

806 - Functional Analysis
Fall 2017 - University of Delaware
Notes by Dr. Constantin Bacuta, Transcribed by Nicholas
Russell

Chapter 1

Banach Spaces

Let X be a vector space over \mathbb{K} , where $\mathbb{K} = \mathbb{R}$ or \mathbb{C} .

Definition 1. A function $d : X \times X \mapsto [0, \infty)$ is said to be a **metric** if

(D1) $d(x, y) \geq 0$, for all $x, y \in X$.

(D2) $d(x, y) = 0$ if and only if $x = y$.

(D3) $d(x, y) = d(y, x)$ for all $x, y \in X$.

(D4) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

Definition 2. A function $\|\cdot\| : X \rightarrow [0, \infty)$ is called a **norm** if

(N1) $\|x\| \geq 0$, for all $x \in X$, $\|x\| = 0$ if and only if $x = 0$.

(N2) $\|\lambda x\| = |\lambda| \|x\|$ for all $\lambda \in \mathbb{K}$, $x \in X$.

(N3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

We know that the norm induces some distance and thus, we call $(X, \|\cdot\|)$ a **normed space**. Defined $d(x, y) = \|x - y\|$ for all $x, y \in X$. We can easily check that this satisfies a distance on X . We also note the following properties:

1. $d(x + z, y + z) = d(x, y)$. This means that d is *invariant under translation*.
2. $d(\lambda x, \lambda y) = |\lambda| d(x, y)$, which means that d is *positively homogeneous*.
3. $B(x, r) = \{y \in X \mid d(x, y) < r\}$ is a convex set, which means that $x, y \in B$, when $\lambda x + (1 - \lambda)y \in B$ for any $\lambda \in [0, 1]$. We define that $B(x, r)$ to be the **open ball** centered at x with radius r , and we define $\overline{B}(x, r) = \{y \in X \mid d(x, y) \leq r\}$ to be the **closed ball** centered at x with radius r .

We note that $x \mapsto \|x\|$ is a continuous function since one can prove that $|\|x\| - \|y\|| \leq \|x - y\|$ and thus, we have continuity since $x_n \rightarrow x$ implies that $\|x_n\| \rightarrow \|x\|$.

Definition 3. A sequence $(x_n)_{n \geq 1} \in X \subseteq \mathbb{K}$, a normed space, **converges** to $x_0 \in X$ if

$$\lim_{n \rightarrow \infty} \|x_n - x_0\| = 0.$$

If $(y_n)_{n \geq 1} \in X$, then the series $\sum_{k=1}^{\infty} y_k$ is convergent if $x_n = \sum_{k=1}^n y_k$ converges to a vector $\mathbf{x}_0 \in X$, i.e.

$$\lim_{n \rightarrow \infty} \|(y_1 + y_2 + \cdots + y_n) - \mathbf{x}_0\| = 0.$$

In this case, $\mathbf{x}_0 = \sum_{n=1}^{\infty} y_n$ is the **sum of the series**.

Definition 4. A sequence $(x_n)_{n \geq 1} \in X$ is a **Cauchy sequence** if for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $m, n \geq N$, then $\|x_n - x_m\| < \epsilon$.

We remark that convergence implies Cauchy, but the opposite is not true in general.

Definition 5. A normed space $(X, \|\cdot\|)$ is **complete** (or a **Banach space**) if every Cauchy sequence is convergent (to an $x \in X$).

We now present a couple of examples:

1. On \mathbb{R}^n , $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, then

$$\|x\| = \|x\|_2 := \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Let $1 \leq p \leq \infty$ and define

$$\|x\|_p := (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$$

if $1 \leq q < \infty$ and

$$\|x\|_{\infty} = \max\{|x_1|, |x_2|, \dots, |x_n|\}.$$

We note that if we fix $x \in \mathbb{R}^n$, $\lim_{p \rightarrow \infty} \|x\|_p = \|x\|_{\infty}$. We can also show that $(\mathbb{R}^n, \|\cdot\|_p)$ is a Banach space.

2. For infinite sequences (ℓ^p spaces, $1 \leq p \leq \infty$), let $x = (x_1, x_2, \dots) = (x_n)_{n \geq 1}$, $x_n \in \mathbb{R}$ (or \mathbb{C}). Then, we can denote

$$\|x\|_p = \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{\frac{1}{p}}$$

and

$$\|x\|_{\infty} = \sup_{k \geq 1} \{|x_k|\}.$$

We will then call the set

$$\ell^p = \{x = (x_n)_{n \geq 1} : \|x\|_p < \infty\} = \left\{ x = (x_n)_{n \geq 1} : \sum_{k=1}^{\infty} |x_k|^p < \infty \right\}$$

Lemma 1. If $p > 1$, $q > 1$, and $\frac{1}{p} + \frac{1}{q} = 1$, then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

(For proof, see P.234 of AB)

Theorem 1. (Holder's Inequality) If $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\sum_{n=1}^{\infty} |x_n \cdot y_n| \leq \|x\|_p \|y\|_q \quad (1.1)$$

for all $x \in \ell^p$, $y \in \ell^q$, where equality holds if and only if there exist α and β independent of n such that $\alpha|x_n|^p = \beta|y_n|^q$, where $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$.

Proof. Assume that $x \neq 0$ and $y \neq 0$, and denote $X = \frac{x}{\|x\|_p}$ and $Y = \frac{y}{\|y\|_q}$. Note then that $\|X\|_p = \|Y\|_q = 1$. Dividing (1) by $\|x\|_p \|y\|_q$, we obtain

$$\sum_{n=1}^{\infty} \left| \frac{x_n}{\|x\|_p} \frac{y_n}{\|y\|_q} \right| \leq 1 \implies \sum_{n=1}^{\infty} |X_n Y_n| \leq 1.$$

Now, we must have that

$$\begin{aligned} \sum_{n=1}^{\infty} |X_n Y_n| &= \sum_{n=1}^{\infty} |X_n| |Y_n| \\ &\leq \sum_{n=1}^{\infty} \left(\frac{1}{p} |X_n|^p + \frac{1}{q} |Y_n|^q \right) \\ &= \frac{1}{p} \sum_{n=1}^{\infty} |X_n|^p + \frac{1}{q} \sum_{n=1}^{\infty} |Y_n|^q \\ &= \frac{1}{p} \|X\|_p^p + \frac{1}{q} \|Y\|_q^q \\ &= \frac{1}{p} + \frac{1}{q} = 1, \end{aligned}$$

there the inequality in line 2 stems from the Lemma 0.1. This proves the formula, and we can see that we have equality if and only if $a^p = b^q$. \square

Theorem 2. (Minkowski's Theorem) If $p > 1$, then

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p \quad (1.2)$$

for all $x, y \in \ell^p$.

Proof. Assume again that $x \neq 0$ and $y \neq 0$. Assuming without loss of generality that $x_n > 0$ and $y_n > 0$, knowing that if $\frac{1}{p} + \frac{1}{q} = 1$, then $q(p-1) = p$, and applying (1), we will have that

$$\begin{aligned}
\|x + y\|_p^p &= \sum_{n=1}^{\infty} |x_n + y_n|^p \\
&\leq (|x_n| + |y_n|)^p \\
&= \sum_{n=1}^{\infty} (x_n + y_n)^p \\
&= \sum_{n=1}^{\infty} x_n (x_n + y_n)^{p-1} + \sum_{n=1}^{\infty} y_n (x_n + y_n)^{p-1} \\
&\leq \left(\sum_{n=1}^{\infty} x_n^p \right)^{\frac{1}{p}} \left(\sum_{n=1}^{\infty} (x_n + y_n)^{(p-1)q} \right)^{\frac{1}{q}} + \left(\sum_{n=1}^{\infty} y_n^p \right)^{\frac{1}{p}} \left(\sum_{n=1}^{\infty} (x_n + y_n)^{(p-1)q} \right)^{\frac{1}{q}} \\
&= \left(\sum_{n=1}^{\infty} (x_n + y_n)^p \right)^{\frac{1}{q}} (\|x\|_p + \|y\|_p).
\end{aligned}$$

So, we will have that

$$\begin{aligned}
\sum_{n=1}^{\infty} (x_n + y_n)^p &\leq \left(\sum_{n=1}^{\infty} (x_n + y_n)^p \right)^{\frac{1}{q}} (\|x\|_p + \|y\|_p) \quad (1.3) \\
\left(\sum_{n=1}^{\infty} (x_n + y_n)^p \right)^{1 - \frac{1}{q}} &\leq \|x\|_p + \|y\|_p \\
\left(\sum_{n=1}^{\infty} (x_n + y_n)^p \right)^{\frac{1}{p}} &\leq \|x\|_p + \|y\|_p \\
\|x + y\|_p &\leq \|x\|_p + \|y\|_p.
\end{aligned}$$

□

Theorem 3. ℓ^p is a Banach space for all $1 \leq p \leq \infty$.

Proof. We have proven that ℓ^p is a normed space, so it just goes to show that it is complete. To this end, let $a^n = (a_k^n)_{k \geq 1}$ is a Cauchy sequence in ℓ^p ($1 \leq p < \infty$). Then, for all $\epsilon > 0$, there exists an N_0 such that for all $m, n \geq N_0$, $\|a^n - a^m\|_p^p < \epsilon^p$. This implies that

$$|a_k^n - a_k^m|^p \leq \sum_{k=1}^{\infty} |a_k^n - a_k^m|^p < \epsilon^p. \quad (1.4)$$

Thus, for all $\epsilon > 0$, there exists that same N_0 (independent of k) such that $|a_k^n - a_k^m| < \epsilon$ for all $m, n \geq N_0$. Then, $(a_k^n)_{n \geq 1}$ (fixing k) is Cauchy in \mathbb{K} , which is complete so $(a_k^n)_{n \geq 1}$ is convergent to $a_k \in \mathbb{K}$ for all $k \in \mathbb{N}$. Let $a = (a_1, a_2, \dots, a_k, \dots)$. Letting $m \rightarrow \infty$ in (4), we have that for all $\epsilon > 0$, there is that $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$,

$$\sum_{k=1}^{\infty} |a_k^n - a_k|^p \leq \epsilon^p \implies \|a^n - a\|_p^p \leq \epsilon^p. \quad (1.5)$$

Choosing $N > N_0$, we have that

$$\|a\|_p \leq \|a^n - a\|_p + \|a^n\|_p < \epsilon^p + \|a^n\|_p < \infty,$$

which implies that $a \in \ell^p$. Therefore, from (5), we have that $a^n \rightarrow a$ in ℓ^p because for all $\epsilon > 0$, there is that N_0 such that for all $n \geq N_0$, $\|a^n - a\| \leq \epsilon$. \square

$\mathbf{L}^p(\Omega)$ spaces

Let $\Omega \subseteq \mathbb{R}^d$ be open and $1 \leq p < \infty$. Then, we define

$$\mathbf{L}^p(\Omega) := \left\{ f : \Omega \mapsto \mathbb{R} : f \text{ measurable and } \int_{\Omega} |f(x)|^p < \infty \right\}.$$

The norm is defined as

$$\|f\|_p = \|f\|_{\mathbf{L}^p(\Omega)} = \left(\int_{\Omega} |f|^p dx \right)^{\frac{1}{p}}.$$

Here, we have that $f \equiv g$ if $f = g$ almost everywhere (a.e.). We also define

$$\mathbf{L}^{\infty}(\Omega) = \{f : \Omega \mapsto \mathbb{R} : \exists M > 0 \text{ s.t. } |f(x)| < M \text{ a.e.}\},$$

where

$$\|f\|_{\infty} := \text{ess sup } |f(x)| := \inf \{M > 0 : |f(x)| < M \text{ a.e.}\}.$$

Exercise: a) Use the ε -characterization of inf to describe $\|f\|_{\infty}$.

b) Justify that if $f \in \mathbf{L}^{\infty}(\Omega)$, then for any $\varepsilon > 0$ the set

$$\{x \in \mathbb{R} : |f(x)| > \|f\|_{\infty} - \varepsilon\}$$

is of positive measure.

Theorem 4. $\mathbf{L}^p(\Omega)$ for $1 \leq p \leq \infty$ are Banach spaces.

A couple of notes on this theorem. The proof is on page 235, you need the Holder and Minkowski's Theorems, and the proof that \mathbf{L}^p is complete is called the Riesz-Fisher Theorem.

Linear Operators

Let X, Y be normed spaces over \mathbb{K} (either \mathbb{R} or \mathbb{C}).

Definition 6. A **linear operator** is a mapping $T : \text{Dom}(T) \subset X \rightarrow Y$ such that

$$T(c_1x_1 + c_2x_2) = c_1Tx_1 + c_2Tx_2, \text{ for all } x_1, x_2 \in \text{Dom}(T), c_1, c_2 \in \mathbb{K}.$$

We also define the **range** of a function to be

$$\text{Range}(T) = R(T) = \{Tx : x \in \text{Dom}(T)\} \subset Y.$$

We also define the **null space** or **kernel** of T to be $\ker(T) = \{x \in \text{Dom}(T) : Tx = 0\}$. We note that T is injective if and only if $\ker(T) = \{0\}$.

Definition 7. Assume that $\text{Dom}(T) = X$ and T is a linear operator from $D(T) = X \rightarrow Y$. Then, we say T is **bounded** if

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

Theorem 5. Suppose $T : X \rightarrow Y$ is a linear operator. The following statements are equivalent:

- (a) T is continuous on X
- (b) T is continuous at 0
- (c) T is bounded
- (d) There exists a $C > 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in X$.

Proof. We will prove that (a) \Leftrightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (b).

(a) \Rightarrow (b): This is clear.

(b) \Rightarrow (a): Let $x_n \rightarrow x \in X$. We need to prove that $Tx_n \rightarrow Tx$. To show this, we note that $x_n \rightarrow x$ implies that $\|x_n - x\| \rightarrow 0$. Since T is continuous at 0, we have that $T(x_n - x) \rightarrow T(0) = 0$, or

$$Tx_n - Tx \rightarrow 0 \iff Tx_n \rightarrow Tx \text{ in } Y.$$

(b) \Rightarrow (c): Assume that T is continuous at 0. We claim that $A = \{\|Tx\| : \|x\| < 1\}$ is bounded, and to prove this, we suppose not. Then, for all $n \geq 1$, there exists $x_n \in X$ such that $\|x_n\| \leq 1$ and $\|Tx_n\| \geq n$. Let $y_n = \frac{1}{n}x_n$. Then, we will have that $\|y_n\| = \frac{1}{n}\|x_n\| \leq \frac{1}{n} \rightarrow 0$, which implies that $y_n \rightarrow 0$. Thus, we must have that $Ty_n \rightarrow T(0) = 0$. We also can see that

$$\|Ty_n\| = \left\| T\left(\frac{1}{n}x_n\right) \right\| = \frac{1}{n}\|Tx_n\| \geq 1,$$

but we have that $Ty_n \rightarrow 0$, which is a contradiction. Thus, we proved that A is bounded.

(c) \Rightarrow (d): Let us assume that T is bounded. Let $M = \sup_{\|x\| \leq 1} \|Tx\| < \infty$. For any $x \in X$, $x \neq 0$, where $x' = \frac{x}{\|x\|}$ satisfies $\|x'\| = 1$, we must have that $\|Tx'\| \leq M$, which means that

$$\left\| T \frac{x}{\|x\|} \right\| \leq M \implies \frac{1}{\|x\|} \|Tx\| \leq M \implies \|Tx\| \leq M\|x\|.$$

(d) \Rightarrow (b): Assume that there exists a $c > 0$ such that $\|Tx\| \leq c\|x\|$. We need to prove that $x_n \rightarrow 0$ implies that $Tx_n \rightarrow 0$. Let $x_n \in x$, $x_n \rightarrow 0$. Then, we have that $\|Tx_n\| \leq c\|x_n\|$ and $\|x_n\| \rightarrow 0$. this must imply that $\|Tx_n\| \rightarrow 0$ and thus, $Tx_n \rightarrow 0$. Thus, T is continuous at 0.

□

We will define $B(X, Y) : \{T : X \rightarrow Y : T \text{ is a linear, continuous operator}\}$ and is the space of all bounded, linear operators. If $T, T_1, T_2 \in B(X, Y)$ $\alpha \in \mathbb{K}$, and $x \in X$, then we will have that

$$(T_1 + T_2)x = T_1x + T_2x, \quad T(\alpha x) = \alpha T(x), \quad \|T\| = \sup_{\|x\| \leq 1} \|Tx\|.$$

Theorem 6. *If X, Y are normed spaces over \mathbb{K} and Y is a Banach space, then $B(X, Y)$ is a Banach space (Theorem 2.12 in book).*

If we have that $Y = \mathbb{K}$, then $B(X, Y) = B(X, \mathbb{K})$ is denoted by X^* , which is the **dual** of X . If, $f \in B(X, \mathbb{K})$, then $f : X \rightarrow \mathbb{K}$ and f is linear. This is called a **functional**. Then, $\|f\|_{X^*} = \sup_{\|x\| \leq 1} |f(x)|$. Even though X does not have to be a Banach space, X^* is a Banach space. If $Y = X$, then $B(X, Y) = B(X) = \{T : X \mapsto X : T \text{ is a linear functional}\}$. Here are a few examples:

1. Let $x = \{f : (0, \pi) \mapsto \mathbb{R} : f \text{ continuous, bounded on } (0, \pi)\}$. Then, we have that $\|f\| < \infty$. We define $T : X \mapsto X$ where $Tf = f'$. Then, we have that

$$\text{Dom}(x) = \{f : (0, \pi) \mapsto \mathbb{R} : f \text{ differentiable, } f' \text{ continuous and differentiable on } (0, \pi)\} \subsetneq X.$$

We note however that $f_n(x) = \sin(nx)$, $x \in (0, \pi) \in \text{Dom}(T)$. Then, $Tf_n(x) = n \cos(nx) \in \text{Dom}(T)$. We note that $\|f_n\| = 1$, and then $\|Tf_n\| = n$. So, we can see that T is not bounded since $\|f_n\| \leq 1$ and $\|T\| = +\infty$, for $f_n \rightarrow f$.

2. We can define $\Lambda_+, \Lambda_- : \ell^p \mapsto \ell^p$ as the shift operators, where

$$\Lambda_+(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$$

$$\Lambda_-(x_1, x_2, \dots) = (x_2, x_3, \dots).$$

We can find that Λ_+ is injective but not surjective, and vice-versa for Λ_- .

Finite-Dimensional Normed Spaces

Lemma 2. *Let X be an N -dimensional normed space over \mathbb{K} . Let $\{u_1, \dots, u_N\}$ be a basis for X . For any $\alpha \in \mathbb{R}^N$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)$, define*

$$T\alpha = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_N u_N \in X$$

(Note that $T : \mathbb{R}^N \mapsto X$). Then, we have that

1. T is invertible and linear
2. T is continuous
3. There exists $c_1, c_2 > 0$ such that

$$c_1 \|\alpha\| \leq \|T\alpha\| \leq c_2 \|\alpha\| \quad (1.1)$$

for all $\alpha \in \mathbb{R}^N$.

Consequently, T^{-1} is continuous and

$$c_1 \|T^{-1}x\| \leq \|x\| \leq c_2 \|T^{-1}x\| \quad (1.2)$$

for all $x \in X$ (i.e. T is an isomorphism).

Proof. For 1, we know that $T\alpha = 0$ implies that $\alpha = 0$ since $\alpha = \alpha_1 u_1 + \dots + \alpha_N u_N$ and $\{u_1, u_2, \dots, u_N\}$ is a basis must imply that $\alpha_1 = \dots = \alpha_N = 0$. Therefore, this implies that T is invertible.

For 3, we can see that

$$\begin{aligned} \|T\alpha\| &= \|\alpha u_1 + \alpha_2 u_2 + \dots + \alpha_N u_N\| \\ &\leq |\alpha_1| \|u_1\| + \dots + |\alpha_N| \|u_N\| \\ &\leq \sqrt{\alpha_1^2 + \dots + \alpha_N^2} \cdot \sqrt{\|u_1\|^2 + \dots + \|u_N\|^2} \\ &= \|\alpha\| c_2, \end{aligned}$$

for all $\alpha \in \mathbb{R}^N$. This also proves 2 as well since we have boundedness. Now, to prove that we can have a c_1 on the other side, we assume by contradiction that (1) does not hold. Thus for any $n \geq 1$, (take $c_1 = \frac{1}{n}$), there must exist an $\alpha^n \in \mathbb{R}^n$ such that

$$\|T\alpha^n\| < \frac{1}{n} \|\alpha^n\|.$$

Now, take $\beta^n = \frac{\alpha^n}{\|\alpha^n\|} \in \mathbb{R}^N$ and thus, $\|\beta^n\| = 1$. Note now that

$$\|T\beta^n\| = \frac{1}{\|\alpha^n\|} \|T\alpha^n\| < \frac{1}{\|\alpha^n\|} \frac{1}{n} \|\alpha^n\| = \frac{1}{n}.$$

This means that $T\beta^n \rightarrow 0$. Therefore, since $\beta^n \in \overline{B}(0, 1) \subset \mathbb{R}^N$, there must exist a convergent subsequence β^{n_k} . Therefore, $\beta^{n_k} \rightarrow \beta$ in \mathbb{R}^N . But, we have

that $\|\beta^{n_k}\| = 1$ which implies that $\|\beta\| = 1$. Since T is continuous, and $\beta^{n_k} \rightarrow \beta$, this must imply that $T\beta^{n_k} \rightarrow T\beta = 0$, which is a contradiction since $\|\beta\| = 1$ means that $\beta \neq 0$. This proves that T is not injective, but this is a contradiction with 2. Thus, 3 must hold.

To prove the rest, all one does is denote $T\alpha = x$ and thus, $\alpha = T^{-1}x$ and it follows quickly. \square

Definition 8. If $\|\cdot\|_1$ and $\|\cdot\|_2$ are 2 norms on X , then they are **equivalent** if there exists $\alpha, \beta > 0$ such that

$$\alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1$$

for all $x \in X$.

Consequence 1. In any finite-dimensional normed space, all norms are equivalent.

Proof. Assume that $\dim(X) = N$ and u_1, \dots, u_N is a basis. Denote $T : \mathbb{R}^N \rightarrow X$ as in the above lemma. Then, $(\alpha_1, \dots, \alpha_N)$ implies that $\alpha_1 u_1 + \dots + \alpha_N u_N \in X$. Then, if we have $(X, \|\cdot\|_1)$, then the lemma tells us that a_1, a_2 such that

$$a_1\|T^{-1}x\|_{\mathbb{R}^N} \leq \|x\|_1 \leq a_2\|T^{-1}x\|_{\mathbb{R}^N}$$

for all $x \in X$. As well, if we have $(X, \|\cdot\|_2)$, we know that there exist b_1, b_2 such that

$$b_1\|T^{-1}x\|_{\mathbb{R}^N} \leq \|x\|_2 \leq b_2\|T^{-1}x\|_{\mathbb{R}^N}$$

for all $x \in X$. Therefore, we can see that

$$\frac{b_1}{a_2}\|x\|_1 \leq \|x\|_2 \leq \frac{b_2}{a_1}\|x\|_1$$

which proves that they are equivalent. \square

Consequence 2. Any finite-dimensional normed space is complete.

Proof. Let $(X, \|\cdot\|)$ of dimension N . Then, from the lemma, we know that there exist C_1, C_2 such that

$$C_1\|T^{-1}x\| \leq \|x\| \leq C_2\|T^{-1}x\|. \quad (1.3)$$

Let (x_n) be a Cauchy sequence in X . Then, from (3), we get that

$$\|T^{-1}x_n - T^{-1}x_m\| \leq \frac{1}{C_2}\|x_n - x_m\| \rightarrow 0$$

as $m, n \rightarrow \infty$. Then we see that $(T^{-1}x_n)_{n \geq 1} \in \mathbb{R}^N$ is a Cauchy sequence. Since \mathbb{R}^N is complete, we must have that $(T^{-1}x_n)_{n \geq 1}$ has a limit, say β as $n \rightarrow \infty$. Then, since T is continuous, we have that $T(T^{-1}x_n) = x_n \rightarrow T\beta \in X$. Therefore, (x_n) is convergent in X and thus, it is complete. \square

Consequence 3. Let $(X, \|\cdot\|)$ be a normed space and $B_1 = \overline{B}(0, 1)$. Then, B_1 is compact if and only if X is finite dimensional.

Proof. First, assume that X is finite dimensional. Then, using the Lemma, $T : \mathbb{R}^n \rightarrow X$ is an isomorphism. Since B_1 is closed and bounded and T^{-1} is continuous, then $T^{-1}(B_1)$ is closed and bounded in \mathbb{R}^n . By Bolzano-Weierstrass, $T^{-1}(B_1)$ must be compact. Since $B_1 = T(T^{-1}(B_1))$, then we must have that B_1 is compact.

For the other direction, we will need to two steps:

- (i) We will need the Riesz Lemma: Suppose $Y \subsetneq X$ closed subspace of X , which is a normed space. Given $\epsilon > 0$, there exists an $x \in X$ with $\|x\| = 1$ and $d(x, Y) = \inf_{y \in Y} \|x - y\| \geq 1 - \epsilon$.
- (ii) We will prove that contrapositive of the statement, which is the following: To do this, we assume that X is infinite dimensional. We pick $e_1 \in X$ such that $\|e_1\| = 1$ and let $E_1 = \text{span}\{e_1\}$. Then, E_1 is finite dimensional, thus E_1 is complete, and therefore, E_1 is closed. Since $E_1 \subsetneq X$, by Riesz Lemma when $\epsilon = \frac{1}{2}$, there exists an $e_2 \in X$ such that $\|e_2\| = 1$ and $\|e_2 - e_1\| > \frac{1}{2}$. Now, we define $E_2 = \text{span}\{e_1, e_2\} \subsetneq X$ and it is a closed subspace. Thus, by induction, we can construct $E_n = \text{span}\{e_1, e_2, \dots, e_n\}$ and $\|e_k\| = 1$ for all $k \in \{1, \dots, n\}$ and $\|e_j - e_k\| \geq \frac{1}{2}$ for all $j \neq k$. Then, $E_n \subset B_1$ but has no convergent subsequence and therefore, we have that B_1 is not compact. \square

Hahn-Banach Theorem, Consequences, and Representation of the Dual

We assume that X is a normed space over \mathbb{K} . We define

$$X^* = B(X, \mathbb{K}) = \{f : X \mapsto \mathbb{K} : f \text{ linear and bounded} \}$$

and we call X^* the dual of X . Then, if $f \in X^*$, we have that

$$\|f\| = \|f\|_{X^*} = \sup_{\|x\| \leq 1} |f(x)| = \sup_{\|x\| \neq 0} \frac{|f(x)|}{\|x\|}.$$

Then, $(X^*, \|\cdot\|_{X^*})$ is a Banach space.

For an example, let $X = \mathbb{R}^n$. For any $a \in \mathbb{R}^n$, define a real functional, f_a . Then, $f_a(x) = a_1x_1 + \dots + a_nx_n$ and thus,

$$\|f_a\| = \sup_{\|x\|} \frac{|f_a(x)|}{\|x\|} = \|a\| = \sqrt{a_1^2 + \dots + a_n^2}$$

We note that $f_a(x) = 0$ implies that $a_1x_1 + \dots + a_nx_n = 0$. Thus, we can see that we can represent the dual of \mathbb{R}^n as \mathbb{R}^n , or $(\mathbb{R}^n)^* = \mathbb{R}^n$.

Definition 9. Assume that X is a real normed space. Then, $P : X \mapsto \mathbb{R}$ is a convex functional if

- (i) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$;

(ii) $p(tx) = tp(x)$ for all $t \in [0, \infty)$ and for all $x \in X$.

We note that any convex functional must be a convex function.

Theorem 7. (Hahn-Banach) Let X be a normed space over \mathbb{R} and let $P : X \mapsto \mathbb{R}$ be a convex functional on X . Consider $V \subseteq X$ subspace and let $f : V \mapsto \mathbb{R}$ be a linear functional such that $f(x) \leq p(x)$ for all $x \in V$. Then, there exists a linear functional $F : X \mapsto \mathbb{R}$ such that

(i) $F(x) = f(x)$ for all $x \in V$;

(ii) $F(x) \leq p(x)$ for all $x \in X$.

Proof. Proof is in the textbook, but the goal is to first extend to $V + \text{span}\{x_0\}$ where $x_0 \notin V$. Then, we can use the Housdorf Maximum Principle to get the result on X . \square

We now will discuss some different consequences of the Hahn-Banach theorem, now starting with the Extension Theorems.

Theorem 8. (Extension Theorem: Real Case) Assume that $V \subset X$ real normed subspace and $f : V \mapsto \mathbb{R}$ is a linear, bounded functional. Then, there exists an $F : X \mapsto \mathbb{R}$ functional such that $F(x) = f(x)$ for all $x \in V$ and $\|F\|_{X^*} = \|f\|_{V^*}$.

Proof. Let $M = \|f\|_{V^*}$ and define $p : X \mapsto \mathbb{R}$, where $p = m\|x\|$. Note that p is a convex functional and $|f(x)| = \|f\|_{V^*}\|x\| = m\|x\| = p(x)$ for all $x \in V$. Then by H-B theorem, there exists an $F : X \mapsto \mathbb{R}$ that extends F to X and $F(x) \leq p(x)$ for all $x \in X$. Now, we need to prove that F is bounded and $\|F\|_{X^*} = \|f\|_{V^*}$. To do this, we note that $F(x) \leq m\|x\|$ for all $x \in X$ and we have that $-F(x) = F(-x) \leq m\|-x\| \leq m\|x\|$ for all $x \in X$. Therefore, this implies that $|F(x)| \leq m\|x\|$ for all $x \in X$ and thus, F is bounded ($F \in X^*$) and $\|F\| \leq \|f\|$. But,

$$\|F\| = \sup_{x \in X, \|x\| \leq 1} |F(x)| \geq \sup_{x \in V, \|x\| \leq 1} |F(x)| = \sup_{x \in V, \|x\| \leq 1} |f(x)| = \|f\|.$$

Therefore, $\|F\|_{X^*} = \|f\|_{V^*}$. \square

Theorem 9. (Extension Theorem: Complex Case) Assume that $V \subset X$ complex normed subspace and $f : V \mapsto \mathbb{C}$ is a linear, bounded functional. Then, there exists an $F : X \mapsto \mathbb{C}$ functional such that $F(x) = f(x)$ for all $x \in V$ and $\|F\|_{X^*} = \|f\|_{V^*}$.

Proof. Let $X_{\mathbb{R}}$ be X considered over a vector space of \mathbb{R} . Let $u : V_{\mathbb{R}} \mapsto \mathbb{R}$, $u(x) = \Re(f(x))$. Then,

$$|u(x)|_{X_{\mathbb{R}}^*} = |\Re(f(x))| \leq |f(x)| \leq \|f\|_{V^*}\|x\|$$

for all $x \in X_{\mathbb{R}}$. Then, $u \in X_{\mathbb{R}}^*$ and $\|u\|_{X_{\mathbb{R}}^*} \leq \|f\|_{V^*}$. If we apply the real case for $u : V_{\mathbb{R}} \mapsto \mathbb{R}$, then there exists $u : X_{\mathbb{R}} \mapsto \mathbb{R}$ that is linear and extends u and

$$\|u\|_{X_{\mathbb{R}}^*} = \|u\|_{V_{\mathbb{R}}^*} \leq \|f\|_{V^*}.$$

Define $F(x) = u(x) - iu(ix)$ for all $x \in X$. Then, this map satisfies all of the requirements, F extends f and F is linear, continuous, and bounded. As well, $\|F\|_{X^*} = \|f\|_{V^*}$. \square

Proposition 1. (Separation of Points) *If $(X, \|\cdot\|)$ is a normed space and $x \neq y$, then there exists an $F \in X^*$ such that $F(x) \neq F(y)$.*

Proof. Let $V = \text{span}\{x - y\}$ supposing that $x \neq y$. Then, define $f : V \mapsto \mathbb{K}$ where $f(t(x - y)) = t$ for all $t \in \mathbb{K}$. Note that $f(x - y) = 1$, which implies that $f(x) - f(y) = 1$ and thus, $f(x) \neq f(y)$. Now, we also note that

$$\frac{|f(t(x - y))|}{\|t(x - y)\|} = \frac{|t|}{|t|\|x - y\|} = \frac{1}{\|x - y\|} > 0$$

for all $x \neq y \in X$. Therefore,

$$\|f\|_{V^*} = \sup_{\|f\| \neq 0} \frac{\|f(t(x - y))\|}{t\|x - y\|} = \frac{1}{\|x - y\|},$$

which implies that f is bounded. Thus, by the extension theorem, there exists an $F : X \mapsto \mathbb{K}$ such that $F(z) = f(z)$ for all $z \in X$ and $\|F\| = \|f\| = \frac{1}{\|x - y\|} < \infty$. Therefore, $F \in X^*$ and $F(x) \neq F(y)$. \square

We note that if $(X, \|\cdot\|)$ is a normed space, Proposition 1 has the following consequences:

- (i) $\phi(x) = \phi(y)$ for all $\phi \in X^*$ implies that $x = y$;
- (ii) $\phi(x) = 0$ for all $x \in X^*$ implies that $x = 0$;
- (iii) For all $x \in X$, $x \neq 0$, there exists a $\phi \in X^*$ such that $\phi(x) \neq 0$.

Proposition 2. *Let $(X, \|\cdot\|)$ be a Banach space and $x_0 \in X$, $x_0 \neq 0$. Then, there exists a $\phi \in X^*$ such that $\phi(x_0) = \|x_0\|$ and $\|\phi\| = 1$.*

Proposition 3. (Separating a point from a convex set) *Let $(X, \|\cdot\|)$ be a real normed space. If $\Omega \subset X$ is an open, convex set containing $\mathbf{0}$ and $x_0 \notin \Omega$, then there exists an $F \in X^*$ such that $F(x) < 1 \leq F(x_0)$ for all $x \in \Omega$.*

Proof. Define $p : X \mapsto \mathbb{R}$.

From this, we see that $a < b$ implies that

$$g_b(x) - g_a(x) = \begin{cases} 1, & x \in [a, b] \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, for all $\epsilon > 0$, there exists a g_ϵ piece wise constant such that $\|g - g_\epsilon\|_{L^2} < \epsilon$. Then, if we take

$$|(f_n, g) - (f, g)| = |(f_n - f, g)| = |(f_n - f, g - g_\epsilon) + (f_n - f, g_\epsilon)| \leq |(f_n - f, g - g_\epsilon)| + |(f_n - f, g_\epsilon)| < \epsilon$$

for large n . Then, we can have weak convergence.

Then, we have three claims to finish the proof:

Claim 1: $f_n \not\rightarrow f$ since $\|f_n - f\|_{\frac{1}{8}}^2$ for all $n \geq 1$.

Claim 2: $T : \mathbf{L}^2([0, \pi]) \mapsto \mathbf{L}^2([0, \pi])$, where

$$(Tg)(x) = \int_0^x g(t) dt$$

is a compact operator. We can prove this using the Arzela-Ascoli Theorem

Claim 3: $Tf_n \rightarrow Tf$ uniformly on $[0, \pi]$ and hence $Tf_n \rightarrow Tf$ strongly in $\mathbf{L}^2([0, \pi])$.
We see that

$$(Tf_n)(x) = \int_0^x \sin^2(nt) dt = \frac{x}{2} - \frac{1}{4n} \sin(2nx)$$

for all $x \in [0, \pi]$. As well, $(Tf)(x) = \frac{x}{2}$. Therefore, we have that

$$|(Tf_n(x) - Tf(x))| = \frac{1}{4n} |\sin(2nx)| \leq \frac{1}{4n}$$

for all $x \in [0, \pi]$, $n \geq 1$. So, we have uniform convergence. Therefore,

$$\int_0^\pi (Tf_n - Tf)^2 dx \leq \int_0^\pi \frac{1}{16n^2} \rightarrow 0$$

as $n \rightarrow \infty$. Therefore, we have strong convergence.

5.6 - Positive Definite Operators and Lax-Milgram Theorem

Let H be a Hilbert space over \mathbb{R} (all of this is true also over \mathbb{C} , but to simplify calculations, we only use \mathbb{R}). Let $A : H \mapsto H$.

Definition 10. We say that A is *(strictly) positive definite (SPD)* if

$$(Au, u) \geq \beta \|u\|^2, \quad \forall u \in H \tag{1.1}$$

for some $\beta \in \mathbb{R}$.

We note that if $H = \mathbb{R}^n$ and A is an $n \times n$ matrix, then (1) implies that $Ax = 0$ implies that $x = 0$. Thus, A is injective and thus, A is invertible.

Theorem 10. (An SPD operator is invertible) Let $A : H \mapsto H$ satisfying (1). Then,

- (i) For all $f \in H$, there exists a unique $u \in H$ such that $Au = f$ ($u = A^{-1}f$);
- (ii) The inverse satisfies $\|A^{-1}\| \leq \frac{1}{\beta}$.

Proof. Let $A : H \mapsto H$ be a linear, bounded operator that satisfies (1). Then, from (1), we get that

$$\beta\|u\|^2 \leq (Au, u) \leq \|Au\|\|u\|$$

which implies that

$$\|Au\| \geq \beta\|u\|, \quad \forall u \in H \quad (1.2)$$

Therefore, we have that A is injective and the range of A is closed. Now, we claim that $\text{range}(A) = H$. If this is untrue, then there exists a $w \in \text{range}(A)^\perp$ such that $w \neq 0$. Then, take $u = w$ in (1), we get that $\beta\|w\| \leq (Aw, w) = 0$. Since $Aw \in \text{range}(A)$ and $w \in \text{range}(A)^\perp$, we get that $\|w\| = 0$, which implies that $w = 0$. Therefore, A is onto and thus invertible, and (i) is proven. For (ii), first note that A^{-1} is also linear and by taking $u = A^{-1}f$, from (2) we get

$$\beta\|A^{-1}f\| \leq \|f\|, \quad \forall f \in H.$$

Therefore, we get that $\|A^{-1}\| \leq \frac{1}{\beta}$. \square

Let H be a Hilbert space over \mathbb{R} and let $a : H \times H \mapsto \mathbb{R}$ be a **bilinear form**, i.e.

$$a(\alpha_1 u_1 + \alpha_2 u_2, v) = \alpha_1 a(u_1, v) + \alpha_2 a(u_2, v)$$

$$a(u, \beta_1 v_1 + \beta_2 v_2) = \beta_1 a(u, v_1) + \beta_2 a(u, v_2)$$

for all $\alpha_1, \beta_1, \alpha_2, \beta_2 \in \mathbb{R}$ and $v_1, v_2, u_1, u_2 \in H$.

Theorem 11. (*Lax-Milgram*) Assume that a is a bilinear form, $a : H \times H \mapsto \mathbb{R}$, and satisfies

(i) $|a(u, v)| \leq M\|u\|\|v\|$ for all $u, v \in H$ and for some $M > 0$ (a is bounded)

(ii) there exists a $\beta > 0$ such that $a(u, u) \geq \beta\|u\|^2$ for all $u \in H$ (a is coercive).

Let $F : H \mapsto \mathbb{R}$ be a bounded linear functional. Then, the problem

$$\text{Find } u \in H \text{ such that } a(u, v) = F(v) \text{ for all } v \in H,$$

has a unique solution, $u \in H$, and $\|u\| \leq \frac{1}{\beta}\|F\|$.

Proof. Since G , which maps $v \mapsto a(u, v)$ is linear and bounded (from (i)) on H , by the Riesz Representation Theorem, there exists a unique $u \in H$ such that

$$G(v) = a(u, v) = (Au, v), \quad \forall v \in H. \quad (1.3)$$

Using linearity of $a(\cdot, \cdot)$ with respect to the first component, one can prove that A is a linear operator. Remembering that we can write

$$\|y\| = \sup_{\|x\| \neq 0} \frac{|(x, y)|}{\|x\|}$$

for all $y \in H$, we have that

$$\|Au\| = \sup_{\|v\| \neq 0} \frac{|a(u, v)|}{\|v\|} = \sup_{\|v\| \neq 0} \leq \sup_{\|v\| \neq 0} \frac{M\|u\|\|v\|}{\|v\|} = M\|u\|,$$

where $M \geq 0$ and $u \in H$. From (ii), the coercivity condition and (3), we get that $(Au, u) = a(u, u) \geq \beta \|u\|^2$. By using the Riesz Representation Theorem for F we can find an $f \in H$ such that $F(v) = (f, v)$ for all $v \in H$. From the previous theorem, the problem $Au = f$ has a unique solution, $u = A^{-1}f$, and $\|u\| \leq \frac{1}{\beta} \|f\| = \frac{1}{\beta} \|F\|$. Since $Au = f$, we have that $(Au, v) = (f, v)$ for all $v \in H$, and thus, $a(u, v) = F(v)$ for all $v \in H$. \square

Chapter 6 - Compact Operators on Hilbert Spaces

If $A : \mathbb{R}^n \mapsto \mathbb{R}^n$, $A = (a_{ij})_{i,j=\overline{1,n}}$, then

$$Ax = \left(\sum_{j=1}^n a_{ij} x_j \right)_{i=\overline{1,n}} \in \mathbb{R}^n, \text{ where } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

We know that

(R1) A injective if and only if A surjective.

(R2) $\dim(\ker(A)) = \dim(\ker(A^T))$ and $\text{range}(A)^\perp = \ker(A^T)$.

Now, what happens if we replace \mathbb{R}^n by H , an infinite dimensional Hilbert space? We will see that results (R1) and (R2) are still valid for $A = I - K$ if K is compact.

6.1 - Fredholm Alternative

We now present information from previous chapters to remind ourselves of key facts that we will use going forward.

Proposition 1. (1) If $T \in B(X, Y)$ where X, Y are Hilbert spaces, then $T^* : Y \mapsto X$ is the Hilbert Adjoint and $(Tx, y) = (x, T^*y)$.

(2) $T^{**} = T$ and $\|T^*\| = \|T\|$.

Lemma 1. For any $T \in B(X, Y)$ where X, Y are Hilbert we have that

$$(i) \ker(T) = (\text{range}(T^*))^\perp$$

$$(ii) \ker(T^*) = (\text{range}(T))^\perp.$$

We note that (ii) does not imply that $\text{range}(T) = \ker(T^*)$. This is because we have that

$$\ker(T)^\perp = \text{range}(T)^{\perp\perp} = \overline{\text{range}(T)}.$$

So, the result is only true if $\text{range}(T)$ is closed.

Theorem 1. (Fredholm's Theorem) Let H be a Hilbert Space and let $K : H \mapsto H$ be a compact operator. Then,

- (i) $\ker(I - K)$ is finite dimensional
- (ii) $\text{range}(I - K)$ is closed
- (iii) $\text{range}(I - K) = \ker(I - K^*)^\perp$
- (iv) $\ker(I - K) = \{0\} \iff \text{range}(I - K) = H$
- (v) $\ker(I - K)$ and $\ker(I - K^*) (= \text{range}(I - K)^\perp)$ have the same dimension

Proof. (i) Assume that $\ker(I - K)$ is infinite dimensional. Let $\{e_1, e_2, \dots, e_n, \dots\}$ be an orthonormal sequence in $\ker(I - K)$. Then $Ke_n = e_n$ and $\|e_n\| = 1$, and $\|Ke_n - Ke_m\| = \|e_n - e_m\| = 2$. But this implies that $(Ke_n)_{n \geq 1}$ does not admit a convergent subsequence, which contradicts K being compact.

- (ii) We proved this for Banach spaces (HW #3, Problem 9). Let $N = \ker(I - K) \subset H$ which is closed. So, we have that $H = N \oplus N^\perp$. Therefore, for all $x \in H$, we have that $x = x_N + x_{N^\perp}$, where $x_N \in N$ and $x_{N^\perp} \in N^\perp$. So, $(I - K)x = (I - K)x_{N^\perp}$. So, $\text{range}(I - K) = \text{range}(I - K)|_{N^\perp}$. We claim that there exists a $c > 0$ such that

$$\|(I - K)x\| \geq c\|x\|, \quad \forall x \in N^\perp. \quad (1.1)$$

To prove this, assume that (1) does not hold. Then, there exists an $(x_n)_{n \geq 1} \subset N^\perp$ such that $\|x_n\| = 1$ and $\|(I - K)x_n\| \geq \frac{1}{n}$ for all $n \geq 1$. Therefore, we see that

$$(I - K)x_n \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (1.2)$$

Since K is compact, Kx_n must admit a convergent subsequence $(Kx_{n_j})_{j \geq 1}$ which converges. From (2), we have that $(x_{n_j} - Kx_{n_j})_{j \geq 1}$ must converge. Because two sequences converging implies the sum of the sequences converges, we have that $x_{n_j} - Kx_{n_j} + Kx_{n_j} = x_{n_j}$ converges. But, since $(x_{n_j}) \subset N^\perp$, which is closed, we have that $x_{n_j} \rightarrow x \in N^\perp$. Since K is compact, we must have that K is continuous. Thus, $Kx_{n_j} \rightarrow Kx$, and thus, $x_{n_j} - Kx_{n_j} \rightarrow x - Kx = (I - K)x = 0$ from (2). Therefore, $x \in N$. But, $x \in N \cap N^\perp$ implies that $x = 0$, which contradicts the fact that $\|x_{n_j}\| = 1$. Therefore, the claim is proven and that implies that $I - K$ is surjective.

- (iii) Apply Lemma 1, part (ii) for $T = I - K$. Then, $\ker(I - K^*) = \text{range}(I - K)^\perp$. Then,

$$\ker(I - K^*)^\perp = \text{range}(I - K)^{\perp\perp} = \overline{\text{range}(I - K)} = \text{range}(I - K).$$

- (iv) (\rightarrow) Assume that $\ker(I - K) = \{0\}$ and $\text{range}(I - K) \neq H$. Define $H_1 = (I - K)H = \text{range}(I - K) \subsetneq H$. Then, apply (ii) for the operator $(I - K)|_{H_1}$ to get that $H_2 = (I - K)H_1 = (I - K)^2H$, and by injectivity, $(I - K)H = H$, but we assumed that this is not true. Therefore, we have

that $H_2 \subsetneq H_1$. By induction, we find that for any $n \geq 1$, $H_n = (I - K)^n H$ is a closed subspace of H and $H \supsetneq H_1 \supsetneq H_2 \supsetneq \cdots \supsetneq H_n \supsetneq \cdots$. Choose $(e_n) \subset H_n \cap H_{n+1}^\perp$ such that $\|e_n\| = 1$. We claim that Ke_n does not admit a convergent subsequence, and in fact for any $m < n$, $\|Ke_n - Ke_m\| \geq 1$. To this end, we see that

$$Ke_m Ke_n = - \underbrace{(I - K)e_m}_{\in H_{m+1}} + \underbrace{(I - K)e_n}_{\in H_{n+1} \rightarrow \in H_{m+1}} + e_m - \underbrace{e_n}_{\in H_{m+1}}.$$

Therefore, $Ke_m - Ke_n = e_m + z_{m+1}$, where $z_{m+1} \in H_{m+1}$. Therefore,

$$\|Ke_m - Ke_n\|^2 = \|e_m\|^2 + \|z_{m+1}\|^2 \geq \|e_m\|^2 = 1.$$

Therefore, (Ke_m) does not have a convergent subsequence, contradicting the compactness of K . Thus, the first direction is proved.

(\leftarrow) Assume that $\text{range}(I - K) = H$. From Lemma 1, $\ker(I - K^*) = \text{range}(I - K)^\perp$. So, $\ker(I - K^*) = \{0\}$. Therefore, $I - K^*$ is injective. Using that K^* is also compact and the forward direction of (iv), we get that $\text{range}(I - K^*) = H$. Using Lemma 1, (i) for $T = I - K$, we get that

$$\ker(I - K) = \text{range}(I - K^*)^\perp = \{0\}.$$

- (v) We first show that $\dim(\ker(I - K)) \geq \dim \text{range}(I - K)^\perp$. To this end, suppose to the contrary that

$$\dim \ker(I - K) < \dim \text{range}(I - K)^\perp. \quad (1.3)$$

Then, there exists a linear map $A : \ker(I - K) \mapsto \text{range}(I - K)^\perp$ which is one-to-one but not onto. We can extend A to a linear map $A : H \mapsto \text{range}(I - K)^\perp$ defined on the whole space H , by requiring that $Au = 0$ if $u \in \ker(I - K)^\perp$. Since the range of A is finite dimensional, the operator A is compact and so is $K + A$. Now we claim that $\ker(I - (K + A)) = \{0\}$. Indeed, consider any vector $u \in H$ and write $u = u_1 + u_2$, where $u_1 \in \ker(I - K)$ and $u_2 \in \ker(I - K)^\perp$. then, we see that

$$(I - K - A)(u_1 + u_2) = (I - K)u_2 - Au_1 \in \text{range}(I - K) \oplus \text{range}(I - K)^\perp. \quad (1.4)$$

Since $(I - K)u_2$ is orthogonal to Au_1 , the sum $(I - K)u_2 + Au_1$ can vanish only if $(I - K)u_2 = 0$ and $Au_1 = 0$. Recalling that the operator $I - K$ is one-to-one on $\ker(I - K)^\perp$ and A is one-to-one on $\ker(I - K)$, we conclude that $u_1 = u_2 = 0$. Now, applying (iv) to the compact operator $K + A$, we obtain $\text{range}(I - (K + A)) = H$. However, this is impossible: by construction, there exists a vector $v \in \text{range}(I - K)^\perp$ with $v \notin \text{range}(A)$. By (4), the equation

$$u - Ku - Au = v$$

has no solution. This contradiction shows that (3) cannot hold, and we have that $\dim \ker(I - K) \geq \dim \text{range}(I - K)^\perp$. To get the other inequality,

we recall that $\text{range}(I - K^*)^\perp = \ker(I - K)$, from the previous step, we deduce that

$$\dim \ker(I - K^*) \geq \dim \text{range}(I - K^*)^\perp = \dim \ker(I - K).$$

Interchanging the roles of K and K^* , we get the opposite inequality and we are done. \square

The big consequence of this theorem is Fredholm Alternative, written out as follows:

Theorem 2. (Fredholm Alternative:) *Let $K : H \mapsto H$ be a compact operator where H is Hilbert and let $f \in H$. Consider the problem*

$$\text{Find } u \in H \text{ such that } u - Ku = f \iff (I - K)u = f. \quad (1.5)$$

Then, one of the following two cases holds (a dichotomy)

- 1) *If $\ker(I - K) = \{0\}$, then from (iv), $I - K$ is one to one and onto, and therefore, (5) has a unique solution, $u = (I - K)^{-1}f$. In addition, there exists a $c > 0$ such that $\|u\| \leq c\|f\|$.*
- 2) *If $\ker(I - K) = \text{span}\{\phi_1, \dots, \phi_n\}$ where ϕ_1, \dots, ϕ_n is a basis and $n \geq 1$, then $\ker(I - K^*) = \text{span}\{\phi_1, \dots, \phi_n\}$ which is a basis too. Then, the problem (5) has a solution if and only if*

$$f \in \text{range}(I - K) \iff f \in \ker(I - K^*)^\perp \iff (f, w) = 0 \forall w \in \ker(I - K^*) \iff (f, \phi_k) = 0 \forall k = 1, \dots, n.$$

In this case, if u is a solution of (5), then $u + \sum_{k=1}^n a_k \phi_k$ is a solution where $a_k \in \mathbb{R}$.

We now present a couple of notes, or consequences of the Fredholm Alternative:

1. If K is compact, this implies that $-K$ is compact. Therefore, $I + K$ satisfies the Fredholm Alternative as well.
2. If $A \in B(H)$ is invertible, then A^* is invertible and $(A^*)^{-1} = (A^{-1})^*$. We can see this by showing that since $A \in B(H)$, then

$$A^{-1}A = A^{-1}A = I \implies A^*(A^{-1})^* = (A^{-1})^*A^* = I^* = I.$$

3. Let $A, K \in B(H)$ be such that A is invertible and K is compact. Then, $A + K$ satisfies the Fredholm Alternative.

Proof. First, we note that

$$A + K = A(I + A^{-1}K). \quad (1.6)$$

Since $A + K$ is injective, then, $I + A^{-1}K$ is injective since A is invertible. Since $A^{-1}K$ is compact, then $I + A^{-1}K$ must be surjective. Therefore, since A is surjective, we have that $A(I + A^{-1}K) = A + K$ is surjective. Next, from (6), we get that

$$A^* + K^* = (I + (A^{-1}K)^*) A^*. \quad (1.7)$$

Then, we see that by the Fredholm Alternative,

$$\dim \ker(A + K) = \dim \ker(I + A^{-1}K) = \dim \ker(I + (A^{-1}K)^*).$$

From (7), we note that $u \in \ker(A^* + K^*)$ implies that $A^*u \in \ker(I + (A^{-1}K)^*)$. Therefore, the function that sends $u \rightarrow A^*u$ from $\ker(A^* + K^*) \mapsto \ker(I + (A^{-1}K)^*)$ is one-to-one. Thus, we have that by the Fredholm Alternative

$$\dim \ker(I + (A^{-1}K)^*) = \dim \ker(A^* + K^*)$$

□

4. If A is invertible and K is compact, then $(A + K)u = f$ is solvable if and only if $f \in \ker(A^* + K^*)^\perp$

General Spectral Theory

Let X be a Banach space over \mathbb{K} ($= \mathbb{R}$ or \mathbb{C}) and let $T \in B(X)$.

Definition 1. We say that $\lambda \in \mathbb{K}$ is a

- 1) **regular point of T** if $\lambda I - T$ is invertible.
- 2) **singular point of T** if $\lambda I - T$ is not invertible.

We denote

$$\rho(T) = \{x \in \mathbb{K} : (\lambda I - T)^{-1} \text{ exists} \} = \{\text{set of all regular points}\}$$

$$\sigma(T) = \{x \in \mathbb{K} : (\lambda I - T)^{-1} \text{ does not exist} \} = \{\text{set of all singular points}\}$$

where $\rho(T)$ is the **resolvent of T** and $\sigma(T)$ is the **spectrum of T** . As well, we say that $R : \rho(T) \rightarrow B(X)$, where $R(\lambda) = (\lambda I - T)^{-1}$ is the **resolvent function**.

We briefly note that $\rho(T) \cup \sigma(T) = \mathbb{K}$ and $\rho(T) \cap \sigma(T) = \emptyset$. Now, the spectrum $\sigma(T)$ can be split into three parts, and we define them in the next definition.

Definition 2. Define the following three sets:

$$\sigma_p(T) = \{\lambda \in \sigma(T) : \lambda I - T \text{ is not injective} \}$$

$$\sigma_c(T) = \{\lambda \in \sigma(T) : \lambda I - T \text{ is injective, not surjective, but } \overline{\text{range}(\lambda I - T)} = X\}$$

$$\sigma_r(T) = \{\lambda \in \sigma(T) : \lambda I - T \text{ is injective, but } \overline{\text{range}(\lambda I - T)} \neq X\}.$$

We say that $\sigma_p(T)$ is the **point spectrum of T** , $\sigma_c(T)$ is the **continuous spectrum of T** , and $\sigma_r(T)$ is the **residual spectrum of T** . We note that $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$.

Definition 3. Any $\lambda \in \sigma_p(T)$ is called an **eigenvalue of T** , i.e. there exists an $x \neq 0$ such that $(\lambda I - T)x = 0$, which is also rewritten usually as $Tx = \lambda x$. Any $k \in \ker(\lambda I - T)$, $x \neq 0$ is an **eigenvector of T corresponding to λ** . Then, $\ker(\lambda I - T)$ is the **corresponding eigenspace**.

If X is finite-dimensional, then we know that $\dim \ker(\lambda I - T) + \dim \text{range}(\lambda I - T) = \dim(X) < \infty$. Therefore, we can see that $\ker(\lambda I - T) = \{0\}$ if and only if $\lambda I - T$ is invertible. The negation of this statement is true, i.e. $\ker(\lambda I - T) \neq \{0\}$ if and only if $\lambda I - T$ is not invertible. Therefore, this implies that $\lambda \in \sigma_p(T)$ if and only if $\lambda \in \sigma(T)$. Thus, if X is finite dimensional, we get that $\sigma(T) = \sigma_p(T)$.

Lemma 2. Let $S, T \in B(X)$. Then,

- (a) If $\|T\| < 1$, then $I - T$ is invertible and $(I - T)^{-1} = \sum_{n=0}^{\infty} T^n = I + T + T^2 + \dots$,
and

$$\|(I - T)^{-1}\| < \frac{1}{1 - \|T\|}.$$

- (b) If T is invertible and $\|T - S\| < \frac{1}{\|T^{-1}\|}$, then S is invertible and

$$\|S^{-1} - T^{-1}\| < \frac{\|T^{-1}\|^2 \|S - T\|}{1 - \|T^{-1}\| \|S - T\|}$$

Theorem 3. Let X be a Banach space and $T \in B(X)$. then,

- (a) $\sigma(T) \subset \{\lambda \in \mathbb{K} : |\lambda| \leq \|T\|\}$;
(b) $\sigma(T)$ is a compact set in \mathbb{K} ;
(c) If $\lambda > \|T\|$, then $\lambda \in \rho(T)$ and

$$R(\lambda) = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}} \quad \text{and} \quad \|R(\lambda)\| < \frac{1}{|\lambda| - \|T\|}.$$

Proof. Let $T \in B(X)$ where X is a Banach space.

- (a) If $|\lambda| > \|T\|$, then $\lambda I - T = \lambda(I - \frac{1}{\lambda}T)$ and $\|\frac{1}{\lambda}T\| = \frac{1}{|\lambda|}\|T\| < 1$. By Lemma 2, we get that $I - \frac{1}{\lambda}T$ is invertible and therefore, $\lambda I - T$ is invertible. Thus, we must have that $\lambda \in \rho(T)$. Therefore, this all implies that if $|\lambda| > \|T\|$, then $\lambda \in \rho(T)$. Therefore, the contrapositive is true, namely $\lambda \in \sigma(T)$ implies that $|\lambda| \leq \|T\|$, which proves that $\sigma(T)$ is a subset of the set above.

- (b) From (a), we can see that $\sigma(T)$ is bounded. It is enough then to show that $\sigma(T)$ is closed, or that $\sigma(T)^c = \rho(T)$ is open. To do this, let $\lambda_0 \in \rho(T)$. Then, $\lambda I - T$ is invertible by definition. We want to apply Lemma 2 to this situation. We want to pick a λ close to λ_0 such that if we denote $S = \lambda I - T$ and $T = \lambda_0 I - T$, we can get

$$\|S - T\| = \|\lambda - \lambda_0\| = |\lambda - \lambda_0| < \frac{1}{\|(\lambda_0 I - T)^{-1}\|}.$$

Therefore, select a $\lambda \in \rho(T)$ such that $|\lambda - \lambda_0| < \frac{1}{\|(\lambda_0 I - T)^{-1}\|}$. Thus, using Lemma 2 part (b), we get that $\lambda I - T$ is invertible, and thus $\lambda \in \rho(T)$. As well, we just proved that $\lambda \in \rho(T)$ implies that $B_{\frac{1}{\|R(\lambda)\|}}(\lambda_0) \subset \rho(T)$. Therefore, $\rho(T)$ is open and finally, $\sigma(T)$ is closed.

- (c) We can see that

$$R(\lambda) = (\lambda I - T)^{-1} = \frac{1}{\lambda} \left(I - \frac{1}{\lambda} T \right)^{-1}.$$

Now, using Lemma 2(b), we can see that

$$\frac{1}{\lambda} \left(I - \frac{1}{\lambda} T \right)^{-1} = \frac{1}{\lambda} \sum_{n=0}^{\infty} \frac{T^n}{\lambda^n} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}.$$

Using this representation, one can readily see that the norm of R satisfies the condition in the problem.

□

We denote $\sigma_c(T) \cup \sigma_r(T) = \sigma_e(T)$ as the **essential spectrum**. We now consider properties regarding the spectral properties in a Hilbert space, making this into a proposition.

Proposition 2. *Let X be a Hilbert space, $\langle \cdot, \cdot \rangle$ be an inner product and let T^* be defined as $\langle Tx, y \rangle = \langle x, T^*y \rangle$ for all $x, y \in X$ where $T \in B(X)$. Then,*

- (a) $\lambda \in \sigma(T)$ if and only if $\bar{\lambda} \in \sigma(T^*)$;
- (b) If $\lambda \in \sigma_p(T)$, then $\bar{\lambda} \in \sigma_p(T^*) \cup \sigma_r(T^*)$;
- (c) If $\lambda \in \sigma_r(T)$, then $\bar{\lambda} \in \sigma_p(T^*)$.

Proof. (a) We will prove that $\lambda \in \rho(T)$ if and only if $\bar{\lambda} \in \rho(T^*)$ and this will imply our result. We observe the following if and only if statements:

$$\begin{aligned} \lambda \in \rho(T) &\iff \lambda I - T \text{ invertible} \\ &\iff \text{there exists } S \in B(X) \text{ such that } (\lambda I - T)S = S(\lambda I - T) = I \\ &\iff S^*(\lambda I - T)^* = S^*(\bar{\lambda} I - T^*) = (\bar{\lambda} I - T^*)S^* = I^* = I \\ &\iff \bar{\lambda} I - T^* \text{ invertible} \\ &\iff \bar{\lambda} \in \rho(T^*). \end{aligned}$$

Therefore, taking the negation of the statement, we have our result.

- (b) If $\lambda \in \sigma_p(T)$, then this implies that $\ker(\lambda I - T) \neq \{0\}$. Thus, we know that $\text{range}(\overline{\lambda I - T^*})^\perp \neq \{0\}$. Since this is not just the 0 set, we know that the perpendicular set of this cannot be the entire space, or $\text{range}(\overline{\lambda I - T^*})^{\perp\perp} \neq X$ which can also be written then as $\text{range}(\overline{\lambda I - T^*}) \neq X$. Therefore, this must imply that $\overline{\lambda} \notin \sigma_c(T^*)$. Therefore, since it is a disjoint union, we must have that $\overline{\lambda} \in \sigma_r(T^*) \cup \sigma_p(T^*)$.
- (c) Let $\lambda \in \sigma_r(T)$. Then, we know that $K = \overline{(\lambda I - T)X}$ is a proper subspace of X . Therefore, this means that $\{0\} \neq K^\perp = \ker(\overline{\lambda I - T^*})$, which implies that $\overline{\lambda}$ is an eigenvalue of T^* . Therefore, $\overline{\lambda} \in \sigma_p(T^*)$. \square

As an example, take $L, R : \ell^2 \mapsto \ell^2$ as the left and right shift operators, namely

$$L(x_1, x_2, \dots) = (x_2, x_3, x_4, \dots) \quad R(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots).$$

We now want to compute $\sigma_p(L), \sigma(L), \sigma_p(R)$, and $\sigma(R)$. First, we note that $\|L\| = \|R\| = 1$, which has been proved in previous homework assignments. From (a) of Theorem 3, we can see that $\sigma(L), \sigma(R) \subset \overline{D(0, 1)}$, where $D(0, 1)$ is the disk of radius 1 with center 0. So, let $\lambda \in \sigma_p(L)$. Then, we must have that

$$(x_2, x_3, \dots, x_n, \dots) = (\lambda x_1, \lambda x_2, \dots, \lambda x_{n-1}, \dots).$$

Therefore, if $x_1 = 0$, we then have that $x_i = 0$ for all i and thus, $x = 0$, which cannot happen since $x \neq 0$. So, assume that $x_1 = 1$. Then, $x_2 = \lambda, x_3 = \lambda^2$, and thus, $x_n = \lambda^{n-1}$ for all n . So, $(1, \lambda, \lambda^2, \dots)$ is an eigenvector provided that $(1, \lambda, \lambda^2, \dots) \in \ell^2$. Therefore, we need that

$$\sum_{n=0}^{\infty} |\lambda^n|^2 < +\infty, \quad \text{i.e.,} \quad \sum_{n=0}^{\infty} |\lambda^2|^n < \infty,$$

which implies that $|\lambda| < 1$. Therefore, all $\lambda \in D(0, 1)$ belongs to $\sigma_p(L)$ and vice versa. So, $\sigma_p(L) = D(0, 1)$ and $D(0, 1) = \sigma_p(L) \subset \sigma(L) \subset \overline{D(0, 1)}$. Since $\sigma(L)$ is compact, it is closed. Therefore, this must imply that $\sigma(L) = \overline{D(0, 1)}$. We have also seen in our homework that $\sigma(R) = \sigma(L^*) = \overline{D(0, 1)} = \overline{D(0, 1)}$. Thus, we need to find $\sigma_p(R)$. If $\lambda \in \sigma_p(R)$, then $Rx = \lambda x$ for some $x \neq 0$. We then have

$$(0, x_1, x_2, \dots) = (\lambda x_1, \lambda x_2, \dots)$$

So, if $\lambda = 0$, then $x = 0$, which is a contradiction. So, if $\lambda \neq 0$, we must have that $x_1 = 0$, which in turn implies that $x_n = 0$ for all n which again is a contradiction. Therefore, we must have that $\sigma_p(R) = \emptyset$.

6.2 - Spectrum of Compact Operators

Theorem 4. *Let H be an infinite-dimensional Hilbert Space and $K : H \mapsto H$ be compact. Then,*

- (a) $0 \in \sigma(K)$;
- (b) $\sigma(K) = \sigma_p(K) \cup \{0\}$ (i.e., the essential supremum is at most 0);
- (c) $\sigma_p(K)$ is either finite or countable and if countable, then $\sigma_p(K) = \{\lambda_n : n \geq 1\}$ and $\lambda_n \rightarrow 0$, i.e. if $\sigma_p(K)$ is infinite, then it accumulates only at 0.

Proof. (a) If $0 \neq \sigma(K)$, then K is invertible. Then, there must exist a $K^{-1} : H \mapsto H$, $K^{-1} \in B(H)$. So, $I = K \circ K^{-1}$, but since K is compact, K^{-1} is bounded, and compositions of a compact operator and a bounded operator are compact, then I must be compact. Therefore, this implies that H must be finite dimensional, which is a contradiction.

- (b) We always have that $\sigma_p(K) \subset \sigma(K)$ and we know from (a) that $0 \in \sigma(K)$. So, $\sigma_p(K) \cup \{0\} \subset \sigma(K)$. Now, for the other inclusion, it is enough to prove that if $\lambda \in \sigma(K)$ and $\lambda \neq 0$, then $\lambda \in \sigma_p(K)$. Assume to the contrary. Then, $\lambda \notin \sigma_p(K)$ and thus, $\ker(\lambda I - K) = \{0\}$. Then by the Fredholm Alternative, $\text{range}(\lambda I - K) = H$. Thus, by the Open Mapping Theorem, this implies that $\lambda I - K$ is invertible, which implies that $\lambda \notin \sigma(K)$, which is a contradiction. Therefore, equality holds.

- (c) To prove this, assume that $(\lambda_n)_{n \geq 1}$ is a sequence of distinct eigenvalues of K where $\lambda_n \rightarrow \lambda$. We now claim that $\lambda = 0$. Since $\lambda_n \in \sigma_p(K)$, for each $n \geq 1$, there exists an eigenvector w_n such that $Kw_n = \lambda_n w_n$. Denote $H_n := \text{span}\{w_1, \dots, w_n\}$ and w_1, \dots, w_n must be linearly independent since $\lambda_1, \dots, \lambda_n$ are distinct. Therefore, H_n has dimension n . We observe that $(K - \lambda_n I)H_n \subset H_{n-1}$ since

$$(K - \lambda_n I) \left(\sum_{k=1}^n \alpha_k w_k \right) = \sum_{k=1}^{n-1} \alpha_k (K - \lambda_n I) w_k + \underbrace{\alpha_n (K - \lambda_n I) w_n}_{=0} = \sum_{k=1}^{n-1} \alpha_k (\lambda_k - \lambda_n) w_k \in H.$$

For $n \geq 2$, choose $e_n \in H_{n-1}^\perp \cap H_n$ such that $\|e_n\| = 1$. Choose $e_1 \in H_1$ such that $\|e_1\| = 1$. Then, we have that $(e_n)_{n \geq 1}$ is an orthonormal sequence and it is bounded since $\|e_n\| = 1$. Then we see that assuming $m < n$, then

$$\begin{aligned} Ke_n - Ke_m &= \underbrace{(Ke_n - \lambda_n e_n)}_{\in H_{n-1}} - \underbrace{(Ke_m - \lambda_m e_m)}_{\in H_{m-1} \subset H_{n-1}} + \lambda_n e_n - \underbrace{\lambda_m e_m}_{\in H_m \subset H_{n-1}} \\ &= \lambda_n e_n + z_{n-1}, \end{aligned}$$

where $z_{n-1} \in H_{n-1}$. Since $\lambda_n e_n \in H_{n-1}$ as well, we see that

$$\|Ke_n - Ke_m\|^2 = |\lambda_n|^2 + \|z_{n-1}\|^2 \geq |\lambda_n|^2.$$

Therefore, this implies that $\|Ke_n - Ke_m\| \geq |\lambda_n| - |\lambda|$. However, if $\lambda \neq 0$, this contradicts K being compact, so $\lambda = 0$. Using a similar argument, we can prove that if $t \geq 0$, then $\{\lambda \in \sigma_p(K) : |\lambda| > t\}$ is finite. Then,

$$\sigma_p(K) \subset \bigcup_{n \geq 1} \left\{ \lambda \in \sigma_p(K) : |\lambda| \geq \frac{1}{n} \right\} \cup \{0\}$$

is a countable union of finite sets, and thus countable. Therefore, $\sigma_p(K)$ is at most countable. \square

6.3 - Self-Adjoint Operators

Definition 4. Let H be a Hilbert Space and $T : H \mapsto H$ be a bounded linear operator. We define T to be **self-adjoint** (or **symmetric**) if $T^* = T$, i.e. $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in H$.

As an example, define $A : \mathbb{C}^2 \mapsto \mathbb{C}^2$, where

$$A \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 2 & i \\ -i & 3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

One can find that this is self-adjoint since $A^* = \overline{A^T} = A$.

Proposition 3. Eigenvalues of self-adjoint operators are real.

Proof. Let (λ, x) be an eigenpair for $T = T^*$. We know that

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

$$\langle x, Tx \rangle = \langle x, \lambda x \rangle = \bar{\lambda} \|x\|^2.$$

By the self-adjointness of the operators, we must have that $\bar{\lambda} \|x\|^2 = \lambda \|x\|^2$, which implies $\lambda = \bar{\lambda}$. Therefore, $\lambda \in \mathbb{R}$. \square

Proposition 4. If T is self-adjoint, then $(Tx, x) \in \mathbb{R}$.

Proof. We note that $(Tx, x) = \overline{(x, Tx)}$. Since $(x, Tx) = (Tx, x)$, we have that $\overline{(x, Tx)} = (x, Tx)$, which implies that $(x, Tx) \in \mathbb{R}$. Thus, $(Tx, x) \in \mathbb{R}$. \square

Now, we will define

$$m = \inf_{\|u\|=1} (Tu, u) = \inf_{u \neq 0} \frac{(Tu, u)}{\|u\|^2}$$

$$M = \sup_{\|u\|=1} (Tu, u) = \sup_{u \neq 0} \frac{(Tu, u)}{\|u\|^2}$$

Lemma 3. (Bounds of the Spectrum of a self-adjoint Operator) Let $T \in B(H)$ where T is self-adjoint. Then,

- (i) $\sigma(T) \subset [m, M]$
- (ii) $m, M \in \sigma(T)$
- (iii) $\|T\| = \max\{-m, M\} = \max\{|m|, |M|\}$

Proof. (i) We will prove that if $\lambda \in \mathbb{R}$ and $\lambda \notin (-\infty, m) \cup (M, \infty)$, then $\lambda \in \rho(T)$. To this end, if $\lambda > M$, then

$$\langle (\lambda I - T)u, u \rangle = \lambda \langle u, u \rangle - \langle Tu, u \rangle \geq (\lambda - M)\|u\|^2 > 0.$$

By the consequence of the Lax-Milgram theorem, $\lambda I - T$ is invertible and there in $B(H)$. Similarly, if $\lambda < m$, then $\langle (T - \lambda I)u, u \rangle \geq (m - \lambda)\|u\|^2$, which means that $T - \lambda I$ is invertible. Thus, $\lambda \in \rho(T)$. Therefore, $\rho(T) \subseteq (-\infty, m) \cup (M, \infty)$ and therefore, $\sigma(T) \subseteq [m, M]$.

(iii) Without loss of generality, assume $M \geq 0$ and $|m| \leq M$ (if not, replace T by $-T$). Note that for any $u, v \in H$ we have that

$$\begin{aligned} 4\langle Tu, v \rangle &= \langle T(u+v), u+v \rangle - \langle T(u-v), u-v \rangle \leq M(\|u+v\|^2 + \|u-v\|^2) \\ &= 2M(\|u\|^2 + \|v\|^2). \end{aligned}$$

Now, replace v by $\frac{\|u\|}{\|Tu\|}Tu$. So, noting that $\|v\| = \|u\|$, we have that

$$2\|u\|\|Tu\| \leq M(\|u\|^2 + \|u\|^2)$$

$$\|Tu\| \leq M\|u\|$$

for all $u \in H$, and thus, $\|T\| \leq M$. Therefore, for all $u \in H$ where $\|u\| = 1$, we have that

$$\langle Tu, u \rangle \leq \|Tu\|\|u\| \leq \|T\|\|u\|\|u\| = \|T\|.$$

Thus, taking the supremum, we have that

$$M = \sup_{\|u\| \leq 1} \langle Tu, u \rangle \leq \|T\|.$$

Thus, $M = \|T\|$.

(ii) Without loss of generality, assume that $M \geq 0$ and $|m| < M$, and we claim that $M \in \sigma(T)$. From the definition of M , we can choose $(u_n) \subset H$ such that $\|u_n\| = 1$ and $\langle Tu_n, u_n \rangle \rightarrow M$. Then,

$$\begin{aligned} \|Tu_n - Mu_n\|^2 &= \|Tu_n\|^2 - 2M\langle Tu_n, u_n \rangle + M^2\|u_n\|^2 \\ &\leq \underbrace{\|T\|^2}_{=M^2}\|u_n\|^2 - 2M \underbrace{\langle Tu_n, u_n \rangle}_{\rightarrow M} + M^2 \rightarrow 0. \end{aligned}$$

Therefore, $\|(MI - T)u_n\| \rightarrow 0$ which implies that $(MI - T)u_n \rightarrow 0$, and thus, $M \in \sigma(T)$. □

A couple of notes regarding this lemma:

(1) If $T \in B(H)$ and $T = T^*$, then

$$\|T\| = \sup_{\|x\|=1} |\langle Tx, x \rangle| = \sup_{x \neq 0} \frac{|\langle Tx, x \rangle|}{\|x\|^2}.$$

We note that if $M > |m|$, then $\|T\| = M$ and if $M < |m|$, then $\|T\| = |m|$.

(2) If $M \geq 0$ and $m = 0$, then $\|T\| = M = \sup_{x \neq 0} \frac{|\langle Tx, x \rangle|}{\|x\|^2}$.

(3) If $T = T^*$, $T \in B(H)$, and $\sigma(T) = \{0\}$, then (iii) tells us that $\|T\| = 0$, which implies that $T \equiv 0$.

Theorem 5. (Hilbert-Schmidt Theorem) *Eigenvectors of compact symmetric operators form an orthonormal basis. More formally, let H be a separable Hilbert space and $T : H \mapsto H$ be a compact self-adjoint operator. Then, there exists a countable orthonormal basis consisting of eigenvectors of T .*

Proof. We remember that if H is finite dimensional and $T^*T = TT^*$ (which is clearly satisfied if T is self-adjoint), we say that T is **normal**. If $\mathbb{K} = \mathbb{C}$, then we have the Compact Spectral Theorem (in Axler), and we know that there exists a finite basis of orthonormal vectors. If $\mathbb{K} = \mathbb{R}$, we have that the Real Spectral Theorem gives us the same result.

So, we will assume that H is infinite-dimensional. Since T is compact, let $(\lambda_n)_{n \geq 1}$ be the sequence of distinct eigenvalues where $\lambda_n \neq 0$ for all $n \geq 1$. Denote $\lambda_0 = 0$, $E_0 = \ker(T)$, and $E_n = \ker(\lambda_n I - T)$. We know that $E_0 \subset H$ and since H is separable, E_0 is separable. Because of this fact, E_0 admits an orthonormal countable basis of eigenvectors corresponding to $\lambda_0 = 0$. Since T is compact, we have that E_n is finite-dimensional and admits an orthonormal basis.

We now claim that $H = E_0 \oplus E_1 \oplus E_2 \cdots \oplus E_n \oplus \cdots$ in the sense that

(i) the spaces E_n are mutually orthogonal, and

(ii) if $F = \text{span} \left(\bigcup_{n=0}^{\infty} E_n \right)$, then F is dense in H .

To show (i), if $u \in E_n$ and $v \in E_m$, where $n \neq m$, we have that

$$\lambda_n \langle u, v \rangle = \langle \lambda_n u, v \rangle = \langle Tu, v \rangle = \langle u, Tv \rangle = \langle u, \lambda_m v \rangle = \lambda_m \langle u, v \rangle$$

This gives us $(\lambda_n - \lambda_m) \langle u, v \rangle = 0$ and since the eigenvalues are distinct, this necessarily implies that $\langle u, v \rangle = 0$ and thus, all E_n 's are orthogonal to each other.

For (ii), we can show that F and F^\perp are invariant subspaces for T , i.e. $T(F) \subset F$ and $T(F^\perp) \subset F^\perp$. To show this, first let $u \in F = \bigcup E_n$. Therefore, $u \in E_m$ for some $m \in \mathbb{N} \cup \{0\}$. Therefore, $Tu = \lambda_m u$ and since E_m is a subspace, $\lambda_m u \in E_m$. Therefore, F is invariant. To show that F^\perp is invariant, let $u \in F^\perp$, which implies that $u \perp E_n$ for all $n \geq 0$. Thus, for all $v \in F$, we have that

$$\langle Tu, v \rangle = \langle \underbrace{u}_{\in F^\perp}, \underbrace{Tv}_{\in F} \rangle = 0$$

since F^\perp and F are, by definition, orthogonal. Therefore, $Tu \in F^\perp$ since it is orthogonal to all vectors $v \in F$. Thus F^\perp is invariant under T . Now, take $T_0 = T|_{F^\perp}$. Since F^\perp is a closed subspace of H , a Hilbert space, F^\perp is also a Hilbert space. Since T is compact, T_0 is compact and since T is self-adjoint, T_0 is self-adjoint since

$$\langle T_0 x, y \rangle = \langle T x, y \rangle = \langle x, T y \rangle = \langle x, T_0 y \rangle$$

for all $x, y \in F^\perp$. We now claim that $\sigma(T_0) = \{0\}$. To this end, suppose not. Then, since T_0 is compact, there must exist a $\lambda \in \sigma(T_0)$, $\lambda \neq 0$ such that λ is an eigenvalue, i.e. $T_0 v = \lambda v$ for some vector $v \in F^\perp$, $v \neq 0$. This also implies that $Tv = \lambda v$ for some $v \neq 0$, and therefore, there must exist an $n \geq 1$ such that $\lambda = \lambda_n$. However, this implies that $v \in E_n \subset F$, which means that $v \in F \cap F^\perp$. Thus, $v = 0$ which is a contradiction. Therefore, $\sigma(T_0) = \{0\}$, which means that by note (3), $T_0 \equiv 0$.

This gives us that $T|_{F^\perp} \equiv 0$, which necessarily implies that $F^\perp \subset E_0 \subset F$. Thus, we must have that $F^\perp = \{0\}$, which gives us that F is dense in H . Therefore, (ii) is proved and thus, we can choose an orthonormal basis for each E_n and union them all together to get a countable orthonormal basis. \square

We now assume that H is a Hilbert and separable space and $T \in B(H)$ is compact and self-adjoint. Let $(e_k)_{k \geq 1}$ be a basis for $\bigcup_{n \geq 1} E_n$, where $E_n = \ker(\lambda_n I - T)$, $E_0 = \ker(T)$ and $(\lambda_n)_{n \geq 1} = \sigma_p(T)$, $\lambda_n \neq 0$. Then, for all $f \in H$, we have that

$$f = \underbrace{P_{E_0}(f)}_{\in E_0} + \underbrace{(I - P_{E_0})(f)}_{\in E_0^\perp} = P_{E_0}(f) + \sum_{k=1}^{\infty} \langle f, e_k \rangle e_k.$$

If we assume further that for some $m \geq 0$, $\langle Tx, x \rangle \geq m \langle x, x \rangle$ and we note that $\langle Tx, x \rangle \leq \|T\| \|x\|^2$, this implies that $\|Tx\| \geq M \|x\|$ for all $x \in H$ and thus, $\text{range}(T)$ is closed. Therefore, $\text{range}(T) = \ker(T)^\perp$ (since T is self-adjoint). Thus, if $f \in \text{range}(T)$, we have that $P_{E_0}(f) = 0$ and thus,

$$f = \sum_{k=1}^{\infty} \langle f, e_k \rangle e_k,$$

noting that the limits start at 1, and not 0. Why is this helpful? Well, assume we wanted to solve $Tx = f$. Then, we can let

$$x = \sum_{k=1}^{\infty} c_k e_k + P_{E_0}(x)$$

where $P_{E_0}(x)$ is in the kernel of T . Therefore,

$$Tx = \sum_{k=1}^{\infty} c_k T e_k = \sum_{k=1}^{\infty} \lambda_k c_k e_k.$$

Thus, equating coefficients, we have that $\lambda_k c_k = \langle f, e_k \rangle$, or

$$c_k = \frac{\langle f, e_k \rangle}{\lambda_k}$$

for all $k \geq 1$. This gives us a representation for x that is easily found via computation.

Extra Class: Applications of Functional Analysis

The Rellick-Kondrator Theorem

Let $\Omega \subset \mathbb{R}^d$ be an open and bounded set and $D(\Omega) = \{f : \Omega \mapsto \mathbb{R} : \text{supp}(f) \text{ is compact}, f \in C^\infty\}$ be the set of smooth functions, where $\text{supp}(f) = \overline{\{x \in \Omega : f(x) \neq 0\}}$. In this section, if we see (\cdot, \cdot) , assume that it is $(\cdot, \cdot)_{\mathbf{L}^2(\Omega)}$ unless otherwise noted.

Theorem 1. (Poincaré Inequality) For any $u \in D(\Omega)$, we have that

$$\left(\int_{\Omega} |u|^2 dx \right)^{\frac{1}{2}} \leq C(\Omega) \left(\int_{\Omega} |\nabla u|^2 dx \right)^{\frac{1}{2}},$$

where $C(\Omega)$ is a constant that depends on Ω .

We will eventually see that $(D(\Omega), \|\nabla \cdot\|_{\mathbf{L}^2}) \subset \ell^2$ is a normed space, but it is not complete. Now, we define H_0^1 as the closure of $(D(\Omega), \|\nabla \cdot\|_{\mathbf{L}^2(\Omega)})$ in $\mathbf{L}^2(\Omega)$, i.e.

$$H_0^1(\Omega) = \left\{ u \in \mathbf{L}^2(\Omega) : \int_{\Omega} |\nabla u|^2 dx < \infty, u|_{\partial\Omega} = 0 \right\}.$$

We note that the inner product on $H_0^1(\Omega)$ is

$$(u, v)_{H_0^1} = (\nabla u, \nabla v), \quad \|u\|_{H_0^1} = \left(\int_{\Omega} |\nabla u|^2 dx \right)^{\frac{1}{2}}.$$

The Poincaré inequality in H_0^1 is

$$\|u\|_{\mathbf{L}^2(\Omega)} \leq C(\Omega) \|u\|_{H_0^1(\Omega)}, \quad \forall u \in H_0^1(\Omega).$$

Therefore, we have that $(H_0^1(\Omega), \|\cdot\|_{H_0^1(\Omega)}) \subset \mathbf{L}^2(\Omega)$ is an embedding, and we can see that convergence in $H_0^1(\Omega)$ implies convergence in $\mathbf{L}^2(\Omega)$, but not vice versa.

Theorem 2. (Rellick-Kondrator Theorem) Define $i : H_0^1(\Omega) \rightarrow \mathbf{L}^2(\Omega)$, where $iu = u$ for all $u \in H_0^1$. Then, i is a compact operator. This implies that bounded sequences in $H_0^1(\Omega)$ have convergent subsequences.

This theorem means that $H_0^1(\Omega)$ is a **compact embedding** into $\mathbf{L}^2(\Omega)$, denoted by $H_0^1(\Omega) \subset\subset \mathbf{L}^2(\Omega)$. Define i^* to be the Hilbert dual of i . Then, $(i^*f, v)_{H_0^1} = (f, iv) = (f, v)$ for all $v \in H$. Then, $(\nabla u, \nabla v) = (f, v)$ for all $v \in H_0^1(\Omega)$. Then, i^*f is the representation of $v \mapsto (f, v)$ in H_0^1 , and thus,

$$|(f, v)| \leq \underbrace{\|f\| \|v\|_{\mathbf{L}^2} \leq \|f\| C(\Omega) \|v\|_{H_0^1}}_{\text{by Poincaré Identity}}.$$

Elliptic Equations (Ch. 9 in AB)

Given $a^{ij}(x), b^i(x), c(x) : \Omega \subset \mathbb{R}^d \mapsto \mathbb{R}$, define

$$Lu = - \sum_{i,j=1}^d (a^{ij}(x)u_{x_i})_{x_j} + \sum_{i=1}^d (b^i(x)u)_{x_i} + c(x)u \quad (1.1)$$

where $u_{x_j} = \frac{\partial u}{\partial x_j}$. Now, the main problem with these elliptic equations is the following: find $u \in H_0^1(\Omega)$ such that $Lu = f$ given $f \in \mathbf{L}^2(\Omega)$. If $d = 3$, the first summation in (1) is diffusion, the second summation is advection, and the last term is the decay term. We can see the answer $u = u(x)$ as the density of the chemical dispersed within a fluid occupying Ω , where f is the source of the chemical.

For the remaining part of this section, we take some assumptions. Assume that $b^i = 0$, $c(x) = 0$ and

$$\sum_{i,j=1}^d a^{ij}(x)\xi_i\xi_j \leq \theta (|\xi_1|^2 + \dots + |\xi_n|^2)$$

for all $x \in \Omega$. Then, we have our **uniform ellipticity** condition,

$$(A\xi, \xi) \geq \theta|\xi|^2.$$

We also assume that a^{ij} is symmetric, i.e. $a^{ij}(x) = a^{ji}(x)$ for all $x \in \Omega$. Now, we introduce the main problem:

(MP) *Given that $f \in \mathbf{L}^2(\Omega)$, find $u \in H_0^1(\Omega)$ such that*

$$Lu = - \sum_{i,j=1}^d (a^{ij}(x)u_{x_i})_{x_j} = f. \quad (1.2)$$

Trying to solve this question, multiplying (2) by $v \in H_0^1(\Omega)$ and integrating both sides, we have that

$$- \sum_{i,j=1}^d \int_{\Omega} (a^{ij}(x)u_{x_i})_{x_j} v = \int_{\Omega} f v,$$

and integrating by parts, we can get that

$$\int_{\Omega} (a^{ij}(x)u_{x_i})_{x_j} v = \int_{\Omega} a^{ij}(x)u_{x_i}v_{x_j} + \underbrace{\int_{\partial\Omega} (a^{ij}(x)u_{x_j})x_i v}_{=0}.$$

Therefore, we can find the *weak formulation* for the main problem, (MP).

(WF) Find $u \in H_0^1(\Omega)$ such that

$$B(u, v) = \sum_{i,j=1}^d \int_{\Omega} a^{ij}(x)u_{x_i}v_{x_j} = \int_{\Omega} f v, \quad \forall v \in H_0^1(\Omega)$$

or $B(u, v) = (f, v)_{\mathbf{L}^2(\Omega)}$ for all $v \in H_0^1(\Omega)$.

The goal is to now show that (WF) has a unique solution and therefore, this will imply that (MP) has a unique solution. To this end, using that $a^{ij} \in \mathbf{L}^\infty(\Omega)$,

$$B(u, v) \leq c \|u\|_{H_0^1(\Omega)} \|v\|_{H_0^1(\Omega)}, \quad \forall u, v \in H_0^1(\Omega),$$

and uniform ellipticity, we have that $B(u, v) \geq \theta \|u\|_{H_0^1(\Omega)}^2$ for all $u \in H_0^1(\Omega)$. Now, define

$$F(v) = \int_{\Omega} f v = (f, v)$$

which is a bounded functional on H_0^1 , i.e. this belongs to $(H_0^1(\Omega))^*$. Then, we can see by the Poincaré Inequality that

$$|(f, v)| = |F(v)| = \left| \int_{\Omega} f v \right| \leq \|f\| \|v\|_{\mathbf{L}^2(\Omega)} \leq \|f\| C(\Omega) \|v\|_{H_0^1(\Omega)}. \quad (1.3)$$

Then, using the Riesz Representation Theorem, there must exist a $u_f \in H_0^1(\Omega)$ such that $(u_f, v)_{H_0^1(\Omega)} = F(v)$, which implies that $(\nabla u_f, \nabla v) = (f, v)$ for all $v \in H_0^1(\Omega)$. As well, by using (3), we can have that

$$\|u_f\| = \|F\|_{(H_0^1(\Omega))^*} \leq C(\Omega) \|f\|_{\mathbf{L}^2(\Omega)}.$$

Therefore, (WF) is equivalently asking the following question:

Find $u \in H_0^1(\Omega)$ such that $B(u, v) = (u_f, v)_{H_0^1(\Omega)} = (f, v)$ for all $v \in H_0^1(\Omega)$.

We have now satisfied all of the hypotheses of Lax-Milgram and thus, by that Lax-Milgram theorem, there exists a unique $u \in H_0^1$ such that $B(u, v) = (u_f, v) = (f, v)$ for all $v \in H_0^1$. As well, we have that

$$\|u\|_{H_0^1(\Omega)} \leq \frac{1}{\theta} \|u_f\|_{H_0^1(\Omega)} \leq \frac{C(\Omega)}{\theta} \|f\|_{\mathbf{L}^2(\Omega)}.$$

Now, we define $T : \mathbf{L}^2(\Omega) \mapsto \mathbf{L}^2(\Omega)$, $Tf = u$, the solution of the (WF) problem. This is a continuous function, and we can take

$$\underbrace{f \mapsto u_f}_{\text{compact}} = \underbrace{i^* f \mapsto u = Tf}_{\text{continuous}}.$$

Since this is a compact operator composed with a continuous function, we can then claim that T is compact. As well, since a^{ij} is symmetric, we have that T is self-adjoint, i.e.

$$(Tf, g) = (f, Tg) \quad \forall f, g \in \mathbf{L}^2(\Omega).$$

We claim that T is injective as well. To this end, assume that $Tf = u = 0$ for some $x \in H_0^1(\Omega)$. Then, we have that

$$0 = B(u, v) = (f, v) \quad \forall v \in H_0^1(\Omega)$$

which implies that

$$\int_{\Omega} f v = 0 \quad \forall v \in H_0^1(\Omega).$$

Using the fact that $H_0^1(\Omega)$ is dense in $\mathbf{L}^2(\Omega)$, we must have that $f \equiv 0$ and we have proven injectivity.

Now, $(u, f) = (Tf, f) = B(u, u)$ since by definition $B(u, v) = (f, v)$ for all $v \in H_0^1(\Omega)$. Thus, taking $v = u$, we have $B(u, u) = (f, u) = (f, Tf) = (Tf, f)$. Thus,

$$(Tf, f) = B(u, u) > c \|u\|_{H_0^1(\Omega)}^2 > 0.$$

Thus, by the Hilbert-Schmidt Theorem, there exists $(\lambda_k)_{k \geq 1} \subset (0, \infty)$ that are eigenvalues, $\lambda_k \rightarrow 0$ and there also exist $\phi_k \in \mathbf{L}^2(\Omega)$ such that $T\phi_k = \lambda_k \phi_k$, i.e. $L\phi_k = \lambda_k^{-1} \phi_k$. Therefore, $(\phi_k)_{k \geq 1}$ is an orthonormal basis for $\mathbf{L}^2(\Omega)$ and thus,

$$f = \sum_{k=1}^{\infty} (f, \phi_k) \phi_k$$

for all $f \in \mathbf{L}^2(\Omega)$. Therefore, we have that

$$u = Tf = \sum_{k=1}^{\infty} \lambda_k (f, \phi_k) \phi_k, \tag{1.4}$$

which gives us the solution to the (MP)! Therefore, for any type of these problems, we just need to find ϕ_k and λ_k and we have the answer (this is not as easy as it sounds, but it is still very feasible in most cases).

We now present an example of how to use this. Let $\Omega = (0, \pi)$ and $Lu = -u_{xx} = -u''$. Then, we want to solve

$$Lu = -u''(x) = f(x), \quad x \in \Omega$$

$$u(0) = u(\pi) = 0$$

where $f \in \mathbf{L}^2(\Omega)$. Now, to do this, we must just solve for

$$\begin{aligned} Lu &= \frac{1}{\lambda^2} u \\ -u'' &= \frac{1}{\lambda^2} u. \end{aligned}$$

However, the answer to this ODE is just sines and cosines. Via calculations, one can find that $\lambda = k$ and therefore, $\lambda_k = \frac{1}{k^2}$ for $k \geq 1$ and $\phi_k(x) = c_k \sin(kx)$. Needing that $\|\phi_k\| = 1$, we have that $c_k = \sqrt{\frac{2}{\pi}}$. Thus, $\phi_k(x) = \sqrt{\frac{2}{\pi}} \sin(kx)$. Thus, we have found our eigenvalues and eigenfunctions, and we have that $\lambda_k \rightarrow 0$ as $k \rightarrow \infty$. Therefore, we have that

$$\begin{aligned} u = Tf &= \sum_{k=1}^{\infty} \lambda_k (f \phi_k) \phi_k \\ &= \sum_{k=1}^{\infty} \frac{1}{k^2} \left(\int_0^{\pi} f(x) \frac{2}{\sqrt{\pi}} \sin(kx) dx \right) \sqrt{\frac{2}{\pi}} \sin(kx) \\ &= \frac{2}{\pi} \sum_{k=1}^{\infty} \underbrace{\left(\frac{1}{k^2} \int_0^{\pi} f(x) \sin(kx) dx \right)}_{\text{Fourier Coefficients}} \sin(kx). \end{aligned}$$