Spectral Theorem For Bounded Self-Adjoint Opeartors on Hilbert Spaces

Lance Remigio

May 18, 2025

Contents

Abstract

This paper presents a generalization of the spectral theorem from finite-dimensional inner product spaces to the broader context of Hilbert spaces. We focus on the diagonalization of bounded self-adjoint linear operators, examining the structural and functional properties necessary for this extension. In particular, we investigate positive operators, their positive square roots, orthogonal projections, and the spectral families associated with such operators. These tools enable us to represent self-adjoint operators via Riemann–Stieltjes integrals. We conclude by applying the spectral theorem to diagonalize convolution operators, thereby revealing their structure as multiplication operators in the frequency domain. This not only simplifies their analysis but also highlights the deep connection between convolution, Fourier analysis, and the spectral properties of self-adjoint operators.

1 Preliminary Results/Notation

We dedicate this section to stating key results that are frequently used throughout the paper. While we will not include full proofs (as proving them will take us outside the scope of the paper), we assume the reader has at least some familiarity with these theorems. Readers may omit this section and skip to the beginning of the paper here introduction.

1.1 Linear Operators

Definition (Linear Operator Kreyszig (2.6-1)). A **Linear Operator** T is an operator such that

(i) The domain $D(T)$ of T is a vector space and the range $R(T)$ lies in a vector space over the same field,

(ii) For all $x, y \in D(T)$ and scalars α, we have

$$T(x+y) = Tx + Ty$$
$$T(\alpha x) = \alpha Tx.$$

Definition (Bounded Linear Operator Kreyszig (2.7-1)). Let X and Y be normed spaces and $T : D(T) \to Y$ a linear operator, where $D(T) \subseteq X$. The operator T is said to be **bounded** if there is a real number c such that for all $x \in D(T)$,

$$\|Tx\| \leq c\|x\|.$$

1.2 Inner Product Spaces and Hilbert Spaces

Definition (Inner Product Space Kreyszig (3.1-1)). An **inner product space** is a vector space X with an inner product defined on X. A **Hilbert space** is a complete inner product space (complete in the metric defined by the inner product). An **inner product** on X is a mapping of $X \times X$ into the scalar field K of X; that is, with every pair of vectors x and y there is associated a scalar which is written

$$\langle x, y \rangle$$

and is called the **inner product** of x and y, such that for all $x, y, z \in V$ and $\alpha \in K$, we have

(IP1) $\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

(IP2) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$

(IP3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$

(IP4) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.

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

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

and a metric on X given by

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

Corollary (Continuity, null space Kreyszig (2.7-10)). Let T be a bounded linear operator. Then:

(a) $x_n \to x$ [where $x_n, x \in D(T)$] implies $Tx_n \to Tx$.

(b) The null space $N(T)$ is closed.

Lemma (Continuity of Inner Product Kreyszig (3.2-2)). If in an inner product space, $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n \rangle \to \langle x, y \rangle$.

Definition (Projection Kreyszig 3.3). Let H be a Hilbert space and let Y be some non-empty subspace of H. A **projection** is a map $P : H \to Y$ such that

$$Px = y$$

where $x \in H$ and $y \in Y$.

Lemma (Null Space Kreyszig (3.3-5)). The orthogonal complement $Y^{\perp}$ of a closed subspace Y of a Hilbert Space H is the null space $N(P)$ of the orthogonal projection P of H onto Y.

Lemma (Equality Kreyszig (3.8-2)). 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 $v_1 = 0$.

Theorem (Self-Adjointness of Product Kreyszig (3.10-4)). The product of two bounded self-adjoint linear operators S and T on a Hilbert Space H is self-adjoint if and only if the operators commute,

$$ST = TS.$$

Theorem (Uniform Boundedness Theorem Kreyszig (4.7-3)). Let (T_n) be a sequence of bounded linear operators $T_n : X \to Y$ from a Banach space X into a normed space Y such that $\|T_n x\|$ is bounded for every $x \in X$, say,

$$\|T_n x\| \leq c_x \qquad (n \in \mathbb{N})$$

Then the sequence of the norms $\|T_n\|$ is bounded, that is, there is a c such that

$$\|T_n\| \leq c. \qquad (n \in \mathbb{N})$$

Definition (Self-Adjoint, Unitary, Normal Operators Kreyszig 3.10-1). A bounded linear operator $T : H \to H$ on a Hilbert space H is said to be **self-adjoint** or **Hermitian** if $T^* = T$, **unitary** if T is bijective i.e $T^* = T^{-1}$, and **normal** if $TT^* = T^*T$.

1.3 Resolvent and Spectrum of T

Suppose $T : D(T) \to H$ where H is a Hilbert space. We associate T with the real eigenvalue λ as

$$T_\lambda = T - \lambda I$$

where I is the identity operator on $D(T)$.

Definition (Resolvent Operator Kreyszig 7.2). Assuming that T_λ contains an inverse, define the **resolvent operator** $R_\lambda(T)$ by

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

In the remainder of the section, we will denote T's resolvent operator as R_λ for convenience. Below, we will give a brief overview of the resolvent and spectrum sets of T.

Definition (Regular value, Resolvent set, Spectrum Kreyszig (7.2-1))**.** Let H be a complex Hilbert Space and $T : D(T) \to H$ be a linear operator with domain $D(T) \subseteq H$. A **regular value** λ of T is a complex number such that

(I) $R_\lambda(T)$ exists,

(II) $T_\lambda(T)$ is bounded,

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

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

Below are the three distinctions usually discussed about the spectrum of T:

- The **point spectrum** $\sigma_p(T)$ is the set such that $R_\lambda(T)$ does not exist.
- The **continuous spectrum** $\sigma_c(T)$ is the set such that $R_\lambda(T)$ exist and satisfies (III) but is not bounded.
- The **resdidual spectrum** $\sigma_r(T)$ is the set such that $R_\lambda(T)$ exists (may or may not be bounded) but does not satisfy (III).

1.4 Definitions and Results from Analysis

Definition (Fourier Transform Folland (7.2-1))**.** Suppose f is an integrable function that is defined on $\mathbb{R}$. The **Fourier Transform** of f is defined by the function $\hat{f}$ on $\mathbb{R}$ where

$$\hat{f}(\xi) = \int e^{-i\xi x} f(x)\ dx.$$

Definition (L^1 space)**.** We denote $L^1([a,b])$ by the set

$$L([a,b]) = \left\{ f[a,b] \to \mathbb{R} : \int_a^b |f(x)|\ dx < \infty \right\}.$$

Definition (L^2 spaces Folland (3.3))**.** We call $L^2([a,b])$ the space of **square-integrable** functions on $[a,b]$, more specifically,

$$L^2([a,b]) = \left\{ f : \int_a^b |f(x)|^2\ dx < \infty \right\}$$

Remark. If $[a,b]$ is replaced by $\mathbb{R}$, then the Riemann integral in the previous two definitions would turn into an improper integral. That is,

$$L^2(\mathbb{R}) = \left\{ f : [a,b] \to \mathbb{R} : \int_{-\infty}^{+\infty} |f(x)|^2\ dx < \infty \right\}$$

and

$$L^1(\mathbb{R}) = \left\{ f : [a,b] \to \mathbb{R} : \int_{-\infty}^{\infty} |f(x)|\ dx < \infty \right\}.$$

Remark. Another way to write $\hat{f}(\xi)$ is $\mathcal{F}[f(x)] = \hat{f}(\xi)$.

Remark.
- The refers to the space where the Fourier transform resides.
- The refers to the space where the function that is being transformed under.
- The parameter $\xi \in \mathbb{R}$ refers to the of the Fourier transform.

Lemma (Riemann-Lebesgue Lemma Folland (7.2-1))**.** If $f \in L^1$, then $\hat{f}(\xi) \to 0$ as $\xi \to \pm\infty$.

For sake of convenience, we use the notation $f \in R[a, b]$ to denote Riemann-integrable functions over a compact interval $[a, b]$ in $\mathbb{R}$.

Theorem (Fubini's Theorem Rudin)**.** Let $f : [a, b] \times [c, d]$ be a bounded function. Suppose f is Riemann-integrable on the rectangle $[a, b] \times [c, d]$. For each fixed $x \in [a, b]$, the function $y \mapsto f(x, y) \in R[c, d]$ and for each fixed $y \in [c, d]$, the function $x \mapsto f(x, y) \in R[a, b]$. Then the iterated integrals exist and satisfy:

$$\int_a^b \Big(\int_c^d f(x, y)\ dy \Big)\ dx = \int_c^d \Big(\int_a^b f(x, y)\ dx \Big)\ dy = \int_{[a,b]\times[c,d]} f(x, y) dx dy.$$

Theorem (Convolution Theorem Rudin (8.14))**.** Let $f, g \in L^1(\mathbb{R})$, and define their **convolution** by

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x - t) g(t)\ dt.$$

Then $f * g \in L^1(\mathbb{R})$ and the Fourier Transform of the convolution satisfies:

$$\mathcal{F}(f * g)(\xi) = \mathcal{F}(f)(\xi) \cdot \mathcal{F}(g)(\xi),$$

where

$$\mathcal{F}(f)(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi}\ dx$$

is the Fourier transform of $f \in L^1(\mathbb{R})$.

Definition (Partition)**.** A partition P of an interval $[a, b] \subseteq \mathbb{R}$ is a finite set of points in $[a, b]$ that includes both a and b. Denote $P = \{x_0, x_1, \ldots, x_n\}$ whose elements are listed in the following increasing order

$$a = x_0 < x_1 < \cdots < x_n = b.$$

Definition (Riemann-Stieltjes Integrable Functions Rudin (Chapter 6))**.** Let $\alpha : [a, b] \to \mathbb{R}$ be an increasing function. A function $f : [a, b] \to \mathbb{R}$ is said to be Riemann-Stieltjes integrable on $[a, b]$ if

(i) f is bounded

(ii) $L(f, \alpha) = U(f, \alpha)$

In this case, the Riemann-Stieltjes integral of f with respect to α, denoted by

$$\int_a^b f\ d\alpha$$

is the common value of $L(f,\alpha)$ and $U(f,\alpha)$. That is,

$$\int_a^b f\ d\alpha = L(f,\alpha) = U(f,\alpha).$$

Remark. $U(f,\alpha) = \inf_{P\in\pi[a,b]} U(f,\alpha,P)$ and $L(f,\alpha) = \sup_{P\in\pi[a,b]} L(f,\alpha,P)$ above are the upper and lower integrals of f where $U(f,\alpha,P)$ and $L(f,\alpha,P)$ are the upper and lower sum with respect to the partition P of $[a,b]$.

2 Introduction

2.1 Motivation

Recall that in a finite-dimensional unitary space $H = \mathbb{C}^n$, any self-adjoint linear operator $T : H \to H$ can be represented as a Hermitian matrix A with respect to some orthonormal basis of H.

This matrix representation has real eigenvalues, as guaranteed by 9.1-1. Since H is n-dimensional, we can label these eigenvalues as $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, corresponding to an orthonormal set of eigenvectors $\beta = \{x_1, x_2, \ldots, x_n\}$. This set β forms a basis for H, and each eigenvector x_j corresponds to the eigenvalue λ_j $(1 \le j \le n)$. These vectors form the columns of the matrix A representing T. As β is an orthonormal basis, any vector $x \in H$ can be uniquely written as a linear combination of the vectors in β. In particular,

$$\begin{aligned} x &= \sum_{j=1}^n \gamma_j x_j \qquad (1) \\ \gamma_j &= \langle x, x_j \rangle, \end{aligned}$$

where the coefficients γ_j are obtained by projecting x onto each basis vector x_j.

Applying the linearity of T to (1), we obtain

$$Tx = \sum_{j=1}^n \gamma_j \lambda_j x_j. \qquad (2)$$

Although T may seem complicated in general, its action on each eigenvector x_j is particularly simple: it scales x_j by λ_j.

Next, define an operator $P_j : H \to H$ by $P_j(x) = \gamma_j x_j$. Each P_j is the orthogonal projection of x onto the eigenspace corresponding to λ_j. Then equation (1) can be written as

$$x = \sum_{j=1}^n P_j x \quad \Rightarrow \quad I = \sum_{j=1}^n P_j, \qquad (3)$$

where I denotes the identity operator on H. Similarly, equation (2) becomes

$$Tx = \sum_{j=1}^n \lambda_j P_j x \quad \Rightarrow \quad T = \sum_{j=1}^n \lambda_j P_j. \qquad (4)$$

This shows that the self-adjoint operator T admits a simple and elegant decomposition in terms of orthogonal projections. However, this nice finite-dimensional representation does not extend directly to infinite-dimensional inner product spaces. As is often the case, difficulties arise from handling the infinite structure.

To generalize this representation, we introduce the concept of a **spectral family of projections** $\{E_\lambda\}_{\lambda\in\mathbb{R}}$, which captures the cumulative effect of projections up to a certain threshold λ. In the finite-dimensional setting, we can define

$$E_\lambda = \sum_{\lambda_j \leq \lambda} P_j. \tag{5}$$

Here, E_λ is the projection of H onto the subspace spanned by those eigenvectors x_j for which $\lambda_j \leq \lambda$.

Definition (Spectral Family or Decomposition of Unity, Kreyszig (9.7-1))**.** A real spectral family (or real *decomposition of unity*) is a one-parameter family $\mathcal{G} = \{E_\lambda\}_{\lambda\in\mathbb{R}}$ of projections E_λ on a Hilbert space H satisfying:

(i) $E_\lambda \leq E_\mu$ for $\lambda < \mu$ (i.e., $E_\lambda E_\mu = E_\mu E_\lambda = E_\lambda$),

(ii) $\lim\limits_{\lambda\to-\infty} E_\lambda x = 0$ for all $x \in H$,

(iii) $\lim\limits_{\lambda\to+\infty} E_\lambda x = x$ for all $x \in H$,

(iv) $\lim\limits_{\mu\to\lambda^+} E_\mu x = E_\lambda x$ for all $x \in H$.

This means a spectral family is a mapping from $\mathbb{R}$ to $B(H)$, the space of all bounded linear operators on H. Note that condition (i) is bidirectional: $E_\lambda \leq E_\mu$ if and only if $E_\lambda E_\mu = E_\mu E_\lambda = E_\lambda$ (see Kreyszig 9.6-1).

We say that $\mathcal{G}$ is a **spectral family on an interval** $[a, b]$ if:

(i) $E_\lambda = 0$ for $\lambda < a$,

(ii) $E_\lambda = I$ for $\lambda \geq b$.

This restriction to a compact interval will prove useful when studying the spectrum of bounded self-adjoint operators, particularly in expressing them as Riemann–Stieltjes integrals. In finite dimensions, the integral representation reduces to a finite sum.

Assuming distinct eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, we find:

$$\begin{aligned}
E_{\lambda_1} &= P_1, \\
E_{\lambda_2} &= P_1 + P_2, \\
&\vdots \\
E_{\lambda_n} &= P_1 + \cdots + P_n.
\end{aligned}$$

Conversely,

$$\begin{aligned}
P_1 &= E_{\lambda_1}, \\
P_j &= E_{\lambda_j} - E_{\lambda_{j-1}} \quad \text{for } 2 \leq j \leq n.
\end{aligned}$$

Since E_λ remains constant on each interval $[\lambda_{j-1}, \lambda_j)$, we can express the projections as

$$P_j = E_{\lambda_j} - E_{\lambda_j - 0}.$$

Thus, equation (4) becomes:

$$T = \sum_{j=1}^{n} \lambda_j P_j = \sum_{j=1}^{n} \lambda_j (E_{\lambda_j} - E_{\lambda_j-0}). \tag{*}$$

This equation (*) gives the *spectral representation* of the self-adjoint operator T on the n-dimensional Hilbert space H. For any $x, y \in H$, we have:

$$\langle Tx, y\rangle = \sum_{j=1}^{n} \lambda_j \langle (E_{\lambda_j} - E_{\lambda_j-0})x, y\rangle.$$

This sum can be written as a Riemann–Stieltjes integral:

$$\langle Tx, y \rangle = \int_{-\infty}^{+\infty} \lambda \, dw(\lambda),$$

where $w(\lambda) = \langle E_\lambda x, y \rangle$. This integral form captures the action of T through the cumulative effect of the spectral family $\{E_\lambda\}$, and serves as the prototype for the spectral representation in infinite-dimensional Hilbert spaces.

2.2 Main Theorem

Theorem (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:

(a) T has the spectral representation

$$T = \int_{m-0}^{M} \lambda \; dE_\lambda \tag{1}$$

where $\mathcal{G} = (E_\lambda)$ is the spectral family associated with T; the integral is to be understood in the sense of uniform convergence (convergence in the norm on $B(H, H)$), for all $x, y \in H$

$$\langle Tx, y \rangle = \int_{m-0}^{M} \lambda \; dw(\lambda) \tag{1*}$$

where $w(\lambda) = \langle E_\lambda x, y \rangle$ where the integral is an ordinary Riemann-Stieltjes integral.

(b) More generally, if p is a polynomial in λ with real coefficients,

$$p(\lambda) = \alpha_n \lambda^n + \alpha_{n-1} \lambda^{n-1} + \cdots + \alpha_0,$$

then the operator $p(T)$ defined by

$$p(T) = \alpha_n T^n + \alpha_{n-1} T^{n-1} + \cdots + \alpha_0 I$$

has the spectral representation

$$p(T) = \int_{m-0}^{M} p(\lambda) \; dE_\lambda \tag{2}$$

and for all $x, y \in H$

$$\langle p(T)x, y \rangle = \int_{m-0}^{M} p(\lambda) \; dw(\lambda) \tag{2*}$$

where $w(\lambda) = \langle E_\lambda x, y \rangle$.

Our goal, for the remainder of this paper, is to develop the integral representation outlined in the spectral theorem above. To guide the reader, we provide an outline of the paper's structure below.

- In section 3, we introduce some fundamental properties of bounded self-adjoint linear operators, laying the groundwork for understanding their spectral behavior and setting the stage for the spectral representation.
- In section 4, we extend this foundation by exploring positive operators and the square roots of operators, highlighting how these concepts help establish the framework for operator convergence—an essential component of the spectral theorem.
- In section 5 , we further enrich our understanding by introducing orthogonal projections. Building on the results of the previous sections, we examine how projections contribute to the convergence of the integral representation and serve as the key link to the spectral resolution of operators.
- Leading up to the proof of the spectral theorem (section 6) , we introduce the notion of spectral families and analyze their properties. This analysis reveals how spectral families encapsulate the

behavior of bounded self-adjoint operators and facilitate their representation via Riemann-Stieltjes integrals.

- Finally, after establishing the proof of the spectral theorem (section 7), we present a general convolution operator on $L^2(\mathbb{R})$ in terms of Riemann-Stieltjes integral in ours spectral theorem by using tools from Fourier analysis. In particular, we will present a more concrete example by looking at the Gaussian Kernel of a Fourier Transform and representing it in terms of its spectral representation.

3 Properties of Bounded Self-Adjoint Linear Operators

In this chapter, we will give a brief review of linear self-adjoint operators and their spectral properties.

First, we will give a brief definition of self-adjoint and some properties as a consequence.

3.1 Basic Properties of Linear Self-adjoint Operators

Definition (Hilbert-Adjoint Operator Kreyszig 9.1)**.** Let $T : H \to H$ be a bounded linear operator on a complex Hilber space H. Then the **Hilbert-adjoint** operator $T^* : H \to H$ is defined to be the operator satisfying

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle \quad \forall x, y \in H. \tag{*}$$

Note that T is said to be **self-adjoint** or **Hermitian** if

$$T = T^*.$$

If the above holds, then it follows from (*) that

$$\langle T(x), y\rangle = \langle x, T(y)\rangle \quad \forall x, y \in H.$$

As our first step towards the representation shown in the last section, we need to make sure that T induces real eigenvalues and orthogonal eigenvectors.

Theorem (Eigenvalues, eigenvectors Kreyszig 9.1-1)**.** Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilber space H. Then:

(a) All the eigenvalues of T (if they exist) are real.

(b) Eigenvectors corresponding to (numerically) different eigenvalues of T are orthogonal.

Proof. (a) Our goal is to show that $\lambda = \overline{\lambda}$ for all eigenvalues λ of T. To this end, Let λ be an eigenvalue of T and x be the corresponding eigenvector. Then $x \neq 0$ and $Tx = \lambda x$. Since T is self-adjoint, it follows that

$$\begin{aligned}\lambda\langle x, x\rangle = \langle \lambda x, x\rangle &= \langle Tx, x\rangle \\ &= \langle x, Tx\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle.\end{aligned}$$

Since $\langle x, x\rangle \neq 0$, then it follows that $\lambda = \overline{\lambda}$, implying that λ must be real.

(b) Our goal is to show that for any two (distinct) eigenvalues λ and μ corresponding to eigenvectors x and y, respectively, we have $\langle x, y\rangle = 0$. To this end, let λ and μ be eigenvalues of T, and let x and y be corresponding eigenvectors. Then it follows that $Tx = \lambda x$ and $Ty = \mu y$. Since T is self-adjoint and μ is real, we have

$$\begin{aligned}\lambda\langle x, y\rangle = \langle \lambda x, y\rangle &= \langle Tx, y\rangle \\ &= \langle x, Ty\rangle = \langle x, \mu y\rangle = \overline{\mu}\langle x, y\rangle = \mu\langle x, y\rangle.\end{aligned}$$

Since we have assumed that $\lambda \neq \mu$, it follows that $\langle x, y\rangle = 0$. Hence, x and y are orthogonal.

■

3.2 Spectrum and Resolvent sets of T and Properties

We will now discuss the consequences of the self-adjointness of T on its spectrum and resolvent sets.

The following theorem gives us the ability to test whether λ is in the spectrum or in its resolvent set.

Theorem (Resolvent Set Kreyszig (9.1-2)). Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H. Then a number λ belongs to the resolvent set $\rho(T)$ of T if and only if there exists a $c > 0$ such that for every $x \in H$,

$$\|T_\lambda(x)\| \geq c\|x\|. \qquad (T_\lambda = T - \lambda I)$$

Proof. ($\Longrightarrow$) Our goal is to show that there exists a $c > 0$ such that for every $x \in H$, we have

$$\|T_\lambda x\| \geq c\|x\| \tag{2}$$

where $T_\lambda = T - \lambda I$. Since $\lambda \in \rho(T)$, it follows that $R_\lambda = T_\lambda^{-1}$ exists and is bounded. That is, $\|R_\lambda\| = k$, where $k > 0$ since $R_\lambda \neq 0$. Now, we have $I = R_\lambda T_\lambda$ such that every $x \in H$, we have

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

which gives us the following inequality

$$\|T_\lambda x\| \geq c|x|$$

where $c = \frac{1}{k}$.

($\Longleftarrow$) Our goal is to show that $\lambda \in \rho(T)$. We need to show that λ satisfies the following three properties:

(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.

(1) Clearly, T is onto by definition. Our goal is to show that for any $x_1, x_2 \in H$ such that $T_\lambda x_1 = T_\lambda x_2$, we have $x_1 = x_2$. Since T is linear and our assumption, it follows that

$$0 = \|T_\lambda x_1 - T_\lambda x_2\| = \|T_\lambda(x_1 - x_2)\| \geq c\|x_1 - x_2\|$$

for some $c > 0$. Since $\|x_1 - x_2\| \geq 0$, it follows from the above inequality that $x_1 = x_2$. Hence, T is injective and thus T is bijective.

(2) Our goal is to show that $T_\lambda(H)$ is dense in H; that is, $\overline{T_\lambda(H)} = H$. It suffices to show via the Projection Theorem in Chapter 3 of [1] Kreyszig that $x_0 \perp \overline{T_\lambda(H)}$ implies $x_0 = 0$. Let $x_0 \perp \overline{T_\lambda(H)}$. Then we have $x_0 \perp \overline{T_\lambda(H)}$. Hence, for all $x \in H$, we have

$$\begin{aligned} 0 = \langle T_\lambda x, x_0 \rangle &= \langle (T - \lambda I)x, x_0 \rangle \\ &= \langle Tx, x_0 \rangle - \lambda \langle x, x_0 \rangle. \end{aligned}$$

Since T is self-adjoint, it follows that

$$\langle x, Tx_0 \rangle = \langle Tx, x_0 \rangle = \langle x, \overline{\lambda} x_0 \rangle.$$

Hence, we have $Tx_0 = \overline{\lambda} x_0$ by 3.8-2. Note that we would need to have $x_0 = 0$ because otherwise if $x_0 \neq 0$, it would mean that $\overline{\lambda}$ is an eigenvalue of T so that $\overline{\lambda} = \lambda$ by (9.1-1) and $Tx_0 - \lambda x_0 = T_\lambda x_0 = 0$, and (2) would imply

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

since $c > 0$ which is absurd. Hence, we can see that $\overline{T_\lambda(H)^\perp} = \{0\}$ where x_0 is an arbitrary vector orthogonal to $T_\lambda(H)$. Hence, it follows that $\overline{T_\lambda(H)} = H$ by the Projection Theorem and so we can conclude that $T_\lambda(H)$ is dense in H.

(3) Finally, we will show that $T_\lambda(H)$ is closed in H. In what follows, we will show $\overline{T_\lambda(H)} = T_\lambda(H)$. Clearly, we can see that $T_\lambda(H) \subseteq \overline{T_\lambda(H)}$. So, it suffices to show that $\overline{T_\lambda(H)} \subseteq T_\lambda(H)$. Let $y \in \overline{T_\lambda(H)}$.

Then there is a sequence (y_n) in $T_\lambda(H)$j such that $y_n \to y$. Note that $y_n \in T_\lambda(H)$ and so $y_n T_\lambda x_n$ for some $x_n \in H$. By (2), we obtain

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

Since $y_n \to y$, it follows that y_n is a Cauchy sequence. From the inequality above, we can see that (x_n) must also be Cauchy when we let $m, n \to \infty$. Since H is complete, $x_n \to x$ for some $x \in H$. Since T is continuous (because it is bounded), we have that

$$y_n = T_\lambda x_n \to T_\lambda x.$$

Since limits are unique, it follows that $y = T_\lambda x$ and so we have $y \in T_\lambda(H)$. Hence, $T_\lambda(H)$ must be closed. As a consequence, we have $T_\lambda(H) = H$ from (2).

This tells us that $R_\lambda = T_\lambda^{-1}$ is defined on all of H, and is bounded, which follows from the [1] Bounded Inverse Theorem (4.12-2). Thus, we see that $\lambda \in \rho(T)$. ■

This is particularly useful because in more general Hilbert spaces since we may have cases where operators may not have eigenvalues at all. The following theorem makes sure that our spectrum does indeed live on $\mathbb{R}$.

Theorem (Spectrum)**.** The spectrum $\sigma(T)$ of a bounded self-adjoint linear operator $T : H \to H$ on a complex Hilber space H is real.

Proof. Using the previous theorem, we will show that for every $\lambda = \alpha + i\beta \in \sigma(T)$ where $\alpha, \beta \in \mathbb{R}$ with $\beta \neq 0$ that $\lambda \in \rho(T)$. Since T is self-adjoint, it follows from 9.1-1 that $\sigma(T) \subseteq \mathbb{R}$. Hence, it suffices to show that there exists a $c > 0$ such that

$$\|T_\lambda x\| \geq c\|x\|. \tag{*}$$

For every $x \neq 0$ in H, we have

$$\langle T_\lambda x, x\rangle = \langle Tx, x\rangle - \lambda\langle x, x\rangle.$$

Since $\langle x, x\rangle$ and $\langle Tx, x\rangle$ are real and so

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

Note that $\overline{\lambda} = \alpha - i\beta$. Subtracting the two quantities above, we can see that

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

which imply that

$$-\langle T_\lambda x, x\rangle = \beta\|x\|^2.$$

Applying the Cauchy-Schwarz inequality, we can see that

$$|\beta|\|x\|^2 = |\Im\langle T_\lambda x, x\rangle| \leq |\langle T_\lambda x, x\rangle| = |\langle T_\lambda x, x\rangle| \leq \|T_\lambda x\|\|x\|.$$

Since $\|x\| \neq 0$, we see that $|\beta|\|x\| \leq \|T_\lambda x\|$. If $\beta \neq 0$, then $\lambda \in \rho(T)$ by 9.1-2. Hence, for $\lambda \in \sigma(T)$ we see that $\beta = 0$, and so λ is real.

■

The next theorem takes things a step further and tells us that not only does the spectrum of T live on the real line, it also (assuming that T is bounded) lives in a compact interval.

Theorem (Spectrum Kreyszig (9.2-1))**.** The spectrum $\sigma(T)$ of a bounded self-adjoint linear operator $T : H \to H$ on a complex Hilbert space H lies in the closed interval $[m, M]$ on the real axis, where

$$m = \inf_{\|x\|=1}\langle T(x), x\rangle \quad \text{and} \quad M = \sup_{\|x\|=1}\langle T(x), x\rangle. \tag{1}$$

Proof. By (9.1-3), we see that $\sigma(T)$ lies on the real axis. We will show that for any real $\lambda = M + c$ where M is defined above lies in the resolvent set $\rho(T)$. For every $x \neq 0$, define $v = \|x\|^{-1}x$ and so $x = \|x\|v$. As a consequence, we have

$$\langle Tx, x\rangle = \|x\|^2 \langle Tv, v\rangle \leq \|x\|^2 \sup_{\|\tilde{v}\|=1} \langle T\tilde{v}, \tilde{v}\rangle = \langle x, x\rangle M.$$

Hence, we see that

$$-\langle Tx, x\rangle \geq -\langle x, x\rangle M$$

and so by the Schwarz Inequality we obtain

$$\begin{aligned} \|T_\lambda x\| \|x\| \geq -\langle T_\lambda x, x\rangle &= -\langle Tx, x\rangle + \lambda\langle x, x\rangle \\ &\geq (-M + \lambda)\langle x, x\rangle \\ &= c\|x\|^2 \end{aligned}$$

where $c = \lambda - M > 0$ by assumption. Since $x \neq 0$, it follows by division of $\|x\|$ on both sides of the above inequality yields

$$\|T_\lambda x\| \geq c\|x\|.$$

By (9.1-2), $\lambda \in \rho(T)$. If we assume $\lambda < m$, then a similar argument will lead to the same result. ■

The following theorem takes this notion a step further and says that $\|T\|$ is determined by the biggest possible inner product $\langle Tx, y\rangle$ for $x, y \in H$ by taking the absolute value of m and M.

Theorem (Norm Kreyszig (9.2-2)). For any bounded self-adjoint linear operator T on a complex Hilbert space H we have

$$\|T\| = \max(|m|, |M|) = \sup_{\|x\|=1} |\langle T(x), x\rangle|.$$

Proof. Our goal is to show that for any bounded self-adjoint linear operator T, we have

$$\|T\| = \max(|m|, |M|) = \sup_{\|x\|=1} |\langle Tx, x\rangle|.$$

Denote $K = \sup_{\|x\|=1} |\langle Tx, x\rangle|$. It suffices to show that $\|T\| \leq K$ and $K \leq \|T\|$. Using the Schwarz Inequality, it follows that

$$K \leq \sup_{\|x\|=1} \|Tx\|\|x\| = \|T\|.$$

Hence, the second inequality is satisfied. Now, we show $\|T\| \leq K$. Observe that if $Tz = 0$ for all z such that $\|z\| = 1$, then $T = 0$ and we are done. Suppose, otherwise that for any z such that $\|z\| = 1$, we have $Tz \neq 0$. Set $v = \|Tz\|^{1/2}z$ and $w = \|Tz\|^{-1/2}Tz$. Then we have

$$\|v\|^2 = \|w\|^2 = \|Tz\|.$$

Furthermore, set

$$y_1 = v + w \quad \text{and} \quad y_2 = v - w.$$

By a straight-forward calculation and the fact that T is self-adjoint, we obtain the following

$$\begin{aligned} \langle Ty_1, y_1\rangle - \langle Ty_2, y_2\rangle &= 2(\langle Tv, w\rangle + \langle Tw, v\rangle) \\ &= 2(\langle Tz, Tz\rangle + \langle T^2 z, z\rangle) \\ &= 4\|Tz\|^2. \end{aligned}$$

Now, for every $y \neq 0$ and $x = \|y\|^{-1}y$, we have $y = \|y\|x$ and so

$$|\langle Ty, y\rangle| = \|y\|^2 |\langle Tx, x\rangle| \leq \|y\|^2 \sup_{\|\tilde{x}\|=1} |\langle T\tilde{x}, \tilde{x}\rangle| = K\|y\|^2.$$

Using the triangle inequality, it follows that

$$\begin{aligned}
|\langle Ty_1, y_1\rangle - \langle Ty_2, y_2\rangle| &\leq |\langle Ty_1, y_1\rangle| + |\langle Ty_2, y_2\rangle| \\
&\leq K(\|y_1\|^2 + \|y_2\|^2) \\
&= 2K(\|v\|^2 + \|w\|^2) \\
&= 4K\|Tz\|.
\end{aligned}$$

Note that the left-hand side of the above inequality is equal to $4\|Tz\|^2$. As a consequence with $\|Tz\| \neq 0$, we have

$$4\|Tz\|^2 \leq 4K\|Tz\| \implies \|Tz\| \leq K.$$

Taking the supremum over all z of norm 1, it follows that $\|T\| \leq K$. ■

A remarkable property given by bounded self-adjoint linear operators is that m and M are contained in the spectrum of T.

Theorem (m and M are spectral values Kreyszig (9.2-3)). Let H and T be as in Theorem 9.2-1 and $H \neq \{0\}$. Then m and M are defined in (1) are spectral values of T.

Proof. Our goal is to show that m, M are contained in $\sigma(T)$. The proof to show that $m \in \sigma(T)$ is directly analogous. By the Spectral Mapping Theorem, the spectrum of $T + kI$, where $K \in \mathbb{R}$ is a constant, can be obtained from that of T via a translation, and so we have

$$M \in \sigma(T) \iff M + k \in \sigma(T).$$

We will show that $\lambda = M$ cannot belong to the resolvent set of T by 9.1-2. Indeed, without loss of generality assume $0 \leq m \leq M$. Using the previous theorem, we see that

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

Using the definition of a supremum, we can find a sequence (x_n) such that $\|x_n\| = 1$ and

$$\langle Tx_n, x_n\rangle = M - \delta_n$$

where $\delta_n \geq 0$ and that $\delta_n \to 0$. Then we have

$$\|Tx_n\| \leq \|T\|\|x_n\| = \|T\| = M.$$

Since T is self-adjoint, we have

$$\begin{aligned}
\|Tx_n - Mx_n\| &= \langle Tx_n - Mx_n, Tx_n - Mx_n\rangle \\
&= \|Tx_n\|^2 - 2M\langle Tx_n, x_n\rangle + M^2\|x_n\|^2 \\
&\leq M^2 - 2M(M - \delta_n) + M^2 \\
&= 2M\delta_n.
\end{aligned}$$

Since $\delta_n \to 0$, it follows from the Squeeze Theorem that

$$\|Tx_n - Mx_n\| \to 0.$$

Hence, there is no positive c such that

$$\|T_M x_n\| = \|Tx_n - Mx_n\| \geq c = c\|x_n\|.$$

Using 9.1-2 , we now have that $M \notin \rho(T)$ and so we must have $M \in \sigma(T)$ which is our desired result. ■

Another remarkable property of T, is that T's residual spectrum is empty. This establishes the fact that the eigenvalues we are considering are all real and, without a doubt, makes our integral presented in the spectral theorem to be well-defined.

Theorem (Residual Spectrum Kreyszig (9.2-4)). The residual spectrum $\sigma_r(T)$ of a bounded self-adjoint linear operator $T : H \to H$ on a complex Hilbert Space H is empty.

Proof. Suppose for sake of contradiction that $\sigma_r(T) \neq \emptyset$ leads to a contradiction. Let $\lambda \in \sigma_r(T)$. By definition, the inverse of T_λ exists, but its domain $D(T_\lambda^{-1})$ is not dense in H. By the Projection Theorem there exist a $y \neq 0$ in H such that y is orthogonal to $D(T_\lambda^{-1})$. However, $D(T_\lambda^{-1})$ is the range of T_λ. Thus,

$$\langle T_\lambda x, y \rangle = 0$$

for all $x \in H$. Since λ is real by 9.1-3 and T is self-adjoint, we obtain

$$\langle x, T_\lambda y \rangle = 0$$

for all x. Since $y \neq 0$, we see that λ is an eigenvalue of T. But this contradicts the assumption that $\lambda \in \sigma_r(T)$, making $\sigma_r(T) \neq \emptyset$ absurd. Hence, it must follow that $\sigma_r(T) = \emptyset$. ■

From the theorems presented in this section, it follows that when T is a bounded self-adjoint linear operator over H, we have

- All are eigenvalues are real
- The spectrum of T $\sigma(T)$ lives on the real line and contained in the compact interval $[m, M]$ where m and M are the infimum and supremum of $\langle Tx, y \rangle$ $(x, y \in H)$.
- The resolvent set $\rho(T)$ is empty, meaning that there is no chance that ou integral representation of T in terms of the spectrum will break.

4 Positive Operators

4.1 Positive Operators

In this section, we study the partial ordering of bounded self-adjoint linear operators. Our goal is to show that the family of projections serves as an integrator in the integral representation, which is a key step toward our desired spectral representation of T. To establish this, we need to demonstrate that the family of projections is monotonic.

First, we will clarify what is meant by the term partial ordering in this context.

Definition (Partial Order Kreyszig (9.3)). We define a **partial order** on the set of bounded self-adjoint linear operators defined over the complex Hilbert Space H by $T_1 \leq T_2$ if and only if $\langle T_1 x, x \rangle \leq \langle T_2 x, x \rangle$.

A key requirement to show monotonicity involves positive operators.

Definition (Positive Operators Kreyszig (9.3-1)). Let H be a complex Hilbert Space. We say that $T : H \to H$ is **positive**; that is, $T \geq 0$ if $\langle Tx, x \rangle \geq 0$.

A few results that follows immediately from the above definition are given below.

Lemma. Let H be a Complex Hilbert Space and let $T_1, T_2 : H \to H$ be two linear operators such that $T_1 \leq T_1$. We say that $T_2 - T_1$ is positive if $T_2 - T_1 \geq 0$.

Lemma. The sum of two positive operators is positive.

In the remainder of this section, we will prove some basic algebraic properties of positive operators which are dependent upon the essential fact that the product of two self-adjoint operators commute.

Theorem (Product Of Positive Operators Kreyszig (9.3))**.** If two bounded self-adjoint linear operators S and T on a Hilbert space H are positive and commute ($ST = TS$), then their product ST is positive.

Proof. Our goal is to show that $\langle STx, x\rangle \geq 0$ for all $x \in H$. Clearly, if $S = 0$, then the result holds. Suppose $S \neq 0$. We will show the following:

(a) If we consider $S_1 = \frac{1}{\|S\|}S$, $S_{n+1} = S_n - S_n^2$ for all $n \in \mathbb{N}$, then we will show via induction that

$$0 \leq S_n \leq I. \tag{*}$$

(b) Then we conclude that $\langle STx, x\rangle \geq 0$ for all $x \in H$.

(a) For $n = 1$, then the inequality in (*) holds. Indeed, using the Cauchy-Schwarz inequality, we have

$$\langle S_1x, x\rangle = \frac{1}{\|S\|}\langle Sx, x\rangle \leq \frac{1}{\|S\|}\|Sx\|\|x\| \leq \|x\|^2 = \langle Ix, x\rangle \Longrightarrow S_1 \leq I.$$

Suppose (*) holds for all $n = k$; that is,

$$0 \leq S_k \leq I \Longrightarrow 0 \leq I - S_k \leq I.$$

Since each S_k is self-adjoint for every $x \in H$ and $y = S_kx$, we obtain

$$\begin{aligned}\langle S_k^2(I - S_k)x, x\rangle &= \langle (I - S_k)S_kx, S_kx\rangle \\ &= \langle (I - S_k)y, y\rangle \geq 0.\end{aligned}$$

By definition, we see that

$$S_k^2(I - S_k) \geq 0 \tag{1}$$

and similarly, we have

$$S_k(I - S_k)^2 \geq 0. \tag{2}$$

Adding (1) and (2) together, we have

$$0 \leq S_k^2(I - S_k) + S_k(I - S_k)^2 = S_k - S_k^2 = S_{k+1}.$$

Hence, we see that $0 \leq S_{k+1}$. And $S_{k+1} \leq I$ follows from $S_k^2 \geq 0$ and $I - S_k \geq 0$ by addition. Indeed, we see that

$$0 \leq I - S_k + S_k^2 = I - S_{k+1}$$

which completes the induction proof of (a).

(b) We now show that $\langle STx, x\rangle \geq 0$ for all $x \in H$. Because $S_1 = \frac{1}{\|S\|}S$, $S_{n+1} = S_n - S_n^2$ for all $n \in \mathbb{N}$, we obtain

$$\begin{aligned}S_1 &= S_1^2 + S_2 \\ &= S_1^2 + S_2^2 + S_3 \\ &= S_1^2 + S_2^2 + \cdots + S_n^2 + S_{n+1}.\end{aligned}$$

Since $S_{n+1} \geq 0$, we have

$$S_1^2 + \cdots + S_n^2 = S_1 - S_{n+1} \leq S_1.$$

Since n is arbitrary, the infinite series

$$\|S_1x\|^2 + \|S_2x\|^2 + \cdots$$

converges. Thus, $\|S_nx\| \to 0$ and $S_nx \to 0$. By (5), we have

$$\Big(\sum_{j=1}^{n} S_j^2\Big)x = (S_1 - S_{n+1})x \to S_1x \qquad (n \to \infty)$$

By positivity and self-adjointness of S_j, we obtain

$$\sum_{j=1}^{n} \|S_j x\|^2 = \sum_{j=1}^{n} \langle S_j x, S_j x\rangle = \sum_{j=1}^{n} \langle S_j^2 x, x\rangle \leq \langle S_1 x, x\rangle. \tag{†}$$

Note that our S_j's commute with T, they are sums and products of $S_1 = \|S\|^{-1} S$, and S and T commute. Thus, with $S = \|S\| S_1$, (†), $T \geq 0$, and the continuity of the inner product, we find that for every $x \in H$ and $y_j = S_j x$, we have

$$\begin{aligned} \langle STx, x\rangle &= \|S\| \langle TSx, x\rangle \\ &= \|S\| \lim_{n\to\infty} \sum_{j=1}^{n} \langle TS_j^2 x, x\rangle \\ &= \|S\| \lim_{n\to\infty} \sum_{j=1}^{n} \langle Ty_j, y_j\rangle \geq 0; \end{aligned}$$

that is, $\langle STx, x\rangle \geq 0$. ■

Just as the validity of algebraic operations with real numbers is foundational in real analysis, the result above ensures that similar operations can be meaningfully performed with bounded self-adjoint linear operators. This is essential for the integral representation in the spectral theorem to be well-defined. Without the assumptions of commutativity and boundedness, these operations can fail, and the integrity of the integral representation may quickly break down.

Theorem (Monotone Sequence Kreyszig (9.3-3))**.** Let (T_n) be a sequence of bounded self-adjoint linear operators on a complex Hilbert space H such that

$$T_1 \leq T_2 \leq \cdots \leq T_n \leq \cdots \leq K \tag{*}$$

where K is a bounded self-adjoint linear operator on H. Suppose that any T_j commutes with K and with every T_m. Then (T_m) is strongly operator convergent i.e $T_n x \to Tx$ for all $x \in H$ and the limit operator T is linear, bounded, self-adjoint and satisfies $T \leq K$.

Proof. We consider $S_n = K - T_n$ and prove:

(a) The sequence $(\langle S_n^2, x\rangle)$ converges for all $x \in H$.

(b) $T_n x \to Tx$ where T is linear and self-adjoint, and is bounded by the uniform boundedness theorem.

(a) Our goal is to show that S_n converges using the Monotone Convergence Theorem. From our definition of S_n, we can see that for all $n \in \mathbb{N}$, T_n is self-adjoint and K is also a self-adjoint operator. Hence, S_n is also a self-adjoint operator. Observe that

$$\begin{aligned} S_m^2 - S_n S_m &= (S_m - S_n) S_m \\ &= (T_n - T_m)(K - T_m). \end{aligned}$$

Let $m < n$. By property (*), it follows that $T_n - T_m$ and $K - T_m$ are both positive. Also,

$$\begin{aligned} S_n S_m - S_n^2 &= S_n (S_m - S_n) \\ &= (K - T_n)(T_n - T_m). \end{aligned}$$

Hence, we see that $T_n - T_m$ and $K - T_m$ commute and so their product must be positive. Therefore, $S_m^2 - S_n S_m \geq 0$ and $S_m^2 - S_n S_m \geq 0$ for $m < n$. As a consequence, $S_m^2 \geq S_n S_m$ and $S_n S_m \geq S_n^2$. Furthermore,

$$S_m^2 \geq S_n S_m \geq S_n^2.$$

Because each S_n is self-adjoint, we must have

$$\begin{aligned} \langle S_m^2 x, x\rangle \geq \langle S_n S_m x, x\rangle &\geq \langle S_n^2 x, x\rangle \\ &= \langle S_n x, S_n x\rangle \\ &= \|S_n x\|^2 \geq 0. \end{aligned}$$

This shows that $(\langle S_n^2 x, x\rangle)$ with fixed x is a monotone decreasing sequence of nonnegative numbers. Hence, the monotone convergence theorem implies that S_n converges.

(b) Our goal is to show that $(T_n x)$ converges. By assumption, each T_n commutes with each T_m and with K. Hence, each S_j commute. From (a), each S_j is self-adjoint. Also, the fact that $-2\langle S_m S_n x, x\rangle \leq -2\langle S_n^2 x, x\rangle$ implies that, for $m < n$, we have

$$\begin{aligned}
\|S_m x - S_n x\|^2 &= \langle (S_m - S_n)x, (S_m - S_n)x\rangle \\
&= \langle (S_m - S_n)^2 x, x\rangle \\
&= \langle S_m^2 x, x\rangle - 2\langle S_m S_n x, x\rangle + \langle S_n^2 x, x\rangle \\
&\leq \langle S_m^2 x, x\rangle - \langle S_n^2 x, x\rangle.
\end{aligned}$$

Since $(S_n^2 x)$ converges for all $x \in H$ from part (a), it follows that $(S_n x)$ is a Cauchy sequence. Since H is a complete normed space, $(S_n x)$ must converge. Note that the limit depends on $x \in H$ and so $T_n x \to Tx$ for every $x \in H$. We see that T is self-adjoint because each T_n is self-adjoint and the inner product is continuous. Because $(T_n x)$ is also convergent, it is also bounded for every $x \in H$. By the uniform boundedness theorem, it follows that T is also bounded. Finally, the convergence of $T_n x$ to Tx implies that T is also bounded. ■

The result above guarantees the monotonicity and countable additivity required for the spectral theorem's integral representation to be well-defined. Moreover, it provides a crucial ingredient for establishing the validity of the Riemann–Stieltjes integral formulation: namely, the monotonicity of bounded self-adjoint operators allows us to construct an appropriate integrator—an essential component in the spectral representation of such operators.

4.2 Square Roots of Positive Operators

Definition (Positive Square Root Kreyszig (9.4-1))**.** Let $T : H \to H$ be a positive bounded self-adjoint linear operator on a complex Hilbert space H. Then a bounded self-adjoint linear operator A is called a **square root** of T if

$$A^2 = T.$$

If, in addition, $A \geq 0$, then A is called a **positive square root** of T and is denoted by

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

That is, $T^{1/2}$ exists and unique.

The next theorem provides a method for decomposing operators, which is especially useful for constructing orthogonal projections from positive operators. It also equips us with additional tools for analyzing the convergence of operator sequences, particularly in establishing both norm and strong operator convergence—topics that will be explored in later sections of this paper. While the proof of the theorem is detailed (at least in the way Kreyszig has done it) and contains some insight into the current structure, it is too long for the scope of this paper and we will skip its proof.

Theorem (Positive Square Root Kreyszig (9.4-2))**.** Every positive bounded self-adjoint linear operator $T : H \to H$ on a complex Hilbert Space H has a positive square root A, which is unique. This operator A commutes with every bounded linear operator on H which commutes with T.

The fact that A commutes with every bounded linear operator on H which commutes with T is critically important in preserving the algebraic structure of operator relationships. In terms of the spectral theorem, it is even more important that our algebraic operations commute because we want to be able to determine if an approximation of an operator is stable.

Overall, the theorems above enables us to ensure that

- our approximations of operators are valid,
- algebraic operations are preserved, and
- our integral representation behave consistently with respect to the operator topology and commutation relations.

5 Projection Operators

5.1 Projection Operators

Proposition. A linear operator $P : H \to H$ is a projection on H is there is a closed subspace Y of H such that Y is the range of P and $Y^{\perp}$ is the null space of P and $P|_Y$ is the identity operator on Y.

The above implies that we can write any $x \in H$ in the following way

$$x = y + z \tag{1}$$

where $y \in Y$ and $z \in Y^{\perp}$. Furthermore, we can write H in terms of Y and $Y^{\perp}$ as a direct sum; that is,

$$H = Y \oplus Y^{\perp}.$$

We can also rewrite (1) in the following way

$$x = y + z = Px + (I - P)x.$$

Another way to characterize the projection on H is the presented in the following theorem below:

Theorem (Projection Kreyszig (9.5-1))**.** A bounded linear operator $P : H \to H$ on a Hilbert Space H is a projection if and only if P is self-adjoint and idempotent (That is, $P^2 = P$).

The above theorem is more or less an instrumental in proving properties that we desire and need in order for our spectral family to represent T properly. Roughly speaking, the theorem helps us, among many other tools in our disposable, to create increasing families of projections that are associated with (E_λ) (family of projections).

Theorem (Positivity, norm Kreyszig (9.5-2))**.** For any projection P on a Hilbert space H,

(i) $\langle Px, x\rangle = \|Px\|^2$

(ii) $P \geq 0$

(iii) $\|P\| \leq 1$ and $\|P\| = 1$ if $P(H) \neq \{0\}$.

In this theorem, we essentially have that

(i) tells us that for any projection P, it projects elements of the Hilbert space H onto some closed subspace of H, thereby preserving the geometry of H. This further tells us that the projection valued measure E_λ satisfy the following property

$$\langle E_\lambda x, x\rangle = \|E_\lambda x\|^2$$

which is a key step in defining our integrator for our integral representation.

(ii) the positivity for any projection P ensures that our integral representation converges meaningfully.

(iii) the spectral projections E_λ are uniformly bounded by 1.

Now, these two theorems gives us a valuable tool for establishing the following seemingly simple facts (but extremely valuable) needed for our integral representation.

Theorem (Product of Projections Kreyszig (9.5-3))**.** In connection with products (composites) of projections on a Hilbert spcae H, the following two statements hold:

(a) $P = P_1P_2$ is a projection on H if and only if the projections P_1 and P_2 commute, that is, $P_1P_2 = P_2P_1$. Then P projects H onto $Y = Y_1 \cap Y_2$, where $Y_j = P_j(H)$.

Theorem (Sum of Projections Kreyszig (9.5-4)). Let P_1 and P_2 be projections on a Hilbert space H. Then:

(a) The sum $P = P_1 + P_2$ is a projection on H if and only if $Y_1 = P_1(H)$ and $Y_2 = P_2(H)$ are orthogonal.

(b) If $P = P_1 + P_2$ is a projection, P projects H onto $Y = Y_1 \oplus Y_2$.

5.2 Properties of Projections

Theorem (Partial Order Kreyszig (9.6-1)). Let P_1 and P_2 be projections defined on a Hilbert space H. Denote by $Y_1 = P_1(H)$ and $Y_2 = P_2(H)$, the subspaces onto which H is projected by P_1 and P_2, and let $N(P_1)$ and $N(P_2)$ be the null spaces of these projections. Then the following conditions are equivalent.

(1) $P_2P_1 = P_1P_2 = P_1$

(2) $Y_1 \subseteq Y_2$

(3) $N(P_1) \supseteq N(P_2)$

(4) $\|P_1x\| \leq \|P_2x\|$ for all $x \in H$

(5) $P_1 \leq P_2$.

This theorem refines our understanding of partial ordering for positive operators by demonstrating not only the monotonicity of projections, but also an analogue of the nested interval property through the inclusion relationships among their corresponding null spaces. More specifically, for a sequence of projections

$$P_1 \leq P_2 \leq P_3 \leq \cdots$$

we have

$$N(P_1) \supseteq N(P_2) \supseteq N(P_3) \supseteq \cdots .$$

Roughly speaking, the nesting of null spaces can be interpreted as becoming progressively more refined as the sequence of projections P_j increases. In fact, there is a theorem that formalizes this idea. However, before we state it, we introduce an important algebraic fact that allows us to take the difference of projections—providing a way to measure how close two projections are to one another. This will ultimately lead us to a notion of the limit of a sequence of projections.

Theorem (Difference of Projections Kreyszig (9.6-2)). Let P_1 and P_2 be projections on a Hilbert space H. Then:

(a) The difference $P = P_2 - P_1$ is a projection on H if and only if $Y_1 \subseteq Y_2$, where $Y_j = P_j(H)$.

(b) If $P = P_2 - P_1$, then P projects H onto Y, where Y is the orthogonal complement of Y_1 in Y_2.

With our observation made earlier, we now see, with the following theorem, that our intuition is correct.

Theorem (Monotone Increasing Sequence Kreyszig (9.6-3)). Let (P_n) be a monotone increasing sequence of projections P_n defined on a Hilbert space H. Then:

(a) (P_n) is strongly operator convergent, say, $P_nx \to Px$ for every $x \in H$, and the limit operator P is a projection defined on H.

(b) P projects H onto

$$P(H) = \overline{\bigcup_{n=1}^{\infty} P_n(H)}.$$

(c) P has the null space

$$N(P) = \bigcap_{n=1}^{\infty} N(P_n).$$

Proof. (a) Let $m < n$. By assumption, we can see that $P_m \leq P_n$ so that we have $P_m(H) \subseteq P_n(H)$ by 9.6-1 and $P_n - P_m$ is a projection by 9.6-2. Hence, for every fixed $x \in H$, we get via 9.5-2 that

$$\begin{aligned}\|P_n x - P_m x\|^2 &= \|(P_n - P_m)x\|^2 \\ &= \langle (P_n - P_m)x, x\rangle \\ &= \langle P_n x, x\rangle - \langle P_m x, x\rangle \\ &= \|P_n x\|^2 - \|P_m x\|^2.\end{aligned}$$

Using 9.5-2, we get $\|P_n\| \leq 1$ and so $\|P_n x\| \leq \|x\|$ for every $n \in \mathbb{N}$. This tells us that $\|P_n x\|$ is a bounded sequence of numbers. Also, $(\|P_n x\|)$ is a monotone sequence by 9.6-1. By the Monotone Convergence Theorem, it follows that $\|P_n x\|$ is a convergent sequence and hence $(\|P_n x\|)$ is a Cauchy sequence. As a consequence, (P_n) is also a Cauchy sequence in H and so using the completeness of H allows us to now say that $P_n x \to Px$ for some Px in H. It is immediate that P must be a projection on H since is linear, self-adjoint and idempotent by 9.5-3

(b) Our goal is to show that

$$P(H) = \overline{\bigcup_{n=1}^{\infty} P_n(H)}.$$

Let $m < n$. Then $P_m \leq P_n$, that is, $P_n - P_m \geq 0$ and so $\langle (P_n - P_m)x, x\rangle \geq 0$ by definition. Letting $n \to \infty$, we obtain $\langle (P - P_m)x, x\rangle \geq 0$ using the continuity of the inner product 3.2-2; that is, we have $P_m \leq P$ and so 9.6-1 gives us $P_m(H) \subseteq P(H)$ for every m. Thus, we have

$$\bigcup P_m(H) \subseteq P(H).$$

Furthermore, for every m and for every $x \in H$, we have

$$P_m x \in P_m(H) \subseteq \bigcup P_m(H).$$

Since $P_m x \to Px$, we see from 1.4-6(a) that $Px \in \overline{\bigcup P_m(H)}$. Hence, we have

$$\bigcup P_m(H) \subseteq P(H) \subseteq \overline{\bigcup P_m(H)}.$$

From 3.3-5 , it follows that $P(h) = N(I - P)$ so that $P(H)$ is closed by $2.7-10(b)$. This implies that from the above containments that (since $P(H)$ is closed)

$$\overline{\bigcup P_m(H)} \subseteq P(H)$$

Together, we see that

$$P(H) = \overline{\bigcup P_m(H)}.$$

(c) We determine $N(P)$. Using $3.3-5$, we see that

$$N(P) = P(H)^{\perp} \subseteq P_n(H)^{\perp}$$

for every $n \in \mathbb{N}$ due to part (b). Hence, we have

$$N(P) \subseteq \bigcap P_n(H)^{\perp} = \bigcap N(P_n).$$

On the other hand, if $x \in \bigcap N(P_n)$, then $x \in N(P_n)$ for every n so that $P_n x = 0$ and so $P_n x \to Px$ implies that $Px = 0$; that is, $x \in N(P)$. Thus, we see that

$$\bigcap N(P_n) \subseteq N(P).$$

Together, we obtain our result that

$$N(P) = \bigcap N(P_n).$$

■

This theorem shows that even without compactness, monotone increasing sequences of orthogonal projections converge strongly to a projection whose range captures the union of the original projection ranges and whose null space is the intersection of the corresponding null spaces. This elegant behavior reflects the well-structured geometry of Hilbert spaces and plays a foundational role in the construction of spectral projections used in the spectral theorem.

In the next section, we will make use of these tools by defining a spectral family along with the essential properties needed to establish the integral representation of the operator T.

6 Spectral Families

6.1 Properties of Spectral Family

Our goal for this section is to prove that our notion of a spectral family $\mathcal{G}$ introduced in the introduction our paper does, indeed, contain all the properties that we desire for an integral representation of T.

To define $\mathcal{G}$, we will use the following operator

$$T_\lambda = T - \lambda I$$

and the positive square root of T_λ^2, which we denote by B_λ; that is,

$$B_\lambda = (T_\lambda^2)^{1/2}$$

as well as

$$T_\lambda^+ = \frac{1}{2}(B_\lambda - T_\lambda)$$

which is denoted as the **positive part** of T_λ.

Definition (Spectral Family $\mathcal{G}$ of T Kreyszig (9.8)). We denote the **Spectral Family $\mathcal{G}$ of T** by $\mathcal{G} = (E_\lambda)_{\lambda \in |R}$ where E_λ is the projection of H onto the null space $N(T_\lambda^+)$ of T_λ^+.

Our goal for the remaining of this section is to prove that indeed $\mathcal{G}$ contains all the properties outlined in definition 9.7-1. Consider the following operators

$$B = (T^2)^{1/2} \qquad \text{(Postive Square Root of } T^2\text{)}$$
$$T^+ = \frac{1}{2}(B + T) \qquad \text{(Positive part of } T\text{)}$$
$$T^- = \frac{1}{2}(B - T) \qquad \text{(Negative part of } T\text{)}$$

and the projection of H onto the null space T^+ which we denote by E i.e

$$E : H \to Y = N(T^+).$$

Subtracting and adding the positive and negative part of T, respectively, we see that

$$T = T^+ - T^-$$
$$B = T^+ + T^-.$$

A natural question to ask at this point is: why do we need to decompose an operator into its positive and negative parts in the first place? Since we've been leveraging the many useful properties of positive operators—such as the existence of square roots, monotonic sequences of projections, and strong operator convergence—it is advantageous to break a more complicated operator into simpler, more manageable components. This decomposition allows us to better understand the operator's structure and behavior. Without it, we would lose access to the essential tools that make the spectral theorem work.

The next lemma says that B, T^+, T^{-1}, and E all contain the necessary properties that we desire in order for our spectral integral representation to work.

Lemma (Operators Related to T Kreyszig (9.8-1))**.** The operators just defined have the following properties

(a) B, T^+, and T^- are bounded and self-adjoint.

(b) B, T^+, and T^- commute with every bounded linear operator that T commutes with; in particular,

$$BT = TB \quad T^+T = TT^+ \quad T^-T = TT^- \quad T^+T^- = T^-T^+.$$

(c) E commutes with every bounded self-adjoint linear operator that T commutes with; in particular,

$$ET = TE \quad EB = BE.$$

(d) Furthermore,

$$\begin{aligned} T^+T^- &= 0 & T^-T^+ &= 0 \\ T^+E = ET^+ &= 0 & T^-E = ET^- &= T^- \\ TE &= -T^- & T(I-E) &= T^+ \\ T^+ &\geq 0 & T^{-1} &\geq 0. \end{aligned}$$

Proof. (a) By definition of B and T, it follows immediately that B, T^+ and T^- are bounded and self-adjoint.

(b) Suppose $TS = ST$. Then we have

$$T^2S = T(TS) = T(ST) = (TS)T = (ST)T = ST^2$$

Hence, $T^2S = ST^2$. From 9.4-2, it also follows that $BS = SB$ when applied T^2. Hence, we have

$$T^+S = \frac{1}{2}(BS + TS) = \frac{1}{2}(SB + ST) = ST^+.$$

Similarly, we have

$$T^-S = \frac{1}{2}(BS - TS) = \frac{1}{2}(SB - ST) = ST^-.$$

Hence, we have $T^+S = ST^+$ and $T^-S = ST^-$.

(c) For every $x \in H$, we have $y = Ex \in Y = N(T^+)$ since E is a projection of H onto Y. Also, we have $ESEx = SEx$ for all $x \in H$; that is, $ESE = SE$. Since projections are self-adjoint by 9.5-1, we see that S is also self-adjoint. Using (6g)-3..9, we obtain

$$ES = E^*S^* = (SE)^* = (ESE)^* = E^*S^*E^* = ESE = SE.$$

(d) We will prove the following

$$\begin{aligned} T^+T^- &= 0 & T^-T^+ &= 0 && \text{(i)} \\ T^+E = ET^+ &= 0 & T^-E = ET^- &= T^- && \text{(ii)} \\ TE &= -T^- & T(I-E) &= T^+ && \text{(iii)} \\ T^+ &\geq 0 & T^{-1} &\geq 0. && \text{(iv)} \end{aligned}$$

From $B = (T^2)^{1/2}$, we have $B^2 = T^2$. Note that $BT = TB$ by (a). Hence, we see that

$$T^+T = T^-T^+ = \frac{1}{2}(B-T)\frac{1}{2}(B+T) = \frac{1}{4}(B^2 + BT - TB - T^2) = 0.$$

This gives us (i).

Next, we prove (ii). By definition of E, we can see that $Ex \in N(T^+)$ and so $T^+Ex = 0$ for all $x \in H$. Because T^+ is self-adjoint, we have $ET^+x = T^+Ex = 0$ by (a) and (c); that is, we have

$$ET^+ = T^+E = 0.$$

Moreover, we have $T^+T^-x = 0$ by (i). Indeed, we have

$$T^+T^-x = T^-T^+x = T^-0 = 0.$$

Next, we prove (iii). Since $T = T^+ - T^-$ and (ii), we can see that

$$TE = (T^+ - T^-)E = T^+E - T^-E = 0 - T^-E = -T^-.$$

Using $T = T^+ + T^-$ again, we get

$$T(I - E) = T - TE = T + T^- = T^+.$$

Finally, we prove (iv). Using $B = T^+ + T^-$, part (ii), and Theorem 9.3-1, we have

$$T^+ = B - T^- = B - EB = (I - E)B \geq 0$$

where $I - E \geq 0$ by 9.5-2.

■

In addition to decomposing T into more manageable components, the lemma above establishes key properties such as the commutativity and orthogonality of these parts. It also clarifies the role of the projection operator E, which, when applied to T, isolates its positive and negative components for deeper insight into the behavior of T.

Roughly speaking, each projection E_λ corresponds to the portion of the Hilbert space associated with the spectrum of T up to the value λ. This structure allows us to approximate T by partitioning the spectrum and summing over the action of these projections on narrower spectral intervals. The orthogonality of projections on disjoint subintervals of the spectrum ensures that their contributions are non-overlapping, mirroring the disjoint behavior seen in the decomposition into positive and negative parts.

As our next step in developing the associated spectral family of T with λ as the parameter, we shall replace our notion of T, B, T^+, T^-, and E by $T_\lambda = T - \lambda I$, $B_\lambda = (T_\lambda^2)^{1/2}$, and the positive part and negative part of T given by

$$T_\lambda^+ = \frac{1}{2}(B_\lambda + T_\lambda)$$
$$T_\lambda^- = \frac{1}{2}(B_\lambda - T_\lambda)$$

and the projection

$$E_\lambda : H \to Y_\lambda = N(T_\lambda^+)$$

of H onto the null space $Y_\lambda = N(T_\lambda^+)$ of T_λ^+. Now, we will restate the lemma we just proved in the following way:

Lemma (Operators Related to T_λ Kreyszig (9.8-2)). The previous lemma remains true if we replace T, B, T^+, T^-, E by $T_\lambda, B_\lambda, T_\lambda^+, T_\lambda^-, E_\lambda$ where $T_\lambda = T - \lambda I$ and $\lambda \in \mathbb{R}$. Moreover, for any real $\kappa, \lambda, \mu, \nu, \tau$, the following operators all commute:

$$T_\kappa \quad B_\lambda \quad T_\mu^+ \quad T_\nu^- \quad E_\tau$$

Proof. The first statement follows from the previous lemma. To show the second statement, we observe that $IS = SI$ and

$$T_\lambda = T - \lambda I = T - \mu I + \mu I - \lambda I = T - \mu I + (\mu - \lambda)I = T_\mu + (\mu - \lambda)I.$$

Hence, we have

$$\begin{aligned} ST = TS &\implies ST_\mu = T_\mu S \\ &\implies ST_\lambda = T_\lambda S \\ &\implies SB_\lambda = B_\lambda S, SB_\mu = B_\mu S \end{aligned}$$

and etc. For $S = T_\kappa$ and so we have $T_\kappa B_\lambda = B_\lambda T_\kappa$ etc. ■

In fact, it can be proven from the lemma above that we can uniquely express T in terms of a spectral family $\mathcal{G} = (E_\lambda)$.

In addition to the properties introduced earlier in this section, we require our spectral family to satisfy the following additional conditions:

(i) If $\lambda < \mu$, then $E_\lambda \leq E_\mu$.

(ii) If $\lambda < m$, then $E_\lambda = 0$.

(iii) If $\lambda \geq M$, then $E_\lambda = I$

(iv) If $\mu \to \lambda + 0$, then $E_\mu x \to E_\lambda x$

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

Theorem (Spectal Family Associated with an Operator Kreyszig (9.8-3))**.** Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H. Furthermore, let E_λ $(\lambda \in \mathbb{R})$ be the projection of H onto the null space $Y_\lambda = N(T_\lambda^+)$ of the positive part T_λ^+ of $T_\lambda = T - \lambda I$. Then $G = (E_\lambda)_{\lambda \in \mathbb{R}}$ is a spectral family on the interval $[m, M] \subseteq \mathbb{R}$ where m and M are given by (1) in Section 9.2.

Proof. Our goal is to show that properties (i) through (iv) listed above.

To prove (i) through (iv), we will use part of 9.8-1 formulated for $T_\lambda, T_\mu, T_\lambda^+$ instead of T, T^+; namely,

(I) $T_\mu^+ T_\mu^- = 0$

(II) $T_\lambda E_\lambda = -T_\lambda^-$, $T_\lambda(I - E_\lambda) = T_\lambda^+$, $T_\mu E_\mu$

(III) $T_\lambda^+ \geq 0$, $T_\lambda^- \geq 0$, $T_\mu^+ \geq 0$ and $T_\mu^{k-} \geq 0$.

(i) Let $\lambda < \mu$. Then

$$T_\lambda = T_\lambda^+ - T_\lambda^- \leq T_\lambda^+$$

because $-T^- \leq 0$ by (III). Hence, we have

$$T_\lambda^+ - T_\mu \geq T_\lambda - T_\mu = (\mu - \lambda)I \geq 0.$$

Note that $T_\lambda^+ - T_\mu$ is self-adjoint and commutes with T_μ^+ by 9.8-2, and $T_\mu^+ \geq 0$ by (III). Using Theorem 9.3-1, we have

$$T_\mu^+(T_\lambda^+ - T_\mu) = T_\mu^+(T_\lambda^+ - T_\mu^+ + T_\mu^+) \geq 0.$$

Here, we have $T_\mu^+ T_\mu^- = 0$ by (I). Hence, we have $T_\mu^+ T_\lambda^+ \geq T_\mu^{+2}$; that is, for all $x \in H$,

$$\langle T_\mu^+ T_\lambda^+ x, x\rangle \geq \langle T_\mu^+ x, x\rangle = \|T_\mu^+ x\|^2 \geq 0.$$

This tell us that $T_\lambda^+ x = 0$ implies $T_\mu^+ x = 0$ for all $x \in H$. Thus, $N(T_\lambda^+) \subseteq N(T_\mu^+)$, and so we have $E_\lambda \leq E_\mu$ by 9.6-1. Here, we have $\lambda < \mu$.

(ii) Let $\lambda < m$. Suppose for contradiction that $E_\lambda \neq 0$. Then $E_\lambda z \neq 0$ for some $z \in H$. We set $x = E_\lambda z$. Then we have

$$E_\lambda x = E_\lambda(E_\lambda z) = E_\lambda^2 z = x.$$

Without loss of generality, suppose $\|x\| = 1$. Hence, we have

$$\begin{aligned}
\langle T_\lambda E_\lambda x, x\rangle &= \langle T_\lambda x, x\rangle \\
&= \langle (T - \lambda I)x, x\rangle \\
&= \langle Tx, x\rangle - \langle \lambda x, x\rangle \\
&= \langle Tx, x\rangle - \lambda\langle x, x\rangle \\
&= \langle Tx, x\rangle - \lambda && (\|x\| = 1) \\
&\geq \inf_{\|\tilde{x}\|=1} \langle T\tilde{x}, \tilde{x}\rangle - \lambda \\
&= m - \lambda > 0.
\end{aligned}$$

But note that this contradicts the fact that $T_\lambda E_\lambda = -T_\lambda^- \leq 0$ from (II) and (III).

(iii) Suppose for contradiction that $\lambda > M$ but $E_\lambda \neq I$, so that $I - E_\lambda \neq 0$. Then $(I - E_\lambda)x = x$ for some x of norm $\|x\| = 1$. Hence, we have

$$\begin{aligned}
\langle T_\lambda (I - E_\lambda)x, x\rangle &= \langle T_\lambda x, x\rangle \\
&= \langle T_\lambda x - T_\lambda E_\lambda x, x\rangle \\
&= \langle T_\lambda x, x\rangle - \langle T_\lambda (E_\lambda x), x\rangle \\
&= \langle T_\lambda x, x\rangle - \langle T_\lambda (0), x\rangle \\
&= \langle T_\lambda x, x\rangle \\
&= \langle (T - \lambda I)x, x\rangle \\
&= \langle Tx, x\rangle - \lambda\langle x, x\rangle \\
&= \langle Tx, x\rangle - \lambda && (\|x\| = 1) \\
&\leq \sup_{\|\tilde{x}\|=1} \langle T\tilde{x}, \tilde{x}\rangle - \lambda \\
&= M - \lambda < 0.
\end{aligned}$$

But this contradicts $T_\lambda(I - E_\lambda) = T_\lambda^+ \geq 0$ which is obtained from (II) and (III). Moreover, $E_M = I$ by the continuity from the right.

(iv) With an interval $\Delta = (\lambda, \mu]$, we associate the operator

$$\Delta E = E_\mu - E_\lambda.$$

Since $\lambda < \mu$, we have $E_\lambda \leq E_\mu$ by (i) and so $E_\lambda(H) \subseteq E_\mu(H)$ by 9.6-1. Moreover, ΔE is a projection by 9.6-2 and $\Delta E \geq 0$ by 9.5-2. Using 9.6-1,

$$E_\mu \Delta E = E_\mu^2 - E_\mu E_\lambda = E_\mu - E_\lambda = \Delta E \tag{*}$$

and

$$(I - E_\lambda)\Delta E = \Delta E - E_\lambda(E_\mu - E_\lambda) = \Delta E. \tag{**}$$

Since ΔE, T_μ^- and T_λ^+ are positive and commute by 9.8-2, we have that $T_\mu^- \Delta E$ and $T_\lambda^+ \Delta E$ are positive by 9.3-1. From (*) and (**), we can see that

$$\begin{aligned}
T_\mu \Delta E &= T_\mu E_\mu \Delta E = -T_\mu^- \Delta E \leq 0 \\
T_\lambda \Delta E &= T_\lambda (I - E_\lambda)\Delta E = T_\lambda^+ \Delta E \geq 0.
\end{aligned}$$

This tells us that $T\Delta E \leq \mu \Delta E$ and $T\Delta E \geq \lambda \Delta E$, respectively. Together, we have

$$\lambda \Delta E \leq T\Delta E \leq \mu \Delta E \tag{\dagger}$$

where $\Delta E = E_\mu - E_\lambda$.

Now, fix λ and let $\mu \to \lambda$ form the right in a monotone fashion. Then $\Delta Ex \to P(\lambda)$ by the analogue of Theorem 9.3-3 for a decreasing sequence. Here $P(\lambda)$ is bounded and self-adjoint. Now, since ΔE

is idempotent (because it is a projection), we can see that $P(\lambda)$ is also idempotent. Thus, $P(\lambda)$ is a projection. Also, $\lambda P(\lambda) = TP(\lambda)$ by (†); that is, $T_\lambda P(\lambda) = 0$. Thus, by using (II) and [9.8-2], we have

$$T_\lambda^+ P(\lambda) = T_\lambda (I - E_\lambda) P(\lambda) = (I - E_\lambda) T_\lambda P(\lambda) = 0.$$

Hence, we have $T_\lambda^+ P(\lambda)x = 0$ for all $x \in H$. This tell us that $P(\lambda)x \in N(T_\lambda^+)$. Also, E_λ is a projection from H onto $N(T_\lambda^+)$. As a consequence, we have

$$E_\lambda P(\lambda)x = P(\lambda)x,$$

that is, $E_\lambda P(\lambda) = P(\lambda)$. However, if we let $\mu \to \lambda + 0$ in (*), then

$$(I - E_\lambda)P(\lambda) = P(\lambda).$$

All together, we have $P(\lambda) = 0$. Since $\Delta Ex \to P(\lambda)x$, we see that (16) holds and so our Spectral family $\mathcal{G}$ is continuous from the right. ■

Remark. Note that (†) is a key inequality that we will use in our proof of the spectral representation of T in the next section.

7 Spectral Theorem

7.1 Spectral Representation

In the previous section, we showed that we can associate a given bounded self-adjoint linear operator T on a complex Hilbert space H with a spectral family $\mathcal{G} = (E_\lambda)$. In this section, we will prove that T can be represented in terms of a Riemann-Stieltjes integral.

Theorem (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:

(a) T has the spectral representation

$$T = \int_{m-0}^{M} \lambda \, dE_\lambda \tag{1}$$

where $\mathcal{G} = (E_\lambda)$ is the spectral family associated with T; the integral is to be understood in the sense of uniform convergence (convergence in the norm on $B(H, H)$), for all $x, y \in H$

$$\langle Tx, y\rangle = \int_{m-0}^{M} \lambda \, dw(\lambda) \tag{1*}$$

where $w(\lambda) = \langle E_\lambda x, y\rangle$ where the integral is an ordinary Riemann-Stieltjes integral.

(b) More generally, if p is a polynomial in λ with real coefficients,

$$p(\lambda) = \alpha_n \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_0,$$

then the operator $p(T)$ defined by

$$p(T) = \alpha_n T^n + \alpha_{n-1} T^{n-1} + \cdots + \alpha_0 I$$

has the spectral representation

$$p(T) = \int_{m-0}^{M} p(\lambda) \, dE_\lambda \tag{2}$$

and for all $x, y \in H$

$$\langle p(T)x, y\rangle = \int_{m-0}^{M} p(\lambda)\ dw(\lambda) \tag{2*}$$

where $w(\lambda) = \langle E_\lambda x, y\rangle$.

Remark. Before we embark on the proof, the $m-0$ written in the notation of the theorem above is there so that we can take into account the contribution made at $\lambda = m$ occurring at $E_m \neq 0$ (and $m \neq 0$). Thus, we can use any $a < m$ so that

$$\int_a^M \lambda\ dE_\lambda = \int_{m-0}^M \lambda dE_\lambda = mE_m + \int_m^M \lambda\ dE_\lambda.$$

Similarly with part (b), we have

$$\int_a^M p(\lambda)\ dE_\lambda = \int_{m-0}^M p(\lambda)\ dE_\lambda = p(m)E_m + \int_m^M p(\lambda)\ dE_\lambda.$$

Proof. (a) We choose a sequence of partitions (P_n) of $[a, b]$ where $a < m$ and $M < b$. Here, we denote our subintervals of $(a, b]$ as

$$\Delta_{nj} = (\lambda_{nj}, \mu_{nj}] \qquad (1 \leq j \leq n)$$

with length $\ell(\Delta_{nj}) = \mu_{nj} - \lambda_{nj}$. Note that $\mu_{nj} = \lambda_{n,j+1}$ for $1 \leq j \leq n-1$. We assume that this sequence of partitions to have the following property:

$$\eta(P_n) = \max_{j\in\mathbb{N}} \ell(\Delta_{nj}) \to 0 \ \text{ as } n \to \infty. \tag{2}$$

Also, in the last section, we have the following inequality with $\Delta = \Delta_{nj}$; that is,

$$\lambda_{nj}\Delta_{nj}E \leq T\Delta_{nj}E \leq \mu_{nj}\Delta_{nj}E. \tag{3}$$

Since $\mu_{nj} = \lambda_{n,j+1}$ for $1 \leq j \leq n-1$ from the fact that whenever $\lambda < \mu$, $E_\lambda \leq E_\mu$ and whenever $\lambda \geq M$, $E_\lambda = I$. Hence, we have

$$T\Big(\sum_{j=1}^n \Delta_{nj}E\Big) = T\Big(\sum_{j=1}^n (E_{\mu_{nj}} - E_{\lambda_{nj}})\Delta_{nj}E\Big) < \varepsilon I = \varepsilon. \tag{4}$$

From the above equality and from (1), we can see that for any $\varepsilon > 0$, there exists an N such that for every $n > N$ and every choice of $\lambda_{nj} \in \Delta_{nj}$, we have

$$\|T - \sum_{j=1}^n \hat{\delta}_{nj}\Delta_{nj}E\| < \varepsilon.$$

Indeed, we can see that for every choice of $\hat{\lambda_{nj}} \in \Delta_{nj}$, we can see that from (3), we have

$$0 \leq T - \sum_{j=1}^n \hat{\lambda}_{nj}\Delta_{nj}E \leq \sum_{j=1}^n (\mu_{nj} - \lambda_{nj})\Delta_{nj}E. \tag{5}$$

Then from the above, we have for every $n > N$

$$\begin{aligned}
\Big\|T - \sum_{j=1}^n \hat{\lambda_{nj}}\Delta_{nj}E\Big\| &\leq \Big\|\sum_{j=1}^n (\mu_{nj} - \lambda_{nj})\Delta_{nj}E\Big\| \\
&\leq \sum_{j=1}^n \|\mu_{nj} - \lambda_{nj}\|\Delta_{nj}E \\
&< \varepsilon \sum_{j=1}^n \Delta_{nj}E \\
&= \varepsilon I = \varepsilon
\end{aligned}$$

Since E_λ is constant for $\lambda < m$ and for $\lambda \geq M$, the particular of $a < m$ and $b > M$ is arbitrary. This proves (1), where (4) shows that the integral is to be understood in terms of uniform operator convergence. Subsequently, this gives us strong operator convergence since the inner product is continuous and the sum in (5) is a Riemann-Stieltjes Sum. Hence, (1) tells us that (1*) holds for every choice of x and y in H.

(b) Let $p(\lambda) = \lambda^r$ where $r \in \mathbb{N}$. For any $\kappa < \lambda \leq \mu < \nu$, we can see from (7) of section 9.7 that

$$\begin{aligned}\Delta_{nj}E \cdot \Delta_{nk} = (E_\lambda - E_\kappa)(E_\mu - E_\nu) &= E_\lambda E_\mu - E_\lambda E_\nu - E_\kappa E_\mu + E_\kappa E_\nu \\ &= E_\lambda - E_\lambda - E_\kappa + E_\kappa \\ &= 0\end{aligned}$$

for $j \neq k$. Also, since $\Delta_{nj}E$ is a projection, $(\Delta_{nj}E)^s = E(\Delta_{nj})$ for every $s \in \mathbb{N}$. Hence, we obtain

$$\Big[\sum_{j=1}^{n} \hat{\lambda_{nj}} \Delta_{nj} E\Big] = \sum_{j=1}^{n} \left(\hat{\lambda_{nj}}\right)^r \Delta_{nj} E. \tag{6}$$

Note that if the sum in (4) is close to T, then the expression in (6) implies that for any $\varepsilon > 0$, there exists an N such that for any $n > N$, we have

$$\Big\| T^r - \sum_{j=1}^{n} \left(\hat{\lambda_{nj}}\right)^r \Delta_{nj} E \Big\| < \varepsilon.$$

This proves (2) and (2*) for $p(\lambda) = \lambda^r$. Now, it follows immediately from this result that the two formulas presented in (2) and (2*) hold for any polynomial with real coefficients.

■

In the proof above, we see that the limit in (2) relies on a key inequality established in the previous section to demonstrate convergence. This result is made possible by the properties discussed in Section 4 and Section 5, specifically those concerning the positivity, monotonicity, and convergence of the orthogonal projections in the spectral family $\mathcal{G} = (E_\lambda)$.

We conclude this paper by exploring an application of the Spectral Theorem to convolution operators on the Hilbert space $L^2(\mathbb{R})$. The spectral theorem provides a powerful framework for understanding such operators by showing that they can be viewed as multiplication operators in the frequency domain. Specifically, through the Fourier transform, convolution operators—which are translation-invariant—are diagonalized, meaning that their action in the spatial domain to pointwise multiplication in the frequency domain. This perspective not only simplifies the analysis of these operators but also reveals their spectral properties in a clear and tractable way, making the Spectral Theorem an essential tool in both theoretical and applied contexts involving convolution.

7.2 Applications

Before exploring the significance of the Spectral Theorem in representing Convolution Operators, we must first define what a convolution is.

Definition (Convolution of two functions (Folland 7.1)). Given f and g defined on $\mathbb{R}$, the **convolution** is defined by the function $f * g$ where

$$(f * g)(x) = \int_{\mathbb{R}} f(x-y)g(y)\ dy,$$

provided that the integral exists.

In general, the function f is referred to as the (Chapter 2 of Stein-Shakarchi [4]) **kernel** of the convolution when it is the fixed function being convolved with another function. However, in the context of the Fourier transform, the term **kernel** refers instead to the complex exponential function $e^{-2\pi i x t}$, which defines the integral transform. That is, in a convolution of the form

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-t)\, g(t)\, dt,$$

the kernel is typically the function f that defines the operation.

On the other hand, the Fourier transform of a function g is given by

$$\hat{g}(x) = \int_{-\infty}^{\infty} g(t)\, e^{-2\pi i x t}\, dt,$$

where the function $e^{-2\pi i x t}$ acts as the kernel of the transform.

Some key properties that Convolution Operators possess, as we will see shortly, are that they are

- Compact (under specific kernel conditions)
- Self-adjoint via the Fourier Transform
- Diagonalizable

7.3 Convolution Operator on $L^2(\mathbb{R})$

Let $k \in L^1(\mathbb{R})$ be a **real-valued, symmetric function**; that is,

$$k(-x) = k(x)$$

for all $x \in \mathbb{R}$. Define the operator T on $L^2(\mathbb{R})$ by:

$$(Tf)(x) = (k * f)(x) = \int_{\mathbb{R}} k(x-y)f(y)\; dy.$$

It is not to hard to see that T is linear given the linearity properties that the Riemann-integral already possesses. In fact, the following theorem from [3] Folland, without a doubt, shows that T is linear and even further, the collection of convolutions forms a vector space over $\mathbb{R}$.

Theorem (Linearity Properties of the Convolution Folland (7.1)). Assuming that the integrals of f, g, and h exists, the convolution obeys the following properties:

(i) $f * (ag + bh) = a(f * g) + b(f * h)$ for any constants a and b;

(ii) $f * g = g * f$;

(iii) $f * (g * h) = (f * g) * h$

Hence, all that is left to show is that T is self-adjoint and bounded. Then we will show that the spectral representation of T can be expressed as a Fourier Transform.

To show that T is bounded, first note that $k \in L^1(\mathbb{R})$. Using Young's inequality, we have

$$\|Tf\|_{L^2} = \|k * f\|_{L^2} \leq \|k\|_{L^1} \cdot \|f\|_{L^2}.$$

Thus, $T \in B(L^2(\mathbb{R}))$ and $M = \|k\|_{L^1}$.

Next, we will show that T is self-adjoint. Using the symmetry of k, we know that for any $f, g \in L^2(\mathbb{R})$, we have

$$\langle Tf, g\rangle = \int_{\mathbb{R}} (k * f)(x)\overline{g(x)}\; dx.$$

Using Fubini's Theorem and symmetry of k, we obtain

$$\begin{aligned}\langle Tf, g\rangle &= \int_{\mathbb{R}}\int_{\mathbb{R}} k(x-y)f(y)\overline{g(x)}\; dydx \\ &= \int_{\mathbb{R}} f(y)\overline{(k * g)(y)dy} \\ &= \langle f, Tg\rangle.\end{aligned}$$

Hence, we have that $T^* = T$ and so we conclude that T is self-adjoint.

Now, for the spectral representation of T. Let

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\; dx$$

be the Fourier Transform of f. Using the Convolution Theorem, we can say that

$$\widehat{k * f}(\xi) = \hat{k}(\xi) \cdot \hat{f}(\xi).$$

In a frequency domain, we can see that T essentially acts a multiplication operator by the function $\hat{k}(\xi)$; that is,

$$\hat{Tf}(\xi) = \hat{k}(\xi)\hat{f}(\xi).$$

Since $k \in L^1(\mathbb{R})$, real-valued, and symmetric, we have that $\hat{k}$ is also real-valued and $\hat{k} \in C_0(\mathbb{R})$ using the Riemann-Lebesgue lemma.

Thus, T is unitarily equivalent to a multiplication operator $M_{\hat{k}}$ on $L^2(\mathbb{R})$, defined by

$$(M_{\hat{k}}f)(\xi) = \hat{k}(\xi)f(\xi).$$

In terms of our spectral theorem, we can represent T in the following way:

$$w(\lambda) = \int_{\{\xi:\hat{k}(\xi)\leq\lambda\}} \hat{f}(\xi)\overline{\hat{g}(\xi)}\ d\xi$$

where $f, g \in L^2(\mathbb{R})$ fixed and $\mu(\lambda)$ is the measure on $\mathbb{R}$ induced by the level sets of $\hat{k}$. The spectral projection in this example is the multiplication by $\chi\{\xi : \hat{k}(\xi) \leq \lambda\}$ in the corresponding Fourier Space. Roughly speaking, $w(\lambda)$ is monotone and defines a scalar measure derived from the projection-valued measure E_λ under the spectral theorem. Thus, we obtain the Riemann-Stieltjes integral:

$$\langle Tf, g\rangle = \int_{m-0}^{M} \lambda\ dw(\lambda)$$

where $m = \inf_{\xi\in[0,1)} \hat{k}(\xi)$, $M = \sup_{\xi\in[0,1)} \hat{k}(\xi)$.

In the next section, we will see an example of how the convolution operator gets represented in terms of its spectrum with respect to the Gaussian Kernel.

Example (Gaussian Function). Define the Gaussian kernel $k : \mathbb{R} \to \mathbb{C}(\text{or}\mathbb{R})$ by $k(x) = e^{-x^2} \in L(\mathbb{R}) \cap L^2(\mathbb{R})$ with the linear operator T defined as before for all $f \in L^2(\mathbb{R})$.

Taking the fourier transform of k. The (unitary) Fourier transform $\mathcal{F}$ on $L^2(\mathbb{R})$ is defined as

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i\xi x}\ dx.$$

Computing this integral, we get

$$\mathcal{F}[e^{-x^2}](\xi) = \sqrt{\pi}e^{-\pi^2\xi^2}.$$

Normalizing this kernel, we obtain

$$k(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \text{ such that } \hat{k}(x) = e^{-2\pi^2\xi^2}.$$

Note that the $k(x)$ above is real-valued and even. With respect to this kernel $T = k * \cdot$ is self-adjoint and in the Fourier space

$$\hat{Tf}(\xi) = \hat{k}(\xi) \cdot \hat{f}(\xi).$$

This tells us that T is unitarily equivalent to the multiplication operator:

$$M_{\hat{k}}f(\xi) = \hat{k}(\xi)f(\xi) = e^{-2\pi^2\xi^2}f(\xi).$$

Using the Spectral Theorem, we can write

$$T = \int_0^1 \lambda\ dE_\lambda$$

since the spectrum of T is the **essential range** of $\hat{k}(\xi) = e^{-2\pi^2\xi^2}$, which is:

$$R(\hat{k}) = (0, 1]$$

where $\xi \in \mathbb{R}$. Also, the scalar-valued spectral measure:

$$w(\lambda) = \langle E_\lambda f, g \rangle,$$

which, in the Fourier domain, becomes: Then:

$$\langle Tf, g \rangle = \int_0^1 \lambda \, dw(\lambda).$$

where

$$w(\lambda) = \int_{\{\xi : \hat{k}(\xi) \leq \lambda\}} \hat{f}(\xi)\overline{\hat{g}(\xi)} \, d\xi.$$

8 Conclusion

In this paper, we have carefully developed the foundational theory leading to the spectral representation of bounded self-adjoint linear operators on Hilbert spaces. Beginning with an examination of self-adjoint operators in finite-dimensional settings, we gradually introduced the necessary mathematical structures—including positive operators, square roots of operators, and orthogonal projections—to establish a framework conducive to generalization.

We then introduced the notion of a spectral family, whose properties allowed us to formulate a rigorous integral representation of our operator T using the Riemann-Stieltjes integral. This spectral representation serves as a powerful analytical tool, capturing the action of T through its interaction with a continuum of projection operators parameterized by the real line.

Finally, we explored an application of the spectral theorem by connecting it to convolution operators within the context of Fourier analysis. This demonstrated not only the theoretical elegance but also the practical significance of the spectral theorem in analyzing and diagonalizing important classes of operators in mathematical physics and signal processing.

The progression from discrete to continuous spectral decomposition exemplifies the power and reach of functional analysis, and underscores the importance of spectral theory in modern mathematical analysis.

References

[1] Erwin Kreyszig (1978) *Introductory Functional Analysis with Applications* John-Wiley & Sons. Inc.

[2] Walter Rudin (1987) *Real and Complex Analysis*, Mcgraw-Hill 3rd ed.

[3] Gerald B. Folland (1992) *Fourier Analysis and its Applications*, Wadsworth & Brooks/Cole pp. 206-207

[4] Elias M. Stein & Rami Shakarchi (2007) *Fourier Analysis: An Introduction* Princeton University Press pp. 44-48