Math 241A Project Notes

Lance Remigio

May 7, 2025

Chapter 1

Preliminaries

Corollary (Continuity, null space)**.** 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)**.** If in an inner product space, $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n \rangle \to \langle x, y \rangle$.

Lemma (Null Space)**.** 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 (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)**.** 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,

$$ST = TS.$$

Chapter 2

Spectral Properties

2.1 Spectral Properties of Bounded Self-adjoint Linear Operators

First, we will recall two relevant definitions from our study of Hilbert Spaces from chapter 3.

> **Definition** (Hilbert-Adjoint Operator). Let $T : H \to H$ be a bounded linear operartor 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

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

In developing the theory for the Spectral Theorem, we first need to make sure that eigenvalues are real so that we are able to respect the geometry induced by inner product spaces.

> **Theorem** (Eigenvalues, eigenvectors). 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.

■

Theorem (Resolvent Set)**.** 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 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 Bounded Inverse Theorem (4.12-2). Thus, we see that $\lambda \in \rho(T)$. ■

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.

■

2.2 Further Spectral Properties

In the previous section, we saw that the spectrum $\sigma(T)$ of a bounded self-adjoint linear operator T is real and that it is a compact set by chapter 4.

Theorem (Spectrum)**.** 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. ∎

> **Theorem** (Norm)**.** 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

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

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

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

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

This theorem is important in the following ways:

- It tells us how the operator norm is governed by the spectrum of T.
- Even in cases where the eigenvalues of T do not exist, we have a direct way to estimate the spectrum and understand the behavior of T.
- It gives us a way to approximate infinite-dimensional operators.

Theorem (m and M are spectral values)**.** 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

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

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. ■

Theorem (Residual Spectrum)**.** 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 $\sigma_r(T) = \emptyset$. ■

Chapter 3

Positive Operators

3.1 Positive Operators

In this section, we develop the framework that enables us to do more analysis on our self-adjoint linear operators. To do this, we will create a partial order on our operators which will be an extension of the ordering defined on the set of real numbers in real analysis.

Definition (Partial Order)**.** 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 $\langle T_1 x, x\rangle \leq \langle T_2 x, x\rangle$.

As we will see at the end of this section, we will develop a framework on determining the convergence of a sequence of bounded self-adjoint operators. Something of which is essential in representing these operators in more general inner product spaces.

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

In what follows, we will mention two facts that follow immediately from the definitions above.

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.

Theorem (Product Of Positive Operators)**.** 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

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

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 $y = S_k x$, we obtain

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

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).

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

$$T_1 \leq T_2 \leq \cdots \leq T_n \leq \cdots \leq K$$

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. ■

■

3.2 Square Roots of a Positive Operator

Theorem (Positive Square Root)**.** 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.

Proof. ■

Theorem (Positive Square Root)**.** 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.

Proof. ■

Chapter 4

Projection Operators

4.1 Projection Operators

Definition (Projections)**.** The map $P : H \to Y$ defined by $x \mapsto y = Px$ is called the projection of H onto Y.

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)**.** 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, $P^2 = P$).

Proof. ($\Longrightarrow$) Suppose that $P_1P_2 = P_2P_1$. Then by Theorem 3.10-4, P is self-adjoint. Also, we see that P is idempotent because

$$P^2 = (P_1P_2)(P_1P_2) = P_1^2P_2^2 = P_1P_2 = P]$$

where P_1 and P_2 are projections on H. By 9.5-1 , P is a projection, and for every $x \in H$, we have

$$Px = P_1(P_2x) = P_2(P_1x).$$

Since P_1 is a projects H onto Y_1, we must have $P_1(P_2x) \in Y_1$. Similarly, we see that $P_2(P_1x \in Y_2)$. Hence, we see that $Px \in Y_1 \cap Y_2$. Note that this projection is onto since

$$Py = P_1P_2y = P_1y = y.$$

($\Longleftarrow$) If $P = P_1P_2$ is a projection defined on H, then P is self-adjoint by 9.5-1. Also, we see that P_1 and P_2 commute by Theorem 3.10-4. ∎

Theorem (Positivity, norm)**.** 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 $P(H) \neq \{0\}$.

Proof. ■

Theorem (Sum of Projections)**.** 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 $Y = Y_1 \oplus Y_2$.

(a) ($\Longrightarrow$) Suppose that $P = P_1 + P_2$ is a projection on H. Our goal is to show that Y_1 and Y_2 are orthogonal; that is, it suffices to show that $P_1P_2 = 0$ by Theorem 9.5-3(b). Observe that P is idempotent. Hence, we have

$$\begin{aligned} P_1 + P_2 &= (P_1 + P_2)^2 \\ &= P_1^2 + P_1P_2 + P_2P_1 + P_2^2 \\ &= P_1 + P_1P_2 + P_2P_1 + P_2. \end{aligned}$$

Hence, we are left with

$$P_1P_2 + P_2P_1 = 0.$$

Multiplying the quantity above by P_2 on the left, we get

$$P_2P_1P_2 + P_2^2P_1 = 0 \Longleftrightarrow P_2P_1P_2 + P_2P_1 = 0.$$

Multiplying the quantity above by P_2 on the right side, we obtain

$$P_2P_1P_2^2 + P_2P_1P_2 = 0 \Longrightarrow 2P_2P_1P_2 = 0.$$

Hence, it follows that

$$P_1P_2 = 0$$

which is our desired result. ($\Longleftarrow$) Suppose $Y_1 \perp Y_2$. To show that $P = P_1 + P_2$ is a projection on H, it suffices to show that $P^2 = P$ and P is self-adjoint. Starting with the first result, since P_1 and P_2 are both idempotent, we have that

$$\begin{aligned} P^2 &= (P_1 + P_2)^2 \\ &= P_1^2 + P_1P_2 + P_2P_1 + P_2^2 \\ &= P_1^2 + P_2^2 \\ &= P_1 + P_2 \\ &= P. \end{aligned}$$

Hence, we see that $P^2 = P$ and so P is idempotent. Also, since P_1 and P_2 are also self-adjoint, we can see immediately that $P = P_1 + P_2$ is also idempotent. Thus, we can conclude based on Theorem 9.5-3 that P is indeed a projection on H.

(b) Suppose $P = P_1 + P_2$ is a projection. Note that Y is the closed subspace onto which P projects to. By definition of P, it follows that for all $x \in H$, we have

$$\begin{aligned} Px &= (P_1 + P_2)x \\ &= P_1x + P_2x. \end{aligned}$$

Since P is an onto projection, we have

$$y = Px = P_1x + P_2x \tag{*}$$

where $P_1x \in Y_1$ and $P_2x \in Y_2$. We will show now that $Y = Y_1 \oplus Y_2$. Let $y \in Y$. From (*), it follows that $y \in Y_1 \oplus Y_2$ (clearly, P being a projection on Y means that $Y_1 \perp Y_2$ and so $Y_1 \cap Y_2 = \{0\}$). Now, let $v \in Y_1 \oplus Y_2$. Then it follows that

$$\begin{aligned}P(v) &= P_1 v + P_2 v \\ &= P_1(y_1) + P_2(y_2) \\ &= Y_1 + Y_2 \\ &= v.\end{aligned}$$

Hence, $v \in Y$ and so we have $Y_1 \oplus Y_2 \subseteq Y$. Thus, we conclude that $Y = Y_1 \oplus Y_2$.

4.2 Further Properties of Projections

Theorem (Partial Order). 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) $P_1 \leq P_2$.

Proof. (1) $\implies$ (4) From 9.5-2, it follows that $\|P_1\| \leq 1$. Using (1), we can see that for all $x \in H$

$$\|P_1x\| = \|P_1P_2x\| \leq \|P_1\|\|P_2x\| \leq \|P_2x\|.$$

Hence, $\|P_1x\| \leq \|P_2x\|$.

(4) $\implies$ (5) Since $\langle P_1x, x\rangle = \|P_1x\|^2$, it follows that for all $x \in H$

$$\langle P_1x, x\rangle = \|P_1x\|^2 \leq \|P_2x\|^2 = \langle P_2x, x\rangle.$$

Thus, we have $P_1 \leq P_2$ by definition.

(5) $\implies$ (3) Assume $P_1 \leq P_2$. Our goal is to show $N(P_2) \subseteq N(P_1)$. Let $x \in N(P_2)$. Then $P_2x = 0$ by definition. Using the fact that $\langle P_1x, x\rangle = \|P_1x\|^2$ and (5) from our theorem, it follows that

$$\|P_1x\|^2 = \langle P_1x, x\rangle \leq \langle P_2x, x\rangle = 0.$$

Thus, we can see that $P_1x = 0$ and so $x \in N(P_1)$ and $N(P_2) \subseteq N(P_1)$.

(3) $\implies$ (2) By 3.3-5, since $N(P_j)$ is the orthogonal complement of Y_j in H, it immediately follows that $Y_1 \subseteq Y_2$.

(2) $\implies$ (1) Assume $Y_1 \subseteq Y_2$. For every $x \in H$, we have $P_1x \in Y_1$. Thus, $P_1x \in Y_2$ by assumption and so we have

$$P_2(P_1x) = P_1x.$$

Thus, $P_2P_1 = P_1$. Note that since P_1 is a projection, P_1 must also be self-adjoint by 9.5-1. Using 3.10-4, we see immediately that

$$P_1 = P_2P_1 = P_1P_2.$$

■

Theorem (Difference of Projections). 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.

Proof. (a) ($\Longrightarrow$) Suppose $P = P_2 - P_1$ is a projection on H. Our goal is to show that $Y_1 \subseteq Y_2$. Our strategy is to prove that

$$P_2P_1 = P_1P_2 = P_1$$

so that we may apply 9.6-1 to obtain our result. By assumption, we see that $P^2 = P$ by 9.5-1 and so we obtain

$$P_2 - P_1 = (P_2 - P_1)^2 = P_2^2 - P_2P_1 - P_1P_2 + P_1^2.$$

On the right-hand side, since P_1 and P_2 are also projections, we have $P_2^2 = P_2$ and $P_1^2 = P_1$ by 9.5-1. And so the above equation enables us to write

$$P_1P_2 = P_2P_1 = 2P_1. \tag{I}$$

Multiplying P_2 on the left-hand side of (I) gives us

$$P_2P_1P_2 + P_2P_1 = 2P_2P_1.$$

Then we have

$$P_2P_1P_2 = P_2P_1. \tag{1}$$

Multiplying P_2 on the right-hand side of (I) gives us

$$P_1P_2 + P_2P_1P_2 = 2P_1P_2.$$

Then we have

$$P_2P_1P_2 = P_1P_2. \tag{2}$$

From (1) and (2), it follows that

$$P_2P_1 = P_1P_2 = P_1$$

which is our desired result.

($\Longleftarrow$) Our goal is to show that P is idempotent and P is self-adjoint by 9.5-1. If $Y_1 \subseteq Y_2$, we see from 9.6-1 that $P_2P_1 = P_1P_2 = P_1$ which further implies that

$$P_1P_2 + P_2P_1 = 2P_1$$

and consequently shows that P is idempotent. Indeed, we see that

$$\begin{aligned} P^2 = (P_2 - P_1)^2 &= P_2^2 - (P_2P_1 + P_1P_2) + P_1^2 \\ &= P_2 - (P_2P_1 + P_1P_2) + P_1 \\ &= P_2 - P_1 \\ &= P. \end{aligned}$$

Hence, $P^2 = P$. Since P_1 and P_2 are self-adjoint, it follows that $P = P_2 - P_1$ is also self-adjoint, and so P is a projection.

(b) Our goal is to show that $Y = V = Y_2 \cap Y_1^{\perp}$. Note that $Y = P(H)$ consists of all vectors of the form

$$y = Px = P_2x - P_1x$$

for all $x \in H$. Suppose $P = P_2 - P_1$ is a projection. Then from part (a), it follows that $Y_1 \subseteq Y_2$ and so we have $P_2P_1 = P_1$. Thus, we obtain from the above equation that

$$P_2y = P_2^2x - P_2P_1x = P_2x - P_1x = y.$$

This tells us that $y \in Y_2$. Similarly, we have

$$P_1y = P_1P_2x - P_1^2x = P_1x - P_1x = 0.$$

This tells us that $y \in N(P_1) = Y_1^{\perp}$ by 3.3 – 5. Together, $y \in V$ where $V = Y_2 \cap Y_1^{\perp}$. Since $y \in Y$ was arbitrary, we have $Y \subseteq V$

Now, we will show that $V \subseteq Y$. Observe that the projection of H onto $Y_1^{\perp}$ is $I - P_1$, and so for every $v \in V$, we can see that

$$v = (I - P_1)y_2.$$

Using $P_2P_1 = P_1$, we obtain from the above equation that, since $P_2y_2 = y_2$, we have

$$\begin{aligned} Pv &= (P_2 - P_1)(I - P_1)y_2 \\ &= (P_2 - P_2P_1 - P_1 + P_1^2)y_2 \\ &= (P_2 - P_2P_1)y_2 \\ &= P_2y_2 - (P_2P_1)y_2 \\ &= (I - P_1)y_2 \\ &= v. \end{aligned}$$

Hence, we have $v \in Y$ and so $V \subseteq Y$ since v is arbitrary. Together, we identify $Y = P(H)$ with $V = Y_2 \cap Y_1^{\perp}$.

■

Theorem (Monotone Increasing Sequence)**.** 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

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

Using 9.5-2, we get $\|P_n\| \leq 1$ and so $\|P_nx\| \leq \|x\|$ for every $n \in \mathbb{N}$. This tells us that $\|P_nx\|$ is a bounded sequence of numbers. Also, $(\|P_nx\|)$ is a monotone sequence by 9.6-1. By the Monotone Convergence Theorem, it follows that $\|P_nx\|$ is a convergent sequence and hence $(\|P_nx\|)$ 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 $P_nx \to Px$ 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

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

Since $P_mx \to Px$, 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

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

■

Chapter 5

Spectral Representation

5.1 Spectral Family

- Recall that our main goal of this chapter is to find a way to represent bounded self-adjoint linear operators on a Hilbert Space in terms of simpler operators which are projections of our Hilbert space.
- Such representations are called a spectral representations.
- We will do this by associating our linear operator T with a spectral family.
- In this section, we will motivate the use of these spectral families.

Recall that in the finite dimensional case that if we let $T : H \to H$ be a self-adjoint linear operator on the unitary space $H = C^n$. Then T is bounded and so we can choose a basis for H and represent T in terms of a Hermitian matrix which we can denote by T.

Suppose that the matrix that represents T consists of distinct eigenvalues which are real by 9.1-1. Furthermore, assume that we have an n number of these eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. This implies that T contains an orthonormal set of n eigenvectors $\{x_1, x_2, \ldots, x_n\}$ where each x_j corresponds to λ_j and represents the column vectors of the matrix representation of T. Since our orthonormal set produces a unique representation of every $x \in H$, we can write the following formulas

$$x = \sum_{j=1}^{n} \gamma_j x_j \tag{1}$$

$$\gamma_j = \langle x, x_j \rangle x^T \overline{x_j}.$$

Applying T to the above formula, we obtain that

$$Tx = \sum_{j=1}^{n} \lambda_j \gamma_j x_j \tag{2}$$

We find that even if T may be a complicated operator to deal with, it acts on each x_j in a relatively simple way.

If we define an operator $P_j : H \to H$ where $x \mapsto \gamma_j \gamma_j x_j$, then each P_j is a (orthonormal) projection of H onto the eigenspace of T corresponding to λ_j. Note that (1) can be written as

$$x = \sum_{j=1}^{n} P_j x \Longrightarrow I = \sum_{j=1}^{n} P_j \tag{3}$$

where I is the identity operator on H. Now, (2) can be written as

$$Tx = \sum_{j=1}^{n} \lambda_j P_j x \Longrightarrow T = \sum_{j=1}^{n} \lambda_j P_j. \tag{4}$$

This gives us the representation of T in terms of our projections P_j for $1 \le j \le n$.

Given how "nice" our representation are for the linear operator T, it is unfortunate that the same representation cannot be applied in infinite dimensional vector spaces. Instead of using projections $P_1, P_2, \dots, P_n$, we will sum such projections; that is, we will have, for any real λ, define

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

The family projections P_j for $j \in \mathbb{N}$ are associated with one-parameter. From the above definition, we can see that for any λ, the operator E_λ is the projection of H onto the subspace V_λ are spanned by the $x_j's$ such that $\lambda_j \le \lambda$.

Definition (Spectral Family or Decomposition of Unity). A real spectral family (or real *decomposition of unity*) is a one-parameter family $\mathscr{G} = (E_\lambda)_{\lambda \in \mathbb{R}}$ of projections E_λ defined on a Hilbert Space H (of any dimension) which depends on a real parameter λ and is such that

(i) $E_\lambda \le E_\mu \Longrightarrow E_\lambda E_\mu = E_\mu E_\lambda = E_\lambda$ for $\lambda < \mu$.

(ii) $\lim_{\lambda \to -\infty} E_\lambda x = 0$.

(iii) $\lim_{\lambda \to +\infty} E_\lambda x = x$.

(iv) $E_{\lambda+0} x = \lim_{\mu \to \lambda+0} E_\mu x = E_\lambda x$ for $x \in H$.

From the definition above, we can interpret a real spectral family as a mapping from $\mathbb{R}$ to $B(H,H)$ where $\lambda \in \mathbb{R}$ maps to $E_\lambda \in B(H,H)$ where $B(H,H)$ is the space of all bounded linear operators from H into H. Note that in the definition above, the implication in (i) is actually bidirectional; that is, the two statements are equivalent by 9.6-1. Also, we can see that $\mathscr{G}$ is called a **spectral family on an interval** on $[a,b]$ if the following conditions hold

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

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

The fact that $\mathscr{G}$ lies in a finite interval will prove to be advantageous when it comes to the spectrum of a bounded self-adjoint linear operator; that is, expressing such an operator in terms of Riemann-Stieltjes Integral. In the finite dimensional case, the integral representation reduces to a finite sum in terms of the spectral family. Indeed, we can still write the representation given in (3) in terms of (4)

Assume that the eigenvalues $\lambda_1, \dots, \lambda_n$ are distinct and that $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. Then we have

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

Conversely, we have

$$\begin{aligned} P_1 &= E_{\lambda_1} \\ P_j &= E_{\lambda_j} - E_{\lambda_{j-1}}. && (2 \le j \le n) \end{aligned}$$

Since E_λ does not change for λ in $[\lambda_{j-1}, \lambda_j)$, we can write

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

and so (4) becomes

$$x = \sum_{j=1}^n P_j x = \sum_{j=1}^n (E_{\lambda_j} - E_{\lambda_j - 0}) x$$

and subsequently (5) becomes

$$Tx = \sum_{j=1}^n \lambda_j P_j x = \sum_{j=1}^n \lambda_j (E_{\lambda_j} - E_{\lambda_j - 0}) x.$$

Dropping x, we have $\Delta E_\lambda = E_\lambda - E_{\lambda-0}$ and so

$$T = \sum_{j=1}^{n} \lambda_j \Delta E_{\lambda_j}. \tag{*}$$

As we can see, (*) is the *spectral representation* of the self-adjoint linear operator T we are after with eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ on the $n-$dimensional Hilbert Space H. This tells us that for any $x, y \in H$

$$\langle Tx, y\rangle = \sum_{j=1}^{n} \lambda_j \langle \Delta E_{\lambda_j} x, y\rangle.$$

We can write the above in terms of a Riemann-Stieltjes integral

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

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

5.2 Properties of Spectral Family

Our goal for this section is to, indeed, prove that our notion of a spectral family $\mathscr{G}$ introduced in the previous section does have all the properties that we desired for a representation of T that was characterized.

To define $\mathscr{G}$, 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 $\mathscr{G}$ of T)**.** We denote the **Spectral Family** $\mathscr{G}$ **of** T by $\mathscr{G} = (E_\lambda)_{\lambda \in |R}$ 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 $\mathscr{G}$ contains all the properties outlined in definition 9.7-1. Consider the following operators

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

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

$$\begin{aligned} T &= T^+ - T^- \\ B &= T^+ + T^-. \end{aligned}$$

Lemma (Operators Related to T). 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,

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

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.

■

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). 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

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

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 that from the lemma above, we can express T in terms of a spectral family $\mathscr{G} = (E_\lambda)$ that is unique.

Theorem (Spectal Family Associated with an Operator). 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

(i) $\lambda < \mu \Longrightarrow E_\lambda \leq E_\mu$

(ii) $\lambda < m \Longrightarrow E_\lambda = 0$

(iii) $\lambda \geq M \Longrightarrow E_\lambda = I$

(iv) $\lambda \geq M \Longrightarrow E_\lambda = I$

(v) $\mu \to \lambda + 0 \Longrightarrow E_\mu x \to E_\lambda x$.

To prove (i) through (v), we will use part of 9.8-1 formulated for T_λ, T_μ, T_λ^+ 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

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

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

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

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

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

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

Now, fix λ and let $\mu \to \lambda$ form the right in a monotone fashion. Then $\Delta E x \to P(\lambda)$ 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 $\Delta E x \to P(\lambda)x$, 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.

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

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

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

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

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

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

$$\lambda_{nj}\Delta_{nj}E \le T\Delta_{nj}E \le \mu_{nj}\Delta_{nj}E. \tag{4}$$

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

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

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

$$\|T - \sum_{j=1}^{n} \hat{\delta}_{nj}\Delta_{nj}E\| < \varepsilon. \tag{5}$$

Add more explanation of the above later. Indeed, we can see that for every choice of $\hat{\lambda_{nj}} \in \Delta_{nj}$, we can see that from (4), we have

$$0 \le T - \sum_{j=1}^{n} \hat{\lambda}_{nj}\Delta_{nj}E \le \sum_{j=1}^{n} (\mu_{nj} - \lambda_{nj})\Delta_{nj}E. \tag{6}$$

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

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

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 (5) 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

$$\Big[\sum_{j=1}^{n} \hat{\lambda_{nj}} \Delta_{nj} E\Big] = \sum_{j=1}^{n} \hat{\lambda_{nj}}^r \Delta_{nj} E. \tag{7}$$

Note that if the sum in (5) is close to T, then the expression in (7) implies that for any $\varepsilon > 0$, there exists an N such that for any $n > N$, we have

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

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.

■

Below are some properties of operators $p(T)$ as a consequence of the above theorem.

Theorem (Properties of $p(T)$). Let T be as in the previous theorem, and let p, p_1, and p_2 be polynomials with real coefficients. Then:

(a) $p(T)$ is self-adjoint.

(b) If $p(\lambda) = \alpha p_1(\lambda) + \beta p_2(\lambda)$, then $p(T) = \alpha p_1(T) + \beta p_2(T)$.

(c) If $p(\lambda) = p_1(\lambda) p_2(\lambda)$, then $p(T) = p_1(T) p_2(T)$.

(d) If $p(\lambda) \geq 0$ for all $\lambda \in [m, M]$, then $p(T) \geq 0$.

(e) If $p_1(\lambda) \leq p_2(\lambda)$ for all $\lambda \in [m, M]$, then $p_1(T) \leq p_2(T)$.

(f) $\|p(T)\| \leq \max_{\lambda \in J} |p(\lambda)|$, where $J = [m, M]$.

(g) If a bounded linear operator commutes with T, it also commutes with $p(T)$.