Spectral Theory of Unbounded Self-adjoint Operators in Hilbert spaces

Timo Weckman
013626675
In this work we present a derivation of the spectral theorem of unbounded spectral operators in a Hilbert space. The spectral theorem has several applications, most notably in the theory of quantum mechanics. The theorem allows a self-adjoint linear operator on a Hilbert space to be represented in terms of simpler operators, projections.

The focus of this work are the self-adjoint operators in Hilbert spaces. The work begins with the introduction of vector and inner product spaces and the definition of the complete inner product space, the Hilbert space. Three classes of bounded linear operators relevant for this work are introduced: self-adjoint, unitary and projection operators. The unbounded self-adjoint operator and its properties are also discussed. For the derivation of the spectral theorem, the basic spectral properties of operators in Hilbert space are presented.

The spectral theorem is first derived for bounded operators. With the definition of basic spectral properties and the introduction of the spectral family, the spectral theorem for bounded self-adjoint operators is presented with a proof. Using Weckens lemma, the spectral theorem can be written for the special class of unitary operators.

Using the spectral theorem for unitary operators, we can write the spectral theorem of unbounded self-adjoint operators. Using the Cayley transform, the unbounded self-adjoint operator is rewritten in terms of bounded unitary operators and the spectral theorem is presented in the most general form. In the last section of the thesis, the application of the above results in quantum mechanics is briefly discussed.
Contents

1 Introduction 1

2 Operators in Hilbert spaces 2
   2.1 Vector spaces ..................................................... 2
   2.2 Inner product spaces .............................................. 5
   2.3 Bounded linear operators on Hilbert space ...................... 5
   2.4 Projection operators ............................................. 12

3 Unbounded operators in Hilbert Space 13

4 Spectral families and spectral theorem for bounded self-adjoint operators 16
   4.1 Basic spectral properties ........................................ 17
   4.2 Spectral family .................................................. 23
   4.3 Spectral theorem for bounded linear operators ................ 25

5 Unitary operators, Cayley transform and the general spectral theorem 30
   5.1 Spectral theorem for unitary operators ........................ 30
   5.2 Cayley transform and the general spectral theorem .......... 34

6 Applications in quantum mechanics 38
1 Introduction

The idea of the spectral theory arose from attempts to generalize the finite dimensional eigenvalue theory to infinite dimensions. The spectral theorem allows a self-adjoint linear operator on a Hilbert space to be represented in terms of simple operators, projections. While spectral theory has many applications, it is of special importance in quantum mechanics. With spectral theorem, the complicated quantum mechanical operator can be constructed starting from projectors and the spectra of operators form the sets of possible measurement outcomes of observables. \[1, 2\]

In this work, the spectral representation of a self-adjoint operator on a Hilbert space is constructed. A Hilbert space is a (finite or infinite dimensional) complete inner product space. Self-adjoint operators are linear operators on a complex Hilbert space that map the Hilbert space into itself and are their own adjoint operators. Self-adjoint operators are a fundamental concept in quantum mechanics, where observables are represented as self-adjoint operators in Hilbert spaces.

In the finite dimensional case the spectral theorem reduces to an eigenvalue theorem in a normed space. If an operator $T$ on an $n$-dimensional space has a set of orthonormal eigenvectors $(x_1, \ldots, x_n)$ corresponding to $n$ different eigenvalues $\lambda_1, \ldots, \lambda_n$, then any vector $x$ has a unique representation

$$x = \sum_{i=1}^{n} \langle x, x_i \rangle x_i \quad (1)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. The usefulness of the spectral representation is evident, since the complicated nature of the original operator $T$ is broken down into simple operations where $T$ only acts on the eigenvectors:

$$Tx = \sum_{i=1}^{n} \lambda_i \langle x, x_i \rangle x_i. \quad (2)$$

While the finite dimensional case may be trivial, the infinite-dimensional extension requires a lot of tools to be rigorously constructed.

This work follows a functional analysis pathway to arrive at the spectral theorem. An alternative approach is to use the theory of Lebesgue measure and integration and to express the spectral theorem using projection-valued measures. The latter approach is often used in the literature, especially when applying spectral theorem in quantum mechanics. The main source for this work has been the excellent *Introductory functional analysis with applications* by Erwin Kreyszig\[3\] with additional insights from other sources\[1, 2, 4\].
2 Operators in Hilbert spaces

In this section the key concepts of the spectral theory are presented, starting with vector spaces and the definitions of completeness and the inner product. Also, some necessary properties of bounded linear operators are discussed and several special classes of bounded operators (self-adjoint, unitary and projection operators) used later in this work are introduced.[3, 4]

2.1 Vector spaces

**Definition 2.1** (Metric space) A metric space is a pair \((X, d)\), where \(X\) is a set and \(d\) is a metric on \(X\). The metric \(d\) is a function defined on \(X \times X\) such that for all \(x, y, z \in X\) we have

\[(M1)\] \(d\) is real-valued, finite and non-negative

\[(M2)\] \(d(x, y) = 0 \iff x = y\)

\[(M3)\] \(d(x, y) = d(y, x)\)

\[(M4)\] \(d(x, y) \leq d(x, z) + d(z, y)\)

**Definition 2.2** (Cauchy sequence, completeness) A sequence \((x_n)_{n \in \mathbb{N}}\) in metric space \(X\) is Cauchy, if for all \(\epsilon > 0\) one can find \(N \in \mathbb{N}\) such that \(d(x_n, x_m) < \epsilon\) for all \(n, m > N\). The space \(X\) is said to be complete if every Cauchy sequence in \(X\) converges.

**Theorem 2.3** (Convergent sequence) Every convergent sequence in a metric space is a Cauchy sequence.

**Proof:** If \(x_n \to x\), then for every \(\epsilon > 0\) there exists an \(N\), such that

\[d(x_n, x) < \frac{\epsilon}{2}\]  \hspace{1cm} (3)

for all \(n > N\). Using triangle inequality we can write for \(m, n > N\) that

\[d(x_n, x_m) < d(x_m, x) + d(x_n, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\]  \hspace{1cm} (4)

and hence, \((x_n)\) is Cauchy. \(\Box\)

In the following, unless stated otherwise, \(X\) denotes a vector space over a field \(\mathbb{K} = \mathbb{R}\) or \(\mathbb{C}\). By a vector space we mean a nonempty set \(X\) of elements \(x, y, \ldots\) together with two
algebraic operations called *vector addition* and *multiplication of vectors by scalars* such that \( x + y \in X \) and \( \alpha x \in X \) \( \forall \alpha \in \mathbb{K} \). Let us recall some basic definitions and properties of vector spaces. A subspace \( Y \) of \( X \) is a vector space such that \( Y \subseteq X \) and for every pair \( x, y \in Y \) and \( \alpha, \beta \in \mathbb{K} \), it follows that \( \alpha x + \beta y \in Y \).

A vector space \( X \) is said to be the direct sum of two subspaces \( Y \) and \( Z \) of \( X \),

\[
X = Y \oplus Z
\]  
(5)

if each \( x \in X \) has a unique representation

\[
x = y + z
\]  
(6)

where \( y \in Y \) and \( z \in Z \). Here \( Z \) is called the algebraic complement of \( Y \) in \( X \) and vice versa.

**Theorem 2.4** (Complete subspace) A subspace \( M \) of a complex metric space \( X \) is itself complete if and only if the set \( M \) is closed in \( X \).

**Proof:** Let \( M \) be complete. Then, for every \( x \in \overline{M} \) there is a sequence \( (x_n) \) in \( M \) which converges to \( x \). Since \( (x_n) \) is Cauchy by theorem 2.3 and \( M \) is complete, \( (x_n) \) converges to a unique limit in \( M \) and hence \( x \in M \). This proves that \( M \) is closed as \( x \) was arbitrary.

Conversely, let \( M \) be closed and \( (x_n) \) a Cauchy sequence in \( M \). Then \( x_n \to x \in X \) since \( X \) is complete, which implies that \( x \in \overline{M} = M \) since \( M \) is closed. This means that any arbitrary Cauchy sequence in \( M \) converges meaning that \( M \) is complete. \( \square \)

**Definition 2.5** A subset \( M \) of a metric space \( X \) is said to be

1. rare in \( X \) if its closure \( \overline{M} \) has no interior points
2. meager in \( X \) if \( M \) is the union of countably many sets each of which is rare in \( X \)
3. nonmeager in \( X \) if \( M \) is not meager in \( X \).

A set of vectors \( x_1, \ldots, x_m \) is called *linearly independent* if equation

\[
\alpha_1 x_1 + \ldots + \alpha_m x_m = 0
\]  
(7)

holds only for a \( m \)-tuple of scalars where \( \alpha_1 = \ldots = \alpha_m = 0 \). Otherwise, the set of vectors is linearly dependent. A vector space \( X \) is said to be *finite dimensional* if there is a positive integer \( n \) such that \( X \) contains a linearly independent set of \( n \) vectors whereas
any set of $n+1$ or more vectors of $X$ is linearly dependent. Then $n$ is called the dimension of $X$. If $X$ is not finite dimensional, it is infinite dimensional.

A set of vectors $x_1, \ldots, x_k \in X$ span a subspace $Y \subset X$, if $Y$ consists of all possible linear combinations of the set $x_1, \ldots, x_k$: $Y = \text{span}(x_1, \ldots, x_k) = \{\alpha_1 x_1 + \ldots, \alpha_k x_k |\alpha_1, \ldots, \alpha_k \in \mathbb{R}\}$.

**Definition 2.6 (Normed space)** Let $X$ be a vector space over field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. A map $N : X \to \mathbb{R}$ is called a norm on $X$ and $(X, N)$ is called normed space if

\begin{itemize}
  \item[(N1)] $N(x) \geq 0$ for all $x \in X$
  \item[(N2)] $N(\lambda x) = |\lambda| N(x)$ for any $\lambda \in \mathbb{K}$ and $x \in X$
  \item[(N3)] $N(x + y) \leq N(x) + N(y)$ for any $x, y \in X$
  \item[(N4)] $N(x) = 0 \Rightarrow x = 0$.
\end{itemize}

**Definition 2.7 (Banach space)** A Banach space is a complete normed space, i.e. it is complete in the metric defined by the norm.

**Definition 2.8** Let $X$ and $Y$ be metric spaces. Then $T : D(T) \to Y$ with domain $D(T) \subset X$ is called an open mapping if for every open set in $D(T)$ the image is an open set in $Y$.

**Theorem 2.9** (Bounded inverse theorem) A bounded linear operator $T$ from a Banach space $X$ onto a Banach space $Y$ is an open mapping. Hence, if $T$ is bijective, $T^{-1}$ is bounded.

The proof of this theorem is omitted.

Since the spectral theorem is closely related to (and in finite dimensions reduces to) the eigenvalue decomposition in normed space, we should define the eigenvalue problem. An eigenvalue $\lambda \in \mathbb{C}$ of a square matrix $A \in \mathbb{R}^{n \times n}$ is number that satisfies the eigenvalue equation:

$$Ax = \lambda x$$

for some vector $x \neq 0$. Vector $x$ is called the eigenvector of $A$. Given an eigenvalue $\lambda$, the eigenvectors corresponding to this eigenvalue, together with the null-vector, span a subspace in $X$ called an eigenspace. The set of all eigenvalues is called the spectrum of $A$. 

4
2.2 Inner product spaces

**Definition 2.10 (Inner product)** An inner product on $X$ is a mapping of $X \times X$ into scalar field $\mathbb{K}$ of $X$, that is for every pair $x, y \in X$ there is an associated scalar $\langle x, y \rangle$ called the inner product. For all vectors $x, y, z \in X$ and scalars $\lambda \in \mathbb{K}$ we have

(I1) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

(I2) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$

(I3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$

(I4) $\langle x, x \rangle \geq 0$, $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

The overline in (I3) denotes complex conjugation. The complex conjugate of a complex number $z = a + ib$ is $\overline{z} = a - ib$, where $i$ is the imaginary unit, $i^2 = -1$.

**Definition 2.11 (Inner product space)** An inner product space is a vector space $X$ with an inner product defined on $X$. Given an inner product in a vector space $X$, the expression $\sqrt{\langle x, x \rangle}$ is a norm on $X$. We shall always consider inner product space as a normed space with this norm.

**Definition 2.12 (Orthogonality)** Vectors $x, y$ in an inner product space $X$ are said to be orthogonal if $\langle x, y \rangle = 0$.

**Definition 2.13 (Hilbert space)** A Hilbert space, $H$, is a complete inner product space (complete in the metric defined by the inner product).

2.3 Bounded linear operators on Hilbert space

In this section some general properties of bounded linear operators are presented and two special classes of bounded linear operators, self-adjoint and unitary operators are introduced. For the rest of the text, unless otherwise stated, $X$ and $Y$ denote normed vector spaces and $H$ denotes a complex Hilbert space.

**Definition 2.14 (Linear mapping)** Let $X$ and $Y$ be vector spaces. A mapping $T : X \rightarrow Y$ is linear, if $T(ax + by) = aTx + bTy$ for all scalars $a, b$ and all $x, y \in X$. Linear mappings are usually called linear operators or just operators.

**Definition 2.15 (Linear functional)** A linear operator with a domain in a vector space and range in a scalar field is called a linear functional.
**Definition 2.16** (Boundedness) Let $T$ be an operator $T : X \to Y$. $T$ is bounded, if there is a number $M > 0$ such that $||Tx|| \leq M||x||$ for all $x$.

**Theorem 2.17** A linear operator is bounded if and only if it is continuous.

*Proof:*  
 a) Suppose an operator $T$ is linear and bounded. For any $x, y \in X$,  
$$||Tx - Ty|| = ||T(x - y)|| \leq M||x - y||.$$  
(9)  
Now, for any $\epsilon > 0$, $||Tx - Ty|| < \epsilon$ whenever $||x - y|| < \frac{\epsilon}{M}$, which means that $T$ is continuous.

b) Now suppose that $T$ is linear and continuous. Taking $\epsilon = 1$ in the definition of continuity at $y = 0$, it follows that there exists $\delta > 0$ such that  
$$||Tx|| < 1 \text{ for } ||x|| < \delta.$$  
(10)  
Hence for any $y$,  
$$||Ty|| = \left( \frac{2||y||}{\delta} \right) |T\left( \frac{y\delta}{2||y||} \right)| \leq \frac{2||y||}{\delta}$$  
(11)  
and thus $T$ is bounded with $M = \frac{2}{\delta}$.  

Given normed spaces $X$ and $Y$, the bounded operators $T : X \to Y$ form a normed space $\mathcal{L}(X, Y)$ when endowed with the operator norm  
$$||T|| = \sup_{||x|| \leq 1} ||Tx||_Y.$$  
(12)  
One can show that $||T||$ coincides with the infimum of the number $M$ in definition 2.16. Also, $\mathcal{L}(X, Y)$ is a Banach space whenever $Y$ is complete. With the operator norm, the set of all bounded linear functionals on $X$ constitutes a normed space called the dual space of $X$ and is denoted by $X'$.

**Definition 2.18** A sequence in normed space is said to be strongly convergent if  
$$\exists x \in X : \lim_{n \to \infty} ||x_n - x|| = 0.$$  
(13)  

**Definition 2.19** A sequence in normed space is said to be weakly convergent if  
$$\exists x \in X : \forall f \in X', \lim_{n \to \infty} f(x_n) = f(x).$$  
(14)
Definition 2.20 (Strong convergence of sequence of operators) A sequence \((T_n)\) of bounded operators \(T_n : X \to Y\) is said to be

1. uniformly operator convergent if \((T_n)\) converges in the norm
2. strongly operator convergent if \((T_n x)\) converges strongly in \(Y\) for every \(x \in X\)
3. weakly operator convergent if \((T_n x)\) converges weakly in \(Y\) for every \(x \in X\)

Theorem 2.21 (Baire’s category theorem) If a metric space \(X \neq \emptyset\) is complete, it is a nonmeager in itself. This means that if \(X \neq \emptyset\) and \(X\) is a union countably many closed subsets, i.e.

\[
X = \bigcup_{k=1}^{\infty} A_k
\]

with every \(A_k\) closed, then at least one \(A_k\) contains a nonempty open subset.

Proof: Suppose instead that the complete metric space \(X \neq \emptyset\) was meager in itself. Then

\[
X = \bigcup_{k=1}^{\infty} M_k
\]

with each \(M_k\) rare in \(X\). If we now construct a Cauchy sequence \((x_k)\) whose limit \(x\) (which exists due to completeness) does not exist in any \(M_k\), we arrive at a contradiction and provide a proof for the theorem.

Consider a subset \(M_1\). By our assumption, \(M_1\) is rare in \(X\) and hence \(\overline{M}_1\) does not contain a nonempty open set. Since \(X\) does contain a nonempty open set (e.g. \(X\) contains itself), this means that \(\overline{M}_1 \neq X\) and hence the complement \(\overline{M}_1^c = X - \overline{M}_1\) of \(\overline{M}_1\) is nonempty and open. We may now choose a point \(x_1 \in \overline{M}_1^c\) and an open ball around it, e.g.

\[
B_1 = B(x_1; r_1) \subset \overline{M}_1^c
\]

with \(r < \frac{1}{2}\). However, with the same logic, \(M_2\) is rare in \(X\) and so \(\overline{M}_2\) does not contain a nonempty open set. Hence, it does not contain the open ball \(B\left(x_1, \frac{r_1}{2}\right)\). This implies that \(\overline{M}_2 \cap B\left(x_1, \frac{r_1}{2}\right)\) is not empty and open and so we may choose an open ball in this set, say,

\[
B_2 = B(x_2, r_2) \subset \overline{M}_2^c \cap B\left(x_1, \frac{r_1}{2}\right)
\]

with \(r_2 < \frac{r_1}{4}\). By induction, we obtain a sequence of balls

\[
B_k = B(x_k; r_k)
\]
with $r_k < 2^{-k}$ such that $B_k \cap M_k = \emptyset$ and

$$B_{k+1} \subset B \left( x_k; \frac{r_k}{2} \right) \subset B_k. \quad (20)$$

Since $r_k < 2^{-k}$, the sequence $(x_k)$ of the centers is Cauchy and converges $x_k \to x \in X$ because $X$ is complete. Also, for every $m$ and $n > m$, we have $B_n \subset B \left( x_m; \frac{r_m}{2} \right)$ and so

$$d(x_m, x) \leq d(x_m, x_n) + d(x_n, x) \leq \frac{r_m}{2} + d(x_n, x) \to \frac{r_m}{2} \quad (21)$$

as $n \to \infty$. Hence, $x \in B_m$ for every $m$. Since $B_m \subset \overline{M_m^c}$, we see that $p \notin M_m$ for every $m$ so that $p \notin \bigcup M_m = X$ which contradicts $x \in X$. Hence, the theorem is proved. $\square$

The following famous result is also known as the uniform boundedness theorem.

**Theorem 2.22** (Banach–Steinhaus) Let $(T_n)_{n=1}^{\infty}$ be a sequence of bounded linear operators from Banach space $X$ into a normed space $Y$ such that $(||T_n x||)_{n=1}^{\infty}$ is bounded for every $x \in X$, say,

$$||T_n x|| \leq c_x \quad (22)$$

where $c_x$ is a real number. Then the sequence of norms $||T_n||$ is bounded, i.e.

$$\exists c > 0 : ||T_n|| \leq c. \quad (23)$$

**Proof:** For every $k \in \mathbb{N}$, let $A_k \subset X$ be the set of all $x$ such that

$$||T_n x|| \leq k \quad (24)$$

for all $n$. We shall show that $A_k$ is closed. For every $x \in \overline{A_k}$, there is a sequence $(x_i) \in A_k$ converging to $x$. Therefore, for every fixed $n$ we have $||T_n x_i|| \leq k$ and we obtain $||T_n x|| \leq k$ because $T_n$ and the norm are continuous. Hence, $x \in A_k$ and $A_k$ is closed.

Each $x \in X$ belongs to some $A_k$ by (22) and hence

$$X = \bigcup_{k=1}^{\infty} A_k.$$

Using Baire’s theorem 2.21 and the fact that $X$ is complete, some $A_k$ contains an open ball, say

$$B_0 = B(x_0; r) \subset A_{k_0}. \quad (25)$$

Let us define $z$ using some $0 \neq x \in X$ by

$$z = x_0 + \gamma x \quad \gamma = \frac{r}{2||x||}. \quad (26)$$
Now \( ||x_0 - z|| < r \), so \( z \in B_0 \). From the definition of \( A_k \), we have that \( ||T_n z|| \leq k_0 \) for all \( n \). Also \( ||T_n x_0|| \leq k_0 \) since \( x_0 \in B_0 \). If we write \( x \) as

\[
x = \frac{1}{\gamma} (z - x_0)
\]

we get for all \( n \) that

\[
||T_n x|| = \frac{1}{\gamma} ||T_n (z - x_0)|| \leq \frac{1}{\gamma} (||T_n z|| + ||T_n x_0||) \leq \frac{4}{r} ||x|| k_0.
\]

Thus, for all \( n \),

\[
||T_n|| = \sup_{||x||=1} ||T_n x|| \leq \frac{4}{r} k_0
\]

which gives us (23) with \( c = \frac{4k_0}{r} \). \( \square \)

**Theorem 2.23** (Inverse operator power series) Let \( T : X \rightarrow X \) be a bounded linear operator on a Banach space \( X \). If \( ||T|| < 1 \), then \( (1 - T)^{-1} \) exists as a bounded linear operator on \( X \) that can be written as a Neumann series

\[
(1 - T)^{-1} = I + T + T^2 + \ldots = \sum_{k=0}^{\infty} T^k.
\]

**Definition 2.24** (Hilbert adjoint operator) Let \( T : H_1 \rightarrow H_2 \) be a bounded linear operator. The Hilbert-adjoint operator \( T^* : H_2 \rightarrow H_1 \) is defined to be the operator satisfying

\[
\langle Tx, y \rangle = \langle x, T^* y \rangle
\]

for all \( x \in H_1 \) and \( y \in H_2 \). If \( T \) is bounded, \( T^* \) is also bounded with a norm \( ||T^*|| = ||T|| \) on \( H \) and \( T^* \) is unique.

Hilbert-adjoint operators have the following general properties (stated here without a proof):

**Theorem 2.25** Let \( H_1 \) and \( H_2 \) be Hilbert spaces, \( S : H_1 \rightarrow H_2 \) and \( T : H_1 \rightarrow H_2 \) bounded linear operators and \( \alpha \) any scalar. Then,

(a) \( \langle T^* y, x \rangle = \langle y, Tx \rangle \) \hspace{1cm} (\( x \in H_1, y \in H_2 \))

(b) \( (S + T)^* = S^* + T^* \)

(c) \( (\alpha T)^* = \overline{\alpha} T^* \)
Theorem 2.26 The product of two bounded self-adjoint linear operators $S$ and $T$ on a Hilbert space $H$ is self-adjoint if and only if $S$ and $T$ commute.

Proof: Let us assume $S$ and $T$ commute, $ST = TS$. By the theorem 2.25(g) and from the fact that $S$ and $T$ are self-adjoint, we have

\[(ST)^* = T^*S^* = TS = ST\]  \hspace{1cm} (32)

so $ST$ is self-adjoint.

Conversely, let us assume that the product is self-adjoint, $(ST)^* = ST$. Then

\[ST = (ST)^* = T^*S^* = TS\]  \hspace{1cm} (33)

and hence $S$ and $T$ commute.

Next we shall present two special operator classes: the self-adjoint and unitary operators. The focus of this work are the (bounded and unbounded) self-adjoint operators on Hilbert space. Unitary operators are surjective and isometric operators that preserve the inner product in Hilbert space and are essential in constructing the spectral theorem for unbounded self-adjoint operators.

Definition 2.27 (Self-adjoint and unitary operators) A bounded linear operator $T : H \rightarrow H$ is said to be

self-adjoint if \hspace{1cm} $T^* = T$

unitary if \hspace{1cm} $T^* = T^{-1}$  \hspace{1cm} (34)

Theorem 2.28 Let $T : H \rightarrow H$ be a bounded self-adjoint linear operator. Then,

a) All eigenvalues of $T$ are real, if they exist 

b) Eigenvectors corresponding to numerically different eigenvalues of $T$ are orthogonal to each other.

Proof: a) Let $\lambda$ be any eigenvalue of $T$ and let $x$ be the corresponding eigenvector. Using the self-adjointness of $T$,

\[\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Tx, x \rangle = \langle x, Tx \rangle = \langle x, \lambda x \rangle = \overline{\lambda} \langle x, x \rangle\]  \hspace{1cm} (35)

\[(T^*)^* = T\]
\[
\|T^*T\| = \|TT^*\| = \|T\|^2
\]

\[T^*T = 0 \iff T = 0\]

\[(ST)^* = T^*S^*\]
Since $\langle x, x \rangle \neq 0$ when $x \neq 0$, the last line may be divided by $\langle x, x \rangle$, resulting in $\lambda = \overline{\lambda}$. Hence $\lambda$ is real.

b) Let $\lambda$ and $\mu$ be two eigenvalues of $T$ and let $x$ and $y$ be the corresponding eigenvectors, respectively. Using the fact that $\lambda$ and $\mu$ are real, one obtains

$$\lambda \langle x, y \rangle = \langle \lambda x, y \rangle = \langle x, Ty \rangle = \langle x, \mu y \rangle = \mu \langle x, y \rangle$$  \hspace{1cm} (36)

If $\lambda \neq \mu$, then the equality holds only if $\langle x, y \rangle = 0$.

\[ \square \]

**Theorem 2.29** Let $T : H \to H$ be a bounded linear operator on a Hilbert space $H$. Then,

1. If $T$ is self-adjoint, $\langle Tx, x \rangle$ is real for all $x \in H$.

2. If $H$ is complex and $\langle Tx, x \rangle$ is real for all $x \in H$, then the operator $T$ is self-adjoint.

**Proof:**

1. If $T$ is self-adjoint, then for all $x \in H$,

$$\langle Tx, x \rangle = \langle x, Tx \rangle = \langle Tx, x \rangle$$  \hspace{1cm} (37)

and for $\langle Tx, x \rangle$ to be its own complex conjugate, it must be real.

2. If $\langle Tx, x \rangle$ is real for all $x$, then

$$\langle Tx, x \rangle = \overline{\langle Tx, x \rangle} = \overline{\langle x, T^*x \rangle} = \langle T^*x, x \rangle.$$  \hspace{1cm} (38)

Therefore

$$0 = \langle Tx, x \rangle - \langle T^*x, x \rangle = \langle (T - T^*)x, x \rangle$$  \hspace{1cm} (39)

and $T - T^* = 0$ since $H$ is complex.

An unitary operator has the following general properties:

**Theorem 2.30** Let the operators $U : H \to H$ and $V : H \to H$ be unitary. Then,

(a) $||Ux|| = ||x||$ (in particular $||U|| = 1$, i.e. $U$ is isometric)

(b) $U^{-1} = U^*$

(c) $U^{-1}$ is unitary

(d) $UV$ is unitary

(e) $U$ is normal i.e. $UU^* = U^*U$
2.4 Projection operators

A third special class of bounded operators are the projection operators. Projection operators are essential in the spectral theory as many operators can be written in the spectral representation using projection operators.

First, it is worth noting that a general Hilbert space $H$ can be written as a direct sum of a closed subspace $Y$ and its orthogonal complement

$$Y^\perp = \{ z \in H | z \perp Y \}$$

which is a set of all vectors orthogonal to $Y$. We recall that given such a $Y$, one can define the orthogonal projection $P$ from $H$ onto $Y$, which has the following properties:

1. $||P|| = 1$
2. $Px \in Y$ for all $x \in H$
3. $Px = x$ for all $x \in Y$
4. $Px = 0 \iff x \in Y^\perp$

These properties imply the following result:

Theorem 2.31 Let $Y$ be any closed subspace of a Hilbert space $H$, then

$$H = Y \oplus Z.$$  \hspace{1cm} (41)

where $Z = Y^\perp = \mathcal{N}(P) = \{ x |Px = 0 \}$. The following theorems related to projections are stated without proof.

Theorem 2.32 (Projection operator) A bounded linear operator $P : H \to H$ is an orthogonal projection (or projection) onto some subspace $Y$ if and only if $P$ is self-adjoint and idempotent ($P^2 = P$).

Theorem 2.33 (Products of projections) Let $P_1$ and $P_2$ be projections onto the subspaces $Y_1$ and $Y_2$ of a Hilbert space $H$. Then $P = P_1P_2$ is a projection on $H$ if and only if the projections $P_1$ and $P_2$ commute.

Theorem 2.34 (Partial order) Let $P_1$ and $P_2$ be projections onto the subspaces $Y_1$ and $Y_2$ of a Hilbert space $H$. The following conditions are equivalent:
\( Y_1 \subset Y_2 \)

\( P_2 P_1 = P_1 P_2 = P_1 \)

\( \mathcal{N}(P_2) \subset \mathcal{N}(P_1) \)

\( ||P_1 x|| \leq ||P_2 x|| \)

**Theorem 2.35** (Difference of projections) Let \( P_1 \) and \( P_2 \) be projections defined on a Hilbert space \( H \). Then the difference \( P = P_2 - P_1 \) is a projection on \( H \) if and only if \( Y_1 \subset Y_2 \).

### 3 Unbounded operators in Hilbert Space

The focus of this section are the linear operators \( T : \mathcal{D}(T) \to H \) whose domain of definition \( \mathcal{D}(T) \) is a subspace of a complex Hilbert space \( H \). Such operators are not necessarily bounded and are referred to as unbounded operators. Unbounded operators have many applications, notably in differential equations and quantum mechanics.

For a *bounded* linear operator \( T \) on \( H \), the self-adjointness was defined by

\[
\langle Tx, y \rangle = \langle x, Ty \rangle.
\]  

(42)

However, the following Hellinger–Toeplitz theorem shows that an *unbounded* linear operator that satisfies (42), cannot be defined on all of \( H \).

**Theorem 3.1** (Hellinger–Toeplitz) If a linear operator \( T \) is defined on all of a complex Hilbert space \( H \) and satisfies condition (42), then \( T \) is bounded.

**Proof:** If the statement were not true, \( H \) would contain a sequence \( (y_n) \) such that

\[
||y_n|| = 1 \quad ||Ty_n|| \to \infty.
\]

Consider a functional \( f_n \) defined by

\[
f_n(x) = \langle Tx, y_n \rangle = \langle x, Ty_n \rangle.
\]  

(43)

Each \( f_n \) is defined on all of \( H \) and is linear. For fixed \( n \), the functional \( f_n \) is bounded. This can be observed from the Schwartz inequality:

\[
||f_n(x)|| = ||\langle x, Ty_n \rangle|| \leq ||Ty_n|| ||x||.
\]  

(44)
Similarly, for every fixed $x$, the sequence $f_n(x)$ is bounded:

$$|f_n(x)| = |\langle Tx, y_n \rangle| \leq ||Tx|| ||y_n|| = ||Tx||$$

(45)

since $||y_n|| = 1$. Hence, based on the uniform boundedness theorem 2.22, $||f_n|| \leq k$ for all $n$ where $||f_n||$ denotes the norm of $f_n : H \to \mathbb{C}$ in the dual space $H'$. This implies that for every $x \in H$,

$$|f_n(x)| \leq ||f_n|| ||x|| \leq k ||x||$$

(46)

and, taking $x = Ty_n$, one gets

$$||Ty_n||^2 = \langle Ty_n, Ty_n \rangle = |f_n(Ty_n)| \leq k ||Ty_n||$$

(47)

and thus $||Ty_n|| \leq k$, which contradicts the assumption $||Ty_n|| \to \infty$ and proves the theorem.

Since the entire Hilbert space cannot be the domain of an unbounded operator, a suitable domain needs to be specified for the existence of an adjoint operator.

**Definition 3.2** Let $T : \mathcal{D}(T) \to H$ be a (unbounded) densely defined linear operator (i.e. $\mathcal{D}(T)$ is dense) in a complex Hilbert space $H$. The Hilbert adjoint operator $T^* : \mathcal{D}(T^*) \to H$ of $T$ is defined in the following way: the domain $\mathcal{D}(T^*)$ consists of all $y \in H$ such that there is a $y^* \in H$ satisfying

$$\langle Tx, y \rangle = \langle x, y^* \rangle \ \forall \ x \in H.$$

(48)

We then define $T^* y = y^*$.

Often unbounded linear operators in applications are closed or at least have a linear extension which is closed.

**Definition 3.3** Let $T : \mathcal{D}(T) \to H$ be a linear operator, where $\mathcal{D}(T) \subset H$. Then $T$ is called a closed linear operator if its graph

$$\mathcal{G}(T) = \{(x, y) | x \in \mathcal{D}(T), y = Tx\}$$

(49)

is closed in $H \times H$, where the norm on $H \times H$ is defined by

$$|| (x, y) || = (||x||^2 + ||y||^2)^{1/2}$$

(50)

and results from the inner product defined by

$$\langle (x_1, y_1), (x_2, y_2) \rangle = \langle x_1, x_2 \rangle + \langle y_1, y_2 \rangle.$$

(51)
**Theorem 3.4** Let $T : \mathcal{D}(T) \to H$ be a linear operator, where $\mathcal{D}(T) \subset H$. Then:

(a) $T$ is closed if and only if
\[
x_n \to x \quad [x_n \in \mathcal{D}(T)] \quad \text{and} \quad Tx_n \to y
\]

(52) together imply that $x \in \mathcal{D}(T)$ and $Tx = y$.

(b) If $T$ is closed and $\mathcal{D}(T)$ is closed, then $T$ is also bounded.

(c) Let $T$ be bounded. Then $T$ is closed if and only if $\mathcal{D}(T)$ is a closed subspace of $H$.

**Proof:**

(a) By definition, $\mathcal{G}(T)$ is closed if and only if $z = (x, y) \in \overline{\mathcal{G}(T)}$ implies that $z \in \mathcal{G}(T)$. But $z \in \overline{\mathcal{G}(T)}$ if and only if there are $z_n = (x_n, Tx_n) \in \mathcal{G}(T)$ such that $z_n \to z$, i.e.
\[
x_n \to x \quad \text{and} \quad Tx_n \to y
\]

and $z \in \mathcal{G}(T)$ if and only if $x \in \mathcal{D}(T)$ and $y = Tx$.

(b) $\mathcal{G}(T)$ and $\mathcal{D}(T)$ are closed by assumption. Hence, $\mathcal{G}(T)$ and $\mathcal{D}(T)$ are complete by theorem 2.4. Let us now consider the mapping
\[
P : \mathcal{G}(T) \to \mathcal{D}(T)
(\mathbf{x}, Tx) \mapsto \mathbf{x}.
\]

(53) $P$ is linear and also bounded since
\[
\|P(\mathbf{x}, Tx)\| = \|\mathbf{x}\| \leq \|\mathbf{x}\| + \|Tx\| = \|(\mathbf{x}, Tx)\|.
\]

(54) $P$ is also bijective and the inverse is given by
\[
P^{-1} : \mathcal{D}(T) \to \mathcal{G}(T)
x \mapsto (x, Tx).
\]

(55) The inverse is also bounded by theorem 2.9, i.e. $\|(x, Tx)\| \leq c\|x\|$ for some $c \in \mathbb{R}$ and for all $x \in \mathcal{D}(T)$. This shows that $T$ is bounded because
\[
\|Tx\| \leq \|Tx\| + \|x\| = \|(x, Tx)\| \leq c\|x\|
\]

(56) for all $x \in \mathcal{D}(T)$. 

15
(c) Let’s assume $D(T)$ is closed. If $(x_n)$ is in $D(T)$ and $x_n \to x$ and is also such that $(Tx_n)$ also converges, then $x \in \overline{D(T)} = D(T)$ since $D(T)$ is closed and $Tx_n \to Tx$ since $T$ is continuous. Hence $T$ is closed by (a).

Conversely, if $T$ is closed, for $x \in \overline{D(T)}$ there is a sequence $(x_n)$ in $D(T)$ such that $x_n \to x$. Since $T$ is bounded,

$$||Tx_n - Tx_m|| = ||T(x_n - x_m)|| \leq ||T|| ||x_n - x_m||.$$  \hspace{1cm} (57)

This shows that $(Tx_n)$ is Cauchy. $(Tx_n)$ converges to, say, $Tx_n \to y$. Since $T$ is closed, $x \in D(T)$ by (a). This means that $D(T)$ is closed since $x \in \overline{D(T)}$ was arbitrary. \hspace{1cm} $\Box$

**Theorem 3.5** (Hilbert-adjoint operator) The Hilbert-adjoint operator $T^*$ is closed.

**Proof:** Consider a sequence $(y_n) \in D(T^*)$ such that

$$y_n \to y_0 \quad \text{and} \quad T^*y_n \to z_0.$$ \hspace{1cm} (58)

By the definition of $T^*$ we have for every $y \in D(T^*)$

$$\langle Ty, y_n \rangle = \langle y, T^*y_n \rangle.$$ \hspace{1cm} (59)

If we now let $n \to \infty$, we obtain for every $y \in D(T)$

$$\langle Ty, y_0 \rangle = \langle y, z_0 \rangle,$$ \hspace{1cm} (60)

since the inner product is continuous. This means that $y_0 \in D(T^*)$ and $z_0 = T^*y_0$. If we now apply the theorem 3.4(a), we may conclude that $T^*$ is closed. \hspace{1cm} $\Box$

4 Spectral families and spectral theorem for bounded self-adjoint operators

Here the spectrum of an operator is defined and discussed. The general properties that a spectrum of an operator has depends on the kind of space on which the operator is defined as well as the kind of operator in question. In this section $H$ denotes a nonempty complex Hilbert space.
4.1 Basic spectral properties

Let $T : \mathcal{D}(T) \rightarrow H$ be a linear operator with a domain $\mathcal{D}(T) \subset H$. In the following we always assume that $\mathcal{D}(T)$ is dense in $H$. $T$ may be associated with $T_\lambda$:

$$T_\lambda = T - \lambda I$$  \hfill (61)

where $\lambda$ is a complex number and $I$ the identity operator in $\mathcal{D}(T)$. If $T_\lambda$ has an inverse operator, the inverse is called the resolvent operator of $T$:

$$R_\lambda(T) = T_\lambda^{-1} = (T - \lambda I)^{-1}$$  \hfill (62)

where $R_\lambda$ is to be understood as the inverse of the restriction of $T$, i.e. $R_\lambda : \mathcal{R}(T) \rightarrow \mathcal{D}(T)$.

The properties of $T_\lambda$ and $R_\lambda$ depend on $\lambda$ and spectral theory focuses on these properties.

**Definition 4.1 (Regular value, resolvent set, spectrum)** Let $T : \mathcal{D}(T) \rightarrow H$ be a linear operator. A regular value $\lambda$ of $T$ is a complex number such that,

(R1) $R_\lambda(T)$ exists

(R2) $R_\lambda(T)$ is bounded

(R3) $R_\lambda(T)$ is defined on a set which is dense in $H$

The resolvent set $\rho(T)$ of $T$ is the set of all regular values $\lambda$ of $T$. The complement $\sigma(T) = \mathbb{C}\setminus\rho(T)$ is called the spectrum of $T$ and $\lambda \in \sigma(T)$ is a spectral value of $T$.

The spectrum $\sigma(T)$ can be partitioned into three disjoint sets:

- The point spectrum or discrete spectrum $\sigma_p(T)$ is the set such that $R_\lambda(T)$ does not exist (R1 not satisfied), since $T_\lambda$ is not an injection. Thus, there exists a $x \neq 0$, $x \in \mathcal{D}(T)$, such that $Tx - \lambda x = 0$ and hence $Tx = \lambda$. The vector $x$ is then an eigenvector and $\lambda \in \sigma_p(T)$ is an eigenvalue of $T$.

- The continuous spectrum $\sigma_c(T)$ is the set such that $R_\lambda(T)$ exists and is defined on a dense set in $X$ but is unbounded (R2 not satisfied).

- The residual spectrum $\sigma_r(T)$ is the set such that $R_\lambda(T)$ exists (bounded or unbounded) but the domain of $R_\lambda(T)$ is not dense in $H$ (R3 not satisfied).

Some of these sets may be empty.
Lemma 4.2 (Domain of $R_{\lambda}$) Let $T : X \to X$ be a linear operator and $\lambda \in \rho(T)$. Let $T$ be either 1) closed, or 2) bounded. Then the resolvent operator $R_{\lambda}$ is bounded and is defined on the whole space $H$.

Proof:

1) Since $T$ is closed, so is $T_{\lambda}$ by theorem 3.4(a). Hence $R_{\lambda}$ is also closed. By the definition of the resolvent set (R2), $R_{\lambda}$ is bounded and its domain $D(T)$ is closed by theorem 3.4(c). Now (R3) implies that $D(T) = \overline{D(T)} = H$.

2) Since $D(T) = H$ is closed, $T$ is closed by theorem 3.4(c) and the proof follows from part 1) of the proof. 

Theorem 4.3 Let $T$ be a bounded linear operator on $H$ and let $\lambda_{0} \in \rho(T)$. If $\lambda \in \mathbb{C}$ satisfies,

$$|\lambda - \lambda_{0}| < \frac{1}{\|R_{\lambda_{0}}\|}$$

then $\lambda \in \rho(T)$ and the resolvent $R_{\lambda}(T)$ has the representation

$$R_{\lambda} = \sum_{k=0}^{\infty} (\lambda - \lambda_{0})^{k} R_{\lambda_{0}}^{k+1}. \quad (64)$$

This series is absolutely convergent for every $\lambda$ in the open disk given by (63). The disk (63) is a subset of $\rho(T)$.

Proof: Let $\rho(T) \neq \emptyset$. For a fixed $\lambda_{0} \in \rho(T)$ and any $\lambda \in \mathbb{C}$,

$$T_{\lambda} = T - \lambda I = T - \lambda_{0} I - (\lambda - \lambda_{0}) I$$
$$= (T - \lambda_{0} I)[I - (\lambda - \lambda_{0})(T - \lambda_{0} I)^{-1}]$$
$$= T_{\lambda_{0}}[I - (\lambda - \lambda_{0})R_{\lambda_{0}}].$$

Since $\lambda_{0} \in \rho(T)$ and $T$ is bounded, $R_{\lambda_{0}}$ exists and is bounded by lemma 4.2. The expression in the square brackets on the right-hand-side can be inverted and written as a Neumann series using theorem 2.23,

$$[I - (\lambda - \lambda_{0})R_{\lambda_{0}}]^{-1} = \sum_{k=0}^{\infty} [(\lambda - \lambda_{0})R_{\lambda_{0}}]^{k}. \quad (65)$$

For this series to converge, it is required that $\|(\lambda - \lambda_{0})R_{\lambda_{0}}\| < 1$, which means that

$$|\lambda - \lambda_{0}| < \frac{1}{\|R_{\lambda_{0}}\|}$$
For every operator satisfying (63), the operator $T_\lambda$ has an inverse $R_\lambda$,

$$R_\lambda = (T_{\lambda_0}[I - (\lambda - \lambda_0)R_{\lambda_0}])^{-1} = [I - (\lambda - \lambda_0)R_{\lambda_0}]^{-1}R_{\lambda_0} = \sum_{k=0}^\infty (\lambda - \lambda_0)^k R_{\lambda_0}^{k+1}.$$  

**Theorem 4.4** The resolvent set $\rho(T)$ of a bounded linear operator $T$ on $H$ is open; hence the spectrum $\sigma(T)$ is closed.

**Proof:** If $\rho(T) = \emptyset$, then $\rho(T)$ would be open (however, as stated by theorem 4.5, the resolvent set cannot be empty). For a fixed $\lambda_0 \in \rho(T)$ and any $\lambda \in \mathbb{C}$ satisfying (63) in theorem 4.3, the operator $T_\lambda$ has an inverse $R_\lambda = \sum_{k=0}^\infty (\lambda - \lambda_0)^k R_{\lambda_0}^{k+1}$.

Thus, for each arbitrary $\lambda_0$, there exists an open neighbourhood of regular values $\lambda \in \rho(T)$ where $T_\lambda$ is invertible and hence, $\rho(T)$ is open. As a consequence, the spectrum $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is closed.  

**Theorem 4.5** The spectrum $\sigma(T)$ of a bounded linear operator $T : H \to H$ is compact and lies in the disk given by

$$|\lambda| \leq ||T|| \quad (66)$$

Hence, the resolvent set $\rho(T)$ of $T$ is not empty.

**Proof:** Let $\lambda \neq 0$. Using the theorem 2.23, the resolvent operator can be expanded as series

$$R_\lambda = (T - \lambda I)^{-1} = -\frac{1}{\lambda} (I - \frac{1}{\lambda} T)^{-1} = -\frac{1}{\lambda} \sum_{k=0}^\infty \left( \frac{1}{\lambda} T \right)^k$$

that converges when

$$\left| \frac{1}{\lambda} T \right| = \frac{||T||}{|\lambda|} < 1 \iff |\lambda| > ||T||. \quad (67)$$

Hence, all $\lambda$ that satisfy $|\lambda| > ||T||$, belong to the resolvent set and, as a consequence, the spectrum lies in the disk given by (66) and $\sigma(T)$ is bounded.  

19
Above it was shown that the eigenvalues of a bounded self-adjoint linear operator are real, if they exist. Moreover, we shall see that the entire spectrum of a bounded self-adjoint linear operator is in fact real (theorem 4.7). To prove this, we shall first show a characterization of the resolvent set $\rho(T)$ of $T$.

**Theorem 4.6** Let $T : H \to H$ be a bounded self-adjoint linear operator on $H$. Then a number $\lambda$ belongs to the resolvent set $\rho(T)$ if and only if there exists a $c > 0$ such that for every $x \in H$:

$$||T_\lambda x|| \geq c||x||.$$  \hfill (68)

*Proof:*

a) If $\lambda \in \rho(T)$, then $R_\lambda = T_\lambda^{-1} : H \to H$ exists and is bounded (say, $||R_\lambda|| = k > 0$). Now, $I = R_\lambda T_\lambda$, so that for every $x \in H$ we have

$$||x|| = ||R_\lambda T_\lambda x|| \leq ||R_\lambda|| ||T_\lambda x|| = k||T_\lambda x||$$

which gives $||T_\lambda x|| \geq c||x||$ when $c = \frac{1}{k}$.

b) Let us assume instead that $||T_\lambda x|| \geq c||x||$ for some $c > 0$ and $x \in H$. Let us prove that then

1. $T_\lambda : H \to T_\lambda(H)$ is bijective,
2. $T_\lambda(H)$ is dense in $H$,
3. $T_\lambda(H)$ is closed in $H$,

so that $T_\lambda(H) = H$ and $R_\lambda = T_\lambda^{-1}$ is bounded by theorem 2.9.

1. We need to show that $T_\lambda x_1 = T_\lambda x_2$ implies that $x_1 = x_2$. This follows from our assumption:

$$0 = ||T_\lambda x_1 - T_\lambda x_2|| = ||T_\lambda(x_1 - x_2)|| \geq c||x_1 - x_2|| \Rightarrow ||x_1 - x_2|| = 0. \hfill (70)$$

Since $x_1$ and $x_2$ were arbitrary, this shows that $T_\lambda : H \to T_\lambda(H)$ is bijective.

2. Let $x_0 \perp \overline{T_\lambda(H)}$. Then $x_0 \perp T_\lambda(H)$ and therefore for all $x \in H$ we have

$$0 = \langle T_\lambda x, x \rangle = \langle Tx, x \rangle - \lambda \langle x, x \rangle. \hfill (71)$$

Since $T$ is self-adjoint, we obtain

$$\langle x, T_\lambda x_0 \rangle = \langle Tx, x_0 \rangle = \langle x, \lambda x_0 \rangle \hfill (72)$$
and so $Tx_0 = \bar{\lambda}x_0$. If $x_0 \neq 0$, we arrive at a contradiction with our initial assumption, since then $\bar{\lambda}$ would be an eigenvalue of $T$ and so $\bar{\lambda} = \lambda$ by theorem 2.28 and $Tx_0 - \lambda x_0 = T_{\lambda}x_0 = 0$, but

$$0 = ||T_{\lambda}x_0|| \geq c||x_0|| > 0$$

(73)
since $c > 0$. This contradiction means that $x_0 = 0$. Thus $T_{\lambda}(H) = \{0\}$ because $x_0$ was an eigenvalue of $T_{\lambda}(H)$. Hence $T_{\lambda}(H) = H$ i.e. $T_{\lambda}(H)$ is dense in $H$.

3. Let $y \in T_{\lambda}(H)$. Now, there exists a sequence $(y_n)$ in $T_{\lambda}(H)$ which converges to $y$. Since $y_n \in T_{\lambda}(H)$, we have $y_n = T_{\lambda}x_n$ for some $x_n \in H$. From our assumption, we get

$$||x_n - x_m|| \leq \frac{1}{c}||T_{\lambda}(x_n - x_m)|| = \frac{1}{c}||y_n - y_m||$$

(74)

and so $(x_n)$ is Cauchy since $(y_n)$ converges. $H$ is complete so that $(x_n)$ converges $x_n \to x$. Since $T$ is continuous, so is $T_{\lambda}$, and $y_n = T_{\lambda}x_n \to T_{\lambda}x$. Since the limit is unique, $T_{\lambda}x = y \in T_{\lambda}(H)$ and hence $T_{\lambda}(H)$ is closed because $y$ was arbitrary. We have $T_{\lambda}(H) = H$ by the point 2. This means that $R_{\lambda} = T_{\lambda}^{-1}$ is defined on all of $H$, is bounded by theorem 2.9. Hence, $\lambda \in \rho(T)$.

**Theorem 4.7** The spectrum of a bounded self-adjoint linear operator $T : H \to H$ is real.

**Proof:** The proof is based on showing that every $\lambda = \alpha + i\beta$ with $\beta \neq 0$ must belong to $\rho(T)$ due to the theorem 4.6 and therefore $\sigma(T) \subset \mathbb{R}$.

For every $x \neq 0$,

$$\langle T_{\lambda}x, x \rangle = \langle Tx, x \rangle - \lambda \langle x, x \rangle$$

(75)

Since $\langle x, x \rangle$ and $\langle Tx, x \rangle$ are real, we get

$$\overline{\langle T_{\lambda}x, x \rangle} = \langle Tx, x \rangle - \bar{\lambda} \langle x, x \rangle$$

(76)

Inserting $\lambda = \alpha + i\beta$ and subtracting, we have

$$\overline{\langle T_{\lambda}x, x \rangle} - \langle T_{\lambda}x, x \rangle = \lambda \langle x, x \rangle - \bar{\lambda} \langle x, x \rangle = 2i\beta||x||^2 = -2i\text{Im} \langle T_{\lambda}x, x \rangle$$

(77)

Taking absolute values from this result, we obtain

$$||\beta|| ||x||^2 = ||\text{Im} \langle T_{\lambda}x, x \rangle|| \leq ||\langle T_{\lambda}x, x \rangle|| \leq ||T_{\lambda}x|| ||x||$$

(78)

Dividing by $||x||$ give $||\beta|| ||x|| \leq ||T_{\lambda}x||$. If $\beta \neq 0$, then $\lambda \in \rho(T)$ by theorem 4.6. Hence for $\lambda \in \sigma(T)$, we must have $\beta = 0$, that is, $\lambda$ is real. \qed
**Theorem 4.8** The spectrum of a bounded self-adjoint linear operator $T : H \to H$ lies in the closed interval $[m, M]$ on the real axis, where

$$
m = \inf_{||x||=1} \langle Tx, x \rangle \
M = \sup_{||x||=1} \langle Tx, x \rangle.
$$

(79)

$m$ and $M$ are the spectral values of $T$.

**Proof:** Here only the proof for the existence of supremum (and infimum) is given but the proof that $m$ and $M$ are spectral values is omitted.

We get from theorem 4.7 that the spectrum $\sigma(T)$ lies in the real axis. It remains to show that for all $c > 0$, $\lambda = M + c \in \rho(T)$.

For every $x \neq 0$ we define $\hat{x} = \frac{x}{||x||}$ and we have

$$\langle Tx, x \rangle = ||x||^2 \langle T\hat{x}, \hat{x} \rangle \leq ||x||^2 \sup_{||\hat{x}||=1} \langle T\hat{x}, \hat{x} \rangle = \langle x, x \rangle M$$

(80)

Also $-\langle Tx, x \rangle \geq -\langle x, x \rangle M$. Now, using the Schwarz inequality,

$$||T\lambda x|| ||x|| \geq -\langle T\lambda x, x \rangle = -\langle Tx, x \rangle + \langle \lambda x, x \rangle \geq (-M + \lambda) \langle x, x \rangle$$

$$= c \langle x, x \rangle = c ||x||^2$$

(81)

where $c = \lambda - M > 0$ by assumption. Hence, we obtain $||T\lambda x|| > c ||x||$ and therefore $x \in \rho(T)$ by theorem 4.6. It is easy to show that the same applies for the infimum.

**Theorem 4.9** The residual spectrum $\sigma_r(T)$ of a bounded self-adjoint linear operator $T : H \to H$ is empty.

**Proof:** The theorem can be proven via an indirect proof. Let $\sigma_r \neq \emptyset$ and $\lambda \in \sigma_r(T)$. Then, by definition $R_\lambda$ exists but its domain is not dense in $H$. Hence, there is a $y \neq 0$ in $H$ that is orthogonal to $D(R_\lambda)$. However, the domain $D(R_\lambda)$ is also the range of $T_\lambda$, so

$$\langle T_\lambda x, y \rangle = 0 \ \forall x \in H$$

(82)

Since $\lambda$ is real and $T_\lambda$ is self-adjoint, we may also write

$$\langle x, T_\lambda y \rangle = 0 \ \forall x \in H$$

(83)

If we now define $x = T_\lambda y$ we get $\langle T_\lambda y, T_\lambda y \rangle = 0$, which means that

$$T_\lambda y = Ty - \lambda y = 0 \Leftrightarrow Ty = \lambda y$$

(84)

and we see that $\lambda$ is an eigenvalue of $T$ (i.e $\lambda \in \sigma_p(T)$), which contradicts the assumption that $\lambda \in \sigma_r(T)$. Hence, we conclude that $\sigma_r = \emptyset$.

Several of the properties of bounded self-adjoint linear operators hold true for unbounded operators as well, e.g. the theorems 4.6 and 4.7 generalize to unbounded operators.
**Theorem 4.10** Let $T : \mathcal{D}(T) \to H$ be a self-adjoint linear operator which is densely defined in $H$. Then a number $\lambda$ belongs to the resolvent set $\rho(T)$ of $T$ if and only if there exists $c > 0$ such that for every $x \in \mathcal{D}(T)$:

$$||T_\lambda x|| \geq c ||x||. \quad (85)$$

**Theorem 4.11** The spectrum $\sigma(T)$ of a self-adjoint linear operator $T : \mathcal{D}(T) \to H$ is real and closed.

The proofs of the previous theorems are very similar to theorems 4.6 and 4.7 and are therefore omitted.

### 4.2 Spectral family

As discussed in the introduction, a bounded self-adjoint linear operator $T$ in a finite dimensional Hilbert space $H$ operating on a vector $x \in H$, can be written as a sum

$$Tx = \sum_{i=1}^{n} \lambda_i \gamma_i x_i, \quad (86)$$

where $\gamma_i = \langle x, x_i \rangle$

and $(x_i)_{i=1}^{n}$ is an orthonormal basis of $H$. However, the operation can be generalized in terms of projection operators. In the finite dimensional case we have an orthogonal projection $P_i : H \to H$ which projects $x$ onto the eigenspace of $T$ corresponding to $\lambda_i$, $x \mapsto \gamma_i x_i$. The transformation can now be written as

$$Tx = \sum_{i=1}^{n} \lambda_i P_i x \quad \Rightarrow \quad T = \sum_{i=1}^{n} \lambda_i P_i \quad (87)$$

which gives us the operator $T$ in terms of projections.

In order to generalize this result for infinite-dimensional Hilbert spaces, we need to take into account the more complicated nature of the spectrum in infinite-dimensions. In a finite dimensional Hilbert space, instead of the projections $P_1, \ldots, P_n$ themselves, we may take the sum of projections, such that for each $\lambda$ we define

$$E_\lambda = \sum_{\lambda_i \leq \lambda} P_i. \quad (88)$$
This is called a one-parameter family of projections. The operator $E_{\lambda}$ is a projection of $H$ onto the subspace $V_{\lambda}$ spanned by all those $x_i$ for which $\lambda_i \leq \lambda$. Thus, if $\lambda \leq \mu$, it follows that

$$V_{\lambda} \subset V_{\mu}. \quad (89)$$

Obviously, as $\lambda$ increases, $E_{\lambda}$ spans larger portion of the Hilbert space, growing from 0 to $I$. The operator $E_{\lambda}$ remains unchanged for an interval that contains no eigenvalues of $T$. Hence the operator $E_{\lambda}$ has the following properties:

$$E_{\lambda}E_{\mu} = E_{\mu}E_{\lambda} = E_{\lambda} \quad \text{if } \lambda < \mu$$

$$E_{\lambda} = 0 \quad \text{if } \lambda < \lambda_1$$

$$E_{\lambda} = I \quad \text{if } \lambda > \lambda_n \quad (90)$$

$$E_{\lambda+0} = \lim_{\mu \to \lambda+0} E_{\mu} = E_{\lambda}$$

where $\mu \to \lambda + 0$ means that $\mu$ approaches $\lambda$ from the right.

Let us now generalise the above ideas for infinite dimensional Hilbert spaces. The spectral family $\mathcal{E}$ of $T$ is formed as a one-parameter family $\mathcal{E} = (E_{\lambda})_{\lambda \in \mathbb{R}}$. A spectral family is also called a resolution of identity and its properties are summarized in the following definition.

**Definition 4.12 (Spectral family)** A real spectral family is a one-parameter family $\mathcal{E} = (E_{\lambda})_{\lambda \in \mathbb{R}}$ of projections $E_{\lambda}$ defined on a Hilbert space $H$ which depends on a real parameter $\lambda$ and has the following properties:

$$E_{\lambda}E_{\mu} = E_{\mu}E_{\lambda} = E_{\lambda} \quad (\lambda < \mu)$$

$$\lim_{\lambda \to -\infty} E_{\lambda}x = 0$$

$$\lim_{\lambda \to +\infty} E_{\lambda}x = x$$

$$E_{\lambda+0} = \lim_{\mu \to \lambda+0} E_{\mu}x = E_{\lambda}x$$

for all $x \in \mathcal{E}$.

$\mathcal{E}$ is called a spectral family on an interval $[a, b]$ if

$$E_{\lambda} = 0 \text{ for } \lambda < a, \quad E_{\lambda} = I \text{ for } \lambda \geq b. \quad (91)$$

Since the spectrum of a bounded self-adjoint linear operator lies on a finite interval (Theorem 4.8), these spectral families are of special relevance to us.

We shall now derive a useful relation related to the spectral family for later use. Given $\lambda < \mu$, let us define an interval $\Delta = (\lambda, \mu]$ and associate the operator

$$E(\Delta) = E_{\mu} - E_{\lambda} \quad (92)$$
with the interval. Since $\lambda < \mu$, we have $E_\lambda \leq E_\mu$ meaning $E_\mu - E_\lambda \geq 0$ and $E_\lambda(H) \subset E_\mu(H)$ by theorem 2.34. Also, by theorem 2.35, $E(\Delta)$ is also a projection.

Using the theorem 2.34 again, we can write

$$E_\mu E(\Delta) = E_\mu^2 - E_\mu E_\lambda = E_\mu - E_\lambda = E(\Delta)$$

(93)

where every $E_\lambda$ is an orthogonal projection on $H$. Also

$$(I - E_\lambda)E(\Delta) = E(\Delta) - E_\lambda E(\Delta) = E(\Delta) - E_\lambda E_\mu + E_\mu^2 = E(\Delta).$$

(94)

Now, $E$ commutes with every self-adjoint linear operator that commutes with $T$. This holds true also for $T_\lambda$. Using the above results, we have

$$T_\mu E(\Delta) = T_\mu E_\mu E(\Delta) \leq 0$$

$$T_\lambda E(\Delta) = T_\lambda(I - E_\lambda)E_\mu E(\Delta) \geq 0$$

(95)

because the $E(\Delta)$ is a positive operator. If $T : H \to H$ is a bounded linear operator, we denote $T \geq 0$, if $\langle x, Tx \rangle \geq 0$ for all $x \in H$. This means that $T$ is a positive operator.

The equation (95) implies that $TE(\Delta) \leq \mu E(\Delta)$ and that $TE(\Delta) \geq \lambda E(\Delta)$ and so

$$\lambda E(\Delta) \leq TE(\Delta) \leq \mu E(\Delta)$$

(96)

From this relation in can be shown[3] that the spectral family $E$ is continuous from the right i.e. the final property in definition 4.12.

4.3 Spectral theorem for bounded linear operators

Before presenting the spectral theorem for bounded operators, we shall first define the Riemann–Stieltjes integral.

**Definition 4.13** (Riemann–Stieltjes integral) Let $x_1, x_2, \ldots, x_n$ be a set of increasing values of $x$ between $a$ and $b$, with $x_{r+1} - x_r < \delta$ and $x_0 = a$, $x_{n+1} = b$. For each interval $\xi_r \in [x_r, x_{r+1}]$, we may form the sum

$$S_n = f(\xi_0)(x_1 - a) + f(\xi_1)(x_2 - x_1) + \ldots + f(\xi_n)(b - x_n).$$

(97)

The limit of this sum (if it exists) as $\delta \to 0$ is called the Riemann integral and is denoted by

$$\int_a^b f(x)dx.$$
If \( f(x) \) and \( g(x) \) are both bounded functions of \( x \), we may form the sum

\[
S_n = f(\xi_0)(g(x_1) - g(a)) + f(\xi_1)(g(x_2) - g(x_1)) + \ldots + f(\xi_n)(g(b) - g(x_n))
\]

with \( \xi_r \) chosen as in (97). If this sum tends to a unique limit when \( \delta \to 0 \), the limit is called a Riemann–Stieltjes integral and is denoted by

\[
\int_a^b f(x)dg(x).
\]

**Theorem 4.14** (Spectral theorem for bounded self-adjoint linear operators) Let \( T : H \to H \) be a bounded self-adjoint linear operator on a complex Hilbert space \( H \). Then there exists a spectral family \( \{E_\lambda : \lambda \in \mathbb{R}\} \) such that for all \( x, y \in H \),

\[
\langle Tx, y \rangle = \int_{m-\delta}^{M} \lambda dw(\lambda) \quad w(\lambda) = \langle E_\lambda x, y \rangle
\]

where the integral is an ordinary Riemann–Stieltjes integral.

If \( p \) is a polynomial in \( \lambda \) with real coefficients,

\[
p(\lambda) = \alpha_n\lambda^n + \alpha_{n-1}\lambda^{n-1} + \ldots + \alpha_0
\]

then for the operator \( p(T) \)

\[
p(T) = \alpha_nT^n + \alpha_{n-1}T^{n-1} + \ldots + \alpha_0
\]

and for any \( x, y \in H \), we may write the inner product \( \langle p(T)x, y \rangle \) as

\[
\langle p(T)x, y \rangle = \int_{m-\delta}^{M} p(\lambda)dw(\lambda) \quad w(\lambda) = \langle E_\lambda x, y \rangle.
\]

**Proof:** Here we do not present the proof of the existence of the spectral family. This can be shown using the Riesz representation theorem and Banach algebra methods.

Let us define a sequence \( (\mathcal{P}_n) \) of partitions of \( (a, b] \) where \( a < m \) and \( M < b \) with every \( \mathcal{P}_n \) partitioning \( (a, b] \) into intervals

\[
\Delta_{nj} = (\lambda_{nj}, \mu_{nj}] \quad j = 1, \ldots, n
\]

of length \( l(\Delta_{nj}) = \mu_{nj} - \lambda_{nj} \). We assume the sequence \( (\mathcal{P}_n) \) to be such that

\[
\delta(\mathcal{P}_n) = \max_j l(\Delta_{nj}) \to 0 \quad \text{as } n \to \infty.
\]
With (96), we have
\[ \lambda_{nj} E(\Delta_{nj}) \leq T E(\Delta_{nj}) \leq \mu_{nj} E(\Delta_{nj}). \] (107)

By summing over \( j \) from 1 to \( n \), we get
\[ \sum_{j=1}^{n} \lambda_{nj} E(\Delta_{nj}) \leq \sum_{j=1}^{n} T E(\Delta_{nj}) \leq \sum_{j=1}^{n} \mu_{nj} E(\Delta_{nj}). \] (108)

Since \( \mu_{nj} = \lambda_{n,j+1} \) for \( j = 1, \ldots, n-1 \), from the definition 4.12 we have
\[ T \sum_{j=1}^{n} E(\Delta_{nj}) = T \sum_{j=1}^{n} (E_{\mu_{nj}} - E_{\lambda_{nj}}) = T(I - 0) = T. \] (109)

The equation (106) requires that for every \( \epsilon > 0 \) there exists \( n \) such that \( \delta(P_n) < \epsilon \). Therefore, we should have
\[ \sum_{j=1}^{n} \mu_{nj} E(\Delta_{nj}) - \sum_{j=1}^{n} \lambda_{nj} E(\Delta_{nj}) = \sum_{j=1}^{n} (\mu_{nj} - \lambda_{nj}) E(\Delta_{nj}) < \epsilon I. \] (110)

Combining this with (108), it follows that for any \( \epsilon > 0 \) there is an \( N \) such that for any \( n > N \) and every choice of \( \hat{\lambda}_{nj} \in \Delta_{nj} \) we have
\[ \left| \left< Tx - \sum_{j=1}^{n} \hat{\lambda}_{nj} E(\Delta_{nj}) x, y \right> \right| < \epsilon \] (111)
for all \( x, y \in E \). Since \( E_{\lambda} \) remains unchanged when \( \lambda < m \) or \( \lambda \geq M \), the choice of \( a < m \) and \( b > M \) is of no consequence. The equation (111) proves the first part of the theorem since the integral in (101) is to be understood as uniform operator convergence which implies strong operator convergence as defined in 2.20. Furthermore, (101) implies (103).

For the theorem on polynomials, the theorem will be proven for \( \lambda^r, r \in \mathbb{N} \). The first of the spectral family properties in definition 4.12 will be used: for any \( \kappa < \lambda \leq \mu < \nu \) we have
\[ (E_{\lambda} - E_{\kappa})(E_{\mu} - E_{\nu}) = E_{\lambda} E_{\mu} - E_{\lambda} E_{\nu} - E_{\kappa} E_{\mu} + E_{\kappa} E_{\nu} \]
\[ = E_{\lambda} - E_{\kappa} - E_{\mu} + E_{\nu} = 0. \]

This means that \( E(\Delta_{nj}) E(\Delta_{nk}) = 0 \) for every \( j \neq k \). Also, since \( E(\Delta_{nj}) \) is a projection, \( E(\Delta_{nj})^s = E(\Delta_{nj}) \) for every \( s \in \mathbb{N} \). Consequently, we obtain
\[ \left[ \sum_{j=1}^{n} \hat{\lambda}_{nj} E(\Delta_{nj}) \right]^r = \sum_{j=1}^{n} \hat{\lambda}_{nj}^r E(\Delta_{nj}). \] (112)
If the sum in (111) is close to $T$, then the expression on the left hand side in (112) will be close to $T^r$ because multiplication of bounded linear operators in continuous. Hence, given $\epsilon > 0$ there will be an $N$ such that for all $n > N$,
\[
\left| \left\langle Tx - \sum_{j=1}^{n} \hat{\lambda}_{nj} E(\Delta_{nj})x, y \right\rangle \right| < \epsilon
\] (113)
for all $x, y \in \mathcal{E}$. This proves the polynomial part of the theorem for $p(\lambda) = \lambda^r$ and the readily follow for any arbitrary polynomial with real coefficients. \[\square\]

**Remark 4.15.** (Spectral representation) It is possible to define a spectral measure $dE_{\lambda}$ and present the operator $T$ as an integral with respect to the spectral measure as
\[
T = \int_{m-0}^{M} \lambda dE_{\lambda}.
\] (114)

This is called the spectral representation of $T$.

The $m-0$ in theorem 4.14 indicates that we need to take into account a contribution at $\lambda = m$ which occurs if $E_m \neq 0$. Then for any $a < m$, we may write
\[
\int_{a}^{M} \lambda dE_{\lambda} = \int_{m-0}^{M} \lambda dE_{\lambda} = mE_{m} + \int_{m-0}^{M} \lambda dE_{\lambda}.
\] (115)

**Theorem 4.16** Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space $H$ and $\mathcal{E} = E_{\lambda}$ the corresponding spectral family. Then $\lambda \mapsto E_{\lambda}$ has a discontinuity at any $\lambda = \lambda_{0}$ if and only if $\lambda_{0}$ is an eigenvalue of $T$. In this case the corresponding eigenspace is
\[
\mathcal{N}(T - \lambda_{0}I) = (E_{\lambda_{0}} - E_{\lambda_{0}-0})(H).
\] (116)

**Proof:** By definition, $\lambda_{0}$ is an eigenvalue of $T$ if and only if $\mathcal{N}(T - \lambda_{0}I) \neq \{0\}$, and so the discontinuity follows directly from (116) so we shall focus on proving (116). Let us write $F_{0} = E_{\lambda_{0}} - E_{\lambda_{0}-0}$. In order to show that the equality holds, we need to show that
\[
F_{0}(H) \subset \mathcal{N}(T - \lambda_{0}I)
\] (117)
and
\[
F_{0}(H) \supset \mathcal{N}(T - \lambda_{0}I).
\] (118)

Using inequality (96) with $\lambda = \lambda_{0} - \frac{1}{n}$ and $\mu = \lambda_{0}$, we get
\[
\left( \lambda_{0} - \frac{1}{n} \right) E(\Delta_{0}) \leq TE(\Delta_{0}) \leq \lambda_{0}E(\Delta_{0})
\] (119)
where $\Delta_0 = (\lambda_0 - 1/n, \lambda_0]$. If we let $n \to \infty$, then $E(\Delta_0) \to F_0$, so that

$$\lambda_0 F_0 \leq TF_0 \leq \lambda_0 F_0 \quad (120)$$

and $TF_0 = \lambda_0 F_0$. This means that $(T - \lambda_0)F_0 = 0$ and hence we have proven (117).

Next, let us consider $x \in N(T - \lambda_0 I)$. We shall prove (118) by showing that $x \in F_0(H)$. By theorem 4.8, if $\lambda_0 \notin [m, M]$ then $\lambda_0 \in \rho(T)$. Hence, $N(T - \lambda_0 I) = \{0\} \subset F_0(H)$ since $F_0(H)$ is a vector space. Now let $\lambda_0 \in [m, M]$. By assumption, $(T - \lambda_0 I)x = 0$ which also means that $(T - \lambda_0 I)^2x = 0$. Now, using (101),

$$\int_a^b (\lambda - \lambda_0)^2 d\lambda \geq 0 = \int_a^b (\lambda - \lambda_0)^2 d\lambda \geq \epsilon^2 \int_a^b d\lambda = \epsilon^2 \langle E_{\lambda_0 - \epsilon} x, x \rangle \quad (122)$$

as well as

$$0 = \int_{\lambda_0 + \epsilon}^{\lambda_0 + \epsilon} (\lambda - \lambda_0)^2 d\lambda \geq \epsilon^2 \int_{\lambda_0 + \epsilon}^{\lambda_0 + \epsilon} d\lambda = \epsilon^2 \langle Ix, x \rangle - \epsilon^2 \langle E_{\lambda_0 + \epsilon} x, x \rangle. \quad (123)$$

Since $\epsilon > 0$, using the properties of projection operators in section 2.4, we obtain,

$$\langle E_{\lambda_0 - \epsilon} x, x \rangle = 0 \quad \Rightarrow \quad E_{\lambda_0 - \epsilon} x = 0 \quad (124)$$

$$\langle x - E_{\lambda_0 + \epsilon} x, x \rangle = 0 \quad \Rightarrow \quad x - E_{\lambda_0 + \epsilon} x = 0 \quad (125)$$

and may write

$$x = (E_{\lambda_0 + \epsilon} - E_{\lambda_0 - \epsilon})x. \quad (126)$$

If we let $\epsilon \to 0$, we obtain $x = F_0x$ because $\lambda \mapsto E_\lambda$ is continuous from the right. This proves (118).

To conclude the work thus far; we can associate a spectral family to any bounded self-adjoint linear operator $T$ on any Hilbert space and this spectral family can be used to represent $T$ as an Riemann–Stieltjes integral. What remains is to extend this to unbounded operators in Hilbert space.

29
Unitary operators, Cayley transform and the general spectral theorem

In this section, the spectral theorem for bounded linear operators in Hilbert space will be expanded to include also unbounded operators. For this the spectral theorem of unitary operators is needed. The spectrum of unbounded self-adjoint operators is similar to bounded operators, for example, the spectrum of unbounded self-adjoint operators is real. Unitary operators are bounded linear operators whose spectrum lie on the unit circle. The spectrum of the unitary operators can then be mapped onto the real line using a Cayley transform and, hence, the spectrum of an unbounded operator may be written using the spectral theorem of unitary operators as the basis.

5.1 Spectral theorem for unitary operators

**Theorem 5.1** If $U : H \to H$ is a unitary linear operator on a complex Hilbert space $H \neq \{0\}$, then the spectrum $\sigma(U)$ is a closed subset of the unit circle and $|\lambda| = 1$ for every $\lambda \in \sigma(U)$.

**Proof:** First we note that

$$||Ux||^2 = \langle Ux, Ux \rangle = \langle x, U^* Ux \rangle = \langle x, Ix \rangle = ||x||^2$$

so $||U|| = 1$. Hence, $|\lambda| \leq 1$ by theorem 4.5. Also $0 \in \rho(U)$ since for $\lambda = 0$ the resolvent operator of $U$ is $U^{-1} = U^*$, which is also unitary. However, according to the theorem 4.3, setting $\lambda_0 = 0$, every $\lambda$ that satisfies $|\lambda| < \frac{1}{||U||} = 1$ belongs to $\rho(T)$. Hence, the spectrum of $U$ must lie on the unit circle. This set is closed based on theorem 4.4.

The spectral theorem for bounded unitary operators is derived by power series method and with a lemma by Wecken[3].

**Lemma 5.2** Let

$$h(\lambda) = \sum_{n=0}^{\infty} \alpha_n \lambda^n$$

be absolutely convergent for all $\lambda$ such that $|\lambda| \leq k$. Suppose that $S \in \mathcal{L}(H, H)$ is self-adjoint and has norm $||S|| \leq k$. Then

$$h(S) = \sum_{n=0}^{\infty} \alpha_n S^n$$
is a bounded self-adjoint linear operator and

\[ ||h(S)|| \leq \sum_{n=0}^{\infty} |\alpha_n|k^n \]  \hfill (130)

If a bounded linear operator commutes with \( S \), it also commutes with \( h(S) \).

**Partial proof:** We shall only conclude here that \( h(S) \) is self-adjoint.

Let \( h_n(\lambda) \) denote the \( n \)th partial sum of the series in (128). Each partial sum \( h_n(S) \) is self-adjoint by theorem 2.26 and so \( \langle h_n(S)x, x \rangle \) are real by theorem 2.29. Since the inner product is continuous, \( \langle h(S)x, x \rangle \) is also real and so the sum \( h(S) \) is also self-adjoint since \( H \) is complex.

**Lemma 5.3** (Wecken) Let \( W \) and \( A \) be bounded self-adjoint linear operators on a complex Hilbert space \( H \). Suppose that \( WA = AW \) and \( W^2 = A^2 \). Let \( P \) be the projection of \( H \) onto the null space \( \mathcal{N}(W - A) \). Then:

(a) If a bounded linear operator commutes with \( W - A \), it also commutes with \( P \).

(b) \( Wx = 0 \) implies \( Px = x \)

(c) We have \( W = (2P - I)A \)

**Proof:**

a) Suppose that \( B \) commutes with \( W - A \). By definition, \( Px \in \mathcal{N}(W - A) \) for every \( x \in H \) and we obtain

\[ (W - A)BPx = B(W - A)Px = 0 \]  \hfill (131)

Thus, we have shown that \( BPx \in \mathcal{N}(W - A) \), which implies that \( P(BPx) = BPx \), i.e.

\[ PBP = BP \]  \hfill (132)

Since \( W - A \) is self-adjoint, we can write

\[ (W - A)B^* = [B(W - A)]^* = [(W - A)B]^* = B^*(W - A) \]  \hfill (133)

and thus \( W - A \) and \( B^* \) also commute and hence we have also \( PB^*P = B^*P \). Finally, since projections are self-adjoint, we have

\[ PBP = (PB^*P)^* = (B^*P)^* = PB \]  \hfill (134)

Combining the above results, we obtain \( PB = BP \).
b) Let \( Wx = 0 \). Now, since \( A^2 = W^2 \) and both \( A \) and \( W \) are self-adjoint, we have
\[
||Ax||^2 = \langle Ax, Ax \rangle = \langle A^2 x, x \rangle = \langle W^2 x, x \rangle = \langle Wx, Wx \rangle = ||Wx||^2 = 0
\]
and hence \( Ax = 0 \). Therefore, \((W - A)x = 0\). Now, if \( x \in \mathcal{N}(W - A) \), then \( Px = x \) since \( P \) is a projection onto \( \mathcal{N}(W - A) \).

c) From the assumptions, we get
\[
(W - A)(W + A) = W^2 - A^2 = 0
\]
which implies that \((W + A)x \in \mathcal{N}(W - A)\) for every \( x \in H \). Since \( P \) projects \( H \) onto \( \mathcal{N}(W - A) \), we obtain
\[
P(W + A)x = (W + A)x
\]
for every \( x \in H \), so \( P(W + A) = W + A \). Based on a), \( P(W - A) = (W - A)P \) and, because \( P \) projects \( H \) onto \( \mathcal{N}(W - A) \), we obtain
\[
2PA = P(W + A) - P(W - A) = W + A
\]
\[
\Rightarrow W = 2PA - A
\]

\[\square\]

**Theorem 5.4** (Spectral theorem for unitary operators) Let \( U : H \rightarrow H \) be a unitary operator on a complex Hilbert space \( H \neq \{0\} \). Then there exists a spectral family \( \mathcal{E} = E_\theta \) on \([-\pi, \pi]\) such that for all \( x, y \in H \) and for every polynomial \( f \),
\[
\langle f(U)x, y \rangle = \int_{-\pi}^{\pi} f(e^{i\theta})dw(\theta)
\]
where \( w(\theta) = \langle E_\theta x, y \rangle \).

**Proof:** We shall show that, given a unitary operator \( U \), there exists a bounded self-adjointed linear operator \( S \) with \( \sigma(S) \subset [-\pi, \pi] \) such that
\[
U = e^{iS} = \cos S + i \sin S
\]
where the functions \( \cos S \) and \( \sin S \) are defined as power series
\[
\cos S = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} S^{2n} \quad \sin S = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} S^{2n+1}.
\]

32
Since $S$ is bounded and self-adjoint, $\cos S$ and $\sin S$ are also bounded and self-adjoint and commute with one another, by lemma 5.2 and we can see that $U$ defined in this manner is unitary

$$UU^* = (\cos S + i \sin S)(\cos S - i \sin S)$$

$$= (\cos^2 S + \sin^2 S) = I. \quad (142)$$

Similarly one can show that $U^*U = I$.

Now, let us write

$$U = V + iW \quad (143)$$

where

$$V = \frac{1}{2} (U + U^*) \quad W = \frac{1}{2i} (U - U^*) \quad (144)$$

From the properties of adjoint operators in theorem 2.25, we can see that $V$ and $W$ are obviously self-adjoint. Since $UU^* = I$, we have $VW = WV$ and also $V^2 + W^2 = I$. Also, since $||U|| = 1$ by definition (theorem 2.30), for $V$ and $W$ we have,

$$||V|| \leq 1, \quad ||W|| \leq 1. \quad (145)$$

Consider

$$g(\lambda) = \arccos \lambda = \frac{\pi}{2} - \arcsin \lambda = \frac{\pi}{2} - \lambda - \frac{1}{6} \lambda^3 + \ldots \quad (146)$$

The Maclaurin series converges for $|\lambda| \leq 1$. Since $||V|| \leq 1$, the lemma 5.2 implies that the operator

$$g(V) = \arccos V = \frac{\pi}{2} - V - \frac{1}{6} V^3 + \ldots \quad (147)$$

exists and is self-adjoint. If we now define

$$A = \sin g(V) \quad (148)$$

we have a power series in $V$ and based on lemma 5.2, $A$ is self-adjoint and commutes with $V$ and $W$. Since $\cos(g(V)) = V$, we have

$$V^2 + A^2 = (\cos^2 + \sin^2) \ (g(V)) = I. \quad (149)$$

which implies that $W^2 = A^2$. Hence, we can apply Weckens lemma 5.3 and conclude that $W = (2P - I)A$, $Wx = 0$ implies that $Px = x$ and $P$ commutes with $V$ and with $g(V)$ since these operators commute with $W - A$.

We can now define

$$S = (2P - I)g(V) = g(V)(2P - I). \quad (150)$$
$S$ is self-adjoint. We may now define that $S$ satisfies (140). We set $\kappa = \lambda^2$ and define $h_1$ and $h_2$ by

\begin{align*}
    h_1(\kappa) &= \cos \lambda = 1 - \frac{1}{2!}\lambda^2 + \ldots \\
    \lambda h_2(\kappa) &= \sin \lambda = \lambda - \frac{1}{3!}\lambda^3 + \ldots
\end{align*}

These functions exist for all $\kappa$. Since $P$ is a projection, we have $(2P-I)^2 = 4P - 4P + I = I$ and so

$$S^2 = (2P-I)^2g(V)^2 = g(V)^2 \quad (152)$$

Therefore

\begin{align*}
    \cos S &= h_1(S^2) = h_1(g(V)^2) = \cos(g(V)) \\
    = V \\
    \sin S &= Sh_2(S^2) \\
    &= (2P-I)g(V)h_2(g(V)^2) \\
    &= (2P-I)\sin g(V) \\
    &= (2P-I)A = W
\end{align*}

and hence $U = e^{iS} = \cos S + i \sin S$. We note that $||S|| \leq \pi$. Since $S$ is self-adjoint and bounded, $\sigma(S)$ is real and from theorem 4.5 we conclude that $\sigma(S) \subset [-\pi, \pi]$. If $(E_\theta)$ is the spectral family of the operator $S$, then theorem 5.4 follows from the general theorem 4.14 for bounded linear operators. \qed

5.2 Cayley transform and the general spectral theorem

Any self-adjoint unbounded operator $T$ can be expressed in terms of unitary operator $U$ using a Cayley transform,

$$U = (T - iI)(T + iI)^{-1}. \quad (154)$$

The inverse $(T + iI)^{-1}$ exists since $T$ is self-adjoint with a real spectrum. The expression is very similar to a Möbius transformation. The spectrum of $T$ lies on the real axis of the complex plane, but the spectrum of the unitary operator is on the unit circle. A Möbius transformation like (154) maps the real axis onto the unit circle. We shall now show that the operator defined by (154) is unitary and hence it has a spectral representation.

**Theorem 5.5** (Cayley transform) The Cayley transformation

$$U = (T - iI)(T + iI)^{-1}.$$ 

of a self-adjoint linear operator $T : \mathcal{D}(T) \to H$ exists on $H$ and is a unitary operator.
Proof: First we observe that $U$ is isometric. For this, consider any $x \in H$ and set $y = (T + iI)^{-1}x$ and with direct calculation we obtain

$$||Ux||^2 = ||(T - iI)y||^2$$

$$= \langle (T - iI)y, (T - iI)y \rangle$$

$$= \langle Ty, Ty \rangle + i \langle Ty, y \rangle - i \langle y, Ty \rangle + \langle iy, iy \rangle$$

$$= \langle (T + iI)y, (T + iI)y \rangle$$

$$= ||(T + iI)y||^2$$

$$= ||(T + iI)(T + iI)^{-1}x||^2$$

$$= ||x||^2.$$

Second, we note that $U$ is surjective. Since $T$ is self-adjoint, the spectrum $\sigma(T)$ is real by theorem 4.7. Therefore, $i$ and $-i$ belong to the resolvent set $\rho(T)$ and, by the definition of the resolvent set, the inverses $(T + iI)^{-1}$ and $(T - iI)^{-1}$ exist on a dense subset of $H$ and are bounded. Since $T = T^*$, theorem 3.5 implies that $T$ is closed. From lemma 4.2, we see that those inverses are defined on all of $H$, that is

$$\mathcal{R}(T \pm iI) = H \quad (155)$$

Since $I$ is defined on all of $H$, we observe that

$$(T + iI)^{-1}(H) = \mathcal{D}(T + iI) = \mathcal{D}(T) = \mathcal{D}(T - iI) \quad (156)$$

and that

$$(T - iI)(\mathcal{D}(T)) = H. \quad (157)$$

Thus, the operator $U$ is a bijection of $H$ onto itself.

Because $U$ is isometric and surjective, it is unitary. \qed

Theorem 5.6 If $T : \mathcal{D}(T) \to H$ is a self-adjoint linear operator and $U$ is as defined in theorem 5.5, then

$$T = i(I + U)(I - U)^{-1} \quad (158)$$

Also, $1$ is not an eigenvalue of $U$.

Proof: Let $x \in \mathcal{D}(T)$ and

$$y = (T + iI)x. \quad (159)$$

Because $(T + iI)^{-1}(T + iI) = I$, acting on the left with $U$ we get

$$Uy = (T - iI)x. \quad (160)$$
By adding and subtracting the two equations, we get
\begin{align}
(I + U)y &= 2Tx \\
(I - U)y &= 2ix.
\end{align} \tag{161a}
\tag{161b}

From (155) we know that \( y \in \mathcal{R}(T + iI) = H \). From (161b) we see that \( I - U \) maps \( H \) onto \( \mathcal{D}(T) \) and also, that
\begin{align}
(I - U)y = 0 \iff 2ix = 0 \\
x = 0 \iff y = (T + iI)x = 0
\end{align}
so \((I - U)y = 0\) if and only if \( y = 0 \). Hence, there exists an inverse \((I - U)^{-1}\) and it is defined on the range of \( I - U \) i.e. \( \mathcal{D}(T) \). Thus, we can invert (161) to get
\[ y = 2i(I - U)^{-1}x \tag{162} \]
and by substituting into (161),
\[ Tx = \frac{1}{2}(I + U)y = i(I + U)(I - U)^{-1}x \quad \forall x \in \mathcal{D}(T) \tag{163} \]

Since the inverse \((I - U)^{-1}\) exists, 1 cannot be an eigenvalue of the Cayley transform \( U \). \( \square \)

**Theorem 5.7** (Spectral theorem for self-adjoint linear operators) Let \( T : \mathcal{D}(T) \to H \) be a self-adjoint linear operator, where \( H \neq \{0\} \) is a complex Hilbert space and \( \mathcal{D}(T) \) is dense in \( H \). Let \( U \) be a Cayley transform of \( T \) as in theorem 5.5 and let \((E_\theta)\) be the spectral family in the spectral representation of \(-U\) as in theorem 5.4. Then for all \( x \in \mathcal{D}(T) \),
\begin{align}
\langle Tx, x \rangle &= \int_{-\pi}^{\pi} \tan \frac{\theta}{2} dw(\theta) \\
&= \int_{-\infty}^{\infty} \lambda d\nu(\lambda)
\end{align} \tag{164}

where \( F_\lambda = E_{2\arctan \lambda} \)

*Proof:* Just as in the proof of theorem 5.4, we note that there exists an operator \( S \) such that \(-U = \cos S + i \sin S\). Since \( \sigma(S) \subset [-\pi, \pi] \), we have \( E_{-\pi-0} = 0 \). Hence if \( E_{-\pi} \neq 0 \), then \(-\pi\) would be an eigenvalue of \( S \). The operator \( U \) would then have the eigenvalue
\[ -\cos(-\pi) + i \sin(-\pi) = 1 \tag{165} \]
which is a contradiction as 1 is not an allowed eigenvalue for $U$ by theorem 5.6 and hence $E_\theta$ is continuous. Similarly, $E_\pi = I$ and if $E_\pi \neq I$, this would also mean that 1 is an eigenvalue of $U$.

Let $x \in H$ and $y = (I - U)x$. Then $y \in D(T)$ since $I - U : H \to D(T)$ as shown in proof of theorem 5.6. Inserting these into (158), we get
\[
Ty = i(I + U)(I - U)^{-1}y = i(I + U)x
\] (166)
Since unitary operators are isometric,
\[
\langle Ty, y \rangle = \langle i(I + U)x, (I - U)x \rangle
\] (167)
\[
 = i \left( \langle x, x \rangle - \langle Ux, x \rangle - \langle x, Ux \rangle - \langle Ux, Ux \rangle \right)
\] (168)
\[
 = i \left( ||x||^2 - \langle Ux, x \rangle - \overline{\langle Ux, x \rangle} - ||Ux||^2 \right)
\] (169)
\[
 = -2i \text{Im} \langle Ux, x \rangle
\] (170)
\[
 = 2 \int_{-\pi}^{\pi} \sin \theta d \langle E_\theta x, x \rangle.
\] (171)
Hence, using the relation $\sin \theta = 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}$,
\[
\langle Ty, y \rangle = 4 \int_{-\pi}^{\pi} \sin \frac{\theta}{2} \cos \frac{\theta}{2} \langle E_\theta x, x \rangle.
\] (172)
Since $E_\theta$ is the spectral family of the operator $S$, they commute and by lemma 5.2. From (139) we obtain
\[
\langle E_\theta y, y \rangle = \langle E_\theta (I - U)x, (I - U)x \rangle
\] (173)
\[
 = \langle (I - U)^* (I - U) E_\theta x, x \rangle
\] (174)
\[
 = \int_{-\pi}^{\pi} (1 + e^{-i\phi})(1 + e^{i\phi})d \langle E_\theta z, x \rangle
\] (175)
where $z = E_\theta x$. Since $E_\theta E_\phi = E_\phi$ when $\phi \leq \theta$ and $E_\phi E_\theta = E_\theta$ when $\phi > \theta$, and $(1 + e^{-i\phi})(1 + e^{i\phi}) = 4 \cos^2 \frac{\phi}{2}$, we get
\[
\langle E_\theta y, y \rangle = \int_{-\pi}^{\pi} 4 \cos^2 \frac{\phi}{2} \langle E_\theta z, x \rangle.
\] (176)
From this and using the continuity of $E_{\theta}$ at $\pm \pi$, we arrive at

$$\int_{-\pi}^{\pi} \tan \theta \frac{d}{2} \langle E_{\theta} y, y \rangle = \int_{-\pi}^{\pi} \tan \theta \left(4 \cos^2 \frac{\theta}{2}\right) d \langle E_{\theta} x, x \rangle$$

$$= 4 \int_{-\pi}^{\pi} \sin \frac{\theta}{2} \cos \frac{\theta}{2} d \langle E_{\theta} x, x \rangle = \langle Ty, y \rangle.$$

This expression is equivalent to (164). The alternative form in the theorem is obtained with substitution $\theta = 2 \arctan \lambda$.

\[\square\]

6 Applications in quantum mechanics

Here the results derived in this thesis and their application to quantum mechanics is discussed. The standard way of interpreting quantum mechanics dictates that observables can be associated to unbounded self-adjoint operators on a suitable Hilbert space and the spectrum of each operator coincides with the values the observable may attain. The spectral theorem allows the construction of complicated operators starting from projections or to decompose operators into projections on the spectrum.[2]

Historically the spectral theorem for self-adjoint operators was first proved by John von Neumann. The development of this theorem is deemed crucial for the Hilbert-space formulation of quantum mechanics and is one of the most important achievements in mathematics and mathematical physics of the 20th century. Von Neumann also built the modern axiomatic notion of an abstract Hilbert space by considering the first two approaches to quantum mechanics by Heisenberg and Schrödinger. [2, 6]

We shall define quantum mechanics using an axiomatic approach suggested first by von Neumann:[7–9]

1. States $\psi$ of a quantum system are nonzero vectors of a complex separable Hilbert space $H$, considered up to a nonzero complex factor. There is a one-to-one correspondence between observables and linear self-adjoint operators in $H$. We consider states as unit vectors in $H$. These state vectors contain the most complete information available of the system.

2. Observables are represented by self-adjoint linear operators on $H$. Each observable $\hat{A}$ is defined maximally, on a dense subset $\mathcal{D}(\hat{A}) \subseteq H$.

3. When an observable is measured on a state $\psi \in H$, the result is always one of the values in the spectrum of $A$. The expectation value of the measurement of $A$ is
computed as a mean value \( \langle \hat{A} \rangle_\psi \) of the operator at state \( \psi \). Observables \( A_1, \ldots, A_n \) are simultaneously measurable if and only if the self-adjoint operators \( \hat{A}_1, \ldots, \hat{A}_n \) mutually commute.

4. There exists a one parameter group \( U_t \) of unitary operators called evolution operators that map an initial state \( \psi_0 \) at the time \( t = 0 \) to the state \( \psi(t) = U_t \psi_0 \) at the time \( t \). The operator \( U_t \) is of the form

\[
U_t = \exp \left( -i \frac{t \hat{H}}{\hbar} \right)
\]

where \( \hbar \) is the Planck constant and \( \hat{H} \) is called the Hamiltonian of the system. If \( \psi_0 \in \mathcal{D}(\hat{H}) \), then \( \psi(t) \) is differentiable and

\[
i \hbar \frac{d\psi(t)}{dt} = H \psi(t).
\]

As stated by the axioms, an observable \( A \) (energy of the system, position or momentum of a particle etc.) is associated with a self-adjoint operator \( \hat{A} \) with the domain \( \mathcal{D}(\hat{A}) \). The dynamics of the system are dictated by the Schrödinger equation (174) and the solutions to this equation, \( \psi \), are states of the system. Quantum mechanics describes the microscopic structure of matter and maps stochastically onto the macroscopic world; while the Schrödinger equation and the dynamics thereof are deterministic, the quantum mechanics may only assign probabilities to different macroscopically observable values. The Born interpretation maps the state vector \( \psi \) to a probability distribution \( P \), stating that the probability of finding the particle in an interval \([x, x + \Delta x]\) at time \( t \) is given by

\[
P(t) = \int_x^{x+\Delta x} \frac{\Psi(x,t)\overline{\Psi(x,t)}}{\int_x^{x+\Delta x} \abs{\Psi(x,t)}^2 dx} \Delta x. \tag{175}
\]

If \( \psi \in \mathcal{D}(\hat{A}) \), the expectation value \( \langle A \rangle_\psi \) of the observable \( A \) exists and is

\[
\langle A \rangle_\psi = \langle \hat{A} \psi, \psi \rangle. \tag{176}
\]

This expectation value corresponds to the average of a large set of experimental measurements measuring \( A \) conducted on a system in state \( \psi \).

Since the observable is expressed as an self-adjoint operator \( \hat{A} \) in a Hilbert space \( H \), we may rewrite this equation using the spectral theorem for unbounded self-adjoint operators, theorem 5.7,

\[
\langle A \rangle_\psi = \langle \hat{A} \psi, \psi \rangle = \int_{-\infty}^{\infty} \lambda \nu(\lambda) \quad \nu(\lambda) = \langle E_\lambda \psi, \psi \rangle. \tag{177}
\]
Now we see that the expectation value of the operator is an integral over the spectrum and that the possible measurement outcomes lie on the spectrum of the operator. Let us consider a dispersion of an observable $A$ as

$$\delta_\psi A = ||\hat{A}\psi - \langle A \rangle_\psi\psi||^2.$$  \hfill (178)

We note that if the state $\psi$ of the system is an eigenstate of the observable, then the expectation value is the eigenvalue of the state and the dispersion of the observable is zero:

$$\langle A \rangle_\psi = \langle \hat{A}\psi, \psi \rangle = \lambda \langle \psi, \psi \rangle = \lambda$$

$$\Rightarrow \delta_\psi A = ||\hat{A}\psi - \langle A \rangle_\psi\psi||^2$$

$$= ||\lambda\psi - \lambda\psi||^2 = 0.$$  

Let us now consider two operators $\hat{A}$ and $\hat{B}$ and let $\psi$ be a vector. How does the ordering of the operators affect our result? Let $\hat{A}_1 = \hat{A} - \langle A \rangle_\psi I$ and $\hat{B}_1 = \hat{B} - \langle B \rangle_\psi I$. We denote uncertainties related to the measurement of $A$ and $B$ as

$$\Delta A = ||\hat{A}_1\psi|| = \sqrt{\delta A_\psi}$$

$$\Delta B = ||\hat{B}_1\psi|| = \sqrt{\delta B_\psi}.$$  

Using the fact that

$$\hat{A}_1\hat{B}_1 - \hat{B}_1\hat{A}_1 = \hat{A}\hat{B} - \hat{B}\hat{A},$$  \hfill (179)

let us consider the expectation value of the commutator of $\hat{A}$ and $\hat{B}$ i.e. $(\hat{A}\hat{B} - \hat{B}\hat{A})$:

$$\left| \langle (\hat{A}\hat{B} - \hat{B}\hat{A})\psi, \psi \rangle \right| = \left| \langle \hat{A}_1\hat{B}_1 - \hat{B}_1\hat{A}_1\psi, \psi \rangle \right|$$

$$= \left| \langle \hat{A}_1\hat{B}_1\psi, \psi \rangle - \langle \hat{B}_1\hat{A}_1\psi, \psi \rangle \right|$$

$$= \left| \langle \hat{B}_1\psi, \hat{A}_1\psi \rangle - \langle \hat{A}_1\psi, \hat{B}_1\psi \rangle \right|$$

$$= 2 \left| \text{Im} \langle \hat{A}_1\psi, \hat{B}_1\psi \rangle \right| \leq 2 \left| \langle \hat{A}_1\psi, \hat{B}_1\psi \rangle \right|$$

$$\leq 2 \left| \langle \hat{A}_1\psi, \psi \rangle \right| \left| \langle \hat{B}_1\psi, \psi \rangle \right|$$

$$= 2\Delta A\Delta B$$

$$\Rightarrow \Delta A\Delta B \geq \frac{1}{2} \left| \langle (\hat{A}\hat{B} - \hat{B}\hat{A})\psi, \psi \rangle \right|$$

where the Schwartz inequality has been used. The last line is known as the Heisenberg uncertainty relation and it sets a fundamental limit to the precision at which the values
of two non-commuting observables, such as position and momentum, can be measured. If one knows the position of the particle to high precision, the uncertainty of particles momentum must be high for this fundamental inequality to be satisfied.

The time-evolution of a quantum mechanical system is given by the Schrödinger equation of (174). Let us have a look the equation for an elementary one-particle system in one dimension.

\[
\frac{d}{dt} \Psi(x,t) = \hat{H} \Psi(x,t)
\]

(180)

where \(x\) is the spatial coordinate of the particle, \(t\) is time, \(\psi(x,t)\) is called the wave function or the state of the system. The dynamics of the system is determined by the Hamiltonian \(\hat{H}\). The Hamiltonian consists of a kinetic and potential energy contributions and for the one-particle system we may write it as

\[
\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x,t),
\]

where \(\hbar\) is the reduced Planck's constant and \(m\) is the mass of the particle. The first term describes the kinetic energy of the particle and \(V(x,t)\) is the external potential acting on the particle. While the kinetic energy is always the same, the potential energy is system specific. If the potential in the Hamiltonian has no explicit time-dependence, i.e. \(V(x,t) = V(x)\), then (180) becomes separable and we may write \(\Psi(x,t) = \psi(x)\phi(t)\) to obtain

\[
\frac{d}{dt} \phi(t) = \hat{H} \phi(t).
\]

(181)

This equation can be solve directly as

\[
\phi(t) = \exp(-it\hat{H})\phi(0).
\]

(182)

In a finite dimensional case the Hamiltonian corresponds to a matrix and in order to compute the matrix exponential and to grasp the dynamics of the system one needs to diagonalize the matrix. However, as we discussed in the introduction, going from finite dimensional system to an infinite dimensional system is not straightforward. Diagonalization of a self-adjoint matrix essentially corresponds to a change in basis, but for infinite-dimensional spaces no good theory of bases exists. However, the spectral theorem allows us to rewrite the operator using orthogonal projections, making the operator exponential feasible.

The spatial part separates into

\[
\hat{H} \psi(x) = \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x) = E \psi(x)
\]

(183)
where $E$ is a constant that corresponds to the total energy of the system. This equation is called the time-independent Schrödinger equation. It suggests that the possible energy levels of the system depend on the spectrum of the Hamiltonian operator. For the Born interpretation to be meaningful, the wave function solutions should be square integrable,

$$\psi(x) \in L^2(\mathbb{R}) \iff \int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty. \quad (184)$$

If $E$ is in the resolvent set of $\hat{H}$, then the resolvent of $\hat{H}$ exists and there exists only the trivial solution considered in $L^2(\mathbb{R})$. If on the other hand $E$ belongs to the point spectrum $\sigma_p(\hat{H})$ or to the continuous spectrum $\sigma_c(\hat{H})$, then there exists nontrivial solutions. The residual spectrum $\sigma_r(\hat{H})$ is empty by theorem 4.9 which can be extended also to unbounded operators. [3]

Let us look at the spectrum of two simple quantum mechanical systems: the harmonic oscillator and the free particle. For the quantum mechanical harmonic oscillator, the potential of the system is defined as

$$V(x) = \frac{1}{2}m\omega^2 x^2$$

where $\omega$ is the angular frequency of the oscillator. The solutions for the harmonic oscillator Schrödinger equation can be written as Hermite functions,

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} \left( \frac{m\omega}{\pi\hbar} \right)^{1/4} \exp \left( -\frac{m\omega x^2}{2\hbar} \right) H_n \left( \sqrt{\frac{m\omega}{\hbar}} x \right), \quad n = 0, 1, 2\ldots \quad (185)$$

where the functions $H_n$ are the Hermite polynomials given by

$$H_n(z) = (-1)^n e^{z^2} \frac{d^n}{dz^n} e^{-z^2}. \quad (186)$$

The energy levels of the system are then given as

$$E_n = \hbar \omega \left( \omega + \frac{1}{2} \right) \quad n = 0, 1, 2, \ldots. \quad (187)$$

These levels correspond to the total mechanical energy that the quantum oscillator may assume. It is also evident from the unbounded nature of the eigenvalues that $\hat{H}$ is also unbounded. The solutions are square integrable only for integer values of $n$. Because of this, the spectrum of the harmonic oscillator is pure point spectrum. In a classical oscillator, the energy levels would not be discrete but vary continuously.
Consider next the quantum mechanical free particle, i.e. with potential function $V(x) = 0$. The solution to the Schrödinger equation is an eigenstate of the kinetic energy operator which corresponds to a simple complex plane wave,

$$\psi(x) = A \exp(-ikx)$$  \hspace{1cm} (188)

where the parameter $k$ is related to the energy,

$$E = \frac{\hbar^2 k^2}{2m}. \hspace{1cm} (189)$$

Unlike in the case of the harmonic oscillator, the energy of the system may now vary continuously with the parameter $k$. This solution is not itself square-integrable: by the Heisenberg uncertainty principle, if the particle is in a momentum eigenstate, then the wave function cannot be localised anywhere. However, we can combine a set of solutions with different $k$ to represent any $\psi \in L^2(\mathbb{R})$ as a wave packet,

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(k) \exp(-ikx) dk$$  \hspace{1cm} (190)

where

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(q) \exp(-ikq) dq.$$  \hspace{1cm} (191)

Due to linearity, the wave packet is also a solution to the Schrödinger equation, although not an eigenstate of the Hamiltonian.

In general, when $E \in \sigma_p(\hat{H})$, such as in the case of a quantum mechanical harmonic oscillator, the solutions $\psi$ are in $L^2(\mathbb{R})$ but when $E \in \sigma_c(\hat{H})$, such as in the case of the free particle, the equation (183) has no nonzero solutions in $L^2(\mathbb{R})$. However, from a set of nonzero solutions it may be possible to integrate a solution in $L^2(\mathbb{R})$. \cite{2, 3}

References


