i}$ which appear on the spectral resolution of A and B, respectively, also commute. In the same way $[A, \wedge_{b_i}] = 0$ and $[B, \wedge_{a_i}] = 0$. In light of this and previous paragraph, if B and A are the commuting observables, then a measurement of B on the system in the state $W = \sum_i w_{a_i} \wedge_{a_i}$ does not alter the state W.

$$W' = \sum_i \wedge_{b_i} W \wedge_{b_i} = \sum_i \wedge_{b_i} \wedge_{b_i} W = \sum_i \wedge_{b_i} W = \sum_i |b_i\rangle\langle b_i| W = \mathbb{I}\, W = W. \quad (3.29)$$

Two such observables are called compatible. Compatible observables have common eigenvectors [1].

Let A be an observable, and $|a_i\rangle$ be any normalized eigenvector of A labelled with its eigenvalue. Then the one dimensional subspace, $h_{a_i} = |a_i\rangle\langle a_i|\mathcal{H} = \wedge_{a_i}\mathcal{H} = \{\wedge_{a_i} x \mid x \in \mathcal{H}\}$, contains many normalized vectors. They are obtained by multiplying the normalized vector $|a_i\rangle$ with a phase factor $e^{i\alpha}$ where $0 \le \alpha < 2\pi$.

We write the energy observable H, namely the Hamiltonian, using its spectral resolution. $H = \sum_i E_i |E_i\rangle\langle E_i|$, where $\{|E_i\rangle\}$ are a set of eigenvectors of H that constitute an orthonormal basis for the Hilbert space $\mathcal{H}$. Clearly $\langle E_i|H|E_j\rangle = E_i \delta_{ij}$ which means H is diagonalized in the basis $\{|E_i\rangle\}$.

We will give an example. Let $A : \mathbb{C}^2 \to \mathbb{C}^2$ be a matrix on the Hilbert space $\mathbb{C}^2$ such that,

$$A = \begin{pmatrix} 1 & -2i \\ 2i & 1 \end{pmatrix} \quad \text{and} \quad A^\dagger = [A^T]^* = \begin{pmatrix} 1 & 2i \\ -2i & 1 \end{pmatrix}^* = \begin{pmatrix} 1 & -2i \\ 2i & 1 \end{pmatrix} = A. \quad (3.30)$$

Since A is Hermitian, its eigenvalues are real and eigenvectors are orthogonal to each other. We will find these eigenvectors and diagonalize the matrix A. Let λ be a eigenvalue of A, then by Definition (2.1.43), $(A-\lambda\mathbb{I})^{-1}$ does not exist, (where $\mathbb{I}$ is identity matrix). Hence, as we know from the linear algebra, the determinant of the matrix $(A-\lambda\mathbb{I})$ is zero, that is $det(A-\lambda\mathbb{I})=0$. Then, by Definition (2.1.43), the vector $x \in ker(A-\lambda\mathbb{I})$ is the eigenvector of A for the eigenvalue λ.

$$\begin{gathered} det\begin{pmatrix} 1-\lambda & -2i \\ 2i & 1-\lambda \end{pmatrix} = 0 \Rightarrow (1-\lambda)^2 - 4 = 0 \Rightarrow \lambda^2 - 2\lambda - 3 = 0 \\ \Rightarrow \lambda_1 = 3 \text{ and } \lambda_2 = -1, \end{gathered} \tag{3.31}$$

$$\begin{aligned} &\text{for } \lambda_1 = 3, \quad (A-3\mathbb{I})\, x_1 = 0 \Rightarrow \begin{pmatrix} -2 & -2i \\ 2i & -2 \end{pmatrix} x_1 = 0 \Rightarrow x_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix} \\ &\text{for } \lambda_2 = -1, \quad (A+\mathbb{I})\, x_2 = 0 \Rightarrow \begin{pmatrix} 2 & -2i \\ 2i & 2 \end{pmatrix} x_2 = 0 \Rightarrow x_2 = \begin{pmatrix} \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}. \end{aligned} \tag{3.32}$$

Note that x_1 and x_2 are normalized; $\|x_1\| = \|x_2\| = 1$ and $\langle x_1|x_2\rangle = 0$. Hence, $\{|x_i\rangle\}$ is an orthonormal basis for $\mathbb{C}^2$. Now write A using its spectral resolution

$$A = \sum_i \lambda_i |x_i\rangle\langle x_i|. \tag{3.33}$$

That yields

$$A = 3\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{pmatrix} - 1\begin{pmatrix} \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & -2i \\ 2i & 1 \end{pmatrix} \tag{3.34}$$

We can obtain the eigenvalues of A as the following,

$$\langle x_1|A|x_1\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & -2i \\ 2i & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix} = 3 = \lambda_1; \tag{3.35}$$

$$\langle x_2|A|x_2\rangle = \begin{pmatrix} \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & -2i \\ 2i & 1 \end{pmatrix} \begin{pmatrix} \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = -1 = \lambda_2. \tag{3.36}$$

A can be diagonalized as:

$$\begin{pmatrix} \langle x_1| \\ \langle x_2| \end{pmatrix} A \begin{pmatrix} |x_1\rangle & |x_2\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & -2i \\ 2i & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{3.37}$$

$$= \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = A_D. \tag{3.38}$$

Note that, if $T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$, then $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = T^{-1}$, so $T^{-1}AT = A_D$.

Note also that, if $A = (a_{ij})$, then $Tr(A) = \sum_i a_{ii} = 1 + 1 = Tr(A_D) = \lambda_1 + \lambda_2 = 3 + (-1) = 2$. Recall that the trace of a matrix is an invariant of the matrix, and equals to the sum of its eigenvalues. This comes from the property (3.16i) of the trace, that is $Tr(A_D) = Tr(T^{-1}AT) = Tr(ATT^{-1}) = Tr(A\mathbb{I}) = Tr(A)$.

Before we study the next example, let $H : \mathcal{H} \to \mathcal{H}$ be a non-Hermitian operator with a discrete spectrum on a separable Hilbert space $\mathcal{H}$. Let $|\xi\rangle$ be an eigenvector of H with an eigenvalue ξ, $H|\xi\rangle = \xi|\xi\rangle$, then $\langle\xi|H^\dagger = \xi^*\langle\xi|$. Now let $|a\rangle$ be an eigenvector of $H^\dagger$ with eigenvalue a, $H^\dagger|a\rangle = a|a\rangle$. Then $\langle\xi|H^\dagger|a\rangle = \xi^*\langle\xi|a\rangle \Rightarrow a\langle\xi|a\rangle = \xi^*\langle\xi|a\rangle$. Hence if vectors $|\xi\rangle, |a\rangle$ are not orthogonal to each other, i.e., $\langle\xi|a\rangle \neq 0$ then $a = \xi^*$. We express this fact as $H^\dagger|\xi^*\rangle = \xi^*|\xi^*\rangle$ and $\langle\xi|\xi^*\rangle \neq 0$. Let $|\mu\rangle$ be another eigenvector of H with an eigenvalue $\mu \neq \xi$. Using the equation $\langle\xi^*|H = \xi\langle\xi^*|$ we have $\langle\xi^*|H|\mu\rangle = \xi\langle\xi^*|\mu\rangle \Rightarrow \mu\langle\xi^*|\mu\rangle = \xi\langle\xi^*|\mu\rangle \Rightarrow (\mu - \xi)\langle\xi^*|\mu\rangle = 0$. Since $\mu \neq \xi$, $\langle\xi^*|\mu\rangle = 0$ which shows orthogonality.

We now will study a non-Hermitian matrix, as an example. This example is also an application of a biorthonormal system, discussed in the Section (2.2), in the bra-ket notation and finite-dimension. Let B be a non-Hermitian matrix such that, eigenvectors $|\psi_i\rangle$ of B and eigenvectors $|\phi_i\rangle$ of $B^\dagger$ constitute bases $\{|\psi_i\rangle\}$ and $\{|\phi_i\rangle\}$ of $\mathbb{C}^2$ respectively. Here $\{|\psi_i\rangle\}$ and $\{|\phi_i\rangle\}$ are not necessarily orthogonal bases. Then we will see that $\{|\psi_i\rangle, |\phi_i\rangle\}$ constitute a (bounded) complete biorthonormal system for $\mathbb{C}^2$. We will write B in its spectral resolution and diagonalize it. For example, let

$$B = \begin{pmatrix} i & 1 \\ 0 & 2 \end{pmatrix} \quad \text{with eigenvalues} \quad \lambda_1 = i, \qquad \lambda_2 = 2, \tag{3.39}$$

and

$$B^\dagger = \begin{pmatrix} -i & 0 \\ 1 & 2 \end{pmatrix} \quad \text{with eigenvalues} \quad \lambda_1^* = -i, \qquad \lambda_2^* = 2. \tag{3.40}$$

We can choose for the eigenvectors:

$$\begin{aligned} |\psi_1\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |\psi_2\rangle = \begin{pmatrix} 1 \\ 2-i \end{pmatrix} \\ |\phi_1\rangle &= \begin{pmatrix} 1 \\ -\frac{1}{2+i} \end{pmatrix}, \quad |\phi_2\rangle = \begin{pmatrix} 0 \\ \frac{1}{2+i} \end{pmatrix}. \end{aligned} \tag{3.41}$$

Note that $\langle\phi_1|\psi_2\rangle = \langle\phi_2|\psi_1\rangle = 0$ and $\langle\phi_1|\psi_1\rangle = \langle\phi_2|\psi_2\rangle = 1$. Hence, $\{|\psi_i\rangle, |\phi_i\rangle\}$ constitute a (bounded) complete biorthonormal system on $\mathbb{C}^2$. Since $\{|\psi_i\rangle\}_{i=1}^2$ is a basis for $\mathbb{C}^2$, for any $x \in \mathbb{C}^2$,

$$\begin{gathered} x = c_1|\psi_1\rangle + c_2|\psi_2\rangle \qquad c_i : \text{constant} \\ \langle\phi_1|x\rangle = \langle\phi_1|c_1\psi_1\rangle + \langle\phi_1|c_2\psi_2\rangle = c_1, \quad \text{and similarly} \quad \langle\phi_2|x\rangle = c_2. \end{gathered} \tag{3.42}$$

Here we recall that, in the bra-ket notation, the inner product is linear in the second entry. Then,

$$x = \langle\phi_1|x\rangle|\psi_1\rangle + \langle\phi_2|x\rangle|\psi_2\rangle \quad \text{or} \quad x = |\psi_1\rangle\langle\phi_1|x\rangle + |\psi_2\rangle\langle\phi_2|x\rangle. \tag{3.43}$$

Hence, any $x \in \mathbb{C}^2$ can be written as,

$$x = \sum_i |\psi_i\rangle\langle\phi_i|x\rangle. \tag{3.44}$$

Similarly, we have

$$x = \sum_i |\phi_i\rangle\langle\psi_i|x\rangle. \tag{3.45}$$

If we omit arbitrary x in (3.44) and (3.45), then we get the spectral resolution of identity, that is,

$$\mathbb{I} = \sum_i |\psi_i\rangle\langle\phi_i| = \sum_i |\phi_i\rangle\langle\psi_i|, \tag{3.46}$$

Now apply B to this arbitrary x,

$$Bx = \sum_i B|\psi_i\rangle\langle\phi_i|x\rangle = \sum_i \lambda_i|\psi_i\rangle\langle\phi_i|x\rangle, \tag{3.47}$$

and omit x. We obtain the spectral resolution of B.

$$B = \sum_i \lambda_i |\psi_i\rangle\langle\phi_i|, \tag{3.48}$$

and this really satisfies,

$$\begin{pmatrix} i & 1 \\ 0 & 2 \end{pmatrix} = i \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2-i} \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 2-i \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2-i} \end{pmatrix}. \tag{3.49}$$

Eigenvalues of B can be obtained by,

$$\langle\phi_1|B|\psi_1\rangle = i \quad \text{and} \quad \langle\phi_2|B|\psi_2\rangle = 2, \tag{3.50}$$

that is,

$$\begin{gathered} \begin{pmatrix} 1 & -\frac{1}{2-i} \end{pmatrix} \begin{pmatrix} i & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2-i} \end{pmatrix} \begin{pmatrix} i \\ 0 \end{pmatrix} = i, \\ \begin{pmatrix} 0 & \frac{1}{2-i} \end{pmatrix} \begin{pmatrix} i & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2-i \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2-i} \end{pmatrix} \begin{pmatrix} 2 \\ 4-2i \end{pmatrix} = 2. \end{gathered} \tag{3.51}$$

Diagonalize the matrix B by,

$$\begin{pmatrix} \langle\phi_1| \\ \langle\phi_2| \end{pmatrix} B \begin{pmatrix} |\psi_1\rangle & |\psi_2\rangle \end{pmatrix} = B_D \tag{3.52}$$

that is,

$$\begin{pmatrix} 1 & -\frac{1}{2-i} \\ 0 & \frac{1}{2-i} \end{pmatrix} \begin{pmatrix} i & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2-i \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & 2 \end{pmatrix} = B_D. \tag{3.53}$$

Note that letting $T := \begin{pmatrix} 1 & 1 \\ 0 & 2-i \end{pmatrix}$, we have $\begin{pmatrix} 1 & -\frac{1}{2-i} \\ 0 & \frac{1}{2-i} \end{pmatrix} = T^{-1}$, and $B_D = T^{-1}BT$.

$Tr(B) = Tr(B_D) = \lambda_1 + \lambda_2$, because of the property (3.16i) of trace.

We can write $B^\dagger$ in the spectral resolution of it, by using (3.48),

$$B^\dagger = \sum_i \lambda_i^* |\phi_i\rangle\langle\psi_i|. \tag{3.54}$$

Finally, assume that energy observable H, with $\mathcal{D}(H) \subseteq \mathcal{H}$ dense in $\mathcal{H}$, is a non-Hermitian operator on $\mathcal{H}$ with a discrete spectrum and eigenvectors $\{|\psi_i\rangle\}$ of H and eigenvectors $\{|\phi_i\rangle\}$ of $H^\dagger$ together constitute a bounded complete biorhonormal system on $\mathcal{H}$. Then H is said to be diagonalizable. Now let eigenvalues $\{E_i\}$ of H be real. Then, the spectral resolution of H is, $H = \sum_{i=1}^{\infty} E_i|\psi_i\rangle\langle\phi_i|$, and that of $H^\dagger$ is, $H^\dagger = \sum_{i=1}^{\infty} E_i|\phi_i\rangle\langle\psi_i|$. Possible energy values can be obtained by $\langle\phi_i|H|\psi_i\rangle = E_i$. Resolution of identity is $\mathbb{I} = \sum_i |\psi_i\rangle\langle\phi_i| = \sum_i |\phi_i\rangle\langle\psi_i|$.

Note that here $|\psi_i\rangle\langle\phi_i|$ and $|\phi_i\rangle\langle\psi_i|$ are projection operators, since $|\psi_i\rangle\langle\phi_i|\psi_i\rangle\langle\phi_i| = |\psi_i\rangle\langle\phi_i|$ and $|\phi_i\rangle\langle\psi_i|\phi_i\rangle\langle\psi_i| = |\phi_i\rangle\langle\psi_i|$.

3.2 *Pseudo-Hermiticity*

Definition 3.2.1. *[11] Let $A : \mathcal{D}(A) \to \mathcal{H}$, with $\mathcal{D}(A) \subseteq \mathcal{H}$, be a densely defined linear operator in the separable Hilbert Space $\mathcal{H}$. If there exist a bounded, Hermitian and invertible linear operator $\eta : \mathcal{H} \to \mathcal{H}$, defined on all of $\mathcal{H}$, and satisfying*

$$A^\dagger = \eta A \eta^{-1}, \tag{3.55}$$

then A is said to be pseudo-Hermitian. In particular, given such an operator η, A is said to be η-pseudo-Hermitian. Note that, if $\eta = \mathbb{I}$ satisfies (3.55), then A will be a Hermitian operator.

Definition 3.2.2. *[11] Let η be a linear operator as above. Then we can define a modified inner product according to*

$$\langle\langle\xi|\rho\rangle\rangle_\eta = \langle\xi|\eta|\rho\rangle \qquad \forall|\xi\rangle, |\rho\rangle \in \mathcal{H}. \tag{3.56}$$

Note that $\langle\langle\ .\ |\ .\ \rangle\rangle$ is not a genuine inner product, because it may be indefinite; we may obtain imaginary numbers as norms by an indefinite inner product.

There are important results of these definitions of pseudo-Hermiticity and modified inner product. Some of the results will be explained below. Others can be found in References [11][12][13][14].

Proposition 3.2.3. *[11] Let $A : \mathcal{D}(A) \to \mathcal{H}$, with $\mathcal{D}(A) \subseteq \mathcal{H}$ dense in $\mathcal{H}$, be a $\eta-$pseudo-Hermitian linear operator in the Hilbert space $\mathcal{H}$. Then A is Hermitian with respect to the inner product (3.56), that is $\langle\langle\xi|A\rho\rangle\rangle_\eta = \langle\langle A\xi|\rho\rangle\rangle_\eta$ where $|\rho\rangle, |\xi\rangle \in \mathcal{D}(A)$.*

Proof. Since A is $\eta-$pseudo-Hermitian then $A^\dagger\eta = \eta A$, and

$$\langle\langle\xi|A\rho\rangle\rangle_\eta = \langle\xi|\eta A\rho\rangle = \langle\xi|A^\dagger\eta\rho\rangle = \langle A\xi|\eta\rho\rangle = \langle\langle A\xi|\rho\rangle\rangle_\eta. \tag{3.57}$$

□

Assume that, $H : \mathcal{D}(H) \to \mathcal{H}$, with $\mathcal{D}(H) \subseteq \mathcal{H}$ dense in $\mathcal{H}$, is the energy observable, namely Hamiltonian, in a quantum mechanical system. Let H has a (real or complex) discrete spectrum, and the eigenvectors $\{|\psi_n\rangle\}$ of H, and the eigenvectors $\{|\phi_n\rangle\}$ of $H^\dagger$, together constitute a bounded complete biorthonormal system $\{|\psi_n\rangle, |\phi_n\rangle\}$ on $\mathcal{H}$. Then we have,

$$H|\psi_n\rangle = E_n|\psi_n\rangle \qquad H^\dagger|\phi_n\rangle = E_n^*|\phi_n\rangle, \tag{3.58}$$

with

$$\langle\phi_m|\psi_n\rangle = \delta_{mn}, \tag{3.59}$$

and,

$$H = \sum_n E_n|\psi_n\rangle\langle\phi_n|, \qquad H^\dagger = \sum_n E_n^*|\phi_n\rangle\langle\psi_n|, \tag{3.60}$$

$$\sum_n |\psi_n\rangle\langle\phi_n| = \sum_n |\phi_n\rangle\langle\psi_n| = \mathbb{I}. \tag{3.61}$$

Proposition 3.2.4. *[11]Let H be a pseudo-Hermitian Hamiltonian with these properties. Then the nonreal eigenvalues of H come in complex conjugate pairs.*

Proof.

$$H(\eta^{-1}|\phi_n\rangle) = \eta^{-1}H^\dagger|\phi_n\rangle = E_n^*(\eta^{-1}|\phi_n\rangle). \tag{3.62}$$

Because η^{-1} is invertible, $\eta^{-1}|\phi_n\rangle \neq 0$ is an eigenvector of H with eigenvalue E_n^*. More generally, η^{-1} maps the eigensubspace associated with E_n to that associated with E_n^*. □

Now write such an H as,

$$H = \sum_{n_0} E_{n_0} |\psi_{n_0}\rangle\langle\phi_{n_0}| + \sum_{n_+} (\ E_{n_+} |\psi_{n_+}\rangle\langle\phi_{n_+}| + E^*_{n_+} |\psi_{n_-}\rangle\langle\phi_{n_-}|\ ) \tag{3.63}$$

where n_0 denote real eigenvalues and $n_\pm$ denote complex eigenvalues with $\pm$ imaginary part. Then we can construct[11] η and η^{-1} as,

$$\eta = \sum_{n_0} |\phi_{n_0}\rangle\langle\phi_{n_0}| + \sum_{n_+} (\ |\phi_{n_-}\rangle\langle\phi_{n_+}| + |\phi_{n_+}\rangle\langle\phi_{n_-}|\ ), \tag{3.64}$$

$$\eta^{-1} = \sum_{n_0} |\psi_{n_0}\rangle\langle\psi_{n_0}| + \sum_{n_+} (\ |\psi_{n_-}\rangle\langle\psi_{n_+}| + |\psi_{n_+}\rangle\langle\psi_{n_-}|\ ). \tag{3.65}$$

If the discrete spectrum is real, then

$$\eta = \sum_{n_0} |\phi_{n_0}\rangle\langle\phi_{n_0}|, \tag{3.66}$$

$$\eta^{-1} = \sum_{n_0} |\psi_{n_0}\rangle\langle\psi_{n_0}|. \tag{3.67}$$

Theorem 3.2.5. *[11] Let H be a non-Hermitian Hamiltonian with a (real or complex) discrete spectrum and bounded complete biorthonormal system of eigenvectors $\{|\psi_n\rangle, |\phi_n\rangle\}$. Then H is pseudo-Hermitian if and only if one of the following conditions hold*

1. *The spectrum of H is real.*

2. *The complex eigenvalues come in complex conjugate pairs.*

Proof. [11] We have already shown in Proposition (3.2.4) that pseudo-Hermiticity of H implies at least one of these conditions. To prove that these conditions are sufficient for the pseudo-Hermiticity of H, write H as in (3.63) and construct η as in (3.64). Then, by construction, H and η satisfy $H^\dagger = \eta H \eta^{-1}$. □

Theorem 3.2.6. *[12] Let $H : \mathcal{H} \to \mathcal{H}$ be a Hamiltonian that acts in a separable Hilbert Space $\mathcal{H}$, has a discrete spectrum, and admits a bounded complete biorthonormal system of eigenvectors $\{|\psi_n\rangle, |\phi_n\rangle\}$. Then the spectrum of H is real if and* *only if there is an invertible linear operator $O : \mathcal{H} \to \mathcal{H}$ such that H is $OO^\dagger$-pseudo-Hermitian.*

Note that $OO^\dagger$ is a bounded and full domain Hermitian linear operator on $\mathcal{H}$. If O is invertible then [5]$(O^{-1})^\dagger = (O^\dagger)^{-1}$ *be a bounded full domain operator on $\mathcal{H}$ (see Definitions (2.1.42) and (2.1.50)). They together imply that $O = OO^\dagger(O^\dagger)^{-1}$ is a bounded linear operator on $\mathcal{H}$.*

Proof. [12] Let $\{|n\rangle\}$ be a complete orthonormal basis of $\mathcal{H}$, i.e.,

$$\langle m|n\rangle = \delta_{mn}, \qquad \sum_n |n\rangle\langle n| = \mathbb{I} \tag{3.68}$$

and $\mathcal{O} : \mathcal{H} \to \mathcal{H}$ and $H_0 : \mathcal{H} \to \mathcal{H}$ be defined by

$$\mathcal{O} := \sum_n |\psi_n\rangle\langle n|, \qquad H_0 := \sum_n E_n|n\rangle\langle n|. \tag{3.69}$$

Then $\mathcal{O}$ is invertible and the inverse is given by

$$\mathcal{O}^{-1} = \sum_n |n\rangle\langle\phi_n| \tag{3.70}$$

since,

$$\begin{aligned} \mathcal{O}\mathcal{O}^{-1} &= (\textstyle\sum_n |\psi_n\rangle\langle n|)(\sum_m |m\rangle\langle\phi_m|) = \sum_n\sum_m |\psi_n\rangle\langle n|m\rangle\langle\phi_m| \\ &= \textstyle\sum_n\sum_m \delta_{mn}|\psi_n\rangle\langle\phi_m| = \sum_n |\psi_n\rangle\langle\phi_n| = \mathbb{I} \end{aligned} \tag{3.71}$$

and

$$\begin{aligned} \mathcal{O}^{-1}\mathcal{O} &= (\textstyle\sum_m |m\rangle\langle\phi_m|)(\sum_n |\psi_n\rangle\langle n|) = \sum_n\sum_m |m\rangle\langle\phi_m|\psi_n\rangle\langle n| \\ &= \textstyle\sum_n\sum_m \delta_{mn}|m\rangle\langle n| = \sum_n |n\rangle\langle n| = \mathbb{I}. \end{aligned} \tag{3.72}$$

Then,

$$\mathcal{O}^{-1}H\mathcal{O} = (\sum_n |n\rangle\langle\phi_n|)(\sum_k E_k|\psi_k\rangle\langle\phi_k|)(\sum_m |\psi_m\rangle\langle m|) = \sum_n E_n|n\rangle\langle n| = H_0 \tag{3.73}$$

Now suppose that the spectrum of H is real so H_0 is Hermitian, that is

$$H_0^\dagger = H_0. \tag{3.74}$$

[5]Ref. [9].

Hence

$$\mathcal{O}^\dagger H^\dagger \mathcal{O}^{-1\dagger} = \mathcal{O}^{-1} H \mathcal{O}, \tag{3.75}$$

or

$$H^\dagger = \mathcal{O}^{-1\dagger}\mathcal{O}^{-1} H \mathcal{O}\mathcal{O}^\dagger = (\mathcal{O}\mathcal{O}^\dagger)^{-1} H (\mathcal{O}\mathcal{O}^\dagger). \tag{3.76}$$

Define $O := \mathcal{O}^{\dagger^{-1}}$. This shows that H is $OO^\dagger$-pseudo-Hermitian. This completes the proof of necessity. Next we suppose that H is $OO^\dagger$-pseudo-Hermitian. Then (3.76) and consequently (3.75) hold. By (3.69) and (3.70), we have

$$\begin{aligned}\mathcal{O}^{-1} H \mathcal{O} &= \Big(\textstyle\sum_n |n\rangle\langle\phi_n|\Big) H \Big(\textstyle\sum_m |\psi_m\rangle\langle m|\Big)\\ &= \textstyle\sum_n\sum_m E_m |n\rangle\langle\phi_n|\psi_m\rangle\langle m| = \sum_n\sum_m E_m |n\rangle\delta_{mn}\langle m|\\ &= \textstyle\sum_n E_n |n\rangle\langle n| = H_0\end{aligned} \tag{3.77}$$

$$\begin{aligned}\mathcal{O}^\dagger H^\dagger \mathcal{O}^{-1\dagger} &= \Big(\textstyle\sum_n |n\rangle\langle\psi_n|\Big) H^\dagger \Big(\textstyle\sum_m |\phi_m\rangle\langle m|\Big)\\ &= \textstyle\sum_n\sum_m E_m^* |n\rangle\langle\psi_n|\phi_m\rangle\langle m| = \sum_n\sum_m E_m^* |n\rangle\delta_{mn}\langle m|\\ &= \textstyle\sum_n E_n^* |n\rangle\langle n| = H_0^\dagger.\end{aligned} \tag{3.78}$$

Therefore, (3.75) implies that H_0 is Hermitian and the eigenvalues E_n are all real. $\square$

Remark: $OO^\dagger : \mathcal{H} \to \mathcal{H}$ is a positive-definite operator on $\mathcal{H}$. Moreover, the modified inner product $\langle\langle\,\cdot\,|\,\cdot\,\rangle\rangle_{OO^\dagger}$ satisfies all the conditions (2.17)-(2.20) of inner product.

Proof. Let x be any vector in $\mathcal{H}$, then

$$\langle x|OO^\dagger x\rangle = \langle xO^\dagger|O^\dagger x\rangle = \|O^\dagger x\|^2 > 0 \quad \text{for} \quad x \neq 0. \tag{3.79}$$

Let $\xi = OO^\dagger$ for simplicity. Now ξ is a bounded, Hermitian, invertible, positive-definite linear operator on $\mathcal{H}$. Then $\langle\langle\,\cdot\,|\,\cdot\,\rangle\rangle_\xi$ satisfies all the conditions (2.17)-(2.20).

Let $x, y, z \in \mathcal{H}$ and $\alpha \in \mathbb{C}$, then

$$\begin{aligned}
& i.\langle\langle z|x+y\rangle\rangle_\xi = \langle z|\xi(x+y)\rangle = \langle \xi(x+y), z\rangle = \langle \xi x, z\rangle + \langle \xi y, z\rangle \\
& \qquad = \langle z|\xi x\rangle + \langle z|\xi y\rangle = \langle\langle z|x\rangle\rangle_\xi + \langle\langle z|y\rangle\rangle_\xi \\
& ii.\langle\langle y|\alpha x\rangle\rangle_\xi = \langle y|\xi\alpha x\rangle = \langle \xi\alpha x, y\rangle = \alpha\langle \xi x, y\rangle = \alpha\langle y|\xi x\rangle = \alpha\langle\langle y|x\rangle\rangle_\xi \\
& iii.\langle\langle y|x\rangle\rangle_\xi = \langle y|\xi x\rangle = \langle \xi x, y\rangle = \langle y, \xi x\rangle^* = \langle \xi y, x\rangle^* = \langle x|\xi y\rangle^* = \langle\langle x|y\rangle\rangle_\xi^* \\
& iv.\langle\langle x|x\rangle\rangle_\xi = \langle x|\xi x\rangle = \langle \xi x, x\rangle > 0 \ \text{ for all } \ x \in \mathcal{H} - \{0\} \\
& \quad \text{since } \xi \text{ is positive-definite and } \langle\langle x|x\rangle\rangle_\xi = 0 \quad \Leftrightarrow \quad x = 0.
\end{aligned} \tag{3.80}$$

Note that, indefinite inner product $\langle\langle \cdot | \cdot \rangle\rangle_\eta$ differs from inner product $\langle\langle \cdot | \cdot \rangle\rangle_\xi$ in the iv^{th} property(2.20). □

So far, we have worked in the Hilbert space $\mathcal{H}$, which is an inner product space, and complete with the metric defined by the inner product $\langle \cdot | \cdot \rangle$. We defined a modified inner product $\langle\langle \cdot | \cdot \rangle\rangle_\xi$ in this space. We will show in the next chapter that $\mathcal{H}$ is complete with the metric defined by the inner product $\langle\langle \cdot | \cdot \rangle\rangle_\xi$. Let us denote the Hilbert space, with inner product $\langle\langle \cdot | \cdot \rangle\rangle_\xi$, by $\mathcal{H}_\xi$. There is an important property for a ξ-pseudo-Hermitian Hamiltonian in the space $\mathcal{H}_\xi$.

Proposition 3.2.7. *Let $\xi = OO^\dagger$ as in the Theorem (3.2.6). Eigenvectors of the ξ-pseudo-Hermitian Hamiltonian H that correspond to the different eigenvalues are orthogonal to each other in the Hilbert space $\mathcal{H}_\xi$.*

Note that by the Theorem (3.2.6), the eigenvalues of H are real.

Proof.

$$\begin{aligned}
& H^\dagger\xi - \xi H = 0, \quad H|\psi_n\rangle = E_n|\psi_n\rangle, \quad \langle\psi_n|H^\dagger = E_n^*\langle\psi_n|, \ \ (E_n^* = E_n) \\
& \qquad 0 = \langle\xi_m|(H^\dagger\xi - \xi H)|\psi_n\rangle \\
& \qquad\ \ = \langle\psi_m|H^\dagger\xi|\psi_n\rangle - \langle\psi_m|\xi H|\psi_n\rangle \\
& \qquad\ \ = E_m\langle\psi_m|\xi|\psi_n\rangle - E_n\langle\psi_m|\xi|\psi_n\rangle \\
& \qquad\ \ = (E_m - E_n)\langle\psi_m|\xi|\psi_n\rangle \\
& \qquad\ \ = (E_m - E_n)\langle\langle\psi_m|\psi_n\rangle\rangle_\xi = 0.
\end{aligned} \tag{3.81}$$

Now if $m \neq n$ then $\langle\langle\psi_m|\psi_n\rangle\rangle_\xi = 0$ which shows orthogonality. □

Note that, we can prove this theorem easily just saying that H is a Hermitian operator acting in $\mathcal{H}_\xi$.

Chapter 4

CONSTRUCTION OF A CONSISTENT QUANTUM MECHANICAL SYSTEM WITH NON-HERMITIAN SET OF OPERATORS

4.1 Quasi-Hermiticity

In the previous chapter, a quantum mechanical system is explained in which the observables are represented by Hermitian operators. Eigenvalues of a Hermitian operator are real. This is very reasonable since they represent the possible values which can be obtained in an experiment. Furthermore, eigenvectors of a Hermitian operator with different eigenvalues are orthogonal to each other. Scholtz *et. al.* [20] state that it is a restrictive point of view to choose the Hilbert space or the inner-product first, and then choose the observables in such a way that they are Hermitian with respect to the primarily chosen inner product. They claim that we can first choose the observables and then find, if possible, a unique inner product with respect to which these operators are Hermitian. In our discussion of the concept of pseudo-Hermiticity, we have shown that under certain general conditions an operator has a real spectrum if and only if it is η-pseudo-Hermitian for a positive-definite η. Here in Quasi-Hermiticity, Scholtz *et.al.* first take a set of bounded operators, which are actually T-pseudo-Hermitian with respect to a positive-definite, Hermitian, bounded and invertible linear operator T, and they give a necessary and sufficient condition, for determining this T uniquely. This T defines a new inner product on Hilbert space namely $\langle \cdot \mid \cdot \rangle_T = \langle \cdot | T \cdot \rangle$. A unique T is required for a consistent quantum mechanical system, since we should have a unique inner product to calculate an expectation value for example. Scholtz *et. al.* claim that if we construct such a consistent quantum mechanical system, then as long as the inner product $\langle \cdot \mid \cdot \rangle_T$ is employed we can use all the usual theorems about Hermitian operators.

We will give the definition of quasi-Hermiticity, and the theorem which shows the necessary and sufficient condition to construct such a consistent quantum mechanical system with a set of quasi-Hermitian operators.

Definition 4.1.1. *[20] Let $\{A_i\}$ be the set of (not necessarily Hermitian) bounded linear operators on the Hilbert space $\mathcal{H}$, with domain $\mathcal{D}(A_i) = \mathcal{D}(A_i^\dagger) = \mathcal{H}$. Then we call the set of operators $\{A_i\}$ quasi-Hermitian if there exists an invetible linear operator $T : \mathcal{H} \to \mathcal{H}$ such that*

$$\mathcal{D}(T) = \mathcal{H} \tag{4.1}$$

$$T^\dagger = T \qquad (Hermiticity) \tag{4.2}$$

$$\langle\phi|T\phi\rangle > 0; \quad \forall\phi \in \mathcal{H} \;\; and \;\; \phi \neq 0 \qquad (positive - definiteness) \tag{4.3}$$

$$\|T\phi\| \leq \|T\|\|\phi\|; \quad \forall\phi \in \mathcal{H} \qquad (boundedness) \tag{4.4}$$

$$TA_i = A_i^\dagger T; \quad \forall i \tag{4.5}$$

We define the new inner product according to[20],

$$\langle\phi|\psi\rangle_T = \langle\phi|T\psi\rangle. \tag{4.6}$$

$\mathcal{H}$ is complete with the metric defined by the inner product $\langle\phi|\psi\rangle_T$. This will be proved in the following. We denote the Hilbert space $\mathcal{H}$, with the inner product $\langle\phi|\psi\rangle_T$, as $\mathcal{H}_T$. The set of operators A_i are Hermitian in $\mathcal{H}_T$[20], because

$$\langle\phi|A_i\psi\rangle_T = \langle\phi|TA_i\psi\rangle = \langle\phi|A_i^\dagger T\psi\rangle = \langle A_i\phi|T\psi\rangle = \langle A_i\phi|\psi\rangle_T. \tag{4.7}$$

4.2 Required Definitions and Theorems

We will generally follow Ref. [9] in this section.

Definition 4.2.1. *(Positive square root)*

Let $T : \mathcal{H} \to \mathcal{H}$ be a positive, bounded, Hermitian linear operator on a complex Hilbert *space $\mathcal{H}$. Then a bounded Hermitian linear operator A is called a square root of T if*

$$A^2 = T. \tag{4.8}$$

If, in addition, A is positive, then A is called a positive square root of T and is denoted by

$$A = T^{1/2}.$$

Theorem 4.2.2. *Every positive bounded Hermitian linear operator $T : \mathcal{H} \to \mathcal{H}$ on a complex Hilbert space $\mathcal{H}$ has a unique positive square root A. This operator A commutes with every bounded linear operator on $\mathcal{H}$ which commutes with T.*

Note that A is also positive bounded and Hermitian linear operator on $\mathcal{H}$.

Remark: Let $T : \mathcal{H} \to \mathcal{H}$ be a positive bounded Hermitian linear operator on a complex Hilbert space $\mathcal{H}$. Then for all $x \in \mathcal{H}$,

$$\|Tx\| \leq \|T\|^{1/2}\langle x|Tx\rangle^{1/2} \tag{4.9}$$

which implies

$$\langle x|Tx\rangle = 0 \Leftrightarrow Tx = 0. \tag{4.10}$$

Note that (4.10) shows that if T is invertible (or just $ker(T) = \{0\}$) then it is positive-definite.

Proof. By using positive square root A of T, for all $x, y \in \mathcal{H}$ we have

$$|\langle y|Tx\rangle| \leq \langle y|Ty\rangle^{1/2}\langle x|Tx\rangle^{1/2} \tag{4.11}$$

since by Schwarz inequality[1]

$$|\langle y|Tx\rangle| = |\langle y|A^2x\rangle| = |\langle Ay|Ax\rangle| \leq \|Ay\|\|Ax\| = \langle Ay|Ay\rangle^{1/2}\langle Ax|Ax\rangle^{1/2}$$
$$= \langle y|Ty\rangle^{1/2}\langle x|Tx\rangle^{1/2}.$$

Now by using (4.11) for all $x \in \mathcal{H}$ we have (4.9) since for $y = Tx$

$$\|Tx\|^2 \leq \langle Tx|T^2x\rangle^{1/2}\langle x|Tx\rangle^{1/2} \leq \|Tx\|^{1/2}\|T^2x\|^{1/2}\langle x|Tx\rangle^{1/2}$$

[1] $|\langle x, y\rangle| \leq \|x\|\|y\|$.

$$\leq \|Tx\|^{1/2}\|T\|^{1/2}\|Tx\|^{1/2}\langle x|Tx\rangle^{1/2} = \langle x|Tx\rangle^{1/2}\|T\|^{1/2}\|Tx\|$$

by using Schwarz inequality. □

Definition 4.2.3. *Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator. Then a subspace $Y \subseteq \mathcal{H}$ is said to be invariant under T if $T(Y) \subseteq Y$.*[2]

Remark 1: A closed subspace Y of $\mathcal{H}$ is invariant under T if and only if its orthogonal complement $Y^{\perp}$ is invariant under $T^{\dagger}$.

Proof. Let y be any element in Y. Then

$$\langle x|Ty\rangle = 0 \;\; \forall x \in Y^{\perp} \quad \Leftrightarrow \quad \langle T^{\dagger}x|y\rangle = 0 \;\; \forall x \in Y^{\perp} \quad \Leftrightarrow \quad T^{\dagger}(Y^{\perp}) \subseteq Y^{\perp}.$$

□

Remark 2: An eigenspace of T is invariant under T.

Definition 4.2.4. *[20]A set of bounded operators $\{A_i\}$ on a Hilbert space $\mathcal{H}$ is termed irreducible if there are no proper subspaces which are invariant under the action of all the operators A_i.*

Definition 4.2.5. *[2] Let $\mathcal{M}$ be a set of bounded operators on a Hilbert space $\mathcal{H}$. Then the set $\mathcal{M}'$ of all bounded operators on $\mathcal{H}$ which commute with each $A_i \in \mathcal{M}$ is called the commutant of $\mathcal{M}$.*

Definition 4.2.6. *[2] Let $\mathcal{M}$ be a set of bounded operators on a Hilbert space $\mathcal{H}$. Then a vector $\psi \in \mathcal{H}$ is called cyclic for $\mathcal{M}$ if the set $Span\{A\psi; A \in \mathcal{M}\}$ is dense in $\mathcal{H}$.*

Theorem 4.2.7. *[2] Let $\mathcal{M}$ be a set of Hermitian bounded operators on the Hilbert space $\mathcal{H}$. The following conditions are equivalent.*

1. *$\mathcal{M}$ is irreducible;*
2. *the commutant $\mathcal{M}'$ of $\mathcal{M}$ consist of multiples of the identity operator;*
3. *every nonzero vector $\psi \in \mathcal{H}$ is cyclic for $\mathcal{M}$ in $\mathcal{H}$, or $\mathcal{M} = 0$ and $\mathcal{H} = \mathbb{C}$.*

[2] $T(Y) = \{Ty \;:\; \forall y \in Y\}$.

4.3 Construction of a Consistent Quantum Mechanical System

Theorem 4.3.1. *[20] Let $\{A_i\}$ be the set of quasi-Hermitian operators in the Hilbert space $\mathcal{H}$ and T be a linear operator with properties given in the Definition (4.1.1). Then T is uniquely determined up to a constant multiple, if and only if the set of operators $\{A_i\}$ is irreducible on $\mathcal{H}$.*

Proof. [20] First note that since T is invertible T^{-1} is bounded and $\mathcal{D}(T^{-1}) = \mathcal{H}$. Moreover T^{-1} is Hermitian, since using (2.67) $\mathbb{I} = \mathbb{I}^\dagger = (TT^{-1})^\dagger = (T^{-1})^\dagger T \Rightarrow T^{-1} = (T^{-1})^\dagger$ and T^{-1} is positive-definite, since $\langle x|T^{-1}x\rangle = \langle TT^{-1}x|T^{-1}x\rangle = \langle Ty|y\rangle = \langle y|Ty\rangle > 0 \ \forall x \in \mathcal{H}-\{0\}$. Now we can introduce the unique positive-definite bounded and Hermitian linear operator S satisfying $T = S^2$ and $T^{-1} = (S^{-1})^2$(see Definition 4.2.1 and Theorem 4.2.2 and Remark of it). We also have $\|S\| = \|T\|^{1/2}$ and $\|S^{-1}\| = \|T^{-1}\|^{1/2}$ by (2.65). Suppose $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence with respect to the norm $\|x\|_T = \langle x|Tx\rangle^{1/2}$. Then

$$
\begin{aligned}
\|x_n - x_m\| = \|S^{-1}S(x_n - x_m)\| \leq \|S^{-1}\|\|S(x_n - x_m)\| \\
&= \|S^{-1}\|\langle S(x_n - x_m)|S(x_n - x_m)\rangle^{1/2} \\
&= \|S^{-1}\|\langle (x_n - x_m)|S^2(x_n - x_m)\rangle^{1/2} \\
&= \|S^{-1}\|\langle (x_n - x_m)|T(x_n - x_m)\rangle^{1/2} \\
&= \|S^{-1}\|\|(x_n - x_m)\|_T
\end{aligned}
$$

Hence $\{x_n\}$ is Cauchy in the norm $\|x\| = \langle x|x\rangle^{1/2}$. But $\mathcal{H}$ is complete with respect to $\|\cdot\|$; hence there exist an $x \in \mathcal{H}$ such that $\|x_n - x\| \to 0$ as $n \to \infty$. Then boundedness of S yields

$$\|x_n - x\|_T = \|S(x_n - x)\| \leq \|S\|\|x_n - x\| \to 0 \quad \text{as} \quad n \to \infty.$$

Hence $\{x_n\}$ converges to $x \in \mathcal{H}$ in the norm $\|\cdot\|_T$. We conclude that $\mathcal{H}_T$, endowed with the inner-product $\langle x|y\rangle_T = \langle x|Ty\rangle$, forms a Hilbert space $\mathcal{H}_T$.

We now prove that T is uniquely defined on $\mathcal{H}$ if and only if the set of observables A_i is irreducible on $\mathcal{H}$.

First we prove that if irreducibility holds, T is unique. To do this we need the Theorem (4.2.7), which asserts that

A set of Hermitian bounded operators on a Hilbert space $\mathcal{H}$ is irreducible if and only if every bounded operator that commutes with every element of this set is proportional to the identity.

By assumption the set of quasi-Hermitian observables A_i are bounded on $\mathcal{H}$. Then they are bounded on $\mathcal{H}_T$, since

$$\begin{aligned}\|A_i x\|_T &= \langle A_i x|A_i x\rangle_T^{1/2} = \langle A_i x|TA_i x\rangle^{1/2} = \langle A_i x|S^2 A_i x\rangle^{1/2}\\ &= \langle SA_i x|SA_i x\rangle^{1/2} = \|SA_i x\| = \|SA_i S^{-1} Sx\|\\ &\leq \|SA_i S^{-1}\|\|Sx\| = \|SA_i S^{-1}\|\ \langle Sx|Sx\rangle^{1/2}\\ &= \|SA_i S^{-1}\|\ \langle x|S^2 x\rangle^{1/2} = \|SA_i S^{-1}\|\ \langle x|Tx\rangle^{1/2}\\ &= \|SA_i S^{-1}\|\ \langle x|x\rangle_T^{1/2}\\ &= \|SA_i S^{-1}\|\ \|x\|_T.\end{aligned} \tag{4.12}$$

Furthermore, they are Hermitian on $\mathcal{H}_T$ by (4.7) and by assumption they are irreducible.

Assume that $\tilde{T}$ satisfies all the properties of T. Then it follows from the property $TA_i = A_i^\dagger T,\ \forall i$, that $[T^{-1}\tilde{T}, A_i] = 0,\ \forall i$. Because,

$$TA_i = A_i^\dagger T \quad\Rightarrow\quad A_i T^{-1} = T^{-1}A_i^\dagger \quad\Rightarrow\quad A_i T^{-1}\tilde{T} = T^{-1}A_i^\dagger \tilde{T}$$

$$\tilde{T}A_i = A_i^\dagger \tilde{T} \quad\Rightarrow\quad T^{-1}\tilde{T}A_i = T^{-1}A_i^\dagger \tilde{T}$$

then $A_i T^{-1}\tilde{T} = T^{-1}\tilde{T}A_i \quad\Rightarrow\quad T^{-1}\tilde{T}A_i - A_i T^{-1}\tilde{T} = 0$, i.e. $[T^{-1}\tilde{T}, A_i] = 0$.

Since, $T^{-1}\tilde{T}$ is bounded on $\mathcal{H}$, it is also bounded on $\mathcal{H}_T$, as it is proved for A_i in (4.12). Hence by theorem just stated, $T^{-1}\tilde{T}$ is proportional to identity and $\tilde{T}$ is, therefore, proportional to T.

To prove the converse we simply have to prove that if the set $\{A_i\}$ is reducible, T is not unique.

Suppose the set $\{A_i\}$ is reducible on $\mathcal{H}$, and therefore also on $\mathcal{H}_T$. Then we can find at least one proper subspace of $\mathcal{H}_T$ which is invariant under the set $\{A_i\}$, and since these operators are Hermitian on $\mathcal{H}_T$, they also leave the orthogonal complement of this subspace in $\mathcal{H}_T$ invariant (see Definition (4.2.3) and Remark 1 of it). Let P denote the Hermitian projection operator on $\mathcal{H}_T$ that projects onto this subspace and Q the Hermitian projection operator that projects onto its orthogonal complement in $\mathcal{H}_T$ (see Definition 3.1.4). Since P and Q are Hermitian on $\mathcal{H}_T$, it follows that,

$$\langle u|Pv\rangle_T = \langle Pu|v\rangle_T \quad \Rightarrow \quad \langle u|TPv\rangle = \langle Pu|Tv\rangle \quad \Rightarrow \quad \langle u|TPv\rangle = \langle u|P^\dagger Tv\rangle,$$

and by the same application for Q,

$$\begin{aligned} TP &= P^\dagger T \\ TQ &= Q^\dagger T. \end{aligned} \tag{4.13}$$

Now since $PQx = QPx = 0$ for all $x \in \mathcal{H}$, $PQ = QP = 0$, then $Q^\dagger P^\dagger = P^\dagger Q^\dagger = 0$. We also have $(P^\dagger)^2 = P^\dagger$ and $(Q^\dagger)^2 = Q^\dagger$ in $\mathcal{H}$. We construct the operator $V = \alpha P + \beta Q$ with α and β positive real numbers. Now we maintain that $\tilde{T} = TV$, which is clearly not proportional to T, satisfies all the conditions stated in Definition (4.1.1);

(4.1) Clearly, $\mathcal{D}(\tilde{T}) = \mathcal{H}$.

(4.2) $\tilde{T}^\dagger = V^\dagger T = \alpha P^\dagger T + \beta Q^\dagger T = \alpha TP + \beta TQ = TV = \tilde{T}$.

(4.3) $\langle\phi|\tilde{T}\phi\rangle = \alpha\langle\phi|TP\phi\rangle + \beta\langle\phi|TQ\phi\rangle = \alpha\langle\phi|P^\dagger T\phi\rangle + \beta\langle\phi|Q^\dagger T\phi\rangle = \alpha\langle\phi|(P^\dagger)^2 T\phi\rangle + \beta\langle\phi|(Q^\dagger)^2 T\phi\rangle = \alpha\langle P\phi|P^\dagger T\phi\rangle + \beta\langle Q\phi|Q^\dagger T\phi\rangle = \alpha\langle P\phi|TP\phi\rangle + \beta\langle Q\phi|TQ\phi\rangle = \alpha\|P\phi\|_T^2 + \beta\|Q\phi\|_T^2 > 0 \ \forall\phi \neq 0$

(4.4) $\tilde{T} = TV$ is clearly bounded since V is bounded.

(4.5) Finally, it follows from the invariance of $P\mathcal{H}$ and $Q\mathcal{H}$ under the set $\{A_i\}$ that $[A_i, V] = 0, \ \forall i$. Since, by Projection Theorem (2.1.24) we can write any $\psi \in \mathcal{H}$ such that $\psi = P\psi + Q\psi$ and $[A_i, \alpha P + \beta Q](P\psi + Q\psi) = \alpha A_i P(P\psi + Q\psi) + \beta A_i Q(P\psi + Q\psi) - \alpha PA_i(P\psi + Q\psi) - \beta QA_i(P\psi + Q\psi) = \alpha A_i P\psi + \beta A_i Q\psi - \alpha PA_i P\psi - \beta QA_i Q\psi = \alpha A_i P\psi + \beta A_i Q\psi - \alpha A_i P\psi - \beta A_i Q\psi = 0$.

Multiply $TA_i = A_i^\dagger T$ from the right with V, $TA_iV = A_i^\dagger TV \ \Rightarrow \ TVA_i = A_i^\dagger TV \ \Rightarrow \tilde{T}A_i = A_i^\dagger \tilde{T}$. □

Chapter 5

THE MOST GENERAL η

In the discussion of the concept of pseudo-Hermiticity, we construct an η for which the Hamiltonian H is η-pseudo-Hermitian. It is stated in Ref.[14] that η is not unique and H is also η'-pseudo-Hermitian for an operator η' which has the same properties as η. We can obtain this kind of an operator η' by a formula $\eta' = A^{\dagger}\eta A$ [14] where $A : \mathcal{H} \to \mathcal{H}$ $(\mathcal{D}(A) = \mathcal{H})$ is a bounded invertible linear operator, commuting with H. Note that if A is the identity operator then $\eta' = \eta$. Besides, in the concept of quasi-Hermiticity, the operator T is defined with the same properties as a positive-definite η. Then each A_i in a set of bounded linear operators $\{A_i\}$ is assumed to be T-pseudo-Hermitian and $\{A_i\}$ is called a set of quasi-Hermitian operators. Theorem (4.3.1) shows that T(if exists) is uniquely determined up to a constant multiple if and only if the set of quasi-Hermitian operators $\{A_i\}$ is irreducible on $\mathcal{H}$.

In Section (5.1), we calculate the most general positive-definite η according to the equation $\eta' = A^{\dagger}\eta A$ for a Hamiltonian H, with a real discrete spectrum, on a N-dimensional Hilbert space $\mathcal{H}$. Analysis is given for $N < \infty$ though it can be generalized to $N = \infty$; separable ∞-dimensional $\mathcal{H}$ under additional conditions. In section (5.2), we choose another observable S such that the quasi-Hermitian set of operators $\{S, H\}$ is irreducible on $\mathcal{H}$. According to Theorem (4.3.1), $\{S, H\}$ determines η' uniquely up to a constant multiple. We assume that η' is determined uniquely as $\eta' = \eta$ by $\{S, H\}$. After that we perturb η and denote the perturbed η as $\tilde{\eta}$. Then by using first order perturbation, we try to find modified S, $\tilde{S}$, such that the irreducible set of quasi-Hermitian operators $\{\tilde{S}, H\}$ determines η' uniquely as $\tilde{\eta}$.

5.1 Calculation of The Most General η

Let H be a non-Hermitian Hamiltonian in N-dimensional Hilbert space $\mathcal{H}$ and discrete spectrum of H is real and non-degenerate.

$$H = \sum_{n=1}^{N} E_n |\psi_n\rangle\langle\phi_n|. \tag{5.1}$$

Then, according to (3.66),

$$\eta = \sum_{n=1}^{N} |\phi_n\rangle\langle\phi_n|. \tag{5.2}$$

The most general η[14]

$$\eta' = A^\dagger \eta A \tag{5.3}$$

where $A : \mathcal{H} \to \mathcal{H}$ $(\mathcal{D}(A) = \mathcal{H})$ is a (bounded) invertible linear operator, commuting with H. We now will determine the operator A. First note that η in (5.2) is unique up to the choice of (bounded) complete biorthonormal system $\{|\psi_n\rangle, |\phi_n\rangle\}$[14]. If the matrix representation of H is written in another basis then another eigenvector set $\{|\tilde{\psi}_n\rangle\}$ of H and (bounded) complete biorthonormal system $\{|\tilde{\psi}_n\rangle, |\tilde{\phi}_n\rangle\}$ will have been chosen. Here the operator A transforms $\{|\psi_n\rangle, |\phi_n\rangle\}$ to all possible $\{|\tilde{\psi}_n\rangle, |\tilde{\phi}_n\rangle\}$. That is, $|\tilde{\phi}_n\rangle := A^\dagger|\phi_n\rangle$ and $|\tilde{\psi}_n\rangle := A^{-1}|\psi_n\rangle$ and so $\langle\tilde{\phi}_n|\tilde{\psi}_m\rangle = \delta_{nm}$[14]. Moreover since A and H commute they possess common eigenvectors[1]. Therefore

$$A := \sum_{n=1}^{N} a_n |\psi_n\rangle\langle\phi_n|, \qquad a_n \in \mathbb{C}. \tag{5.4}$$

One can show that H is pseudo-Hermitian with respect to η'. Noting that $A^\dagger$ commutes with $H^\dagger$ since A commutes with H we have

$$\begin{aligned}\eta' H \eta'^{-1} &= A^\dagger \eta A H A^{-1} \eta^{-1} {A^\dagger}^{-1} = A^\dagger \eta H A A^{-1} \eta^{-1} {A^\dagger}^{-1} = A^\dagger \eta H \eta^{-1} {A^\dagger}^{-1} \\ &= A^\dagger H^\dagger {A^\dagger}^{-1} = H^\dagger A^\dagger {A^\dagger}^{-1} = H^\dagger.\end{aligned} \tag{5.5}$$

Then $A^\dagger$ and the most general positive-definite η, η', is determined by

$$A^\dagger = \sum_{n=1}^{N} a_n^* |\phi_n\rangle\langle\psi_n| \quad \text{and} \quad \eta' = A^\dagger \eta A = \sum_{n=1}^{N} |a_n|^2 |\phi_n\rangle\langle\phi_n|. \tag{5.6}$$

Now, supposing that a_n are smooth functions of E_n, A can be written as a function of H, such that,

$$\begin{aligned} A = f(H) &= \textstyle\sum_{k=0}^{\infty} \frac{1}{k!} \frac{d^k f(0)}{dx^k} x^k \Big|_{x=H} = \sum_{k=0}^{\infty} c_k H^k \\ &= \textstyle\sum_{k=0}^{\infty} c_k \Big(\sum_{n=1}^{N} E_n^k |\psi_n\rangle\langle\phi_n| \Big) \\ &= \textstyle\sum_{n=1}^{N} \Big(\sum_{k=0}^{\infty} c_k E_n^k \Big) |\psi_n\rangle\langle\phi_n|, \end{aligned} \tag{5.7}$$

which is a Taylor expansion of f. Since $\sum_{k=0}^{\infty} c_k E_n^k = f(E_n)$ by definition,

$$a_n = f(E_n) = \sum_{k=0}^{\infty} c_k E_n^k. \tag{5.8}$$

We write an equivalence relation such that, two functions f, g would be equivalent, if they take the value a_n for E_n, that is $f(E_n) = g(E_n) = a_n,\ n = 1, 2, \cdots N$,

$$\mathcal{F} = \{ f : [0, \infty) \to \mathbb{C},\ \text{analytic and } f(E_n) = a_n,\ n = 1, 2, \cdots N \}.$$

Then $\{a_n\} \longleftrightarrow \mathcal{F}_{/\sim}$, that is there exists one-to-one correspondence between all possible $\{a_n\}$ and equivalence classes of $\mathcal{F}$.

Now we transform η to the form $\eta = OO^{\dagger}$ [1], where $O : \mathcal{H} \to \mathcal{H}$ ($\mathcal{D}(O) = \mathcal{H}$) is a (bounded) invertible linear operator;

$$O := \sum_{n=1}^{N} |\phi_n\rangle\langle n| \quad \text{ then } \quad O^{-1} = \sum_{n=1}^{N} |n\rangle\langle\psi_n|, \tag{5.9}$$

where $\{|n\rangle\}$ be an orthonormal basis for $\mathcal{H}$. By using this form of η we obtain the Hermitian operator $O^{\dagger} H O^{\dagger^{-1}}$ such that

$$\begin{aligned} H^{\dagger} = \eta H \eta^{-1} = OO^{\dagger} H O^{\dagger^{-1}} O^{-1} \quad &\Rightarrow \quad O^{-1} H^{\dagger} O = O^{\dagger} H O^{\dagger^{-1}} \\ &\Rightarrow \quad (O^{\dagger} H O^{\dagger^{-1}})^{\dagger} = O^{\dagger} H O^{\dagger^{-1}} \end{aligned} \tag{5.10}$$

[1] Let $B : \mathcal{H} \to \mathcal{H}$ ($\mathcal{D}(B) = \mathcal{H}$) be a Hermitian, positive-definite operator on the N-dimensional Hilbert space $\mathcal{H}$, such that $B = \sum_{n=1}^{N} \alpha_n |\phi_n\rangle\langle\phi_n|$, where $\alpha_n > 0$ and $\{|\phi_i\rangle\}$ be any basis on $\mathcal{H}$. Let $\{|n\rangle\}$ be an orthonormal basis for $\mathcal{H}$. Since $\langle n|m\rangle = \delta_{mn}$, $B = \sum_{n=1}^{N} \alpha_n |\phi_n\rangle\langle n|n\rangle\langle\phi_n| = \sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_n |\phi_n\rangle\langle n|m\rangle\langle\phi_m|$. Write α_n as $\alpha_n = b_n b_n^*$ where $b_n \in \mathbb{C}$ and * denote complex conjugation. Then $B = (\sum_{n=1}^{N} b_n |\phi_n\rangle\langle n|)(\sum_{m=1}^{N} b_m^* |m\rangle\langle\phi_m|) = CC^{\dagger}$, where $C := \sum_{n=1}^{N} b_n |\phi_n\rangle\langle n|$ which is a non-Hermitian operator on $\mathcal{H}$.

and indeed

$$O^{\dagger}HO^{\dagger^{-1}} = \Big(\sum_{n=1}^{N}|n\rangle\langle\phi_n|\Big)\Big(\sum_{m=1}^{N}E_m|\psi_m\rangle\langle\phi_m|\Big)\Big(\sum_{k=1}^{N}|\psi_k\rangle\langle k|\Big) = \sum_{n=1}^{N}E_n|n\rangle\langle n| \quad (5.11)$$

where H transformed to a Hermitian operator

$$H_0 = \sum_{n=1}^{N}E_n|n\rangle\langle n| \quad (5.12)$$

with a similarity transformation.

Then we express A in terms of

$$|\psi_n\rangle = O^{\dagger^{-1}}|n\rangle \quad \text{and} \quad |\phi_n\rangle = O|n\rangle, \quad (5.13)$$

and obtain

$$A = \sum_{n=1}^{N}a_n|\psi_n\rangle\langle\phi_n| = \sum_{n=1}^{N}a_nO^{\dagger^{-1}}|n\rangle\langle n|O^{\dagger} = O^{\dagger^{-1}}\Big(\sum_{n=1}^{N}a_n|n\rangle\langle n|\Big)O^{\dagger} = O^{\dagger^{-1}}A_{H_0}O^{\dagger}. \quad (5.14)$$

Moreover

$$A_{H_0} = \sum_{n=1}^{N}a_n|n\rangle\langle n| = \sum_{n=1}^{N}\Big(\sum_{k=0}^{\infty}c_kE_n^k\Big)|n\rangle\langle n| = \sum_{k=0}^{\infty}c_k\Big(\sum_{n=1}^{N}E_n^k|n\rangle\langle n|\Big) = \sum_{k=0}^{\infty}c_kH_0^k \quad (5.15)$$

which means

$$A_{H_0} = f(H_0). \quad (5.16)$$

The most general η, $\eta' = A^{\dagger}\eta A$, by using equations

$$A = O^{\dagger^{-1}}f(H_0)O^{\dagger} \quad \text{and} \quad A^{\dagger} = Of(H_0)^{\dagger}O^{-1}, \quad (5.17)$$

is in the form

$$\eta' = A^{\dagger}\eta A = Of(H_0)^{\dagger}O^{-1}OO^{\dagger}O^{\dagger^{-1}}f(H_0)O^{\dagger} = Of(H_0)^{\dagger}f(H_0)O^{\dagger}. \quad (5.18)$$

Let

$$\mathcal{Z} = f(H_0)^{\dagger}f(H_0) \quad (5.19)$$

then

$$\mathcal{Z} = \Big(\sum_{k=0}^{\infty}c_kH_0^k\Big)^{\dagger}\Big(\sum_{m=0}^{\infty}c_mH_0^m\Big) = \Big(\sum_{k=0}^{\infty}c_k^*H_0^k\Big)\Big(\sum_{m=0}^{\infty}c_mH_0^m\Big) = \sum_{n=0}^{\infty}d_nH_0^n \quad (5.20)$$

where $d_n = \sum_{k=0}^{n} c_k^* c_{n-k} = c_0^* c_n + c_1^* c_{n-1} + \cdots + c_{n-1}^* c_1 + c_n^* c_0 \in \mathbb{R}$.

Finally, using

$$\eta' = O\mathcal{Z}O^\dagger = O\Big(\sum_{n=0}^{\infty} d_n H_0^n\Big)O^\dagger = \sum_{n=0}^{\infty} d_n O H_0^n O^\dagger = \sum_{n=0}^{\infty} d_n OO^\dagger {O^\dagger}^{-1} H_0^n O^\dagger \tag{5.21}$$

and

$$({O^\dagger}^{-1} H_0 O^\dagger)^n = {O^\dagger}^{-1} H_0 O^\dagger {O^\dagger}^{-1} H_0 O^\dagger {O^\dagger}^{-1} H_0 O^\dagger \cdots = {O^\dagger}^{-1} H_0^n O^\dagger, \tag{5.22}$$

we obtain

$$\eta' = \sum_{n=0}^{\infty} d_n OO^\dagger ({O^\dagger}^{-1} H_0 O^\dagger)^n = \sum_{n=0}^{\infty} d_n OO^\dagger H^n = \sum_{n=0}^{\infty} d_n \eta H^n = \eta \sum_{n=0}^{\infty} d_n H^n. \tag{5.23}$$

One can also check that

$$\begin{aligned}
{\eta'}^\dagger &= \Big(\textstyle\sum_{n=0}^{\infty} d_n(\eta H^n)\Big)^\dagger = \sum_{n=0}^{\infty} d_n(\eta H^n)^\dagger = \sum_{n=0}^{\infty} d_n(H^\dagger H^\dagger \cdots H^\dagger \eta)\\
&= \textstyle\sum_{n=0}^{\infty} d_n(H^\dagger \eta\eta^{-1} H^\dagger \eta\eta^{-1} \cdots H^\dagger \eta)\\
&= \textstyle\sum_{n=0}^{\infty} d_n[(H^\dagger\eta)\eta^{-1}(H^\dagger\eta)\eta^{-1}\cdots(H^\dagger\eta)]\\
&= \textstyle\sum_{n=0}^{\infty} d_n[(\eta H)\eta^{-1}(\eta H)\eta^{-1}\cdots(\eta H)] = \sum_{n=0}^{\infty} d_n \eta H^n\\
&= \eta \textstyle\sum_{n=0}^{\infty} d_n H^n = \eta'.
\end{aligned} \tag{5.24}$$

If we define

$$\xi := \sum_{n=0}^{\infty} d_n H^n, \tag{5.25}$$

then η' satisfies

$$\eta' = \eta\xi. \tag{5.26}$$

It is easy to check that

$$\eta' H {\eta'}^{-1} = \eta\xi H \xi^{-1}\eta^{-1} = \eta H \xi\xi^{-1}\eta^{-1} = \eta H \eta^{-1} = H^\dagger, \tag{5.27}$$

since H and ξ commute.

5.2 Determining The Second Observable by First Order Perturbation

Let $H,\ \eta,\ \eta' = \eta\xi$ and ξ are as above, that is,

$$H = \sum_{n=1}^{N} E_n |\psi_n\rangle\langle\phi_n| \qquad (E_n \in \mathbb{R}),$$

$$\eta = \sum_{n=1}^{N} |\phi_n\rangle\langle\phi_n|$$

and

$$\eta' = \eta\xi = \eta'^{\dagger} \quad \text{where} \quad \xi = \sum_{n=0}^{\infty} d_n H^n \ , \quad d_n \in \mathbb{R}.$$

Now let S be an operator such that the quasi-Hermitian set of operators $\{S, H\}$ is irreducible on $\mathcal{H}$. Then according to Theorem (4.3.1), one of η' is determined uniquely up to a constant multiple by the set $\{S, H\}$. Assume that η' is determined as $\eta' = \eta$ ($\xi = \mathbb{I}$) by the set $\{S, H\}$. Now we perturb ξ around $\mathbb{I}$ and examine the alteration of S by using first order perturbation. We express perturbed η as

$$\tilde{\eta} = \eta\tilde{\xi} = \eta(\mathbb{I} + \gamma K) \tag{5.28}$$

where $\gamma << 1$ and $K = K(H) = \sum_{n=1}^{\infty} d_n H^n$. Here H is $\tilde{\eta}$-pseudo Hermitian but S is not since $\{S, H\}$ determines η' uniquely as η. However we will try to find a modified S, denoted as $\tilde{S}$, such that the irreducible set of quasi-Hermitian operators $\{\tilde{S}, H\}$ determines η' uniquely as $\tilde{\eta}$. The observable S changes as $\tilde{S} = S + \gamma W + \mathcal{O}(\gamma^2)$, where $\mathcal{O}(\gamma^2)$ is terms of order two or higher in γ. We want to find W and for this purpose we write the equation

$$\tilde{\eta}\tilde{S} = \tilde{S}^{\dagger}\tilde{\eta} \tag{5.29}$$

which implies

$$\eta(\mathbb{I} + \gamma K)(S + \gamma W) = (S^{\dagger} + \gamma W^{\dagger})\eta(\mathbb{I} + \gamma K) + \mathcal{O}(\gamma^2). \tag{5.30}$$

Noting that $\eta S = S^{\dagger}\eta$ we then obtain

$$W - \eta^{-1}W^{\dagger}\eta = [S, K] + \mathcal{O}(\gamma^2). \tag{5.31}$$

By denoting $W^{\sharp} = \eta^{-1}W^{\dagger}\eta$, we have

$$W - W^{\sharp} = [S, K] + \mathcal{O}(\gamma^2). \tag{5.32}$$

Now let

$$W = A + (iB) \tag{5.33}$$

where

$$A = \frac{W + W^{\sharp}}{2} \quad \text{and} \quad B = \frac{W - W^{\sharp}}{2i}. \tag{5.34}$$

Consequently, B can be calculated by

$$B = \frac{1}{2i}[S, K] + \mathcal{O}(\gamma^2) \tag{5.35}$$

where

$$\tilde{S} = S + \gamma(A + iB) + \mathcal{O}(\gamma^2). \tag{5.36}$$

In view of the identities: $\eta A \eta^{-1} = A^{\dagger}$ and $\eta(iB)\eta^{-1} = -(iB)^{\dagger}$, we write $\tilde{S}$ in the form

$$\tilde{S} = (S + \gamma A) + \gamma iB + \mathcal{O}(\gamma^2). \tag{5.37}$$

Here $(S + \gamma A)$ is η-pseudo-Hermitian part of $\tilde{S}$ and the condition $\tilde{\eta}\tilde{S} = \tilde{S}^{\dagger}\tilde{\eta}$ leaves A arbitrary. This is very reasonable since we can always add such an η-pseudo-Hermitian operator A to S. However iB, which is not η-pseudo-Hermitian, is fixed by the condition $\tilde{\eta}\tilde{S} = \tilde{S}^{\dagger}\tilde{\eta}$.

Chapter 6

EXAMPLES

In this chapter we study two examples which are related to previous chapters. In Example 1, we follow some consequences obtained in Section (5.1). We calculate η' according to (5.6) for a general diagonalizable 2×2 complex matrix H with real eigenvalues. Recall that η' is the most general (positive-definite) η for H. Then we choose a general 2×2 complex matrix S as the second operator and require it to be η'-pseudo Hermitian. Hence we obtain the form of S as a second operator and provide the existence of $T = \eta'$ such that $\{S, H\}$ is quasi-Hermitian with. If $det([S, H]) \neq 0$[1] for a fixed S then $\{S, H\}$ is an irreducible set. Finally if we fix one such S, then we observe that η' is determined uniquely up to a constant multiple which verifies the uniqueness theorem (4.3.1). In Example 2, we just apply the conclusion of Section (5.2) to a particular two-dimensional system.

6.1 Example 1

Let H be a diagonalizable 2×2 complex matrix, $H = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, with real eigenvalues. The trace of this matrix, $Tr(H) = a + d$, which is equal to the sum of the eigenvalues of H, is real. Let $h := \frac{a+d}{2}$ and $M = H - h\mathbb{I}$. Then M is a traceless matrix, since

$$M = H - h\mathbb{I} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \frac{a+d}{2} & 0 \\ 0 & \frac{a+d}{2} \end{pmatrix} = \begin{pmatrix} \frac{a-d}{2} & b \\ c & -\frac{a-d}{2} \end{pmatrix}. \qquad (6.1)$$

[1]If $\{S, H\}$ have any common eigenvector which makes the set reducible then $det([S, H]) = 0$[16].

Clearly the eigenvectors of M and H are common. Let $|E\rangle$ be any eigenvector of H and M and $H|E\rangle = E|E\rangle$. Eigenvalues of M are real, since,

$$H|E\rangle = M|E\rangle + h|E\rangle \;\Rightarrow\; E|E\rangle = M|E\rangle + h|E\rangle \;\Rightarrow\; M|E\rangle = (E-h)|E\rangle. \quad (6.2)$$

Similarly the eigenvectors of $M^\dagger$ and $H^\dagger$ are common. Let $\{|\phi_i\rangle\}_{i=1}^2$ be eigenvectors of $M^\dagger$. Obviously, if we construct the most general (positive-definite) $\eta' = \sum_{i=1}^2 \alpha_i|\phi_i\rangle\langle\phi_i|$ $(\alpha_i > 0$ in $\mathbb{R})$ for M according to (5.6), then H will be pseudo-Hermitian with respect to η'. Note that $\eta' H \eta'^{-1} = h\mathbb{I} + \eta' M \eta'^{-1} = h\mathbb{I} + M^\dagger = H^\dagger$ where h is real.

Therefore, we may use M instead of H without loss of generality.

Let $\mathcal{E}$ be an eigenvalue of M. Then, because M is traceless, $-\mathcal{E}$ must be also an eigenvalue of M. If $\mathcal{E} = 0$, then either $M = 0$ which would mean $H = \alpha\mathbb{I}$ and is a trivial case, or M is not diagonalizable which violates our assumption. So assume $\mathcal{E} \neq 0$, then we have

$$M = \begin{pmatrix} f & b \\ c & -f \end{pmatrix} = T \begin{pmatrix} \mathcal{E} & 0 \\ 0 & -\mathcal{E} \end{pmatrix} T^{-1}, \quad (6.3)$$

and

$$det(M) = -(f^2 + bc) = det\Big(T \begin{pmatrix} \mathcal{E} & 0 \\ 0 & -\mathcal{E} \end{pmatrix} T^{-1}\Big)$$

$$= det(T) det \begin{pmatrix} \mathcal{E} & 0 \\ 0 & -\mathcal{E} \end{pmatrix} (1/det(T)) = det(T)(-\mathcal{E}^2)(1/det(T)) = -\mathcal{E}^2.$$

Hence,

$$det(M) = -(f^2 + bc) = -\mathcal{E}^2 \leq 0 \quad (6.4)$$

and $\mathcal{E} = \sqrt{f^2 + bc}$.

The eigenvalues of any pseudo-Hermitian matrix M' are real or come in complex conjugate pairs according to the Theorem (3.2.5). In addition, let M' be a traceless 2×2 matrix. Then, the sum of the eigenvalues of M' is zero. In the case that the eigenvalues are real, they will be equal to $\mathcal{E}'$ and $-\mathcal{E}'$ where $\mathcal{E}'$ is real. Therefore $det(M') = -\mathcal{E}'^2$ is negative. However, if the eigenvalues come in complex conjugate pairs, the eigenvalues will be equal to ir and $-ir$, where r is real, and in this case $det(M') = r^2$ is positive. Hence, the determinant of a traceless, pseudo-Hermitian, 2×2 matrix is negative if and only if it has real eigenvalues.

Any traceless 2×2 matrix can be written as a linear combination of Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.5}$$

Note that all, 2×2 complex matrices, can be written as a linear combination of the matrices $\{\sigma_i, \mathbb{I}\}$.

We will write M using Pauli matrices, and find the eigenvectors of $M^\dagger$ to construct η'.

$$M = \begin{pmatrix} c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = c_1\sigma_1 + c_2\sigma_2 + c_3\sigma_3 \ , \qquad c_i \in \mathbb{C}. \tag{6.6}$$

In the view of the fact that the determinant of M is negative we choose c_i as follows. First we introduce [10],

$$\begin{pmatrix} \tilde{c}_1 & \tilde{c}_2 & \tilde{c}_3 \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix}, \tag{6.7}$$

where $\theta, \phi \in \mathbb{C}$; $Re(\theta) \in [0, \pi)$, $Re(\phi) = [0, 2\pi)$. Then, since

$$det([\tilde{c}_1\sigma_1 + \tilde{c}_2\sigma_2 + \tilde{c}_3\sigma_3]) = det \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix} = -1, \tag{6.8}$$

we can set $c_i = \mathcal{E}\tilde{c}_i$ where $det(M) = -\mathcal{E}^2$. Then

$$M = \mathcal{E} \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}, \qquad \mathcal{E} \in \mathbb{R}, \text{ and } \theta, \phi \in \mathbb{C}. \tag{6.9}$$

Using the identities:

$$(\cos\theta)^* = \Big(\frac{e^{i\theta} + e^{-i\theta}}{2}\Big)^* = \frac{e^{-i\theta^*} + e^{i\theta^*}}{2} = \frac{e^{i\theta^*} + e^{-i\theta^*}}{2} = \cos\theta^*, \tag{6.10}$$

$$(\sin\theta)^* = \Big(\frac{e^{i\theta}-e^{-i\theta}}{2i}\Big)^* = \frac{e^{-i\theta^*}-e^{i\theta^*}}{-2i} = \frac{e^{i\theta^*}-e^{-i\theta^*}}{2i} = \sin\theta^*, \tag{6.11}$$

we have

$$M^\dagger = \mathcal{E}\begin{pmatrix} \cos\theta^* & \sin\theta^* e^{-i\phi^*} \\ \sin\theta^* e^{i\phi^*} & -\cos\theta^* \end{pmatrix}. \tag{6.12}$$

Eigenvectors of $M^\dagger$ are

$$|\phi_1\rangle = \begin{pmatrix} \cos\frac{\theta^*}{2} \\ \sin\frac{\theta^*}{2} e^{i\phi^*} \end{pmatrix} \tag{6.13}$$

and

$$|\phi_2\rangle = \begin{pmatrix} \sin\frac{\theta^*}{2} \\ -\cos\frac{\theta^*}{2} e^{i\phi^*} \end{pmatrix}. \tag{6.14}$$

The most general η, $\eta' = \sum_{i=1}^{2}\alpha_i|\phi_i\rangle\langle\phi_i|$, where $\alpha_i \in \mathbb{R}^+$, is given by

$$\eta' = \begin{pmatrix} \alpha_1|\cos\frac{\theta}{2}|^2 + \alpha_2|\sin\frac{\theta}{2}|^2 & \alpha_1\sin\frac{\theta}{2}\cos\frac{\theta^*}{2}e^{-i\phi} - \alpha_2\sin\frac{\theta^*}{2}\cos\frac{\theta}{2}e^{-i\phi} \\ \alpha_1\sin\frac{\theta^*}{2}\cos\frac{\theta}{2}e^{i\phi^*} - \alpha_2\sin\frac{\theta}{2}\cos\frac{\theta^*}{2}e^{i\phi^*} & \alpha_1|\sin\frac{\theta}{2}|^2 e^{i(\phi^*-\phi)} + \alpha_2|\cos\frac{\theta}{2}|^2 e^{i(\phi^*-\phi)} \end{pmatrix}$$

Let,

$$\beta := |\cos\frac{\theta}{2}|^2, \quad \gamma := |\sin\frac{\theta}{2}|^2, \qquad \zeta := \sin\frac{\theta}{2}\cos\frac{\theta^*}{2},$$

then,

$$\eta' = \begin{pmatrix} \alpha_1\beta + \alpha_2\gamma & e^{-i\phi}(\alpha_1\zeta - \alpha_2\zeta^*) \\ e^{i\phi^*}(\alpha_1\zeta^* - \alpha_2\zeta) & e^{i(\phi^*-\phi)}(\alpha_1\gamma + \alpha_2\beta) \end{pmatrix}, \tag{6.15}$$

where $\alpha_1 > 0,\ \alpha_2 > 0,\ \beta \geq 0,\ \gamma \geq 0 \in \mathbb{R}$ and $\zeta, \phi \in \mathbb{C}$.

Note that $\eta' = (\eta')^\dagger$, as expected. Let $\alpha_1/\alpha_2 := \alpha$. Then

$$\eta' = \alpha_2\begin{pmatrix} \alpha\beta + \gamma & e^{-i\phi}(\alpha\zeta - \zeta^*) \\ e^{i\phi^*}(\alpha\zeta^* - \zeta) & e^{i(\phi^*-\phi)}(\alpha\gamma + \beta) \end{pmatrix}, \ \alpha \neq 0,\ \beta,\ \gamma \text{ are positive and } \zeta,\ \phi \in \mathbb{C}. \tag{6.16}$$

We first choose M and calculate the most general η' as above. Since we fix θ and ϕ by determining M, (in addition to the coefficient α_2) there exists one independent variable α in the expression for η'. Now we choose another traceless operator $\tilde{M}$ generally and require it to be η'-pseudo-Hermitian.

Let,

$$\tilde{M} = \begin{pmatrix} r & s \\ t & -r \end{pmatrix} \tag{6.17}$$

and note that there exist six independent real variables, $Re(s)$, $Im(s)$, $Re(t)$, $Im(t)$, $Re(r)$, $Im(r)$ in $\tilde{M}$. Here, by requiring $\tilde{M}$ to be η'-pseudo-Hermitian, we require $\tilde{M}$ to be diagonalizable and to have real eigenvalues and hence negative determinant by Theorem (3.2.6) and (5.9).

Then $\tilde{M}^\dagger \eta' = \eta' \tilde{M}$ implies,

$$r^*(\alpha\beta + \gamma) + t^* e^{i\phi^*}(\alpha\zeta^* - \zeta) = r(\alpha\beta + \gamma) + t e^{-i\phi}(\alpha\zeta - \zeta^*), \tag{6.18}$$

$$r^* e^{-i\phi}(\alpha\zeta - \zeta^*) + t^* e^{i(\phi^* - \phi)}(\alpha\gamma + \beta) = s(\alpha\beta + \gamma) - r e^{-i\phi}(\alpha\zeta - \zeta^*), \tag{6.19}$$

$$s^*(\alpha\beta + \gamma) - r^* e^{i\phi^*}(\alpha\zeta^* - \zeta) = r e^{i\phi^*}(\alpha\zeta^* - \zeta) + t e^{i(\phi^* - \phi)}(\alpha\gamma + \beta), \tag{6.20}$$

$$s^* e^{-i\phi}(\alpha\zeta - \zeta^*) - r^* e^{i(\phi^* - \phi)}(\alpha\gamma + \beta) = s e^{i\phi^*}(\alpha\zeta^* - \zeta) - r e^{i(\phi^* - \phi)}(\alpha\gamma + \beta). \tag{6.21}$$

Note that (6.19) and (6.20) are complex conjugates of each other and we can obtain two real equations from them; in equations (6.18) and (6.21), the left hand side of the equation is complex conjugate of the right hand side of it. Hence we can obtain two more real equations. If these four real equations is written for $Re(r), Im(r), Re(s)$ and $Im(s)$, then determinant of the coefficient matrix is zero. This shows that at most three of these equations are independent. Equation (6.18) yields

$$\begin{aligned}
& t^* e^{i\phi^*}(\alpha\zeta^* - \zeta) - t e^{-i\phi}(\alpha\zeta - \zeta^*) = (r - r^*)(\alpha\beta + \gamma) \\
\Rightarrow \quad & \Big[t^* e^{i\phi^*}(\alpha\zeta^* - \zeta)\Big] - \Big[t^* e^{i\phi^*}(\alpha\zeta^* - \zeta)\Big]^* = 2iIm(r)(\alpha\beta + \gamma) \\
\Rightarrow \quad & 2iIm(t^* e^{i\phi^*}(\alpha\zeta^* - \zeta)) = 2iIm(r)(\alpha\beta + \gamma).
\end{aligned} \tag{6.22}$$

Let

$$\lambda := e^{i\phi^*}(\alpha\zeta^* - \zeta) \quad \text{and} \quad \tau := \alpha\beta + \gamma. \tag{6.23}$$

Then by (6.22) and (6.23) we have

$$Im(t^*\lambda) = Im(r)\tau. \tag{6.24}$$

Now Equation (6.19) yields

$$\begin{aligned} & s(\alpha\beta+\gamma) - t^* e^{i(\phi^*-\phi)}(\alpha\gamma+\beta) = (r+r^*)e^{-i\phi}(\alpha\zeta-\zeta^*) \\ \Rightarrow \quad & s(\alpha\beta+\gamma) - t^* e^{2Im(\phi)}(\alpha\gamma+\beta) = 2Re(r)e^{-i\phi}(\alpha\zeta-\zeta^*). \end{aligned} \tag{6.25}$$

Let

$$\sigma := e^{2Im(\phi)}(\alpha\gamma+\beta). \tag{6.26}$$

Then by (6.25), (6.26) and (6.23) we have

$$s\tau - t^*\sigma = 2Re(r)\lambda^*. \tag{6.27}$$

Finally Equation (6.21) yields

$$\begin{aligned} & se^{i\phi^*}(\alpha\zeta^*-\zeta) - s^*e^{-i\phi}(\alpha\zeta-\zeta^*) = (r-r^*)e^{i(\phi^*-\phi)}(\alpha\gamma+\beta) \\ \Rightarrow \quad & \left[se^{i\phi^*}(\alpha\zeta^*-\zeta)\right] - \left[se^{i\phi^*}(\alpha\zeta^*-\zeta)\right]^* = 2iIm(r)e^{2Im(\phi)}(\alpha\gamma+\beta) \\ \Rightarrow \quad & 2iIm(s\lambda) = 2iIm(r)\sigma \end{aligned} \tag{6.28}$$

Then by (6.28), (6.23) and (6.26) we have

$$Im(s\lambda) = Im(r)\sigma. \tag{6.29}$$

Briefly, from (6.24), (6.27) and (6.29) we obtain the system of equations

$$Im(\lambda t^*) = \tau Im(r) \tag{6.30}$$

$$\tau s - \sigma t^* = 2\lambda^* Re(r) \tag{6.31}$$

$$Im(\lambda s) = \sigma Im(r) \tag{6.32}$$

where

$$\lambda = e^{i\phi^*}(\alpha\zeta^*-\zeta) \qquad \sigma = e^{2Im(\phi)}(\alpha\gamma+\beta) \qquad \tau = \alpha\beta+\gamma. \tag{6.33}$$

Recall that

$$\beta := |\cos\frac{\theta}{2}|^2, \quad \gamma := |\sin\frac{\theta}{2}|^2, \quad \zeta := \sin\frac{\theta}{2}\cos\frac{\theta^*}{2}. \tag{6.34}$$

Before solving the system of equations (6.30)-(6.32), we use Euler's formula to write ζ in the form[16]

$$\zeta = \frac{1}{2}[\sin(Re(\theta)) + i\sinh(Im(\theta))]. \tag{6.35}$$

This shows that $\theta = 0 \Leftrightarrow \zeta = 0$[2]. Note that, in the above system of equations, there are six independent real variables $Re(s)$, $Im(s)$, $Re(t)$, $Im(t)$, $Re(r)$, $Im(r)$ and three variables λ, σ, τ depending on α. Here σ and τ are nonzero, hence we study the cases $\lambda = 0$ and $\lambda \neq 0$ separately.

Case 1: $\lambda = 0$;

1. $\zeta = 0$ ($\Leftrightarrow \theta = 0$)

   or

2. $\frac{\zeta}{\zeta^*} = \alpha \in \mathbb{R}$ and in this case ζ and $\theta \in \mathbb{R}$ and $\alpha = 1$.

Equations (6.30)-(6.32) turn into

$$Im(r) = 0 \tag{6.36}$$

$$\tau s - \sigma t^* = 0. \tag{6.37}$$

According to (6.37)

$$s = \frac{\sigma}{\tau} t^*. \tag{6.38}$$

Therefore;

$$\tilde{M} = \begin{pmatrix} r & s(t) \\ t & -r \end{pmatrix}, \qquad r \in \mathbb{R}. \tag{6.39}$$

Note that there are three independent real variables $Re(t), Im(t), Re(r)$ in $\tilde{M}$ where s is a function of t and $Im(r) = 0$.

Case 2: $\lambda \neq 0$;

1. $\zeta, \theta \neq 0$ and $\alpha \neq 1$,

   or

2. ζ and θ are not real.

[2] Recall that $Re(\theta) \in [0, \pi)$.

The right hand side of the equation (6.31) is real, hence $Im(\tau s - \sigma t^*) = 0$ which is a verification of (6.30) and (6.32). Moreover $Re(\tau s - \sigma t^*) = 2\lambda^* Re(r)$ and this yields

$$\tau Re(\lambda s) - \sigma Re(\lambda t^*) = 2|\lambda|^2 Re(r). \tag{6.40}$$

Then using the system of equations (6.30), (6.40) and (6.32) we obtain

$$Re(t) = -\Big(\frac{i\tau}{(\lambda^* - \lambda)}\Big) Im(r) - \Big(\frac{i(\lambda^* + \lambda)}{(\lambda^* - \lambda)}\Big) Im(t) \tag{6.41}$$

$$s = \Big(\frac{\sigma(\lambda^* + \lambda)}{|\lambda|^2 - \lambda^2} + \frac{i\sigma}{\lambda}\Big) Im(r) + \Big(\frac{i\sigma(\lambda^* - \lambda)}{2\tau\lambda} - \frac{i\sigma(\lambda^* + \lambda)^2}{2\tau(|\lambda|^2 - \lambda^2)}\Big) Im(t) + \Big(\frac{2\lambda^*}{\tau}\Big) Re(r). \tag{6.42}$$

In this case,

$$\tilde{M} = \begin{pmatrix} r & s \\ t & -r \end{pmatrix}, \tag{6.43}$$

involves three independent real variables: $Re(r), Im(r), Im(t)$.

Now we check that if we choose an $\tilde{M}$ satisfying (6.30)-(6.32) and $det[\tilde{M}, M] \neq 0$ then α is fixed and η' is determined uniquely up to a constant multiple, which is the conclusion of Theorem (4.3.1). For Case 1; if $\theta \neq 0$ then $\alpha = 1$ and if $\theta = 0$ then (6.38) is a linear equation for α and has a unique solution. For Case 2; (6.41) is a linear equation for α and therefore has a unique solution.

Recall that $M = H - h\mathbb{I}$ for a general 2×2 complex matrix H with real eigenvalues where $h \in \mathbb{R}$. Let $\tilde{M} = S - \tilde{h}\mathbb{I}$ with $\tilde{h} \in \mathbb{R}$. Then all the results that we have obtained for $\{\tilde{M}, M\}$ are valid for $\{S, H\}$.

6.2 Example 2

We apply the conclusion of Section (5.2) to the Hamiltonian[15]

$$H = \begin{pmatrix} D+1 & D-1 \\ -D+1 & -D-1 \end{pmatrix} \qquad D \in \mathbb{R}^+. \tag{6.44}$$

Let S be another operator such that $S = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}$, $a \in \mathbb{R}$. One can see that H and S do not have common eigenvectors hence they are irreducible on $\mathbb{R}^2$. Note that they are diagonalizable. Here, an η can be determined uniquely up to a constant multiple by the irreducible set of quasi-Hermitian operators $\{S, H\}$. Now we perturb η according to (5.28) for $K = H$, that is $\tilde{\eta} = \eta(\mathbb{I} + \gamma H)$, $\gamma << 1$. Then modified S is

$$\tilde{S} = S + \gamma W + \mathcal{O}(\gamma^2) = (S + \gamma A) + \gamma B + \mathcal{O}(\gamma^2)$$

where

$$B = \frac{1}{2}[S, H] + \mathcal{O}(\gamma^2) = \begin{pmatrix} a - aD & -a - aD \\ a + aD & -a + aD \end{pmatrix} + \mathcal{O}(\gamma^2) \tag{6.45}$$

and A is arbitrary. Note that we make a formal difference by setting $W = A + B$ where $A = \frac{W+W^\sharp}{2}$ and $B = \frac{W-W^\sharp}{2}$ instead of (5.33) and (5.34) where $W^\sharp = \eta^{-1}W^\dagger\eta$ since we work in real plane. Recall that η-pseudo-Hermitian part of W namely A is leaved arbitrary, however B which is not η-pseudo-Hermitian is fixed.

Appendix A

We will generally follow Ref. [9] in the appendix.

A.1 Proofs of Theorems in Chapter 2.

Proposition 2.1.7. Let X be a normed vector space. Then for any $x, y \in X$ we have

$$| \; \|x\| - \|y\| \; | \leq \|x - y\|. \tag{2.7}$$

Proof. By triangle inequality we have

$$\|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\|,$$

then

$$\|x\| - \|y\| \leq \|x - y\|.$$

By the same way we have

$$\|y\| - \|x\| \leq \|y - x\|$$

and

$$\|y\| - \|x\| \leq \|y - x\| = \|(-1)(x - y)\| = \|x - y\|$$

which shows

$$\|x\| - \|y\| \geq -\|x - y\|.$$

Hence

$$-\|x - y\| \leq \|x\| - \|y\| \leq \|x - y\| \quad \Leftrightarrow \quad | \; \|x\| - \|y\| \; | \leq \|x - y\|.$$

$\square$

Theorem 2.1.12. Let M be a nonempty subset of a normed vector space X, then $x \in \overline{M}$ if and only if there is a sequence $\{x_n\}$ in M such that $x_n \to x$.

Proof. ($\Rightarrow$) Let $x \in \overline{M}$. If $x \in M$, a sequence of that type is $(x, x, \cdots)$. If $x \notin M$, it is a limit point of M. Hence for each $n = 1, 2, \cdots$ the ball $B(x; \frac{1}{n})$ contains an $x_n \in M$, and $x_n \to x$, because $\frac{1}{n} \to 0$ as $n \to \infty$.
($\Leftarrow$) If $\{x_n\}$ is in M and $x_n \to x$, then $x \in M$ or every neighborhood of x contains points $x_n \neq x$, so that x is a limit point of M. Hence $x \in \overline{M}$, by the definition of the closure. □

Proposition 2.1.16. Every convergent sequence is Cauchy, but converse is not generally true.

Proof. Let X be a normed vector space and $\{x_n\}$ be a convergent sequence in X such that $x_n \to x$. Hence for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$\forall\ n \geq N, \quad d(x_n, x) = \|x_n - x\| < \varepsilon/2.$$

Now let $m > N$, then by using triangle inequality

$$\|x_m - x_n\| \leq \|x_m - x\| + \|x - x_n\| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Therefore for all $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $m, n \geq N$ $d(x_m - x_n) = \|x_m - x_n\| < \varepsilon$, i.e., $\{x_n\}$ is Cauchy. □

Theorem 2.1.19. Every closed subspace M of a Hilbert Space $\mathcal{H}$ is again a Hilbert Space.

Proof. Let M be any closed subspace of a Hilbert space $\mathcal{H}$. We need to show that for every Cauchy sequence in M, the limit of the sequence is in M. Let $\{x_n\}$ be any Cauchy sequence in M. Since $\mathcal{H}$ is complete $x_n \to x$ in $\mathcal{H}$. Hence $\{x_n\} \subset M$ is a convergent sequence in $\mathcal{H}$. Then since M is closed, $x \in M$ by definition of closure. Consequently, every Cauchy sequence is convergent in M and so M is a Hilbert space. □

Proposition 2.1.22. [8] Let M be a closed subspace of $\mathcal{H}$. Then the subset $M^\perp$ is a closed subspace of $\mathcal{H}$ and is therefore a Hilbert Space. Furthermore $M \cap M^\perp = \{0\}$.

Proof. Obviously for all $x, y \in M^\perp$ and scalar α, $x + \alpha y \in M^\perp$ since $\forall z \in M$ $\langle x + \alpha y, z\rangle = \langle x, z\rangle + \alpha\langle y, z\rangle = 0$ and hence $M^\perp$ is a subspace.

Let $\{x_n\} \subset M^\perp$ be any convergent sequence in $\mathcal{H}$ and $\{x_n\} \to x$. For every $y \in M$ and for all $n \in \mathbb{N}$ $\langle x_n, y\rangle = 0$. Then $\langle x, y\rangle = \langle lim_{n\to\infty} x_n, y\rangle = lim_{n\to\infty}\langle x_n, y\rangle = 0$ by continuity of inner product. We obtain that $\langle x, y\rangle = 0$ for all $y \in M$ and so $x \in M^\perp$ which shows closedness of $M^\perp$ in $\mathcal{H}$.

Now take any nonzero x from $\mathcal{H}$ and let M be any closed subspace of $\mathcal{H}$. If $x \in M^\perp$ then for all $y \in M$ $\langle x, y\rangle = 0$. However $x \notin M$, since $\langle x, x\rangle > 0$ for a nonzero x. Moreover if $x \in M$ then $x \notin M^\perp$ by the same reason. There is no vector in a space other than zero that inner product with all vectors in the space is zero. Hence $M \cap M^\perp = \{0\}$. □

Proposition 2.1.23: [8] If M is a subspace of $\mathcal{H}$ which is not closed, then we can still define the orthogonal complement of M as in Definition (2.1.21) and $M^\perp$ is still a closed subspace of $\mathcal{H}$. Moreover $(M^\perp)^\perp = \overline{M}$ and $M^\perp = \overline{M}^\perp$.

Proof. Let M be any subspace of a Hilbert space $\mathcal{H}$, which is not closed. Then $M^\perp$ is closed subspace of $\mathcal{H}$ by continuity of inner product (it is obvious by above proof, since we do not use the property of closedness of M there). For the second claim, we first show that $\overline{M} \subseteq (M^\perp)^\perp$. Let $y \in \overline{M}$ and $y_n \to y$ where $\{y_n\} \subset M$. Then for any $x \in M^\perp$, $\langle x, y_n\rangle = 0$ and $\langle x, y\rangle = 0$ by continuity of inner product. Since x was arbitrary $\forall x \in M^\perp$ $\langle x, y\rangle = 0$ which means $y \in (M^\perp)^\perp$ by definition. Now we show that $(M^\perp)^\perp \subseteq \overline{M}$. Let $y \in (M^\perp)^\perp$ then for all $x \in M^\perp$ $\langle x, y\rangle = 0$ by definition. Again by definition $x \in M^\perp$ if and only if $\forall m \in M$ $\langle x, m\rangle = 0$. This shows that $y \in M$ or at least $y \in \overline{M}$ by continuity of inner product. Consequently $(M^\perp)^\perp = \overline{M}$.

For the last claim let $M^\perp = N$. N is the closed subspace of $\mathcal{H}$. Since $(M^\perp)^\perp = \overline{M}$, we have $N^\perp = \overline{M}$ and $(N^\perp)^\perp = \overline{N} = N$. Hence $N = \overline{M}^\perp$, i.e., $M^\perp = \overline{M}^\perp$. □

The Projection Theorem:

Proposition A.1.1. *[8] Let M be a closed subspace of a Hilbert space $\mathcal{H}$. For any $x \in \mathcal{H}$ there exists a unique vector $y \in M$ such that if $d \equiv \inf_{z\in M} \|x - z\|$, the distance from x to M, then $d = \|x - y\|$.*

Proof. [8] If $x \in M$, then choose $y = x$.

If $x \notin M$, then there exist a sequence $\{y_n\}$, $y_n \in M$, such that

$$\lim_{n\to\infty} \|x - y_n\| = d$$

by definition of d.[1] We will show that this implies that $\{y_n\}$ is a Cauchy Sequence. Consequently, there exists $y \in M$ such that

$$y = \lim_{n\to\infty} y_n.$$

Then, by continuity,

$$d = \lim_{n\to\infty} \|x - y_n\| = \|x - y\|.$$

To show that $\{y_n\}$ is a Cauchy Sequence, begin with the parallelogram law; $(\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2)$

$$\|y_n - y_m\|^2 = \|(y_n - x) - (y_m - x)\|^2 = 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - \|y_n + y_m - 2x\|^2.$$

Now as $d = inf_{y\in M}\|x - y\|$, it follows that

$$\|(1/2)(y_n + y_m) - x\| \geq d$$

since $y_n + y_m \in M$. Then we get

$$\|y_n - y_m\|^2 \leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4d^2$$

and so as $m, n \to \infty$, right side of this equation approaches $4d^2 - 4d^2 = 0$.

Now we show that y is unique[9]:

We assume that $y \in M$ and $y_0 \in M$ are both satisfy

$$\|x - y\| = d \quad \text{and} \quad \|x - y_0\| = d$$

and show that then $y_0 = y$. By the parallelogram law,

$$\|y - y_0\|^2 = \|(y - x) - (y_0 - x)\|^2 = 2\|y - x\|^2 + 2\|y_0 - x\|^2 - \|(y - x) + (y_0 - x)\|^2$$

$$= 2d^2 + 2d^2 - 2^2\|(1/2)(y + y_0) - x\|^2.$$

[1] Because for $A \subseteq \mathbb{R},\ A \neq \emptyset$ and bounded, $\beta = inf A \Rightarrow \exists\{y_n\} \subset A$ such that $x_n \to \beta$.

Since $\frac{1}{2}(y + y_0) \in M$ we have

$$\|(1/2)(y + y_0) - x\| \geq d.$$

Hence

$$\|y - y_0\|^2 \leq 4d^2 - 4d^2 = 0,$$

then $\|y - y_0\| = 0$ and $y = y_0$. $\square$

Note that as a geometric interpretation, here y is the projection of x on M. It will be clearer if one applies this for any $a \in \mathbb{R}^2$ where M be the $x-$axis.

Theorem 2.1.24. [8] (The Projection Theorem)
Let M be a closed subspace of $\mathcal{H}$. Then any $x \in \mathcal{H}$ can be written uniquely as $x = y + z$ with $y \in M$ and $z \in M^{\perp}$.

Proof. [8] Let y be the projection of x on M, and set $z \equiv x - y$. We show that $z \in M^{\perp}$.

Let $t \in \mathbb{R}$ and $m \in M$.

According to the Proposition (A.1.1) $d = \|x - y\| = \inf_{n \in M} \|x - n\| \leq \|x - (y + tm)\|$, since $y + tm \in M$. Then by using parallelogram law ($\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$) and the equation $Re\langle u, v\rangle = 1/4(\|u + v\|^2 - \|u - v\|^2)$,

$$d^2 \leq \|x-(y+tm)\|^2 = \|(x-y)-tm\|^2 = (1/2)\|(x-y)-tm\|^2+(1/2)\|(x-y)-tm\|^2$$

$$= \left[\|x - y\|^2 + t^2\|m\|^2 - (1/2)\|(x - y) + tm\|^2\right] + (1/2)\|(x - y) - tm\|^2$$

$$= \|x - y\|^2 + t^2\|m\|^2 - \left[(1/2)\|(x - y) + tm\|^2 - (1/2)\|(x - y) - tm\|^2\right]$$

$$= \|x - y\|^2 + t^2\|m\|^2 - 2tRe\langle x - y, m\rangle = d^2 + t^2\|m\|^2 - 2tRe\langle x - y, m\rangle$$

for all t, so
$t^2\|m\|^2 - 2tRe\langle x - y, m\rangle \geq 0$ for all $t \in \mathbb{R}$
which implies that $\langle x - y, m\rangle = \langle z, m\rangle = 0$ so $z \in M^{\perp}$ since $m \in M$ was arbitrary.
To prove uniqueness[9], let

$$x = y + z = y_1 + z_1$$

where $y, y_1 \in M$ and $z, z_1 \in M^\perp$. Then $y - y_1 = z_1 - z$. Since $y - y_1 \in M$ and $y - y_1 = z_1 - z \in M^\perp$, $y - y_1 \in M \cap M^\perp = \{0\}$ by Proposition (2.1.22). Hence $y = y_1$, and $z = z_1$. □

Proposition 2.1.25: $\mathcal{H}^\perp = \{0\}$.

Proof. By Projection Theorem (2.1.24) any $x \in \mathcal{H}$ can be written uniquely as $x = y+z$ with $y \in \mathcal{H}$ and $z \in \mathcal{H}^\perp$. Since x and y in $\mathcal{H}$ then $(x-y) = z \in \mathcal{H}$. Here we also have $z \in \mathcal{H}^\perp$. Then by Proposition (2.1.22) $z \in \mathcal{H} \cap \mathcal{H}^\perp = \{0\}$. Therefore $\mathcal{H}^\perp = \{0\}$. □

Proposition 2.1.26. For any subset $W \neq \emptyset$ of a Hilbert space $\mathcal{H}$, the space M of all finite linear combinations of elements of W, $M = spanW$, is dense in $\mathcal{H}$ if and only if $W^\perp = \{0\}$.

Proof. ($\Rightarrow$) Let $M = spanW$ is dense in $\mathcal{H}$ and $x \in W^\perp$. Here $x \in \overline{M} = \mathcal{H}$ and so there is a sequence $\{x_n\}$ in M such that $x_n \to x$. Since $x \in W^\perp$ and $W^\perp \perp M$ [2] we have $\langle x_n, x\rangle = 0$ for all $n \in \mathbb{N}$. Then continuity of inner product implies that $\langle x_n, x\rangle \to \langle x, x\rangle = \|x\|^2 = 0$ which means $x = 0$. Therefore, $W^\perp = \{0\}$ since x was arbitrary in $W^\perp$.

($\Leftarrow$) Let $W^\perp = \{0\}$ and $M = spanW$. Now assume that $x \in M^\perp$ which means $x \perp M$, then $x \perp W$ and so $x \in W^\perp$. Hence $x = 0$ and since x was arbitrary in $M^\perp$, $M^\perp = \{0\}$. Then by Proposition (2.1.23) and Proposition (2.1.25), $(M^\perp)^\perp = \{0\}^\perp \Rightarrow \overline{M} = \mathcal{H}$ which means M is dense in $\mathcal{H}$. □

Proof of the Theorem 2.1.31:

Proposition A.1.2. *Let $\{x_1, \cdots, x_n\}$ be a linearly independent set of vectors in a normed vector space X (of any dimension). Then there is a number $c > 0$ such that for every choice of scalars $\alpha_1, \cdots, \alpha_n$ we have*

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq c(|\alpha_1| + \cdots + |\alpha_n|) \qquad (c > 0) \qquad \text{(A.1)}$$

[2]See Definition (2.1.20).

Proof. We write $s = |\alpha_1| + \cdots + |\alpha_n|$. If $s = 0$, all α_i are zero, so that (A.1) holds for any c. Let $s > 0$. Then (A.1) is equivalent to the inequality

$$\|\beta_1 x_1 + \cdots + \beta_n x_n\| \geq c \tag{A.2}$$

where $\beta_j = \alpha_j / s$ and $\sum_{j=1}^{n} |\beta_j| = 1$.

Hence it suffices to prove the existence of a $c > 0$ such that (A.2) holds for every $n-$tuple of scalars $\beta_1, \cdots, \beta_n$ with $\sum_{j=1}^{n} |\beta_j| = 1$. Suppose that this is false. Then there exist a sequence $\{y_m\}$ of vectors

$$y_m = \beta_1^{(m)} x_1 + \cdots + \beta_n^{(m)} x_n \qquad (\sum_{j=1}^{n} |\beta_j^{(m)}| = 1)$$

such that

$$\|y_m\| \to 0 \quad \text{as} \quad m \to \infty.$$

Now obviously $|\beta_j^{(m)}| \leq 1$. Hence for each fixed j the sequence

$$\{\beta_j^{(m)}\} = (\beta_j^{(1)}, \beta_j^{(2)}, \cdots)$$

is bounded. Consequently, by the Bolzano-Weierstrass Theorem, $\{\beta_1^{(m)}\}$ has a convergent subsequence. Let β_1 denote the limit of that subsequence, and let $\{y_{1,m}\}$ denote the corresponding subsequence of $\{y_m\}$. By the same argument $\{y_{1,m}\}$ has a subsequence $\{y_{2,m}\}$ for which the corresponding subsequence of scalars $\beta_2^{(m)}$ converges; let β_2 denote the limit. Continuing in this way, after n steps we obtain a subsequence $\{y_{n,m}\} = (y_{n,1}, y_{n,2}, \cdots)$ of $\{y_m\}$ whose terms are of the form

$$y_{n,m} = \sum_{j=1}^{n} \gamma_j^{(m)} x_j \qquad\qquad (\sum_{j=1}^{n} |\gamma_j^{(m)}| = 1)$$

with scalars $\gamma_j^{(m)}$ satisfying $\gamma_j^{(m)} \to \beta_j$ as $m \to \infty$. Hence, as $m \to \infty$,

$$y_{n,m} \to y = \sum_{j=1}^{n} \beta_j x_j$$

where $\sum |\beta_j| = 1$, so that not all β_j can be zero. Since $\{x_1, \cdots, x_n\}$ is a linearly independent set, we thus have $y \neq 0$. On the other hand, $y_{n,m} \to y$ implies $\|y_{n,m}\| \to$

$\|y\|$, by the continuity of the norm. Since $\|y_m\| \to 0$ by assumption and $\{y_{n,m}\}$ is a subsequence of $\{y_m\}$, we must have $\|y_{n,m}\| \to 0$. Hence $\|y\| = 0$, so that $y = 0$. This contradicts $y \neq 0$ and the lemma is proved. □

Theorem 2.1.31. Every finite dimensional inner product space is a separable Hilbert space.

This is based on the theorem that; every finite dimensional subspace Y of a normed vector space X is complete. In particular, every finite dimensional normed vector space is complete.

Proof. We consider an arbitrary Cauchy sequence $\{y_m\}$ in Y and show that it is convergent in Y. Let $dimY = n$ and $\{e_1, \cdots, e_n\}$ be any basis for Y. Then each y_m has a unique representation of the form

$$y_m = \alpha_1^{(m)} e_1 + \cdots + \alpha_n^{(m)} e_n.$$

Since $\{y_m\}$ is a Cauchy sequence, for every $\varepsilon > 0$ there exist an $N \in \mathbb{N}$ such that $\|y_m - y_r\| < \varepsilon$ for all $m, r \geq N$. From this and Proposition (A.1.2) we have for some $c > 0$

$$\varepsilon > \|y_m - y_r\| = \| \sum_{j=1}^{n} (\alpha_j^{(m)} - \alpha_j^{(r)}) e_j \| \geq c \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j^{(r)}|$$

where $m, r \geq N$. Division by $c > 0$ gives

$$|\alpha_j^{(m)} - \alpha_j^{(r)}| \leq \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j^{(r)}| < \frac{\varepsilon}{c} \qquad (m, r \geq N).$$

This shows that each of the n sequences

$$\{\alpha_j^{(m)}\} = (\alpha_j^{(1)}, \alpha_j^{(2)}, \cdots) \qquad j = 1, 2, \cdots, n$$

is Cauchy in $\mathbb{R}$ or $\mathbb{C}$ and so converges. Let $\alpha_j^{(m)} \to \alpha_j$. Using these n limits $\alpha_1, \cdots, \alpha_n$, we define

$$y = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

Clearly, $y \in Y$. Furthermore,

$$\|y_m - y\| = \| \sum_{j=1}^{n} (\alpha_j^{(m)} - \alpha_j) e_j \| \leq \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j| \|e_j\|.$$

Here since $\alpha_j^{(m)} \to \alpha_j$ then $\|y_m - y\| \to 0$, that is, $y_m \to y$. This shows that $\{y_m\}$ is convergent in Y. Since $\{y_m\}$ was an arbitrary Cauchy sequence in Y, this proves that Y is complete. $\square$

Proof of the Theorem 2.1.32:

Theorem A.1.3. *[8](Pythagorean Theorem)*
Let $\{x_n\}_{n=1}^N$ be an orthonormal set for a separable Hilbert space $\mathcal{H}$, then for all $x \in \mathcal{H}$

$$\|x\|^2 = \sum_{n=1}^{N} |\langle x, x_n\rangle|^2 + \|x - \sum_{n=1}^{N} \langle x, x_n\rangle x_n\|^2. \tag{A.3}$$

Proof. [8] Write $x = y + z$, where $y = \sum_{n=1}^N \langle x, x_n\rangle x_n$ and $z = x - y$. We check that $\langle y, z\rangle = 0$:

$$\begin{aligned}
\langle y, z\rangle &= \Big\langle \sum_{n=1}^{N} \langle x, x_n\rangle x_n,\ x - \sum_{m=1}^{N} \langle x, x_m\rangle x_m \Big\rangle \\
&= \Big\langle \sum_{n=1}^{N} \langle x, x_n\rangle x_n,\ x \Big\rangle - \Big\langle \sum_{n=1}^{N} \langle x, x_n\rangle x_n,\ \sum_{m=1}^{N} \langle x, x_m\rangle x_m \Big\rangle \\
&= \sum_{n=1}^{N} \langle x, x_n\rangle \langle x_n, x\rangle - \sum_{n,m=1}^{N} \langle x, x_n\rangle \langle x, x_m\rangle^* \langle x_n, x_m\rangle \\
&= \sum_{n=1}^{N} \langle x, x_n\rangle \langle x, x_n\rangle^* - \sum_{n,m=1}^{N} \langle x, x_n\rangle \langle x, x_m\rangle^* \delta_{mn} \\
&= \sum_{n=1}^{N} |\langle x, x_n\rangle|^2 - \sum_{n=1}^{N} |\langle x, x_n\rangle|^2 = 0.
\end{aligned}$$

Consequently, $\|x\|^2 = \langle y + z, y + z\rangle = \|y\|^2 + \|z\|^2$. $\square$

An immediate consequence of the Pythagorean theorem is the following inequality, known as *Bessel's inequality*[8]. For any orthonormal set $\{x_n\}_{n=1}^N$ and any vector $x \in \mathcal{H}$, we have

$$\sum_{n=1}^{N} |\langle x, x_n\rangle|^2 \leq \|x\|^2.$$

Theorem 2.1.32.[8] Let $B = \{x_n\}$ be an orthonormal basis for an infinite-dimensional separable Hilbert space $\mathcal{H}$. Then for any $x \in \mathcal{H}$

$$x = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n$$

where the sum converges absolutely and

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2.$$

Conversely, if $\{c_n\}$ is a collection of complex numbers satisfying $\sum_{n=1}^{\infty} |c_n| < \infty$ then $\sum_{n=1}^{\infty} c_n x_n$ converges absolutely to a vector in $\mathcal{H}$.

Proof. [8] Let $x \in \mathcal{H}$ and $x \neq 0$. Consider the sequence $\{\langle x, x_n \rangle\}$ in $\mathbb{C}$. By Pythagorean Theorem (A.1.3) we have

$$\sum_{n=1}^{m} |\langle x, x_n \rangle|^2 \leq \|x\|^2 \tag{A.4}$$

for any m. Since $\sum_{n=1}^{m} |\langle x, x_n \rangle|^2$ is a monotonically increasing sequence, it converges, by (A.4), as $m \to \infty$ and

$$\sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2 \leq \|x\|^2. \tag{A.5}$$

Now define a sequence of vectors in $\mathcal{H}$ by

$$x^{(m)} = \sum_{n=1}^{m} \langle x, x_n \rangle x_n. \tag{A.6}$$

Then (A.6) is a Cauchy sequence, since for $m > k$ we have

$$\|x^{(m)} - x^{(k)}\|^2 = \| \sum_{n=k+1}^{m} \langle x, x_n \rangle x_n \|^2 = \sum_{n=k+1}^{m} |\langle x, x_n \rangle|^2$$

which, by (A.5), converges to zero as $m, k \to \infty$. Consequently, $\exists x' \in \mathcal{H}$ such that $x' = lim_{m \to \infty} x^{(m)}$. By construction $\langle x', x_n \rangle = \langle x, x_n \rangle$, and so for all n

$$\langle x' - x, x_n \rangle = 0$$

which shows that $x' - x$ is orthogonal to the orthonormal basis $\{x_n\}$. But by maximality of the orthonormal basis, we must have $x' - x = 0$, for otherwise $\{x' - x, x_1, x_2, \cdots\}$

would be an orthonormal basis properly containing $\{x_n\}$. Consequently, $x = x'$ and the sum

$$x = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n$$

converges where

$$\|x\|^2 = \lim_{n\to\infty} \|x^{(n)}\|^2 = \sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2$$

by continuity of the norm. □

Proposition 2.1.35. Let X and Y be two vector spaces and $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, be a linear operator. Then if T^{-1} exists, it is a linear operator.

Proof. First we show that the range $\mathcal{R}(T)$ of any linear operator T is a linear space.

Take any $y_1, y_2 \in \mathcal{R}(T)$ and show that $\alpha y_1 + \beta y_2 \in \mathcal{R}(T)$ for any scalars α and β. Since $y_1, y_2 \in \mathcal{R}(T)$, we have $y_1 = Tx_1,\ y_2 = Tx_2$ for some $x_1,\ x_2 \in \mathcal{D}(T)$, and $\alpha x_1 + \beta x_2 \in \mathcal{D}(T)$ because $\mathcal{D}(T)$ is a vector space. The linearity of T yields

$$T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 = \alpha y_1 + \beta y_2.$$

Hence $\alpha y_1 + \beta y_2 \in \mathcal{R}(T)$. Since $y_1, y_2 \in \mathcal{R}(T)$ were arbitrary and so were the scalars, this proves that $\mathcal{R}(T)$ is a vector space.

Now we will show that T^{-1} is a linear operator. Assume that T^{-1} is exist. The domain of T^{-1} is $\mathcal{R}(T)$ which is a vector space as shown above. We consider any $x_1,\ x_2 \in \mathcal{D}(T)$ and their images

$$y_1 = Tx_1 \quad \text{and} \quad y_2 = Tx_2.$$

Then

$$x_1 = T^{-1}y_1 \quad \text{and} \quad x_2 = T^{-1}y_2.$$

T is linear so for any scalars α, β we have

$$\alpha y_1 + \beta y_2 = \alpha T x_1 + \beta T x_2 = T(\alpha x_1 + \beta x_2).$$

Since $x_i = T^{-1}y_i$, we have

$$T^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha T^{-1} y_1 + \beta T^{-1} y_2$$

which shows linearity of T^{-1}. □

Theorem 2.1.36. Let X and Y be two vector spaces and $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, be a linear operator. Then $T^{-1} : R(T) \to \mathcal{D}(T)$ exists if and only if

$$Tx = 0 \quad \Rightarrow \quad x = 0,$$

i.e., ker(T)= $\{0\}$.

Proof. ($\Leftarrow$) Suppose that $Tx = 0$ implies $x = 0$. Let $Tx_1 = Tx_2$. Since T is linear,

$$T(x_1 - x_2) = Tx_1 - Tx_2 = 0,$$

so that $x_1 - x_2 = 0$ by hypothesis. Hence $Tx_1 = Tx_2$ implies $x_1 = x_2$, and so T^{-1} exists.

($\Rightarrow$) If T^{-1} exist, then $Tx_1 = Tx_2$ implies $x_1 = x_2$. From that with $x_2 = 0$ we obtain,

$$Tx_1 = T0 = 0 \quad \Rightarrow \quad x_1 = 0.$$

□

Theorem 2.1.39. Let X and Y be two normed vector spaces. A linear operator $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, is continuous at a point $x_0 \in \mathcal{D}(T)$ if and only if

$$x_n \to x_0 \quad \text{implies} \quad Tx_n \to Tx_0.$$

Proof. Assume T is continuous on x_0. Then, for all $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\|x - x_0\| < \delta \quad \text{implies} \quad \|Tx - Tx_0\| < \varepsilon.$$

Let $x_n \to x_0$. Then there is an $N \in \mathbb{N}$ such that for all $n > N$ we have

$$\|x_n - n_0\| < \delta.$$

Hence for all $n > N$,

$$\|Tx_n - Tx_0\| < \varepsilon.$$

By definition, this means that $Tx_n \to Tx_0$.

Conversely, we assume that

$$x_n \to x_0 \quad \text{implies} \quad Tx_n \to Tx_0$$

and prove that then T is continuous at x_0. Suppose this is false. Then there is an $\varepsilon > 0$ such that for every $\delta > 0$ there is an $x \neq x_0$ satisfying

$$\|x - x_0\| < \delta \quad \text{but} \quad \|Tx - Tx_0\| \geq \varepsilon.$$

In particular, for $\delta = 1/n$ there is an x_n satisfying

$$\|x_n - x_0\| < 1/n \quad \text{but} \quad \|Tx_n - Tx_0\| \geq \varepsilon.$$

Clearly $x_n \to x_0$ but (Tx_n) does not converge to Tx_0. This contradicts $Tx_n \to Tx_0$ and proves the theorem. □

Proposition 2.1.40. Let X and Y be two normed vector spaces and $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, be a linear operator. If for each sequence $\{x_n\} \subset \mathcal{D}(T)$ with $\lim_{n\to\infty} x_n = 0$, we have $\lim_{n\to\infty} Tx_n = 0$ then T is continuous.

Proof. Take any sequence $\{x_n\}$ from $\mathcal{D}(T)$ such that $x_n \to x \in \mathcal{D}(T)$. Let $\{y_n\}$ be a sequence such that $\{x_n - x\} = \{y_n\}$. Then $\{y_n\}$ be in $\mathcal{D}(T)$ since $\mathcal{D}(T)$ be a linear vector space and $y_n \to 0$. Now by assumption $Ty_n \to 0$ then

$$Ty_n = T(x_n - x) = Tx_n - Tx \to 0 \Rightarrow Tx_n \to Tx.$$

Since $\{x_n\}$ was arbitrary T is continuous linear operator on $\mathcal{D}(T)$. □

Theorem 2.1.41. Let X and Y be two normed vector spaces and $T : \mathcal{D}(T) \to Y$, with $\mathcal{D}(T) \subseteq X$, be a linear operator. Then the following two statements are equivalent:

1. T is continuous;

2. T is bounded .

Proof. **(2)** $\Rightarrow$ **(1)** We have $\|Tx\| \leq c\|x\|$ and we want to show that T is continuous i.e., for any $\{x_n\} \to 0$, $\{Tx_n\} \to 0$. Then since

$$\lim_{n\to\infty} \|Tx_n\| \leq c \lim_{n\to\infty} \|x_n\| = 0$$

T is continuous linear operator.

(1) $\Rightarrow$ (2) Suppose T is continuous but unbounded. This means that there exists a sequence $\{x_n\}$ in $\mathcal{D}(T)$, with $\|x_n\| < M$, such that $\|Tx_n\| \geq n\|x_n\|$. Let $y_n \equiv x_n\|Tx_n\|^{-1}$. Then $\{y_n\} \subset \mathcal{D}(T)$ since $\mathcal{D}(T)$ is a linear vector space and $\|y_n\| = \|x_n\|/\|Tx_n\| \leq \frac{1}{n}$. Here $y_n \to 0$ but $\|Ty_n\| = 1$, i.e., $y_n \to 0$ but $\lim_{n\to\infty} Ty_n \neq 0$. This contradicts the continuity of T, so T must be bounded. $\square$

The Riesz's Theorem:

Proposition A.1.4. *[8] For any bounded linear functional $f : \mathcal{H} \to K$, with $\mathcal{D}(f) = \mathcal{H}$ and $K = \mathbb{R}$ or $K = \mathbb{C}$, on a Hilbert Space $\mathcal{H}$, kernel of f,*

$$ker(f) = \{x \in \mathcal{H} | f(x) = 0\},$$

is a closed subspace of codimension one (by codimension of a subspace M in $\mathcal{H}$, we mean dim $M^{\perp}$ in $\mathcal{H}$) provided $f \neq 0$.

Proof. [8] $ker(f)$ is a subset of $\mathcal{H}$, and so vector addition and multiplication by scalars is defined in this set. $ker(f)$ is a vector space, because obviously for any $x, y \in ker(f)$, $x + y \in ker(f)$ and for any scalar α, $\alpha x \in ker(f)$. To show $ker(f)$ is closed, let $\{x_n\} \subset ker(f)$ with $x_n \to x$ in $\mathcal{H}$. Then $f(x) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = 0$ by continuity of f. Note that f is continuous by the Theorem (2.1.41), since it is a bounded linear (operator) functional. Hence $x \in ker(f)$, and since $\{x_n\}$ is arbitrary, $ker(f)$ is closed. For the codimension, assume $f \neq 0$, then $ker(f) \neq \mathcal{H}$. By Projection Theorem (2.1.24), there exists $x_0 \neq 0 \in \mathcal{H}$ such that $x_0 \in ker(f)^{\perp}$ and so $f(x_0) \neq 0$. Then by linearity of f for any $x \in \mathcal{H}$ we have

$$f(x - f(x)f(x_0)^{-1}x_0) = f(x) - f(x) = 0.$$

Hence $(x - f(x)f(x_0)^{-1}x_0) \in ker(f)$ for all $x \in \mathcal{H}$. But if y is any vector in $ker(f)^{\perp}$ where $x_0 \in ker(f)^{\perp}$ then $(y - f(y)f(x_0)^{-1}x_0) \in ker(f)^{\perp}$. Therefore by (2.1.22)

$$y - f(y)f(x_0)^{-1}x_0 = 0$$

which implies that $ker(f)^{\perp}$ is one dimensional. $\square$

Theorem 2.1.47. (Riesz's Theorem)
Every bounded linear functional $f : \mathcal{H} \to K$ on a Hilbert space $\mathcal{H}$ with $\mathcal{D}(f) = \mathcal{H}$ can be represented in terms of the inner product as

$$f(x) = \langle x, z \rangle$$

where z is uniquely determined by f, and satisfies

$$\|z\| = \|f\|.$$

Proof.

1. f has a representation $f(x) = \langle x, z \rangle$.

2. z is unique.

3. $\|z\| = \|f\|$.

If $f = 0$, then take $z = 0$ so $f(x) = \langle x, z \rangle$ and $\|z\| = \|f\|$.
Let $f \neq 0$;

1. [8]What properties z must have if a representation $f(x) = \langle x, z \rangle$ exist.
   $i)$ $z \neq 0$, otherwise $f = 0$.
   $ii)$ $\langle y, z \rangle = 0 \ \ \forall y \in Null(f)$. Then by Proposition (A.1.4) there exists a nonzero $z \in Null(f)^{\perp}$ and we can set

$$z = \frac{\overline{f(z_0)}}{\|z_0\|^2}\, z_0$$

   for a nonzero $z_0 \in Null(f)^{\perp}$. Now given any $x \in \mathcal{H}$,

$$f(x - f(x)f(z_0)^{-1}z_0) = 0, \quad \text{hence} \quad (x - f(x)f(z_0)^{-1}z_0) \in Null(f).$$

   Hence, by computation:

$$\langle x, z \rangle = \frac{f(z_0)}{\|z_0\|^2}\langle x, z_0 \rangle = \frac{f(z_0)}{\|z_0\|^2}\Big\langle x + (-f(x)f(z_0)^{-1}z_0 + f(x)f(z_0)^{-1}z_0), z_0 \Big\rangle$$

$$= \frac{f(z_0)}{\|z_0\|^2}\Big[\ \Big\langle x - f(x)f(z_0)^{-1}z_0, z_0 \Big\rangle + \Big\langle f(x)f(z_0)^{-1}z_0, z_0 \Big\rangle \ \Big]$$

$$\begin{aligned}
&= \frac{f(z_0)}{\|z_0\|^2}\Big[\, 0 + \Big\langle f(x) f(z_0)^{-1} z_0, z_0 \Big\rangle \,\Big] \\
&= \frac{f(z_0)}{\|z_0\|^2}\left[f(x) f(z_0)^{-1} \|z_0\|^2 \right] \\
&= f(x)
\end{aligned}$$

2. [9]Let for all $x \in \mathcal{H}$, $f(x) = \langle x, z_1 \rangle = \langle x, z_2 \rangle$ then $\langle x, z_1 - z_2 \rangle = 0$ for all x. Choosing the particular $x = z_1 - z_2$ we have $\langle z_1 - z_2, z_1 - z_2 \rangle = \|z_1 - z_2\|^2 = 0$. Hence $z_1 = z_2$.

3. [9]Let $f \neq 0$. Then $z \neq 0$. By the equation $f(x) = \langle x, z \rangle$, $f(z) = \langle z, z \rangle = \|z\|^2 \leq \|f\| \; \|z\|$ and so $\|z\| \leq \|f\|$. Moreover $|f(x)| = |\langle x, z \rangle| \leq \|x\| \; \|z\|$ by Schwarz inequality ($|\langle x, y \rangle| \leq \|x\| \|y\|$) and by taking the sup over $\|x\| = 1$, $\|f\| \leq \|z\|$. Hence $\|z\| = \|f\|$.

□

Proposition 2.1.49. Dual space $\mathcal{H}'$ of $\mathcal{H}$ and $\mathcal{H}$ are isomorphic normed vector spaces. One can express this more briefly by saying that dual space $\mathcal{H}'$ is $\mathcal{H}$.

Proof. By Theorem (2.1.47), any f in $\mathcal{H}'$ is defined uniquely by a $z \in \mathcal{H}$ in $f(x) = \langle x, z \rangle$. Conversely for any $z \in \mathcal{H}$ we can define a bounded linear functional on $\mathcal{H}$ such that $\langle x, z \rangle = f(x)$, $x \in \mathcal{H}$. Hence there is a one-to-one correspondence between the elements of $\mathcal{H}'$ and $\mathcal{H}$. These elements are also isometric by the same theorem. Hence $\mathcal{H}'$ and $\mathcal{H}$ are isomorphic normed vector spaces. □

Proof of the Theorem 2.1.51:

Definition A.1.5. *(Sesquilinear form)*
Let X and Y be vector spaces over the same field $K (= \mathbb{R}$ or $\mathbb{C})$. Then the sesquilinear form (or sesquilinear functional) h on $X \times Y$ [3] *is a mapping*

$$h : X \times Y \rightarrow K$$

[3]Set of all ordered pairs (x, y), where $x \in X$ and $y \in Y$.

such that for all $x, x_1, x_2 \in X$ *and* $y, y_1, y_2 \in Y$ *and all scalars* α, β,

$$h(x_1 + x_2, y) = h(x_1, y) + h(x_2, y) \tag{A.7}$$

$$h(x, y_1 + y_2) = h(x, y_1) + h(x, y_2) \tag{A.8}$$

$$h(\alpha x, y) = \alpha h(x, y) \tag{A.9}$$

$$h(x, \beta y) = \beta^* h(x, y). \tag{A.10}$$

Hence h *is linear in the first argument and conjugate linear in the second one. If* X *and* Y *are real* ($K = \mathbb{R}$)*, then (A.10) is simply*

$$h(x, \beta y) = \beta h(x, y)$$

and h *is called bilinear since it is linear in both arguments.*

If X *and* Y *are normed vector spaces and if there is a real number* c *such that for all* x, y

$$|h(x, y)| \leq c\|x\|\|y\|,$$

then h *is said to be bounded, and the number*

$$\|h\| = \sup_{x \in X - \{0\},\ y \in Y - \{0\}} \frac{|h(x, y)|}{\|x\|\|y\|} = \sup_{\|x\|=1,\ \|y\|=1} |h(x, y)| \tag{A.11}$$

is called the norm of h.

Lemma A.1.6. *If* $\langle v_1, w\rangle = \langle v_2, w\rangle$ *for all* w *in an inner product space* X*, then* $v_1 = v_2$*. In particular,* $\langle v_1, w\rangle = 0$ *for all* $w \in X$ *implies* $v_1 = 0$.

Proof. By assumption, for all w,

$$\langle v_1 - v_1, w\rangle = \langle v_1, w\rangle - \langle v_1, w\rangle = 0.$$

For $w = v_1 - v_2$ we obtain $\|v_1 - v_2\|^2 = 0$. Hence $v_1 - v_2 = 0$ and $v_1 = v_2$. In particular, $\langle v_1, w\rangle = 0$ with $w = v_1$ gives $\|v_1\|^2 = 0$, so that $v_1 = 0$. □

Theorem A.1.7. *(Riesz representation)*

Let $\mathcal{H}_1, \mathcal{H}_2$ *be Hilbert spaces and*

$$h : \mathcal{H}_1 \times \mathcal{H}_2 \to K$$

a bounded sesquilinear form. Then h *has a representation*

$$h(x, y) = \langle Sx, y \rangle \tag{A.12}$$

where $S : \mathcal{H}_1 \to \mathcal{H}_2$ *is a bounded linear operator.* S *is uniquely determined by* h *and has norm*

$$\|S\| = \|h\|. \tag{A.13}$$

Proof. We consider $h(x, y)^*$ which is linear in second argument because of complex conjugation. To make Riesz Theorem (2.1.47) applicable, we keep x fixed and define

$$f_x(y) := h(x, y)^* = \langle y, z \rangle.$$

Hence z is uniquely determined by f_x depending on x and

$$h(x, y) = \langle z, y \rangle. \tag{A.14}$$

Then (A.14) with variable x defines an operator

$$S : \mathcal{H}_1 \to \mathcal{H}_2 \quad \text{given by} \quad z = Sx.$$

Substituting $z = Sx$ in (A.14), we have (A.12).

Now we show that S is linear, bounded and $\|S\| = \|h\|$. S is linear since

$$\langle S(\alpha x_1 + \beta x_2), y \rangle = h(\alpha x_1 + \beta x_2, y) = \alpha h(x_1, y) + \beta h(x_2, y)$$

$$= \alpha \langle Sx_1, y \rangle + \beta \langle Sx_2, y \rangle = \langle \alpha Sx_1 + \beta Sx_2, y \rangle$$

for all $y \in \mathcal{H}_2$ and hence by lemma (A.1.6),

$$S(\alpha x_1 + \beta x_2) = \alpha Sx_1 + \beta Sx_2.$$

S is bounded since

$$\|h\| = \sup_{x \neq 0 \in \mathcal{H}_1,\ y \neq 0 \in \mathcal{H}_2} \frac{|\langle Sx, y \rangle|}{\|x\| \|y\|} \geq \sup_{x \neq 0 \in \mathcal{H}_1,\ (Sx \neq 0)} \frac{|\langle Sx, Sx \rangle|}{\|x\| \|Sx\|} = \sup_{x \neq 0 \in \mathcal{H}_1} \frac{\|Sx\|}{\|x\|} = \|S\|$$

by (A.11) and (A.12). Now we just show that $\|h\| \leq \|S\|$ for obtaining (A.13). By Schwarz inequality ($|\langle x, y \rangle| \leq \|x\| \|y\|$):

$$\|h\| = \sup_{x \neq 0 \in \mathcal{H}_1,\ y \neq 0 \in \mathcal{H}_2} \frac{|\langle Sx, y \rangle|}{\|x\| \|y\|} \leq \sup_{x \neq 0 \in \mathcal{H}_1,\ y \neq 0 \in \mathcal{H}_2} \frac{\|Sx\| \|y\|}{\|x\| \|y\|} = \sup_{x \neq 0 \in \mathcal{H}_1} \frac{\|Sx\|}{\|x\|} = \|S\|.$$

S is unique. Assume that there is a linear operator $T : \mathcal{H}_1 \to \mathcal{H}_2$ such that for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$ we have

$$h(x, y) = \langle Sx, y\rangle = \langle Tx, y\rangle.$$

Then for all $x \in \mathcal{H}_1$, $Sx = Tx$ by Lemma (A.1.6) and hence $S = T$ by definition. $\square$

Theorem 2.1.51. (Existence)

The Hilbert-adjoint operator $T^\dagger$ of T in Definition (2.1.50) exists, is unique and is a bounded linear operator with norm

$$\|T^\dagger\| = \|T\|. \tag{2.61}$$

Proof. The formula

$$h(y, x) = \langle y, Tx\rangle \tag{A.15}$$

defines a sesquilinear form on $\mathcal{H}_2 \times \mathcal{H}_1$ since inner product is sequilinear by definition and T is linear. Then h is bounded because

$$|h(y, x)| = |\langle y, Tx\rangle| \leq \|y\|\|Tx\| \leq \|T\|\|x\|\|y\| \tag{A.16}$$

by Schwartz inequality($|\langle x, y\rangle| \leq \|x\|\|y\|$). Equation (A.16) also implies that $\|h\| \leq \|T\|$. Next we show that $\|h\| \geq \|T\|$.

$$\|h\| = \sup_{y\neq 0\in\mathcal{H}_1,\ x\neq 0\in\mathcal{H}_2} \frac{|\langle y, Tx\rangle|}{\|x\|\|y\|} \geq \sup_{(Tx\neq 0),\ x\neq 0\in\mathcal{H}_2} \frac{|\langle Tx, Tx\rangle|}{\|x\|\|Tx\|} = \sup_{x\neq 0\in\mathcal{H}_2} \frac{\|Tx\|}{\|x\|} = \|T\|.$$

Hence,

$$\|h\| = \|T\|.$$

By Theorem (A.1.7) there exists a unique bounded linear operator, denote as $T^\dagger$, for h such that

$$h(y, x) = \langle T^\dagger y, x\rangle \tag{A.17}$$

where $T^\dagger : \mathcal{H}_2 \to \mathcal{H}_1$ has the norm $\|T^\dagger\| = \|h\|$. Then we have

$$\|T^\dagger\| = \|h\| = \|T\|$$

which proves (2.61). Moreover by (A.15) and (A.17) we obtain $\langle y, Tx\rangle = \langle T^\dagger y, x\rangle$ and that gives (2.56) by taking complex conjugates. $\square$

Theorem 2.1.52. Let $\mathcal{H}_1$, $\mathcal{H}_2$ be Hilbert spaces, $S : \mathcal{H}_1 \to \mathcal{H}_2$ and $T : \mathcal{H}_1 \to \mathcal{H}_2$, with $\mathcal{D}(S) = \mathcal{D}(T) = \mathcal{H}_1$, bounded linear operators and α any scalar. Then we have

1. $\langle T^\dagger y, x\rangle = \langle y, Tx\rangle \quad (x \in \mathcal{H}_1, y \in \mathcal{H}_2),$ (2.62)

2. $(\alpha T)^\dagger = \alpha^* T^\dagger,$ (2.63)

3. $(T^\dagger)^\dagger = T,$ (2.64)

4. $\|T^\dagger T\| = \|TT^\dagger\| = \|T\|^2,$ (2.65)

5. $T^\dagger T = 0 \quad \Leftrightarrow \quad T = 0,$ (2.66)

6. $(ST)^\dagger = T^\dagger S^\dagger \quad$ (assuming $\mathcal{H}_2 = \mathcal{H}_1$), (2.67)

where * denotes the complex conjugate of the scalar.

Proof. We need following lemma for the proof;
*Lemma:*Let X and Y be two inner product spaces and $Q : X \to Y$ a bounded linear operator. Then, $Q = 0$ if and only if $\langle Qx, y\rangle = 0$ for all $x \in X$ and $y \in Y$.
Proof: $Q = 0$ means $Qx = 0$ for all x and implies $\langle Qx, y\rangle = \langle 0, y\rangle = 0$ for all y. Conversely, $\langle Qx, y\rangle = 0$ for all x and y implies $Qx = 0$ for all x by lemma (A.1.6) and shows $Q = 0$ by definition.

1. From (2.56) we have,

$$\langle T^\dagger y, x\rangle = \langle x, T^\dagger y\rangle^* = \langle Tx, y\rangle^* = \langle y, Tx\rangle.$$

2. By using (2.62),

$$\langle (\alpha T)^\dagger y, x\rangle = \langle y, (\alpha T)x\rangle = \langle y, \alpha(Tx)\rangle = \alpha^*\langle y, Tx\rangle = \alpha^*\langle T^\dagger y, x\rangle = \langle \alpha^* T^\dagger y, x\rangle.$$

   Then let $Q = (\alpha T)^\dagger - \alpha^* T^\dagger$ and apply the above lemma.

3. By using (2.62) and (2.56), for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$ we have,

$$\langle (T^\dagger)^\dagger x, y\rangle = \langle x, T^\dagger y\rangle = \langle Tx, y\rangle,$$

   then (2.64) follows from the above lemma with $Q = (T^\dagger)^\dagger - T$.

4. Note that $T^\dagger T : \mathcal{H}_1 \to \mathcal{H}_1$ and $TT^\dagger : \mathcal{H}_2 \to \mathcal{H}_2$. By Schwarz inequality $(|\langle x, y\rangle| \leq \|x\|\|y\|)$

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^\dagger Tx, x\rangle \leq \|T^\dagger Tx\|\|x\| \leq \|T^\dagger T\|\|x\|^2.$$

Then $\|Tx\|^2/\|x\|^2 \leq \|T^\dagger T\|$ and $\|Tx\|/\|x\| \leq \sqrt{\|T^\dagger T\|}$ moreover $\sup_{x\neq 0\in\mathcal{H}_1}(\|Tx\|/\|x\|) \leq \sqrt{\|T^\dagger T\|}$ and hence $\|T\| \leq \sqrt{\|T^\dagger T\|}$ which implies

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

By using the inequality $\|T_1T_2\| \leq \|T_1\|\|T_2\|$ for two bounded operators (since $\|T_1T_2x\| \leq \|T_1\|\|T_2x\| \leq \|T_1\|\|T_2\|\|x\|$, $\|T_1T_2x\|/\|x\| \leq \|T_1\|\|T_2\|$ and by taking supremum over $x \neq 0$ we obtain related inequality) and (2.61)

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

Hence $\|T^\dagger T\| = \|T\|^2$. Replace T by $T^\dagger$ and apply (2.64) then by (2.61)

$$\|T^{\dagger\dagger}T^\dagger\| = \|T^\dagger\|^2 \Rightarrow \|TT^\dagger\| = \|T\|^2.$$

5. From (2.65), we easily see (2.66).

6. Repeated application of (2.56) gives

$$\langle x, (ST)^\dagger y\rangle = \langle (ST)x, y\rangle = \langle Tx, S^\dagger y\rangle = \langle x, T^\dagger S^\dagger y\rangle.$$

Then $\langle (ST)^\dagger y, x\rangle = \langle T^\dagger S^\dagger y, x\rangle$ for all $x \in \mathcal{H} = \mathcal{H}_1 = \mathcal{H}_2$ and so by (A.1.6) $(ST)^\dagger y = T^\dagger S^\dagger y$ for all $y \in \mathcal{H}$ which is (2.67) by definition.

□

Proposition 2.1.54. The adjoint $T^\dagger$ of a linear operator $T : \mathcal{D}(T) \to \mathcal{H}$ with $\mathcal{D}(T) \subseteq \mathcal{H}$ and $\mathcal{H}$ a Hilbert space is not well-defined unless T is densely defined, i.e., $\mathcal{D}(T)$ is dense in $\mathcal{H}$.

Proof. First note that $\mathcal{D}(T^\dagger)$ consists of all $y \in \mathcal{H}$ such that there is a $\tilde{y} \in \mathcal{H}$ satisfying $\langle Tx, y\rangle = \langle x, \tilde{y}\rangle$ for all $x \in \mathcal{D}(T)$ and $T^\dagger$ is defined in terms of that $\tilde{y}$ by $\tilde{y} = T^\dagger y$.[4] Assume that $\mathcal{D}(T)$ is not dense in $\mathcal{H}$, i.e., $\overline{D(T)} \neq \mathcal{H}$. Hence the orthogonal complement of $D(T)$ contains a nonzero y_1 such that $\langle x, y_1\rangle = 0$ for all $x \in \mathcal{D}(T)$. Then

$$\langle x, \tilde{y}\rangle = \langle x, \tilde{y}\rangle + \langle x, y_1\rangle = \langle x, \tilde{y} + y_1\rangle,$$

which shows non-uniqueness. On the other hand, if $\mathcal{D}(T)$ is dense in $\mathcal{H}$ then $\mathcal{D}(T)^\perp = \{0\}$. Hence $\langle x, y_1\rangle = 0$ for all $x \in \mathcal{D}(T)$ is just satisfied for $y_1 = 0$. □

Hellinger-Toeplitz Theorem:

Definition A.1.8. *(Category)*

A subset M of a normed vector space X is said to be

(a) rare (or nowhere dense) in X if its closure $\overline{M}$ has no interior points,

(b) meager (or of the first category) in X if M is the union of countably many sets each of which is rare in X,

(c) nonmeager (or of the second category) in X, if M is not meager in X.

Theorem A.1.9. *(Baire's Category Theorem) If a normed vector space $X \neq \emptyset$ is complete, it is nonmeager in itself.*

Hence if $X \neq \emptyset$ is complete and

$$X = \bigcup_{k=1}^{\infty} A_k \qquad (A_k \text{ closed}) \tag{A.18}$$

then at least one A_k contains a nonempty open subset.

Proof. Assume by contradiction that the complete normed vector space $X \neq \emptyset$ is meager in itself. Then

$$X = \bigcup_{k=1}^{\infty} M_k \tag{A.19}$$

with each M_k rare in X. We will construct a Cauchy sequence $\{p_k\}$ whose limit p (which exist by completeness) is in no M_k, hence contradicting the representation (A.19).

[4]See Definition 2.1.53.

By assumption, M_1 is rare in X, so that, by definition, $\overline{M}_1$ does not contain a nonempty open set. But X does (for instance, X itself). This implies $\overline{M}_1 \neq X$. Hence the complement $\overline{M}_1^{\,C} = X \backslash \overline{M}_1$ of $\overline{M}_1$ is not empty and open. We choose a point p_1 in $\overline{M}_1^{\,C}$ and an open ball about it, say,

$$B_1 = B(p_1; \varepsilon_1) \subset \overline{M}_1^{\,C}, \quad \varepsilon_1 < \frac{1}{2}.$$

By assumption, M_2 is rare in X, so that $\overline{M}_2$ does not contain a nonempty open set. Hence it does not contain the open ball $B(p_1; \frac{1}{2}\ \varepsilon_1)$. This implies that $\overline{M}_2^{\,C}$ is not empty and open, so that we may choose an open ball in this set, say,

$$B_2 = B(p_2; \varepsilon_2) \subset \overline{M}_2^{\,C} \cap B(p_1; \frac{1}{2}\ \varepsilon_1), \quad \varepsilon_2 < \frac{1}{2}\ \varepsilon_1.$$

We obtain a sequence of balls

$$B_k = B(p_k; \varepsilon_k), \quad \varepsilon_k < 2^{-k}$$

by induction such that $B_k \cap M_k = \emptyset$ and

$$B_{k+1} \subset B(p_k; \frac{1}{2}\ \varepsilon_k) \subset B_k \quad \text{k=1,2,}\cdots.$$

Since $\varepsilon_k < 2^{-k}$, the sequence $\{p_k\}$ of the centers is Cauchy and converges, say, $p_k \to p \in X$ because X is complete by assumption. Also, for every m and $n > m$ we have $B_n \subset B(p_m; \frac{1}{2}\varepsilon_m)$, so that

$$\|p_m - p\| \leq \|p_m - p_n\| + \|p_n - p\| < \frac{1}{2}\ \varepsilon_m + \|p_n - p\| \to \frac{1}{2}\ \varepsilon_m$$

as $n \to \infty$. Hence $p \in B_m$ for every m. Since $B_m \subset \overline{M}_m^{\,C}$, we now see that $p \notin M_m$ for every m, so that $p \notin \bigcup_{m=1}^{\infty} M_m = X$. This contradicts $p \in X$. Baire's theorem is proved. □

Theorem A.1.10. *(Uniform Boundedness Theorem)*
Let $\{T_n\}$ be a sequence of bounded linear operators $T_n : \mathcal{H} \to \mathcal{H}$, with $\mathcal{D}(T_n) = \mathcal{H}$, such that $\{\|T_n x\|\}$ is bounded for every $x \in \mathcal{H}$, say,

$$\|T_n x\| \leq c_x \quad n = 1, 2, \cdots \tag{A.20}$$

where c_x is a real number. Then the sequence of the norms $\|T_n\|$ is bounded, that is, there is a c such that

$$\|T_n\| \le c \quad n = 1, 2, \cdots. \tag{A.21}$$

Proof. For every $k \in \mathbb{N}$, let $A_k \subset \mathcal{H}$ be the set of all x such that

$$\|T_n x\| \le k \quad \text{ for all } n.$$

Now we show that A_k is closed. Let $\{x_j\}$ be any sequence in A_k such that $x_j \to x$ in $\mathcal{H}$, then we have $\|T_n x_j\| \le k$ for all j and for all n. Then by continuity of norm and continuity of T_n [5] we have $\|T_n x\| \le k$ and hence $x \in A_k$.

By (A.20), each $x \in \mathcal{H}$ belongs to some A_k. Hence

$$\mathcal{H} = \bigcup_{k=1}^{\infty} A_k.$$

Since $\mathcal{H}$ is complete, Baire's Theorem (A.1.9) implies that some A_k contains an open ball, say,

$$B_0 = B(x_0; r) \subset A_{k_0}. \tag{A.22}$$

Let a nonzero $x \in \mathcal{H}$ be arbitrary. We set

$$z = x_0 + \gamma x \qquad \gamma = \frac{r}{2\|x\|}. \tag{A.23}$$

Then $\|z - x_0\| < r$, so that $z \in B_0$. By (A.22) and from the definition of A_{k_0} we have $\|T_n z\| \le k_0$ for all n. Also $\|T_n x_0\| \le k_0$ for all n since $x_0 \in B_0$. From (A.23) we obtain

$$x = \frac{1}{\gamma}(z - x_0).$$

Then for any nonzero $x \in \mathcal{H}$

$$\|T_n x\| = \frac{1}{\gamma}\|T_n(z - x_0)\| \le \frac{1}{\gamma}(\|T_n z\| + \|T_n x_0\|) \le \frac{1}{\gamma}(k_0 + k_0) = \frac{1}{\gamma}2k_0 = \frac{2\|x\|}{r}2k_0 = \frac{4}{r}\|x\|k_0$$

for all n. Hence

$$\|T_n\| = \sup_{\|x\|=1} \|T_n x\| \le \frac{4}{r}k_0$$

for all n, which is of the form (A.21) with $c = \frac{4}{r}k_0$. $\square$

[5] T_n are continuous by Theorem (2.1.41) since they are bounded.

Theorem 2.1.56. (Hellinger-Toeplitz Theorem)
If a linear operator $T : \mathcal{H} \to \mathcal{H}$ is defined on all of a complex Hilbert Space $\mathcal{H}$, and satisfies $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in \mathcal{H}$, then T is bounded.

Proof. Otherwise $\mathcal{H}$ contain a sequence $\{y_n\}$ such that

$$\|y_n\| = 1 \quad \text{and} \quad \|Ty_n\| \to \infty.$$

We consider the functional f_n defined by

$$f_n(x) = \langle Tx, y_n\rangle = \langle x, Ty_n\rangle$$

where $n = 1, 2, \cdots$. Each f_n is defined on all of $\mathcal{H}$ and is linear. For each fixed n, the functional f_n is bounded since the Schwarz inequality $(|\langle x, y\rangle| \leq \|x\|\|y\|)$ gives

$$|f_n(x)| = |\langle x, Ty_n\rangle| \leq \|Ty_n\| \ \|x\|.$$

Moreover, for every fixed $x \in \mathcal{H}$, the sequence $\{f_n(x)\}$ is bounded. Indeed, using Schwarz inequality and $\|y_n\| = 1$, we have

$$|f_n(x)| = |\langle Tx, y_n\rangle| \leq \|Tx\|$$

From this and Uniform Boundedness Theorem (A.1.10), we conclude that $\{\|f_n\|\}$ is bounded, say, $\|f_n\| \leq k$ for all n. This implies that for every $x \in \mathcal{H}$ we have

$$|f_n(x)| \leq \|f_n\| \ \|x\| \leq k\|x\|$$

and, taking $x = Ty_n$, we arrive at

$$\|Ty_n\|^2 = \langle Ty_n, Ty_n\rangle = |f_n(Ty_n)| \leq k\|Ty_n\|.$$

Hence $\|Ty_n\| \leq k$, which contradicts our initial assumption $\|Ty_n\| \to \infty$ and prove the theorem. □

Theorem 2.1.58. If a normed space X is finite dimensional, then every linear operator on X is bounded and linear operators on finite dimensional vector spaces can be represented in terms of matrices.

Proof. Let $dimX = n$ and $\{e_1, \cdots, e_n\}$ a basis for X. We take any $x = \sum_{j=1}^{n} \xi_j e_j$ and consider any linear operator T on X. Since T is linear,

$$\|Tx\| = \|\sum_{j=1}^{n} \xi_j Te_j\| \leq \sum_{j=1}^{n} |\xi_j| \|Te_j\| \leq \max_k \|Te_k\| \sum_{j=1}^{n} |\xi_j|.$$

To the last sum apply Proposition (A.1.2) with $\alpha_j = \xi_j$ and $x_j = e_j$. Then we obtain

$$\sum_{j=1}^{n} |\xi_j| \leq \frac{1}{c} \|\sum_{j=1}^{n} \xi_j e_j\| = \frac{1}{c}\|x\|.$$

Hence T is bounded, since

$$\|Tx\| \leq \gamma \|x\|, \qquad \text{where} \qquad \gamma = \frac{1}{c} \max_k \|Te_k\|.$$

Let $T : X \to Y$ and Y is a vector space over the same field as X and $\{b_1, \cdots, b_r\}$ is a basis for Y. Vectors in bases are arranged in a definite order which is kept fixed. Every $x \in X$ has a unique representation as

$$x = \xi_1 e_1 + \cdots + \xi_n e_n \tag{A.24}$$

and since T is linear, x has the image

$$y = Tx = T(\sum_{j=1}^{n} \xi_j e_j) = \sum_{j=1}^{n} \xi_j Te_j. \tag{A.25}$$

Since the representation (A.24) is unique; T is uniquely determined if the images $y_j = Te_j$ of the n basis vectors $e_1, \cdots, e_n$ are prescribed.

Now y and $y_j = Te_j$ have unique representation in Y such that

$$y = \sum_{k=1}^{r} \eta_k b_k$$

$$Te_j := \sum_{k=1}^{r} \tau_{kj} b_k.$$

Substitution into (A.25) gives

$$y = \sum_{k=1}^{r} \eta_k b_k = \sum_{j=1}^{n} \xi_j Te_j = \sum_{j=1}^{n} \xi_j \sum_{k=1}^{r} \tau_{kj} b_k = \sum_{k=1}^{r} (\sum_{j=1}^{n} \tau_{kj} \xi_j) b_k.$$

Since b_k's are linearly independent, coefficients of each b_k on the left and on the right must be the same, that is

$$\eta_k = \sum_{j=1}^{n} \tau_{kj}\xi_j \qquad k = 1, \cdots, r. \tag{A.26}$$

In these chosen bases, matrix representation of T is,

$$T = (\tau_{kj}).$$

That is, since x has a unique representation in the basis $\{e_1, \cdots, e_n\}$ as $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ and y has a unique representation in the basis $\{b_1, \cdots, b_r\}$ as $\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_r \end{pmatrix}$ we have

$$\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_r \end{pmatrix} = \begin{pmatrix} \tau_{11} & \cdots & \tau_{1n} \\ \vdots & \ddots & \vdots \\ \tau_{r1} & \cdots & \tau_{rn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

for

$$y = Tx.$$

□

Lemma 2.2.7. [19] (Landau)

Let $c = (c_1, c_2, \cdots)$ be any sequence of scalars $c_\alpha \in K$. If

$$\left| \sum_{\alpha=1}^{\infty} c_\alpha a_\alpha^* \right| \leq k\|a\| \tag{2.101}$$

for some $k \in \mathbb{R}$ and for all $a = (a_1, a_2, \cdots)$ in $\mathcal{H}_0$[6] then $c = (c_1, c_2, \cdots)$ is in $\mathcal{H}_0$.

Proof. [7] First define c^n such that

$$c^n := (c_1, c_2, c_3, \cdots, c_n, 0, 0, \cdots)$$

[6] See Definition (2.2.5).

[7] Thanks to Prof. Ali Ülger for his suggestion for the proof.

which is in $\mathcal{H}_0$ obviously, since finite sum $\sum_{\alpha=1}^{n} |c_\alpha|^2 < \infty$. Then define the sequence of linear functionals $f_n : \mathcal{H}_0 \to \mathbb{C}$, with $\mathcal{D}(f_n) = \mathcal{H}_0$, according to

$$f_n(a) = \sum_{\alpha=1}^{n} c_\alpha a_\alpha^*.$$

Then by Schwarz inequality ($|\langle x, y\rangle| \leq \|x\|\|y\|$)

$$\Big|\sum_{\alpha=1}^{n} c_\alpha a_\alpha^*\Big| \leq \sqrt{(\sum_{\alpha=1}^{n} |c_\alpha|^2)(\sum_{\alpha=1}^{n} |a_\alpha|^2)}.$$

Now define $k_n := (\sum_{\alpha=1}^{n} |c_\alpha|^2)^{1/2}$. Noting that $(\sum_{\alpha=1}^{n} |a_\alpha|^2)^{1/2} \leq \|a\|$ we obtain

$$|f_n(a)| = \Big|\sum_{\alpha=1}^{n} c_\alpha a_\alpha^*\Big| \leq k_n\|a\|$$

for all $a \in \mathcal{H}_0$. Hence $\{f_n\}$ is a sequence of bounded linear functionals. Observe that $f_n \to \sum_{\alpha=1}^{\infty} c_\alpha a_\alpha^*$ as $n \to \infty$, and for every $a \in \mathcal{H}_0$

$$\lim_{n\to\infty} |f_n(a)| \leq k\|a\|$$

by (2.101) and continuity of the norm. Then by Uniform Boundedness Theorem (A.1.10) there is a $l \in \mathbb{R}$ such that

$$\|f_n\| \leq l \quad \text{for all } n \in \mathbb{Z}^+.$$

Next, note that

$$\|c^n\|^2 = f_n(c^n) \leq \|f_n\|\|c^n\| \quad \Rightarrow \quad \|c^n\| \leq \|f_n\|.$$

Furthermore,

$$\|f_n\| = sup_{a\neq 0\in\mathcal{H}_0}\frac{|f_n(a)|}{\|a\|} = sup_{a\neq 0\in\mathcal{H}_0}\frac{|\langle c^n, a\rangle|}{\|a\|} \leq sup_{a\neq 0\in\mathcal{H}_0}\frac{\|c^n\|\|a\|}{\|a\|} = \|c^n\|,$$

by Schwarz inequality. Hence

$$\|f_n\| = \|c^n\|$$

and

$$\lim_{n\to\infty} \|c^n\| = \|c\| = (\sum_{\alpha=1}^{\infty} |c_\alpha|^2)^{1/2} \leq l.$$

This shows that

$$\sum_{\alpha=1}^{\infty} |c_\alpha|^2 < \infty$$

and $c \in \mathcal{H}_0$. □

A.2 Proofs of Theorems in Chapter 4.

Positive Square Root of a Positive Operator:

Let $T_1 : \mathcal{H} \to \mathcal{H}$ and $T_2 : \mathcal{H} \to \mathcal{H}$ be two bounded Hermitian linear operators on a Hilbert space $\mathcal{H}$. Note that, in the following

$$\langle T_1 x, x\rangle \leq \langle T_2 x, x\rangle \ \ \forall\, x \in \mathcal{H} \ \text{ written as } \ T_1 \leq T_2,$$

$$\langle Tx, x\rangle \geq 0 \ \ \forall\, x \in \mathcal{H} \ \text{ written as } \ T \geq 0 \ \text{ and}$$

$$T_1 \leq T_2 \quad \Leftrightarrow \quad 0 \leq T_2 - T_1.$$

Theorem A.2.1. *(Product of positive operators)*
If two bounded Hermitian linear operators S and T on a Hilbert space $\mathcal{H}$ are positive and commute, then their product ST is positive.

Proof. We must show that $\langle STx, x\rangle \geq 0$ for all $x \in \mathcal{H}$. If $S = 0$, this holds. Let $S \neq 0$. We proceed in two steps (a) and (b):

(a) We consider

$$S_1 = \frac{1}{\|S\|} S, \qquad S_{n+1} = S_n - S_n^2 \quad (n = 1, 2, \cdots) \tag{A.27}$$

and prove by induction that

$$0 \leq S_n \leq \mathbb{I}. \tag{A.28}$$

(b) We prove that $\langle STx, x\rangle \geq 0$ for all $x \in \mathcal{H}$.

The details are as follows:

(a) For $n = 1$ the inequality (A.28) holds. Since $0 \leq S$ implies $0 \leq S_1$ and

$$\langle S_1 x, x\rangle = \frac{1}{\|S\|}\langle Sx, x\rangle \leq \frac{1}{\|S\|}\|Sx\|\|x\| \leq \|x\|^2 = \langle \mathbb{I}x, x\rangle$$

shows that $S_1 \leq \mathbb{I}$ by using Schwarz inequality ($|\langle x, y\rangle| \leq \|x\|\|y\|$) and the inequality $\|Sx\| \leq \|S\|\|x\|$.

Now assume that (A.28) holds for $n = k$, that is,

$$0 \leq S_k \leq \mathbb{I}, \qquad \text{and hence} \qquad 0 \leq \mathbb{I} - S_k \leq \mathbb{I}.$$

Then since S_k is Hermitian, for every $x \in \mathcal{H}$ we have

$$\langle S_k^2(\mathbb{I} - S_k)x, x\rangle = \langle(\mathbb{I} - S_k)S_k x, S_k x\rangle = \langle(\mathbb{I} - S_k)y, y\rangle \geq 0$$

where we set $y = S_k x$. By definition, this proves

$$S_k^2(\mathbb{I} - S_k) \geq 0.$$

Similarly,

$$S_k(\mathbb{I} - S_k)^2 \geq 0.$$

By addition and simplication,

$$0 \leq S_k^2(\mathbb{I} - S_k) + S_k(\mathbb{I} - S_k)^2 = S_k^2 - S_k^3 + S_k^3 - 2S_k^2 + S_k = S_k - S_k^2 = S_{k+1}.$$

Hence $0 \leq S_{k+1}$ and since $S_k^2 \geq 0$ and $\mathbb{I} - S_k \geq 0$ we have

$$0 \leq \mathbb{I} - S_k + S_k^2 = \mathbb{I} - S_{k+1}$$

which implies $S_{k+1} \leq \mathbb{I}$. This completes the inductive proof of (A.28).

(*b*) We now show that $\langle STx, x\rangle \geq 0$ for all $x \in \mathcal{H}$. From (A.27) we obtain successively

$$S_1 = S_1^2 + S_2 = S_1^2 + S_2^2 + S_3 = \cdots = S_1^2 + S_2^2 + \cdots + S_n^2 + S_{n+1}.$$

Since $S_{n+1} \geq 0$ we have

$$S_1^2 + \cdots + S_n^2 = S_1 - S_{n+1} \leq S_1. \tag{A.29}$$

Then

$$\sum_{j=1}^{n} \|S_j x\|^2 = \sum_{j=1}^{n} \langle S_j x, S_j x\rangle = \sum_{j=1}^{n} \langle S_j^2 x, x\rangle \leq \langle S_1 x, x\rangle$$

by Hermiticity of S_j.

Since n is arbitrary, the infinite series $\|S_1 x\|^2 + \|S_2 x\|^2 + \cdots$ converges. Hence $\|S_n x\| \to 0$ and $S_n x \to 0$. By (A.29),

$$\Big(\sum_{j=1}^{n} S_j^2\Big)x = (S_1 - S_{n+1})x \;\to\; S_1 x \quad (n \to \infty). \tag{A.30}$$

All the S_j's commute with T since they are sums and products of $S_1 = \|S\|^{-1}S$ and S and T commute. Using $S = \|S\|S_1$, (A.30), inequality $T \geq 0$ and the continuity of the inner product, we obtain for every $x \in \mathcal{H}$

$$\langle STx, x\rangle = \|S\|\langle TS_1x, x\rangle = \|S\| \lim_{n\to\infty} \sum_{j=1}^{n} \langle TS_j^2x, x\rangle = \|S\| \lim_{n\to\infty} \sum_{j=1}^{n} \langle Ty_j, y_j\rangle \geq 0$$

where we set $y_j = S_jx$. □

Definition A.2.2. *(Monotone Sequence)*

A monotone sequence (T_n) of bounded Hermitian linear operators T_n on a Hilbert space $\mathcal{H}$ is either monotone increasing, that is,

$$T_1 \leq T_2 \leq T_3 \leq \cdots$$

or monotone decreasing, that is,

$$T_1 \geq T_2 \geq T_3 \geq \cdots.$$

Theorem A.2.3. *(Monotone Sequence)*

Let (T_n) be a sequence of bounded Hermitian linear operators on a complex Hilbert space $\mathcal{H}$ such that

$$T_1 \leq T_2 \leq \cdots \leq T_n \leq \cdots \leq K \tag{A.31}$$

where K is a bounded Hermitian linear operator on $\mathcal{H}$. Suppose that any T_j commutes with K and with every T_m. Then (T_n) is strongly operator convergent ($T_nx \to Tx$ for all $x \in \mathcal{H}$) and the limit operator T is linear, bounded and Hermitian and satisfies $T \leq K$.

Proof. Let $S_n = K - T_n$ and prove:

(a) The sequence $(\langle S_n^2x, x\rangle)$ converges for every $x \in \mathcal{H}$.

(b) $T_nx \to Tx$ where T is linear and Hermitian. Moreover T is bounded by Uniform Boundedness Theorem (A.1.10).

The details are as follows:

(a) Clearly, S_n is Hermitian. We have

$$S_m^2 - S_nS_m = (S_m - S_n)S_m = (T_n - T_m)(K - T_m).$$

Let $m < n$. Then $T_n - T_m$ and $K - T_m$ are positive by (A.31). Since these operators commute, their product is positive by Theorem (A.2.1). Hence $S_m^2 - S_n S_m \geq 0$, that is, $S_m^2 \geq S_n S_m$ for $m < n$. Similarly,

$$S_n S_m - S_n^2 = S_n(S_m - S_n) = (K - T_n)(T_n - T_m) \geq 0,$$

so that $S_n S_m \geq S_n^2$. Together,

$$S_m^2 \geq S_n S_m \geq S_n^2 \qquad (m < n).$$

By definition, using the Hermiticity of S_n, we have

$$\langle S_m^2 x, x\rangle \geq \langle S_n S_m x, x\rangle \geq \langle S_n^2 x, x\rangle = \|S_n x\|^2 \geq 0. \tag{A.32}$$

This shows that $(\langle S_n^2 x, x\rangle)$ with fixed x is a monotone decreasing sequence of nonnegative numbers. Hence it converges.

(b) We show that $(T_n x)$ converges. By assumption, every T_n commutes with every T_m and with K. Hence the S_j's all commute. These operators are Hermitian. Since $-2\langle S_m S_n x, x\rangle \leq -2\langle S_n^2 x, x\rangle$ by (A.32), where $m < n$, we obtain

$$\|S_m x - S_n x\|^2 = \langle (S_m - S_n)x, (S_m - S_n)x\rangle = \langle (S_m - S_n)^2 x, x\rangle$$

$$= \langle S_m^2 x, x\rangle - 2\langle S_m S_n x, x\rangle + \langle S_n^2 x, x\rangle \leq \langle S_m^2 x, x\rangle - \langle S_n^2 x, x\rangle.$$

From this and the convergence proved in part (a) we see that $(S_n x)$ is Cauchy. It converges since $\mathcal{H}$ is complete. Since $T_n = K - S_n$, $(T_n x)$ is also converges. Clearly the limit depends on x, so that we can write $T_n x \to Tx$ for every $x \in \mathcal{H}$. Hence this defines an operator $T : \mathcal{H} \to \mathcal{H}$ which is linear. T is Hermitian because T_n is Hermitian and the inner product is continuous. Since $(T_n x)$ converges, it is bounded for every $x \in \mathcal{H}$. Then Uniform Boundedness Theorem (A.1.10) implies that T is bounded. Finally, $T \leq K$ since $T_n \leq K$. □

Theorem 4.2.2. Every positive bounded Hermitian linear operator $T : \mathcal{H} \to \mathcal{H}$ on a complex Hilbert space $\mathcal{H}$ has a unique positive square root A. This operator A commutes with every bounded linear operator on $\mathcal{H}$ which commutes with T.

Proof. We proceed in three steps:

(a) We show that if the theorem holds under the additional assumption $\langle Tx, x\rangle \leq \langle \mathbb{I}x, x\rangle$ for all $x \in \mathcal{H}$, (i.e $T \leq \mathbb{I}$), it also holds without that assumption.

(b) We obtain the existence of the operator $A = T^{1/2}$ from $A_n x \to Ax$, where $A_0 = 0$ and

$$A_{n+1} = A_n + \frac{1}{2}(T - A_n^2), \qquad n = 0, 1, \cdots, \tag{A.33}$$

and we also prove the commutativity stated in the theorem.

(c) We prove uniqueness of the positive square root.

(a) If $T = 0$, we can take $A = T^{1/2} = 0$. Let $T \neq 0$. By the Schwarz inequality $(|\langle x, y\rangle| \leq \|x\|\|y\|)$,

$$\langle Tx, x\rangle \leq \|Tx\|\|x\| \leq \|T\|\|x\|^2.$$

Dividing by $\|T\| \neq 0$ and set $Q = (1/\|T\|)T$, we obtain

$$\langle Qx, x\rangle \leq \|x\|^2 = \langle \mathbb{I}x, x\rangle.$$

Assuming that Q has a unique positive square root $B = Q^{1/2}$ we see that a square root of $T = \|T\|Q$ is $\|T\|^{1/2}B$ since

$$(\|T\|^{1/2}B)^2 = \|T\|B^2 = \|T\|Q = T.$$

Also, it is easy to see that the uniqueness of B implies uniqueness of the positive square root of T.

Hence if we can prove the theorem under the additional assumption $T \leq \mathbb{I}$, we are done.

(b) We consider (A.33). Since $A_0 = 0$, we have $A_1 = \frac{1}{2}T$, $A_2 = T - \frac{1}{8}T^2$, etc. Each A_n is a polynomial in T. Hence the A_n's are Hermitian and all commute, and they also commute with every operator that T commutes with. We now prove

$$A_n \leq \mathbb{I} \qquad n = 0, 1, \cdots; \tag{A.34}$$

$$A_n \leq A_{n+1} \qquad n = 0, 1, \cdots; \tag{A.35}$$

$$A_n x \to Ax, \qquad A = T^{1/2}; \tag{A.36}$$

$$ST = TS \quad \Rightarrow \quad AS = SA \tag{A.37}$$

where S is a bounded linear operator on $\mathcal{H}$.

Proof of (A.34):

We have $A_0 \leq \mathbb{I}$. Let $n > 0$. Since $\mathbb{I} - A_{n-1}$ is Hermitian, $(\mathbb{I} - A_{n-1})^2 \geq 0$ [8]. Also $\langle Tx, x\rangle \leq \langle \mathbb{I}x, x\rangle \ \ \forall x \in \mathcal{H} \Rightarrow \ \langle (\mathbb{I} - T)x, x\rangle \geq 0 \ \ \forall x \in \mathcal{H}$, i.e. $\mathbb{I} - T \geq 0$. From this and (A.33) we obtain (A.34):

$$0 \leq \frac{1}{2}(\mathbb{I} - A_{n-1})^2 + \frac{1}{2}(\mathbb{I} - T) = \mathbb{I} - A_{n-1} - \frac{1}{2}(T - A_{n-1}^2)$$

$$= \mathbb{I} - A_{n-1} - (A_n - A_{n-1}) = \mathbb{I} - A_n.$$

Proof of (A.35):

We use induction. (A.33) gives $0 = A_0 \leq A_1 = \frac{1}{2}T$. We show that if $A_{n-1} \leq A_n$ for any fixed n then $A_n \leq A_{n+1}$. From (A.33) we calculate directly

$$A_{n+1} - A_n = A_n + \frac{1}{2}(T - A_n^2) - A_{n-1} - \frac{1}{2}(T - A_{n-1}^2) = A_n - A_{n-1} - \frac{1}{2}(A_n^2 - A_{n-1}^2)$$

$$= (A_n - A_{n-1})[\mathbb{I} - \frac{1}{2}(A_n + A_{n-1})]$$

Here $A_n - A_{n-1} \geq 0$ by assumption and $[\mathbb{I} - \frac{1}{2}(A_n + A_{n-1})] \geq 0$ by (A.34). Hence $A_{n+1} - A_n \geq 0$ by Theorem (A.2.1).

Proof of (A.36):

(A_n) is monotone by (A.35) and $A_n \leq \mathbb{I}$ by (A.34). Hence Theorem (A.2.3) implies the existence of a bounded Hermitian linear operator A such that $A_n x \rightarrow Ax$ for all $x \in \mathcal{H}$. Since $(A_n x)$ converges, (A.33) gives

$$A_{n+1}x - A_n x = \frac{1}{2}(Tx - A_n^2 x) \rightarrow 0$$

as $n \rightarrow \infty$. Hence $Tx - A^2 x = 0$ for all x, that is, $T = A^2$. Also $A \geq 0$ because $0 = A_0 \leq \cdots \leq A_n \leq A_{n+1} \leq \cdots$ by (A.35), that is, $\langle A_n x, x\rangle \geq 0$ for all $n \in \mathbb{N}$ and for every $x \in \mathcal{H}$, which implies $\langle Ax, x\rangle \geq 0$ for every $x \in \mathcal{H}$ by continuity of the inner product.

[8] If $T : \mathcal{H} \rightarrow \mathcal{H}$ is a bounded Hermitian linear operator then T^2 is positive since $\langle T^2 x, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2 \geq 0 \ \forall x \in \mathcal{H}$.

Proof of (A.37):

Obviously if T commutes with S then any A_n also commutes with S. That is $A_nSx = SA_nx$ for all $x \in \mathcal{H}$. Letting $n \to \infty$, we obtain (A.37).

(c) Let both A and B be positive square roots of T. Then $A^2 = B^2 = T$. Also $BT = BB^2 = B^2B = TB$, so that $AB = BA$ by (A.37). Let $x \in \mathcal{H}$ be arbitrary and $y = (A - B)x$. Here $\langle Ay, y\rangle \geq 0$ and $\langle By, y\rangle \geq 0$ because $A \geq 0$ and $B \geq 0$. Using $AB = BA$ and $A^2 = B^2$, we obtain

$$\langle Ay, y\rangle + \langle By, y\rangle = \langle (A + B)y, y\rangle = \langle (A^2 - B^2)x, y\rangle = 0.$$

Hence $\langle Ay, y\rangle = \langle By, y\rangle = 0$. Since $A \geq 0$ and A is Hermitian, it has itself a positive square root C, that is, $C^2 = A$ and C is Hermitian. We then obtain

$$0 = \langle Ay, y\rangle = \langle C^2y, y\rangle = \langle Cy, Cy\rangle = \|Cy\|^2$$

and $Cy = 0$. Also $Ay = C^2y = C(Cy) = 0$. Similarly, $By = 0$. Hence $(A-B)y = 0$. Using $y = (A - B)x$, we have for all $x \in \mathcal{H}$

$$\|Ax - Bx\|^2 = \langle (A - B)^2x, x\rangle = \langle (A - B)y, x\rangle = 0.$$

This shows that $Ax - Bx = 0$ for all $x \in \mathcal{H}$ and proves that $A = B$. □

Theorem 4.2.7. [2] Let $\mathcal{M}$ be a set of Hermitian bounded operators on the Hilbert space $\mathcal{H}$. The following conditions are equivalent.

1. $\mathcal{M}$ is irreducible;

2. the commutant $\mathcal{M}'$ of $\mathcal{M}$ consist of multiples of the identity operator;

3. every nonzero vector $\psi \in \mathcal{H}$ is cyclic for $\mathcal{M}$ in $\mathcal{H}$, or $\mathcal{M} = 0$ and $\mathcal{H} = \mathbb{C}$.

Proof. [2]

(1) $\Rightarrow$ **(3)** Assume there is a nonzero ψ such that $Span\{A\psi; A \in \mathcal{M}\}$ is not dense in $\mathcal{H}$. The orthogonal complement of this set then contains at least one nonzero vector and is invariant under $\mathcal{M}$ (unless $\mathcal{M} = \{0\}$ and $\mathcal{H} = \mathbb{C}$), and this contradicts condition (1).

(3) $\Rightarrow$ **(2)** If $T \in \mathcal{M}'$ then $T^\dagger \in \mathcal{M}'$ ($AT = TA \Rightarrow T^\dagger A = AT^\dagger \ \forall A \in \mathcal{M}$) and, furthermore $T + T^\dagger \in \mathcal{M}'$ and $(T - T^\dagger)/i \in \mathcal{M}'$ (they are Hermitian). Thus if $\mathcal{M}' \neq \mathbb{C}\mathbb{I}$ then there is a Hermitian operator $S \in \mathcal{M}'$ such that $S \neq \lambda\mathbb{I}$ for any $\lambda \in \mathbb{C}$. As all bounded functions of S must also be in the commutant, one deduces that the spectral projectors of S (projection to its eigenspace) also commute with $\mathcal{M}$. But if E is any such projector and ψ a vector in the range of E then $\psi = E\psi$ cannot be cyclic ($AE = EA \Rightarrow A\psi = E(A\psi) \ \forall A \in \mathcal{M}$) and condition (3) is false.

(2) $\Rightarrow$ **(1)** If condition (1) is false then there exist a closed subspace $\mathcal{Y}$ of $\mathcal{H}$ which is invariant under $\mathcal{M}$. But then (the projection to subspace $\mathcal{Y}$) $P_{\mathcal{Y}} \in \mathcal{M}'$ and condition (2) is false.

$\square$

BIBLIOGRAPHY

[1] A. Bohm, "Quantum Mechanics: Foundations and Applications." Springer-Verlag New York, Inc., (1993).

[2] O. Bratteli, D.W. Robinson, "Operator Algebras and Quantum Statistical Mechanics." Vol. 1, Springer-Verlag, New York, (1987).

[3] G. Chen, D.A. Church, B.-G. Englert, and M.S. Zubairy, CRM Proceedings δ Lecture Notes **33**, 79 (2003).

[4] F. Çallialp, N. Kuruoğlu, "Lineer Cebir." Ondokuzmayis Üniversitesi Yayinlari, Samsun, (1990).

[5] J.W. Dettman, "Mathematical Methods in Physics and Engineering." McGraw-Hill, Inc., 1969.

[6] J.B. Fraleigh, "A Firs Course in Abstract Algebra." 5^{th} edition, Addison-Wesley, (1994).

[7] H.B. Geyer, F.G. Scholtz, I. Snyman, Czechoslovak Journal of Physics **54**, 1069 (2004).

[8] P.D. Hislop, I.M. Sigal, "Introduction to Spectral Theory." Springer-Verlag New York, Inc., 1996.

[9] E. Kreyszig, "Introductory Functional Analysis with Applications." John Wiley δ Sons, Inc., 1989.

[10] A. Kvitsimsky and S. Putterman, J. Math. Phys. **32**, 1403 (1991).

[11] A. Mostafazadeh, J. Math. Phys. **43**, 205 (2002).

[12] A. Mostafazadeh, J. Math. Phys. **43**, 2814 (2002).

[13] A. Mostafazadeh, J. Math. Phys. **43**, 3944 (2002).

[14] A. Mostafazadeh, J. Math. Phys. **44**, 974 (2003).

[15] A. Mostafazadeh, Class. Quantum Grav. **20**, 155 (2003).

[16] A. Mostafazadeh and Seher Özçelik, ”Explicit Realization of Pseudo-Hermitian Quantum Mechanics for Two-Level Systems.”

[17] M. Pedersen, ”Functional Analysis in Applied Mathematics and Engineering.” Chapman and Hall/CRC Press, (2000).

[18] J.J. Sakurai, ”Modern Quantum Mechanics Revised Edition.” Addison-Wesley Publishing Company, Inc., (1994).

[19] W. Schmeidler, ”Linear Oprators in Hilbert Space.” Academic Press Inc., 1965.

[20] F. G. Scholtz, H. B. Geyer, and F. J. W. Hahne, Annals of Physics **213**, 74 (1992)

[21] M.S. Swanson, J. Math. Phys. **45**, 585 (2004).

[22] L.N. Trefethen, D. Bau, ”Numerical Linear Algebra.” Society for Industrial and Applied Mathematics, Philadelphia, (1997).