MATH 806: Functional Analysis

Dr. David Colton

August 29, 2011

Contents

1 Normed Linear Spaces (Kreyszig 58)

Definition 1.1. *A normed linear space is a vector space X over $\mathbb{R}$ or $\mathbb{C}$ and a function $\|\cdot\|$ from X to $\mathbb{R}$, which satisfies*

(1) $\|x\| \geq 0$, *for all* $x \in X$,

(2) $\|x\| = 0$, *if and only if* $x = 0$,

(3) $\|\alpha x\| = |\alpha|\|x\|$, *for all* $\alpha \in \mathbb{R}(or\ \mathbb{C})$ *and for all* $x \in X$,

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

The norm is a continuous function, i.e. if $x \to y$ (i.e. $\|x-y\| \to 0$), then $\|x\| \to \|y\|$ (Kreyszigr, 60).

Definition 1.2. *A bounded linear transformation (bounded liner operator) from a normed linear space $\langle X_1, \|\cdot\|_1\rangle$ to another normed linear space $\langle X_2, \|\cdot\|_2\rangle$ is a function $T: X_1 \to X_2$, such that*

(1) $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$, *for all* $x, y \in X_1$, *with* $\alpha, \beta \in \mathbb{R}(or\ \mathbb{C})$,

(2) *for some* $c > 0$, $\|Tx\|_2 \leq c\|x\|_1$, *for all* $x \in X_1$.

The smallest c is called the norm of T, written as $\|T\|$. Thus,

$$\|T\| = \sup_{\|x\|_1=1} \|Tx\|_2 = \sup_{\|x\|_1 \neq 0} \frac{\|Tx\|_2}{\|x\|_1}.$$

Example: The space $C([0,1])$, with either the norm $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$ or $\|f\|_1 = \int_0^1 |f(x)|\,dx$ are both normed linear spaces.

Theorem 1.3. *Let T be a linear transformation between two normed linear spaces. Then the following are equivalent:*

(a) *T is continuous,*

(b) *T is continuous at 0,*

(c) *T is bounded.*

Proof. (a) $\Rightarrow$ (b) is trivial. Suppose (b) is true, so that T is continuous at the origin. Then there exists a $\delta > 0$, such that $\|x\| < \delta$ implies $\|Tx\| < 1$, for some $\epsilon > 0$, with $\epsilon = 1$. Then if $\|x\| = 1$, we have $\left\|\frac{1}{2}\delta x\right\| < \delta$. This implies $\|Tx\| = \left\|T\left(\frac{\delta x}{2}\right)\right\| < 1$. Hence, $\|Tx\| < \frac{2}{\delta}$. Therefore, for all x, we have

$$\|Tx\| = \|x\| \left\|T\left(\frac{x}{\|x\|}\right)\right\| < \frac{2}{\delta}\|x\|.$$

This implies (c). Now suppose (c) is true. Let $M = \sup_{\|x\|=1} \|Tx\|$. Then for $x \neq y$, we have

$$\left\|T\left(\frac{x-y}{\|x-y\|}\right)\right\| \leq M.$$

Hence, $\|Tx - Ty\| \leq M\|x-y\|$ so that T is continuous. □

Note: In $C([0,1])$ with either $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$ or $\|f\|_1 = \int_0^1 f(x)\,dx$ the operator

$$(Tf)(x) = \int_0^1 k(x,y)f(y)\,dy,$$

for $k(x,y)$ continuous on $[0,1]\times[0,1]$ defines a bounded linear operator into itself.

Definition 1.4. *If T is a linear operator and $Tx_1 = Tx_2$ implies $x_1 = x_2$, then there exists a linear operator T^{-1}, such that if $Tx = y$, then $x = T^{-1}y$ is called the inverse of T.*

Example: If $Tx = y$, and $\|Tx\| \le c\|x\|$, then $T^{-1}y = x$ and we want solutions such that $\|T^{-1}y\| \le M\|y\|$. Stable solutions occur when T^{-1} is bounded. (In finite dimensions, any linear operator is bounded.)

Definition 1.5. *We say $\langle X, \|\cdot\|\rangle$ is complete if every Cauchy sequence converges to an element of X. Every normed space has a completion, i.e. can be embedded as a dense subset of a complete normed space. (Theorem 2.3-2 of Kreyszigr)*

Example: The completion of $\mathbb{Q}$ is $\mathbb{R}$. The completion of $\langle C([0,1]), \|f\|_1\rangle$ is L (the set of Lebesgue integrable functions).

Theorem 1.6. *Let $C([a,b])$ be the space of continuous functions on the closed interval $[a,b]$ with the norm*

$$\|f\|_\infty = \sup_{a\le x\le b} |f(x)|,$$

then $C([a,b])$ is complete.

Proof. Let $\{f_n\}$ be a Cauchy sequence. Then for any fixed x in $[a,b]$,

$$|f_n(x) - f_m(x)| \le \|f_n(x) - f_m(x)\| \to 0, \qquad \text{as } m,n\to\infty$$

so $\{f_n\}$ is a Cauchy sequence of real numbers. Since the real numbers are complete, for each x there is a number $f(x)$ with $f_n(x) \to f(x)$. Given $\epsilon > 0$, find N such that for $m,n \ge N$, $\|f_n(x) - f_m(x)\|_\infty < \epsilon$. Then

$$\begin{aligned}\sup_{a\le x\le b} |f(x) - f_N(x)| &= \sup_{a\le x\le b} \lim_{n\to\infty} |f_n(x) - f_N(x)| \\ &\le \sup_{a\le x\le b} \sup_{n\ge N} |f_n(x) - f_N(x)| \\ &\le \sup_{a\le x\le b} \sup_{n\ge N} \|f_n(x) - f_N(x)\| \\ &= \sup_{n\ge N} \|f_n(x) - f_N(x)\|_\infty < \epsilon.\end{aligned}$$

Thus, if we can show that $f \in C([a,b])$, we can conclude $\|f - f_n\|_\infty \to 0$ so that $f_n \to f$ in $C([a,b])$. Fix $x \in [a,b]$ and $\epsilon > 0$. We want to find a $\delta > 0$ such that $|x-y| < \delta$ that implies $|f(x) - f(y)| < \epsilon$. Pick n such that $\|f_n - f\| < \frac{\epsilon}{3}$. Then since each f_n is continuous, pick $|x-y| < \delta$, which implies $|f_n(x) - f_n(y)| < \frac{\epsilon}{3}$. Then when $|x-y| < \delta$, we have

$$|f(x) - f(y)| \le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Hence, $f \in C([a,b])$. □

---

2 Hilbert Spaces (Kreyszig 128)

Definition 2.1. *A complex vector space X is called an inner product space if there is a complex valued function $\langle \cdot, \cdot \rangle$ on $X \times X$ that satisfies the following conditions for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{C}$:*

- $\langle x, x \rangle \geq 0$, *and* $\langle x, x \rangle = 0$ *if and only if* $x = 0$,
- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$,
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

The function $\langle \cdot, \cdot \rangle$ is called an inner product. Note that

$$\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle,$$

$$\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle.$$

Example: Let $\mathbb{C}^n$ denote the space of all n-tuples of complex numbers where $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ in $\mathbb{C}^n$. We have $\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$.

Example: Let X denote complex valued continuous functions on $[a, b]$ where for $f, g \in X$, the inner product is given by $\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} \, dx$.

Definition 2.2. *Two elements x, y in an inner product space X are said to be orthogonal if $\langle x, y \rangle = 0$. A collection $\{x_\alpha\}_{\alpha \in A}$ of elements in X is called an orthonormal set if $\langle x_\alpha, x_\alpha \rangle = 1$ for all $\alpha \in A$ and $\langle x_\alpha, x_\beta \rangle = 0$ for all $\alpha \neq \beta \in A$. (Note, A may be uncountable)*

We introduce the shorthand $\|x\| = \sqrt{\langle x, x \rangle}$. We will see shortly that $\|x\|$ is a norm.

Theorem 2.3. *Let $\{x_n\}_{n=1}^N$ be an orthonormal set in an inner product space X. Then for all $x \in X$, we have*

$$\|x\|^2 = \sum_{n=1}^{N} |\langle x_n, x \rangle|^2 + \left\| x - \sum_{n=1}^{N} \langle x_n, x \rangle x_n \right\|^2.$$

Proof. We write x as

$$x = \sum_{n=1}^{N} \langle x_n, x \rangle x_n + \left(x - \sum_{n=1}^{N} \langle x_n, x \rangle x \right).$$

Noting that $\sum_{n=1}^{N} \langle x_n, x \rangle x_n$ and $x - \sum_{n=1}^{N} \langle x_n, x \rangle x_n$ are orthogonal, we have that

$$\begin{aligned} \langle x, x \rangle &= \left\| \sum_{n=1}^{N} \langle x_n, x \rangle x_n \right\|^2 + \left\| x - \sum_{n=1}^{N} \langle x_n, x \rangle x_n \right\|^2 \\ &= \sum_{n=1}^{N} |\langle x_n, x \rangle|^2 + \left\| x - \sum_{n=1}^{N} \langle x_n, x \rangle x_n \right\|^2. \end{aligned}$$

□

2.1 Bessel's Inequality (Kreyszig 157)

Corollary 2.4. *Let $\{x_n\}_{n=1}^N$ be an orthonormal set in an inner product space X. Then for all $x \in X$, we have*

$$\|x\|^2 \geq \sum_{n=1}^N |\langle x_n, x\rangle|^2.$$

2.2 Cauchy-Schwarz Inequality (Kreyszig 14)

Corollary 2.5. *If x and y are elements of an inner product space then*

$$|\langle x, y\rangle| \leq \|x\|\,\|y\|.$$

Proof. The case $y = 0$ is trivial, so suppose $y \neq 0$. The element $\frac{y}{\|y\|}$ by itself forms an orthonormal set so applying Bessel's inequality to any $x \in X$ gives

$$\|x\|^2 \geq \left|\left\langle x, \frac{y}{\|y\|}\right\rangle\right|^2 = \frac{|\langle x, y\rangle|^2}{\|y\|^2} \qquad \Rightarrow \qquad \|x\|\,\|y\| \geq |\langle x, y\rangle|.$$

□

Exercise: For x, y on an inner product space X, show that the parallelogram law

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2,$$

is valid.

Theorem 2.6. *Every inner product space X is a normed linear space with norm $\|x\| = \sqrt{\langle x, x\rangle}$.*

Proof. Since X is a vector space, we need only verify that $\|\cdot\|$ is a norm. All the properties required follow immediately, except the triangle inequality. Suppose $x, y \in X$, then

$$\begin{aligned}
\|x+y\|^2 &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\
&= \langle x, x\rangle + 2\mathrm{Re}\langle x, y\rangle + \langle y, y\rangle \\
&\leq \langle x, x\rangle + 2|\langle x, y\rangle| + \langle y, y\rangle \\
&\leq \langle x, x\rangle + 2\langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2} + \langle y, y\rangle \\
&= (\|x\| + \|y\|)^2,
\end{aligned}$$

by the Schwarz Inequality. □

Definition 2.7. *A complete inner product space is called a Hilbert space.*

Definition 2.8. *Two Hilbert spaces, H_1 and H_2, are said to be isomorphic if there is a linear operator called U from H_1 onto H_2 such that $\langle U_x, U_y\rangle_{H_2} = \langle x, y\rangle_{H_1}$ for all $x, y \in H_1$. Such an operator is called unitary.*

Example: The vector space of all continuous functions on $[0, 1]$ is an inner product space with inner product defined by

$$\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\, dx.$$

Unfortunately, this inner product space is not complete. Consider the sequence $a_n = \frac{1}{2} + \frac{1}{n}$, which is a Cauchy sequence, since for $n > m$ we have that $\|f_n - f_m\|^2 = \int_0^1 |f_n(x) - f_m(x)|^2\, dx = \frac{(n-m)^2}{n^2m^2} \to 0$ as $m, n \to \infty$. However, $\{f_n\}$ does not converge to a continuous function. The completion of this inner product space is a Hilbert space which we denote by $L^2([0, 1])$.

Example: We define ℓ^2 to be the set of sequences $\{x_n\}_{n=1}^{\infty}$ of complex numbers such that

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty,$$

with inner product

$$\langle \{x_n\}, \{y_n\}\rangle = \sum_{n=1}^{\infty} x_n\overline{y_n}.$$

It follows from section 1.2-3 of Kreyszigr that ℓ^2 is a Hilbert space.

3 The Riesz Representation Theorem (Kreyszig 188)

Let M be a closed subspace of a Hilbert space H and let $M^\perp$ be the set of vectors in H which are orthogonal to M. $M^\perp$ is called the orthogonal complement of M. It is easily verified that $M^\perp$ is a closed linear subspace of H and hence $M^\perp$ is also a Hilbert space such that M and $M^\perp$ have only the zero element in common. Consider:

- $\|\langle x_n - x, m\rangle\| \le \|x_n - x\|\,\|m\| \quad \Rightarrow \quad \langle x_n, m\rangle \to \langle x, m\rangle$ if $x_n \to x$.
- $\langle x_n, m\rangle = 0$ for all $m \in M$, $x_n \to x$ implies $\langle x, m\rangle = 0$, which implies M is closed.
- A closed subspace in a Hilbert space H implies that it is complete. That is, if $\|x_n - x_m\| \to 0$, then $x_n \to x \in H$, and $M^\perp$ is closed so that $x \in M^\perp$. (Theorem 1.4-7 in Kreyszigr)

Lemma 3.1. *Let H be a Hilbert space, and let M be a subspace of H. Suppose $x \in H$. Then, there exists in M a unique element z closest to x.*

Proof. Let $d = \inf_{y\in M} \|x - y\|$. Choose a sequence $\{y_n\}$, $y_n \in M$ such that $\|x - y_n\| \to d$. Then, by the parallelogram law, we have

$$\begin{aligned}\|y_n - y_m\|^2 = \|(y_n - x) - (y_m - x)\|^2 &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - \| -2x + y_n + y_m\|^2 \\ &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\left\|x - \frac{1}{2}(y_n + y_m)\right\|^2 \\ &\le 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4d^2 \\ &\to 2d^2 + 2d^2 - 4d^2 = 0,\end{aligned}$$

since $\frac{1}{2}(y_n+y_m) \in M$ as $n, m \to \infty$. Hence, $\{y_n\}$ is a Cauchy sequence, and since M is closed $\{y_n\}$ converges to an element $z \in M$. Therefore, $\|x - z\| = d$. To show uniqueness, assume that $d = \|x - z\| = \|x - z_0\|$. Then by the parallelogram law,

$$\begin{aligned}\|z - z_0\|^2 = \|(z - x) - (z_0 - x)\|^2 &= 2\|z - x\|^2 + 2\|z_0 - x\|^2 - \|(z - x) + (z_0 - x)\|^2 \\ &= 2d^2 + 2d^2 - 4\left\|\frac{1}{2}(z + z_0) - x\right\|^2 \\ &\le 2d^2 + 2d^2 - 4d^2 = 0,\end{aligned}$$

which implies $z = z_0$. □

Theorem 3.2. *(Projection) Let H be a Hilbert space, and M a closed subspace of H, then every x in H can be uniquely written as $x = z + w$, where $z \in M$ and $w \in M^\perp$.*

Proof. Let $x \in H$. Then by the lemma, there exists a unique element $z \in M$ closest to x. Define $w = x - z$. Then clearly, $x = z + w$. We must show $w \in M^\perp$. Let $y \in M$ and $t \in \mathbb{R}$. Then if $d = \|x - z\|$, we have

$$d^2 \le \|x - (z + ty)\|^2 = \|w - ty\|^2 = d^2 - 2t\mathrm{Re}\langle w, y\rangle + t^2\|y\|^2.$$

This implies $-2t\mathrm{Re}\langle w, y\rangle + t^2\|y\|^2 \ge 0$ for all $t \in \mathbb{R}$. Hence, dividing by t, we must have $\mathrm{Re}\langle w, y\rangle = 0$. A similar argument using $i\,t$ instead of t shows that $\mathrm{Im}\langle w, y\rangle = 0$. Hence, $\langle w, y\rangle = 0$. To prove uniqueness, assume $x = z + w = z_0 + w_0$ where $z, z_0 \in M$ and $w, w_0 \in M^\perp$. then

$$z - z_0 = w_0 - w \quad \Rightarrow \quad z - z_0 \in M \cap M^\perp = \{0\}.$$

Hence, $z - z_0 = w_0 - w = 0$ and thus, $z = z_0$ and $w_0 = w$, as desired. □

We introduce the notation, $H = M \oplus M^{\perp}$. The linear space of all bounded linear transformations from a Hilbert space, H, into another Hilbert space $\hat{H}$ is denoted by $\mathcal{L}(H, \hat{H})$.

Definition 3.3. *The space $\mathcal{L}(H, \mathbb{C})$ is called the Dual space of H, and is denoted by H'. The elements of H' are called bounded linear functionals.*

3.1 Riesz Representation Theorem

Theorem 3.4. *For each $f \in H'$ there exists a unique $z \in H$ such that $f(x) = \langle x, z \rangle$, for all $x \in H$. In addition, $\|f\| = \|z\|$.*

Proof. Let N be the set of all $x \in H$ such that $f(x) = 0$. By the continuity of f, N is a closed subspace of H. If $N = H$, then $f(x) = \langle x, 0 \rangle$ for all x and we are done. Hence, assume $N \subset H$. Then by the Projection theorem, there is a non-zero element $z_0 \in N^{\perp}$. Define

$$z = \frac{\overline{f(z_0)} z_0}{\|z_0\|^2}.$$

We will show that z has the right properties. First, if $x \in N$, then $\langle x, z \rangle = 0$, since $z \in N^{\perp}$. Furthermore, if $x = \alpha z_0$, we have

$$f(x) = \alpha f(z_0) = \left\langle \alpha z_0, \frac{\overline{f(z_0)} z_0}{\|z_0\|^2} \right\rangle = \langle x, z \rangle.$$

Since $f(\cdot)$ and $\langle \cdot, z \rangle$ are linear and agree on N and z_0, they must agree on the space spanned by N and z_0. But N and z_0 span the whole space, H, since every element $y \in H$ can be written $y = \left(y - \frac{f(y)}{f(z_0)} z_0 \right) + \frac{f(y)}{f(z_0)} z_0$, and

$$f\left(y - \frac{f(y)}{f(z_0)} z_0 \right) = 0.$$

Thus, $f(x) = \langle x, z \rangle$ for all $x \in H$. Now, assume $f(x) = \langle x, z \rangle = \langle x, z_0 \rangle$. Then

$$\|z_0 - z\|^2 = \langle z_0 - z, z_0 - z \rangle = f(z_0 - z) - f(z_0 - z) = 0.$$

Hence, $z_0 = z$. To prove that $\|f\|_{H'} = \|z\|_H$, consider the definition. By definition, we have

$$\|f\|_{H'} = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\|=1} |\langle x, z \rangle| \leq \sup_{\|x\|=1} \|x\| \, \|z\| = \|z\|.$$

Also, we have

$$\|f\|_{H'} = \sup_{\|x\|=1} |f(x)| \geq \left| f\left(\frac{z}{\|z\|} \right) \right| = \left\langle \frac{z}{\|z\|}, z \right\rangle = \|z\|.$$

$\square$

4 Orthonormal Sets (Kreyszig 152)

Definition 4.1. *If S is an orthonormal set in a Hilbert space H and no other orthonormal set contains S as a proper subset, then S is called an orthonormal basis or a complete orthonormal set for H.*

Theorem 4.2. *Every Hilbert space $H \neq \{0\}$ has an orthonormal basis.*

Proof. Consider the collection S of orthonormal sets in H. We order S by inclusion. That is, we say $S_1 \leq S_2$ if $S_1 \subseteq S_2$. With this definition, S is partially ordered. There may be sets such that neither $S_1 \leq S_2$ or $S_2 \leq S_1$ holds. It is also non-empty since if $v \in H$, $v \neq 0$, then $\left\{\frac{v}{\|v\|}\right\}$ is an orthonormal set. Now, let $\{S_\alpha\}$ be any totally ordered subset of S (i.e. $S_{\alpha_1}, S_{\alpha_2}$ in $\{S_\alpha\}$ implies $S_{\alpha_1} \leq S_{\alpha_2}$ or $S_{\alpha_2} \leq S_{\alpha_1}$). Then $\cup_{\alpha\in A} S_\alpha$ is an orthonormal set which contains each S_α and is thus an upper bound for $\{S_\alpha\}_{\alpha\in A}$. Hence, by Zorn's lemma (4.1-6 in Kreyszigr), S has a maximal element (i.e. an orthonormal set not properly contained in any other orthonormal system). □

Remark: From now on we'll always assume $H \neq \{0\}$.

Theorem 4.3. *Let H be a Hilbert space, and $S = \{\phi_\alpha\}_{\alpha\in A}$ an orthonormal basis. Then for each $y \in H$,*

$$y = \sum_{\alpha\in A} \langle y, \phi_\alpha\rangle \phi_\alpha \qquad \text{and} \qquad \|y\|^2 = \sum_{\alpha\in A} |\langle y, \phi_\alpha\rangle|^2.$$

The sum on the right side converges (independent of order) to $y \in H$.

Proof. Bessel's inequality implies that for any finite subset $A' \subset A$,

$$\sum_{\alpha\in A} |\langle y, \phi_\alpha\rangle|^2 \leq \|y\|^2.$$

For any integer m, $|\langle y, \phi_\alpha\rangle| > \frac{1}{m}$ for a finite number of α. This implies $\langle y, \phi_\alpha\rangle \neq 0$ for at most a countable number of α's in A, which we order in some way, $\alpha_1, \alpha_2, \ldots$. Furthermore, since

$$\sum_{j=1}^{N} |\langle y, \phi_\alpha\rangle|^2$$

is monotone increasing and bounded it converges to a finite limit as $N \to \infty$. Let $y_n = \sum_{j=1}^n \langle y, \phi_{\alpha_j}\rangle \phi_{\alpha_j}$. Then for $n > m$, we have

$$\|y_m - y_n\|^2 = \left\| \sum_{j=m+1}^{n} \langle y, \phi_{\alpha_j}\rangle \phi_{\alpha_j} \right\|^2 = \sum_{j=m+1}^{n} |\langle y, \phi_{\alpha_j}\rangle|^2.$$

Therefore, $\{y_n\}$ is a Cauchy sequence and converges to an element $y' \in H$. But

$$\langle y - y', \phi_{\alpha_\ell}\rangle = \lim_{n\to\infty} \left\langle y - \sum_{j=1}^{n} \langle y, \phi_{\alpha_j}\rangle \phi_{\alpha_j}, \phi_{\alpha_\ell} \right\rangle = \langle y, \phi_{\alpha_\ell}\rangle - \langle y, \phi_{\alpha_\ell}\rangle = 0,$$

and if $\alpha \neq \alpha_\ell$ for some ℓ we have (since $\langle y, \phi_\alpha \rangle = 0$ and $\langle \phi_{\alpha_\ell}, \phi_\alpha \rangle = 0$),

$$\langle y - y', \phi_\alpha \rangle = \lim_{n\to\infty} \left\langle y - \sum_{j=1}^{n} \langle y, \phi_{\alpha_j} \rangle \phi_{\alpha_j}, \phi_\alpha \right\rangle = 0.$$

Hence, $y - y'$ is orthonormal to all the ϕ_α's in S and since S is a complete orthonormal set, we must have that $y - y' = 0$. (Otherwise adding $\frac{y-y'}{\|y-y'\|}$ to S would give us a bigger orthonormal set, which contradicts the fact that we had a maximal element.) Therefore, $y = \lim_{n\to\infty} \sum_{j=1}^{n} \langle y, \phi_j \rangle \phi_j$. Furthermore,

$$\begin{aligned} 0 &= \lim_{n\to\infty} \left\| y - \sum_{j=1}^{n} \langle y, \phi_{\alpha_j} \rangle \phi_{\alpha_j} \right\|^2 \\ &= \lim_{n\to\infty} \left(\|y\|^2 - \sum_{j=1}^{n} |\langle y, \phi_{\alpha_j} \rangle|^2 \right) \\ &= \|y\|^2 - \sum_{\alpha\in A} |\langle y, \phi_\alpha \rangle|^2, \end{aligned}$$

since the norm is continuous. □

4.1 Parseval's Relation (Kreyszig 170)

Definition 4.4. *The above equality is called Parseval's relation and $\langle \phi_\alpha, y \rangle$ are called the Fourier coefficients.*

Definition 4.5. *A normed linear space X which has a countable dense subset M (i.e. $\overline{M} = X$) is said to be separable.*

Theorem 4.6. *A Hilbert space H is separable if and only if it has a countable orthonormal basis S. If there are $N < \infty$ elements of S, then H is isomorphic to $\mathbb{C}^N$. If there are countably many elements in S, then H is isomorphic to ℓ^2.*

Proof. ($\Rightarrow$) Suppose H is separable and let $\{x_n\}$ be a countable dense set. By throwing out some of the x_n's, we can get a sub-collection of independent vectors whose span is the same as the set $\{x_n\}$ and is thus dense. Applying the Gram-Schmidt procedure (Kreyszigr 157) to this sub-collection, we obtain a countable complete orthonormal set.
($\Leftarrow$) Conversely, if $\{\phi_n\}$ is a complete orthonormal set for a Hilbert space H, then it follows by Theorem 4.3, that the set of finite linear combinations of the ϕ_n with rational coefficients is dense in H. Since this set is countable, H is countable. This completes the first part of the proof.
Suppose H is separable and let $\{\phi_n\}_{n=1}^{\infty}$ be a complete orthonormal set. We define a map $U : H \to \ell^2$ by

$$U : x \to \{\langle x, \phi_n \rangle\}_{n=1}^{\infty}.$$

Theorem 4.3 implies that this map is well defined and onto (i.e. $x = \sum a_n \phi_n$ implies $\langle x, \phi_n \rangle = \langle \sum a_n \phi_n, \phi_n \rangle = a_n$, and $\sum |a_n|^2 < \infty$. This implies onto.) We need to show that U is unitary. We first note the polarization identity:

$$\begin{aligned} \mathrm{Re}\langle x, y \rangle &= \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) \\ \mathrm{Im}\langle x, y \rangle &= \frac{1}{4}(\|x + iy\|^2 - \|x - iy\|^2). \end{aligned}$$

Then from Theorem 4.3, we have that

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, \phi_n \rangle|^2 = \|Ux\|^2,$$

and by the polarization identities, U preserves the inner product.
The proof that H is isomorphic to $\mathbb{C}^N$ if S has N elements is similar. $\square$

5 Banach Spaces (Kreyszig 58)

Definition 5.1. *A complete normed linear space is called a Banach space.*

Example: Every Hilbert space is a Banach space with respect to the norm $\|x\| = \sqrt{\langle x, x\rangle}$.

Example: For $p > 1$ the Banach space $L^p[a,b]$ is the completion of the normed space which consists of all continuous functions on $[a,b]$ with norm

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, dx \right)^{1/p}.$$

The fact that $\|f\|_p$ is a norm follows from the Minkowski Inequality (Kreyszigr 14), which shows

$$\|f+g\|_p \le \|f\|_p + \|g\|_p.$$

5.1 Important Banach Spaces

Example: Let $a = \{a_n\}_{n=1}^\infty$ be a sequence of complex numbers and define

$$\ell^\infty = \{a : \|a\|_\infty = \sup_n |a_n| < \infty\}$$
$$c = \{a : \lim_{n\to\infty} a_n < \infty\}$$
$$\ell^p = \left\{ a : \|a\|_p = \left(\sum_{n=1}^\infty |a_n|^p \right)^{1/p} < \infty \right\}, \qquad 1 \le p < \infty.$$

The space ℓ^∞ and c are Banach spaces with the $\|\cdot\|_\infty$ norm (Sections 1.5-2 and 1.5-3 of Kreyszigr) and ℓ^p is a Banach space with the $\|\cdot\|_p$ norm (Section 1.5-4 of Kreyszigr).

The set of all bounded linear operators from one normed linear space X to another normed linear space Y is denoted $B(X,Y)$. We can define a norm on $B(X,Y)$ by defining

$$\|A\| = \sup_{x\in X, x\ne 0} \frac{\|Ax\|_Y}{\|x\|_X}.$$

This norm is called the operator norm.

Theorem 5.2. *If Y is complete, then $B(X,Y)$ is a Banach space.*

Proof. Since any finite linear combination of bounded operators is again a bounded operator, $B(X,Y)$ is a vector space. It is also easy to see that $\|\cdot\|$ is a norm. For example, the triangle inequality is proved by the following computation

$$\|A+B\| = \sup_{x\ne 0} \frac{\|(A+B)x\|}{\|x\|} \le \sup_{x\ne 0} \frac{\|Ax\| + \|Bx\|}{\|x\|} \le \sup_{x\ne 0} \frac{\|Ax\|}{\|x\|} + \frac{\|Bx\|}{\|x\|} = \|A\| + \|B\|.$$

To show completeness of $B(X,Y)$, we must show that if $\{A_n\}_{n=1}^\infty$ is a Cauchy sequence in the operator norm, then there is a bounded linear operator A such that $\|A_n - A\| \to 0$. Let $\{A_n\}_{n=1}^\infty$ be a Cauchy sequence. Then for all $x \in X$, $\{A_n x\}_{n=1}^\infty$ is a Cauchy sequence in Y, because

$$\|A_n x - A_m x\| \le \|A_n - A_m\| \, \|x\|,$$

for $n, m \geq N$, for some N large enough. Since Y is complete, $\{A_n x\}_{n=1}^{\infty}$ converges to an element $y \in Y$ so define $Ax = y$. It is easy to check that A is a linear operator. From the triangle inequality, it follows

$$| \|A_n\| - \|A_m\| | \leq \|A_n - A_m\|,$$

which implies $\{\|A_n\|\}_{n=1}^{\infty}$ is a Cauchy sequence of real numbers converging to a real number c. Thus

$$\|Ax\|_Y = \lim_{n\to\infty} \|A_n x\|_Y \leq \lim_{n\to\infty} \|A_n\| \, \|x\|_X = c\|x\|_X,$$

which implies A is a bounded linear operator. We must still show that $A_n \to A$, in the operator norm. Since $\|(A - A_m)x\| = \lim_{n\to\infty} \|(A_m - A_n)x\|$, we have

$$\frac{\|(A - A_n)x\|}{\|x\|} \leq \lim_{m\to\infty} \|A_m - A_n\|,$$

which implies

$$\|A - A_m\| = \sup_{x\neq 0} \frac{\|(A - A_m)x\|}{\|x\|} \leq \lim_{m\to\infty} \|A_m - A_n\|,$$

which is arbitrarily small for large enough n. □

Definition 5.3. *A bounded linear operator from a normed linear space X onto a normed linear space Y is called an isomorphism if it is norm preserving.*

Example: All separable infinite dimensional Hilbert spaces are isomorphic to ℓ^2 (Theorem 4.6).

Definition 5.4. *Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a normed linear space X are called equivalent if there are positive constants a and b such that for all $x \in X$,*

$$a\|x\|_1 \leq \|x\|_2 \leq b\|x\|_1.$$

Theorem 5.5. *On a finite dimensional vector space, all norms are equivalent.*

Proof. Theorem 2.4-5 of Kreyszigr. □

---

6 The Dual Space (Kreyszig 120)

In the previous section we proved that if X is a normed linear space and Y is a Banach space, then the set $B(X,Y)$ of bounded linear transformations from X to Y is also a Banach space. In the special case when $Y = \mathbb{C}$, the space $B(X,Y)$ is denoted X' and is called the Dual space of X. The elements of X' are called bounded linear functionals on X. Convergence in X' is convergence with respect to the operator norm, i.e. if $f \in X'$, then

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

Example: Suppose $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(\mathbb{R})$ (c.f Kreyszigr 62), then Holder's inequality (Kreyszigr 14) says that

$$\|fg\|_1 \leq \|f\|_p \|g\|_q, \qquad \text{where} \qquad \|f\|_p = \left(\int_{-\infty}^{\infty} |f(x)|^p\right)^{1/p},$$

and hence

$$\int_{-\infty}^{\infty} f(x)\overline{g(x)}\, dx,$$

makes sense. Let $g \in L^q(\mathbb{R})$ be fixed and define

$$G(f) = \int_{-\infty}^{\infty} f(x)\overline{g(x)}\, dx,$$

for each $f \in L^p(\mathbb{R})$. By Holder's inequality, $G(\cdot)$ is a bounded linear functional on $L^p(\mathbb{R})$ with norm less than or equal to $\|g\|_q$. It can be shown that the norm is equal to $\|g\|_q$ and every bounded linear functional on L^p is of the form $G(\cdot)$ for some $g \in L^q(\mathbb{R})$.

Example: For $1 < p < \infty$, ℓ^q is the dual space of ℓ^p, where $\frac{1}{p} + \frac{1}{q} = 1$ (2.7-10 of Kreyszigr).

Since the dual of X' of a Banach space is itself a Banach space, it also has a dual space, denoted X'' called the second dual of X. We will show that X can be regarded in a natural way as a subset of X''.

6.1 The Second Dual Space (Kreyszig 239)

Theorem 6.1. *Let X be a Banach space. For each $x \in X$, let $g_x(\cdot)$ be the linear functional on X' which assigns to each $f \in X'$, the number $g_x(f) = f(x)$. Then the map $C : x \mapsto g_x$ is an isomorphism of X onto a (possibly proper) subspace of X''.*

Proof. Since $|g_x(f)| = |f(x)| \leq \|f\|_X \|x\|_X$ and

$$g_x(\alpha f + \beta h) = \alpha f(x) + \beta h(x) = \alpha g_x(f) + \beta g_x(h),$$

for any $f, h \in X'$, we have that g_x is a bounded linear functional on X', such that

$$\|g_x\| \leq \|x\|_X.$$

By a corollary to the Hahn-Banach theorem (which we will prove soon), given x, we can find $f \in X'$ such that $\|f\|_{X'} = 1$, and $f(x) = \|x\|_X$. This implies

$$\|g_x\|_{X''} = \sup_{f\in X', \|f\|=1} |g_x(f)| \geq \|x\|_X.$$

Hence, $\|g_x\|_{X''} = \|x\|_X$. Therefore, C is an isomorphism of X onto its range. □

Definition 6.2. *If the map C is surjective, then X is said to be reflexive.*

Theorem 6.3. *(Hahn-Banach, Real) Let X be a real vector space, P is a real-valued function defined on X satisfying $P(x+y) \leq P(x) + P(y)$ and $P(\alpha x) = \alpha P(x)$ for all $x, y \in X$ and $\alpha \geq 0$. Suppose that f is a linear functional defined in a subspace $Y \subset X$, which satisfies $f(x) \leq P(x)$ for all $x \in Y$. Then there is a linear function $\hat{f}$ defined on X satisfying $\hat{f}(x) \leq P(x)$ for all $x \in X$ and such that $\hat{f}(x) = f(x)$ for all $x \in Y$.*

Proof. The idea is the following: First, we will show that if $z \in X$ but $z \notin Y$, then we can extend f to a functional having the right properties spanned by z and Y. We then use Zorn's lemma to show that the process can be continued to extend f to the whole space.
Let $\hat{Y}$ denote the subspace spanned by Y and z. The extension of f to $\hat{Y}$, call it $\hat{f}$, was specified as soon as we defined $\hat{f}(z)$, since for $y \in Y$, we have

$$\hat{f}(cz + y) = c\hat{f}(z) + f(y).$$

Suppose that $y_1, y_2 \in Y$, $\alpha, \beta > 0$, then

$$\begin{aligned}
\beta f(y_1) + \alpha f(y_2) &= f(\beta y_1 + \alpha y_2) \\
&= (\alpha+\beta) f\left(\frac{\beta}{\alpha+\beta} y_1 + \frac{\alpha}{\alpha+\beta} y_2\right) \\
&\leq (\alpha+\beta) P\left(\frac{\beta}{\alpha+\beta}(y_1 - \alpha z) + \frac{\alpha}{\alpha+\beta}(y_2 + \beta z)\right) \\
&\leq \beta P(y_1 - \alpha z) + \alpha P(y_2 + \beta z).
\end{aligned}$$

Hence, for $\alpha, \beta > 0$ and $y_1, y_2 \in Y$, we have the following

$$\frac{1}{\alpha}\left[-P(y_1 - \alpha z) + f(y_1)\right] \leq \frac{1}{\beta}\left[P(y_2 + \beta z) - f(y_2)\right].$$

We can therefore find a real number a such that

$$\sup_{y \in Y, \alpha > 0} \left[\frac{1}{\alpha}\left(-P(y - \alpha z) + f(y)\right)\right] \leq a \leq \inf_{y \in Y, \alpha > 0} \left[\frac{1}{\alpha}\left(P(y + \beta z) - f(y_2)\right)\right].$$

We now define $\hat{f}(z) = a$. It can be easily verified that $\hat{f}(x) \leq P(x)$, for all $y \in Y$:

$$\hat{f}(cz + y) = c\hat{f}(z) + f(y) = ca + f(y).$$

Assume $c > 0$ ($c < 0$ can be done similarly), then

$$\begin{aligned}
ca + f(y) &\leq \frac{c}{\alpha}\left(P(y_1 + \beta z) - f(y_1)\right)_+ f(y) \qquad \text{for all } y_1 \in Y \\
&= P\left(\frac{c}{\alpha} y_1 + cz\right) - \frac{c}{\alpha} f(y_1) + f(y) \\
&= P(y + cz),
\end{aligned}$$

for $y_1 = \frac{\alpha}{c} y$. This shows that f can be extended one dimension at a time. By Zorn's lemma we can continue this process and extend f to all of X (see a and b Kreyszigr 214). ☐

Theorem 6.4. *(Hahn-Banach, Complex) Let X be a complex vector space, P a real-valued function defined on X satisfying $P(x+y) \le P(x)+P(y)$ and $P(\alpha x) = |\alpha|P(x)$ for all $x, y \in X$ and for any α. Suppose that f is a linear functional defined in a subspace $Y \subset X$, which satisfies $|f(x)| \le P(x)$ for all $x \in Y$. Then there is a linear function $\hat{f}$ defined on X satisfying $|\hat{f}(x)| \le P(x)$ for all $x \in X$ and such that $\hat{f}(x) = f(x)$ for all $x \in Y$.*

Proof. Let $\ell(x) = \mathrm{Re}\{f(x)\}$. Then ℓ is a real linear functional on Y (i.e. ℓ is linear over the vector space consisting of elements of Y and real scalars), and since $\ell(ix) = \mathrm{Re}\{f(ix)\}$, we have

$$\mathrm{Re}\{f(ix)\} = \Re\{if(x)\} = -\Im\{f(x)\}.$$

We see that $f(x) = \ell(x) - i\ell(ix)$. Since ℓ is linear on the vector space Y with scalar multiplication restricted to the real numbers, ℓ has a real linear extension to all of X such that $L(x) \le P(x)$ (by Theorem 6.3), where $L(x)$ is real valued. And now define

$$\hat{f}(x) = L(x) - iL(ix).$$

Then $\hat{f}(x)$ clearly extends $f(x)$ and is real linear. Moreover,

$$\hat{f}(ix) = L(ix) - iL(-x) = L(ix) + iL(x) = i\hat{f}(x),$$

so that $\hat{f}$ is complex linear. To complete the proof, we need only show that $|\hat{f}(x)| \le P(x)$. To this end, first note that $P(\alpha x) = P(x)$ if $|\alpha| = 1$. If we set $\theta = \arg\{\hat{f}(x)\}$ and use the fact that $\Re\{\hat{f}(x)\} = L(x)$, we see that

$$\begin{aligned} |\hat{f}(x)| &= e^{-i\theta}\hat{f}(x) && \text{since } \hat{f}(x) = |\hat{f}(x)|e^{i\theta} \\ &= \hat{f}(e^{-i\theta}x) \\ &= L(e^{-i\theta}x) && \text{since } \hat{f}(e^{-i\theta}x) \text{ is real} \\ &\le P(e^{-i\theta}x) \\ &= P(x), \end{aligned}$$

since $|e^{-i\theta}| = 1$. □

6.2 Hahn-Banach Corollaries

Corollary 6.5. *Let X be a normed space, Y a subspace of X and f an element of Y'. Then there exists an extension $\hat{f} \in X'$ extending f and satisfying $\|\hat{f}\|_{X'} = \|f\|_{Y'}$.*

Proof. Choose $P(x) = \|f\|_{Y'}\|x\|$ and apply the Hahn-Banach theorem. □

Corollary 6.6. *Let $x_0 \neq 0$ be an element of a normed linear space X. Then there exists $\hat{f} \in X'$ (bounded linear functional) such that $\|\hat{f}\| = 1$ and $\hat{f}(x_0) = \|x_0\|$.*

Proof. Let Y be the subspace consisting of all scalar multiples of x_0 and define $f(ax_0) = a\|x_0\|$ as a linear functional. By Corollary 6.5, we can construct $\hat{f}$ with $\|\hat{f}\| = \|f\|$ extending f to all of X, but $\|f\| = \|\hat{f}\| = 1$ and $f(x_0) = \hat{f}(x_0) = \|x_0\|$. □

Corollary 6.7. *For every x in a normed linear space X we have that*

$$\|x\| = \sup_{f\in X', f\neq 0} \frac{|f(x)|}{\|f\|}$$

Hence, if $f(x_0) = 0$ for all $f \in X'$, then $x_0 = 0$.

Proof. Using the $\hat{f}$ from Corollary 6.6, we have, writing x for x_0,

$$\sup_{f\in X'} \frac{|f(x)|}{\|f\|} \geq \frac{|\hat{f}(x)|}{\|\hat{f}\|} = \frac{\|x\|}{1} = \|x\|,$$

and from $|f(x)| \leq \|f\|\,\|x\|$ it follows

$$\sup_{f\neq 0} \frac{|f(x)|}{\|f\|} \leq \|x\|.$$

□

Corollary 6.8. *Let Y be a proper closed subspace of a normed linear space X and suppose x_0 is an element of X, whose distance from Y is d. Then there exists $f \in X'$ such that $\|f\| = 1$, $f(x_0) = d$, and $f(y) = 0$ for all $y \in Y$.*

Proof. Lemma 4.6-7 of Kreyszigr. □

Theorem 6.9. *Let X be a normed linear space. If X' is separable, then X is separable.*

Proof. Let $\{f_n\}$ be a dense set in X'. Choose $\{x_n\}$ in X, $\|x_n\| = 1$, such that $\|f_n(x_n)\| \geq \frac{1}{2}\|f_n\|$. Let D be the set of all finite linear combinations of the $\{x_n\}$ with rational coefficients. Since D is countable, it is sufficient to show that D is dense in X. Suppose D is not dense in X, then there exists a $y \in X \setminus \overline{D}$, such that

$$\inf_{x\in\overline{D}} \|y - x\| > 0 \qquad \text{(since } \overline{D} \text{ is closed)},$$

and a linear functional $f \in X'$ such that $f(y) \neq 0$, but $f(x) = 0$ for all $x \in \overline{D}$ by Corollary 6.8. Let $\{f_{n_k}\}$ be a subsequence of $\{f_n\}$ which converges to f. Then

$$\begin{aligned} 0 \longleftarrow \|x_{n_k}\|\,\|f - f_{n_k}\| &\geq |(f - f_{n_k})(x_{n_k})| \\ &= |f_{n_k}(x_{n_k})| \qquad \text{(since } f \text{ vanishes for all } x_{n_k} \in \overline{D}) \\ &\geq \frac{1}{2}\|f_{n_k}\|. \end{aligned}$$

Hence, $\|f_{n_k}\| \to 0$ as $k \to \infty$. This implies $f = 0$, which is a contradiction since $f(y) \neq 0$. Thus, $\overline{D}$ is dense in X, and so X is separable. □

---

7 The Baire Category Theorem and Its Applications (Kreyszig 247)

Definition 7.1. *A set S in a metric space M is called nowhere dense if $\overline{S}$ has no interior points.*

Theorem 7.2. *(Baire Cetagory) A complete metric space is never the union of a countable number of nowhere dense sets.*

Proof. Suppose X is a complete metric space and assume $X = \cup_{n=1}^{\infty} M_n$ with each M_n a nowhere dense space. We will construct a Cauchy sequence $\{x_n\}$ which stays away from each M_n so that its limit point x (which is in X by completion) is in no way a part of any M_n, thereby contradicting the statement that $X = \cup_{n=1}^{\infty} M_n$.
Since M_1 is nowhere dense, $\overline{M}_1$ does not contain a non-empty open set, but X does (i.e. X itself). This implies, $M_1 \neq X$. Hence, we can find $x_1 \notin \overline{M}_1$. Pick an open ball B_1 about x_1 such that $B_1 \cap M_1 = \emptyset$ and so that the radius of B_1 is smaller than 1. Since, M_2 is nowhere dense, we can find $x_2 \in B_1 \setminus \overline{M}_2$. Let B_2 be an open ball around x_2 such that $\overline{B}_2 \subset B_1$ (nested balls), and $B_2 \cap M_2 = \emptyset$ and B_2 has a radius smaller than $\frac{1}{2}$. Proceeding inductively, we pick $x_n \in B_{n-1} \setminus \overline{M}_n$ and choose an open ball B_n about x_n such that $\overline{B}_n \subset B_{n-1}$, and $B_n \cap M_n = \emptyset$ and B_n has radius smaller than 2^{1-n}. Now $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence since $n, m \geq N$ implies $x_n, x_m \in B_N$ which implies

$$d(x_n, x_m) < 2^{1-N} + 2^{1-N} = 2^{2-N} \to 0 \qquad \text{(diameters)},$$

as $N \to \infty$. Hence, $\{x_n\} \to x$ for some $x \in X$, by completeness. That is, let $x = \lim_{n\to\infty}\{x_n\}$. Then for each $N > 1$, $x_n \in B_n$, for $n > N$, we have that $x \in \overline{B}_N \subset B_{N-1}$. So, $x \notin M_{N-1}$ for any N, which is a contradiction. □

7.1 Uniform Boundedness Theorem (Kreyszig 249)

Theorem 7.3. *Let X be a Banach space. Let $\mathcal{F}$ be a family of bounded linear transformations from X to some normed linear space Y. Suppose that for each $x \in X$, $\{\|Tx\| : T \in \mathcal{F}\}$, is bounded. Then $\{\|T\| : T \in \mathcal{F}\}$ is bounded.*

Proof. Let $A_n := \{x : \|Tx\| \leq n \text{ for all } T \in \mathcal{F}\}$. By hypothesis, each x is in some A_n. That is, $X = \cup_{n=1}^{\infty} A_n$. Moreover, each A_n is closed (since T is continuous). By the Baire Category theorem, some A_n has a non-empty interior, i.e. $B_0 = B(x_0; r) \subseteq A_{n_0}$ for some n_0, x_0, and r. Let $x \in X$, $x \neq 0$, and set $z = x_0 + \gamma x$, where $\gamma = \frac{r}{2\|x\|}$. Then $\|z - x_0\| < r$ which implies $z \in B_0$ which gives $\|Tz\| \leq n_0$ for all $T \in \mathcal{F}$. We also have $\|Tx_0\| \leq n_0$ since $x_0 \in B_0$. Hence,

$$\|Tx\| = \frac{1}{\gamma}\|T(z - x_0)\| \leq \frac{1}{\gamma}(\|Tz\| + \|Tx_0\|) \leq \frac{1}{\gamma}(n_0 + n_0) = \frac{2n_0}{\gamma} = \frac{4n_0}{r}\|x\|.$$

□

"Recall" from MATH 616 that the Fourier Series of a given periodic function $x(t)$ of period 2π is of the form

$$\frac{1}{2}a_0 + \sum_{m=1}^{\infty} \left(a_m \cos mt + b_m \sin mt\right), \tag{7.1}$$

where

$$a_m = \frac{1}{\pi}\int_0^{2\pi} x(t)\, \cos\, mt\, dt, \qquad b_m = \frac{1}{\pi}\int_0^{2\pi} x(t)\, \sin\, mt\, dt.$$

It is well known that continuity and the existence of the right hand and left hand derivatives at a point t_0 is sufficient for convergence of (**??**) at $t = t_0$. The following theorem shows that continuity alone is not enough for convergence.

Theorem 7.4. *There exists real valued continuous functions whose Fourier Series diverge at a given point* t_0.

Proof. Let X be the normed space of all real valued continuous functions of period 2π with norm $\|x\| = \max |x(t)|$. It follows that X is a Banach Space. Without loss of generality, let $t_0 = 0$. Let $f_n(x)$ be the value at $t = 0$ of the n^{th} partial sum of the Fourier Series of $x(t)$. That is,

$$\begin{aligned} f_n(x) &= \frac{1}{2}a_0 t + \sum_{m=1}^{n} a_m \\ &= \frac{1}{\pi}\int_0^{2\pi} x(t)\left[\frac{1}{2} + \sum_{m=1}^{n} \cos mt\right] dt \\ &= \frac{1}{2\pi}\int_0^{2\pi} x(t) q_n(t)\, dt, \end{aligned}$$

where $q_n(t) = \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{1}{2}t}$ (Kreyszig 252). From

$$|f_n(t)| \le \frac{1}{2\pi}\max|x(t)|\int_0^{2\pi} |q_n(t)|\, dt = \frac{\|x\|}{2\pi}\int_0^{2\pi} |q_n(t)|\, dt,$$

we see that f_n is a bounded linear functional on X such that

$$\|f_n\| \le \frac{1}{2\pi}|\int_0^{2\pi} |q_n(t)|\, dt.$$

We first show that the above inequality is in fact an equality. Define

$$y(t) = \begin{cases} 1 & \text{if} \quad q_n(t) \ge 0 \\ -1 & \text{if} \quad q_n(t) < 0 \end{cases}.$$

Let $x(t)$ be be a continuous function of norm one, such that $\epsilon > 0$ implies

$$\frac{1}{2\pi}\left|\int_0^{2\pi} [x(t) - y(t)]\, q_n(t)\, dt\right| \le \epsilon.$$

The we have,

$$\frac{1}{2\pi}\left|\int_0^{2\pi} x(t)q_n(t)\, dt - \int_0^{2\pi} y(t)q_n(t)\, dt\right| = \left|f_n(x) - \frac{1}{2\pi}\int_0^{2\pi} |q_n(t)|\, dt\right| \le \epsilon.$$

The absolute value in the second expression comes from the definition of $y(t)$. Since $\epsilon > 0$ was arbitrary, and $\|x\| = 1$, we can conclude that

$$\|f_n\| = \frac{1}{2\pi}\int_0^{2\pi} |q_n(t)|\, dt.$$

We next show that the sequence $\{f_n(t)\}$ is unbounded. Using the fact that

$$\left|\sin\frac{1}{2}t\right| < \frac{1}{2}t, \qquad \text{for } t \in (0, 2\pi],$$

we see that

$$\begin{aligned}
\|f_n\| &= \frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{1}{2}t}\right| dt \\
&> \frac{1}{\pi}\int_0^{2\pi}\left|\frac{\sin\left(n+\frac{1}{2}\right)t}{t}\right| dt \\
&= \frac{1}{\pi}\int_0^{(2n+1)\pi}\frac{|\sin v|}{v}\, dv \\
&= \frac{1}{\pi}\sum_{k=0}^{2n}\int_{k\pi}^{(k+1)\pi}\frac{|\sin v|}{v}\, dv \\
&\geq \frac{1}{\pi}\sum_{k=0}^{2n}\frac{1}{(k+1)\pi}\int_{k\pi}^{(k+1)\pi}|\sin v|\, dv \\
&= \frac{2}{\pi^2}\sum_{k=0}^{2n}\frac{1}{k+1} \to \infty,
\end{aligned}$$

as $n \to \infty$, which implies $\{\|f_n\|\}$ is unbounded. Since X is a Banach space, by the Uniform Boundedness Principle, we can conclude that there must be $x \in X$ such that $\{|f_n(x)|\}$ is unbounded. Here, the Fourier Series for this x diverges at $t = 0$.

□

7.2 Open Mapping Theorem (Kreyszig 286)

Theorem 7.5. *Let $T : X \to Y$ be a bounded linear transformation of the Banach space X onto the Banach space Y. Then if M is an open set in X, $T[M]$ is open in Y.*

Proof. We first make a series of remarks which will simplify the proof. First, we need only show that for every neighborhood N of $x \in X$, $T[N]$ is a neighborhood of Tx. Since $T[x+N] = Tx + T[N]$, we need only show this for $x = 0$. Secondly, since neighborhoods contain balls, it is sufficient to show that for r sufficiently small $T[B_r^X] \supset B_{r'}^Y$ for some r' where

$$B_r^X = \{x \in X : \|x\| < r\}.$$

Since $T[B_r^X] = rT[B_1^X]$, we need only show that $T[B_r^X]$ is a neighborhood of zero for some r. Finally, it is sufficient to show that $T[B_r^X]$ has a non-empty interior for some r. To see this, suppose $T[B_r^X]$ contains the ball $\{y : \|y - y_0\| < \epsilon\}$ where $Tx_0 = y_0$. Then if $\|x\| < r_0 = \frac{\epsilon}{\|T\|}$ we have that $T(x_0 + x)$ is a neighborhood of y_0, i.e. $T[B_{r_0}^X]$ is a neighborhood of zero. We now proceed to the proof of the theorem. Since T is onto, setting $B_n = B_n^X$, we have that

$$Y = \cup_{n=1}^{\infty} T[B_n].$$

This implies $\overline{T[B_n]} = n\overline{T[B_1]}$ has a nonempty interior for some $n = n_0$ by the Baire Category theorem. That implies $\overline{T[B_1]}$ has a non-empty interior. By scaling and translating we can suppose that B_ϵ for some $\epsilon > 0$ is contained in $\overline{T[B_1]}$. We will show that $\overline{T[B_1]} \subset T[B_2]$ which will complete the proof (since $T[B_2]$ will contain a non-empty interior). Let $y \in \overline{T[B_1]}$, choose $x \in B_1$ such that $y - Tx_1 \in B_{\epsilon/2} \subset \overline{T[B_{1/2}]}$. Now choose x_2 in $B_{1/2}$ such that $y - Tx_1 - Tx_2 \in B_{\epsilon/4}$. By induction, choose $x_n \in B_{2^{1-n}}$ such that $y - \sum_{i=1}^{n} Tx_i \in B_{\epsilon 2^{-n}}$. Then we have

$$x = \sum_{i=1}^{\infty} x_i$$

exists, is in B_2 (since $\|x\| \leq \sum_{i=1}^{\infty} \|x_i\| < \sum_{i=1}^{\infty} 2^{1-i} = 2$) and

$$y = \sum_{i=1}^{\infty} Tx_i = Tx.$$

Hence, $y \in T[B_2]$. □

7.3 Bounded Inverse Theorem (Kreyszig 286)

Corollary 7.6. *A continuous linear bijection $T : X \to Y$ of one Banach space onto another has a continuous inverse.*

Proof. This follows from the fact that if T is a bijection, then $T^{-1} : Y \to X$ exists and T^{-1} is continuous if and only if for all open sets $M \subset X$, we have that $T[M]$ is open (Theorem 1.3-4 of Kreyszigr). □

Definition 7.7. *Let T be a mapping of a normed linear space X into a normed space Y. The graph of T denoted by $\Gamma(T)$ is defined as $\Gamma(T) := \{(x, y) : (x, y) \in X \times Y, y = Tx\}$ where $X \times Y$ is the normed linear space defined by*

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2),$$

$$\alpha(x, y) = (\alpha x, \alpha y),$$

$$\|(x, y)\| = \|x\| + \|y\|.$$

It can easily be verified that if X and Y are Banach spaces then $X \times Y$ is a Banach space (Theorem 4.13-2 of Kreyszigr).

7.4 Closed Graph Theorem (Kreyszig 292)

Theorem 7.8. *Let X and Y be Banach spaces and T a linear map X into Y. Then T is bounded if and only if the graph of T is closed in $X \times Y$.*

Proof. Suppose $\Gamma(T)$ is closed then since T is linear, $\Gamma(T)$ is a subspace of $X \times Y$. By assumption $\Gamma(T)$ is closed and is thus a Banach space in the norm

$$\|(x, Tx)\| = \|x\| + \|Tx\|.$$

Consider the bounded linear maps $P_1 : (x, Tx) \to x$ and $P_2 : (x, Tx) \to Tx$. P_1 is bijective ($P_1^{-1} : x \to (x, Tx)$) and so by the bounded inverse theorem P_1^{-1} is bounded. But $T = P_2 P_1^{-1}$, hence, T is bounded. The converse is trivial, because if T is bounded, then $\Gamma(T)$ is closed by the continuity of T. □

Corollary 7.9. *(Hellmyer-Toeplitz) Let A be an everywhere defined linear operator on a Hilbert space, H, with $\langle x, Ay \rangle = \langle Ax, y \rangle$ for all $x, y \in H$. Then A is bounded.*

Proof. We will prove that $\Gamma(A)$ is closed. Suppose $(x_n, Ax_n) \to (x, y)$. We need to show that $(x, y) \in \Gamma(A)$, i.e. $y = Ax$. But for every $z \in H$, we have

$$\langle z, y \rangle = \lim_{n\to\infty} \langle z, Ax_n \rangle = \lim_{n\to\infty} \langle Az, x_n \rangle = \langle Az, x \rangle = \langle z, Ax \rangle.$$

Hence, $y = Ax$ and $\Gamma(A)$ is closed. □

---

8 The Adjoint Operator

Definition 8.1. *Let X and Y be Banach spaces and T a bounded linear operator from X to Y. The Banach space adjoint of T, denoted $T^{\times}$, is the bounded linear operator from Y' to X' defined by*

$$(T^{\times}g)(x) = g(Tx)$$

for all $g \in Y'$, $x \in X'$

Theorem 8.2. *Let X and Y be a Banach space. The map $T \to T^{\times}$ is an isomorphism of $B(X,Y)$ to $B(Y',X')$.*

Proof. The map $T \to T^{\times}$ is linear. The fact that $T^{\times}$ is bounded and the map is an isomorphism, follows from the following

$$\|T\| = \overbrace{\sup_{\|x\|=1} \|Tx\| = \sup_{\|x\|=1}\left(\sup_{\|g\|=1} |g(Tx)|\right)}^{\text{Hahn-Banach Corollary 6.7}} = \sup_{\|g\|=1}\left(\sup_{\|x\|=1} |(T^{\times}g)(x)|\right) = \sup_{\|g\|=1} \|T^{\times}g\| = \|T^{\times}\|,$$

where $g \in Y'$. ☐

For the rest of this course, we are only interested in the adjoint of T when T is a bounded transformation of a Hilbert space H into itself, i.e. $T \in B(H,H) = B(H)$.
The Banach space adjoint of T is then a mapping of H' to H'. Let $C : H \to H'$ be the map which assigns to each $y \in H$ the bounded linear functional $\langle \cdot, y\rangle$ in H'. Then C is the conjugate linear isomorphism which is surjective with norm one by the Riesz Representation Theorem. Now define $T^* : H \to H$ by

$$T^* = C^{-1}T^{\times}C.$$

Then from Theorem 8.2, we have $\|T^*\| = \|T\|$, because

$$\|T^*\| \le \|C^{-1}\|\|T^{\times}\|\|C\| = \|T^{\times}\| = \|T\| \qquad \text{and} \qquad \|T\| = \|T^{\times}\| \le \|C^{-1}\|\|T^*\|\|C\| = \|T^*\|.$$

Furthermore,

$$\langle Tx, y\rangle = Cy(Tx) = (T^{\times}Cy)(x) = (CT^*y)(x) = \langle x, T^*y\rangle,$$

where Cy is the linear functional associated with Tx from Definition 8.1.

Example: Let $k(x,y)$ be continuous on $[0,1]\times[0,1]$ and define $T : L^2[0,1] \to L^2[0,1]$ by

$$(Tf)(x) = \int_0^1 k(x,y)\, f(y)\, dy.$$

Then it is easily seen that T is bounded because

$$|(Tf)(x)|^2 \le M \int_0^1 |f(y)|^2\, dy, \qquad \text{where } M = \max\left\{ k(x,y)^2 \,\middle|\, (x,y) \in [0,1]\times[0,1]\right\},$$

and

$$(T^*f)(x) = \int_0^1 \overline{k(x,y)}\, f(y)\, dy.$$

Theorem 8.3. *Let X be an inner product space and $T \in B(X)$, then if $\langle Tx, x\rangle = 0$ for every $x \in X$, then $T = 0$.*

Proof. Let $v = \alpha x + y$, where $x, y \in X$, then $0 = \langle T(\alpha x + y), (\alpha x + y)\rangle$. Hence,

$$\begin{aligned} 0 &= |\alpha|^2 \langle Tx, x\rangle + \langle Ty, y\rangle + \alpha\langle Tx, y\rangle + \overline{\alpha}\langle Ty, x\rangle \\ &= \alpha\langle Tx, y\rangle + \overline{\alpha}\langle Ty, x\rangle. \end{aligned}$$

Then for $\alpha = 1$, we have $\langle Tx, y\rangle + \langle Ty, x\rangle = 0$. And for $\alpha = i$, we have $\langle Tx, y\rangle - \langle Ty, x\rangle = 0$. Hence, $\langle Tx, y\rangle = 0$ for all $x, y \in X$, which implies $\|Tx\| = 0$ (set $Tx = y$), and so $Tx = 0$ for all $x \in X$. This means $T = 0$. □

Theorem 8.4. *Let H be a Hilbert space and $S, T \in B(H)$, then*

(a) $(TS)^* = S^*T^*$

(b) $(T^*)^* = T$ *(in particular, $T \to T^*$ maps $B(H)$ onto $B(H)$)*

(c) *if T has a bounded inverse T^{-1}, then so does T^* and the inverse of the adjoint is the adjoint of the inverse, $(T^*)^{-1} = (T^{-1})^*$*

(d) $\|T^*T\| = \|T\|^2$.

Proof. First off, (a) and (b) follow from $\langle Tx, y\rangle = \langle x, T^*y\rangle$. Since $T^{-1}T = I = TT^{-1}$, we have from (a) that $T^*(T^{-1})^* = I^* = I = (T^{-1})^*T^*$, which implies (c). Finally (d) follows from $\|T^*T\| \le \|T\|\|T^*\| = \|T\|^2$ and,

$$\overbrace{\|T^*T\| \ge \sup_{\|x\|=1} |\langle x, T^*Tx\rangle|}^{\text{Schwarz Inequality}} = \sup_{\|x\|=1} \langle Tx, Tx\rangle = \sup_{\|x\|=1} \|Tx\|^2 = \|T\|^2.$$

□

Definition 8.5. *A bounded linear operator T on a Hilbert space is called self-adjoint or Hermitian if $T = T^*$.*

Definition 8.6. *If $P \in B(H)$ and $P^2 = P$, then P is called a projection. If in addition P is self-adjoint, P is called an orthogonal projection (main interest).*

Example: Let $x = z + w$ where $z \in M$ and $w \in M^\perp$ for some closed subspace $M \subset H$. Define $P : x \to z$. Then $P^2x = P(z) = z = Px$, implies $P^2 = P$; hence, P is a projection. Furthermore, for $x, y \in H$, $x = z + w$, $y = z' + w'$, we have that

$$\langle x, P^*y\rangle = \langle Px, y\rangle = \langle z, y\rangle = \langle z, z' + w'\rangle = \langle z, z'\rangle = \langle z, Py\rangle = \langle x + w, Py\rangle = \langle x, Py\rangle.$$

Hence, $P = P^*$ so that P is an orthogonal projection.

Theorem 8.7. *If T is a self-adjoint operator on a Hilbert space H, then*

$$\|T\| = \sup_{\|x\|=1} |\langle Tx, x\rangle|. \tag{8.1}$$

Proof. For all $x \in H$ such that $\|x\| = 1$, we have the following

$$|\langle Tx, x\rangle| \leq \|Tx\|\|x\| \leq \|T\|\|x\|^2 = \|T\| \quad \Rightarrow \quad \|T\| \geq \sup_{\|x\|=1} |\langle Tx, x\rangle|.$$

Let m denote the right hand side of (**??**), i.e. $|\langle Tu, u\rangle| \leq m\|u\|^2$ for all $u \in H$. Then for every $x, y \in H$,

$$\langle T(x \pm y), (x \pm y)\rangle = \langle Tx, x\rangle \pm 2\Re\langle Tx, y\rangle + \langle Ty, y\rangle,$$

since the self-adjoint property of T implies

$$\langle Tx, y\rangle + \langle Ty, x\rangle = \langle Tx, y\rangle + \overline{\langle Ty, x\rangle} = \langle Tx, y\rangle + \overline{\langle Tx, y\rangle} = 2\Re\langle Tx, y\rangle.$$

Here, we have

$$4\Re\langle Tx, y\rangle = \langle T(x+y), (x+y)\rangle - \langle T(x-y), (x-y)\rangle \leq M\left(\|x+y\|^2 + \|x-y\|^2\right) = 2m\left(\|x\|^2 + \|y\|^2\right).$$

Now replace x by λx where $|\lambda| = 1$ and $\lambda\langle Tx, y\rangle \geq 0$. Then we have

$$|\langle Tx, y\rangle| \leq \frac{m}{2}\left(\|x\|^2 + \|y\|^2\right).$$

Suppose $Tx \neq 0$ and choose $y = \frac{\|x\|}{\|Tx\|}Tx$. Since $\|y\| = \|x\|$, w have from the above inequality

$$\|Tx\| \leq m\|x\|.$$

This holds trivially when $Tx = 0$, and hence is true for all $x \in H$, i.e. $\|T\| \leq m = \sup_{\|x\|=1} |\langle Tx, x\rangle|$. This finishes the proof. □

Remark: The theorem is not true if T is not self-adjoint. To see this, consider $T : \mathbb{C}^2 \to \mathbb{C}^2$ defined by $T(x_1, x_2) = (0, x_1)$. Then $\|T\| = 1$, but $\sup_{\|x\|=1} |\langle Tx, x\rangle| = x_1x_2 \leq \frac{1}{2}$.

---

9 Strong and Weak Convergence

Definition 9.1. *A sequence $\{x_n\}$ in a normed linear space X is said to be strongly convergent if there exists in X an x, such that*

$$\lim_{n\to\infty} \|x_n - x\| = 0, \qquad \text{and we write } x_n \to x.$$

Definition 9.2. *A sequence $\{x_n\}$ in a normed linear space X is said to be weakly convergent if there exists $x \in X$ such that for every $f \in X'$, we have that*

$$\lim_{n\to\infty} f(x_n) = f(x), \qquad \text{and we write } x_n \rightharpoonup x.$$

Theorem 9.3. *Let $\{x_n\}$ be a weakly convergent sequence in an normed linear space X, such that $\{x_n\} \rightharpoonup x \in X$. Then*

- (a) *the weak limit x is unique,*
- (b) *every subsequence of $\{x_n\}$ converges weakly to x,*
- (c) *the sequence $\{\|x_n\|\}$ is bounded.*

Proof. For part (a), suppose $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$. Then $f(x_n) \to f(x)$ and $f(x_n) \to f(y)$ for every $f \in X'$. Since $\{f(x_n)\}$ is a sequence of numbers, its limit is unique. This implies $f(x) = f(y)$, which gives $f(x-y) = 0$ since f is linear, for all $f \in X'$. This means $x - y = 0$ by Corollary 6.7. Hence, $x = y$. (In a Hilbert space, we can use the Riesz Representation)

Part (b) follows from the fact that $\{f(x_n)\}$ is a convergent sequence of numbers, either complex or real and hence every subsequence has the same limit.

For part (c), since $\{f(x_n)\}$ is a convergent sequence of numbers, it is bounded, $|f(x_n)| \le C_f$. Now define $g_n \in X''$ by $g_n(f) = f(x_n)$ for $f \in X'$. Then for all n, $|g_n(f)| = |f(x_n)| \le C_f$, so that $\{|g_n(f)|\}$ is bounded for all $f \in X'$ for fixed f. Since X' is complete, by the Uniform Boundedness Theorem, $\|g_n\|$ is bounded. But

$$\|g_n\| = \sup_{f\in X',\, f\neq 0} \frac{|g_n(f)|}{\|f\|} = \sup_{f\in X',\, f\neq 0} \frac{|f(x_n)|}{\|f\|} = \|x_n\|,$$

by Corollary 6.7 of the Hahn-Banach Theorem. This implies that $\|x_n\|$ is bounded. □

Theorem 9.4. *Let $\{x_n\}$ be a sequence in a normed linear space X, then*

- (a) *strong convergence implies weak convergence, and the limits are the same,*
- (b) *the converse of (a) is in general not true.*

Proof. For part (a), let $x_n \to x$. Then for every $f \in X'$, we have $|f(x_n) - f(x)| = |f(x_n - x)| \le \|f\|\|x_n - x\| \to 0$. Hence, $x_n \to x$.

For part (b), let H be an infinite dimensional Hilbert space with an orthonormal basis $\{\phi_n\}$, with $f \in H'$, so that $f(x) = \langle x, z\rangle$, for some $z \in H$. By Parseval's Relation, for all $z \in H$, we have

$$\sum_{n=1}^{\infty} |\langle \phi_n, z\rangle|^2 = \|z\|^2.$$

This means $\langle \phi_n, z\rangle \to 0$, for all $z \in H$. Hence, $\phi_n \to 0$. But for $n \neq m$, we have $\|\phi_n - \phi_m\|^2 = \langle(\phi_n - \phi_m), (\phi_n - \phi_m)\rangle = 2$, which implies $\{\phi_n\}$ is not strongly convergent. □

Theorem 9.5. *Every bounded sequence in a Hilbert space contains a weakly convergent subsequence.*

Proof. Let $\{\phi_n\}$ be a bounded sequence, $\|\phi_n\| \leq c$. Then for each integer m the sequence $\langle \phi_m, \phi_n \rangle$ is bounded for all fixed m and for all n. Hence, by the Bolzano-Weierstrass Theorem, and a diagonalization process (c.f. Kreyszigr 408), we can select a subsequence $\{\phi_{n(k)}\}$ such that $\langle \phi_m, \phi_{n(k)} \rangle$ converges as $k \to \infty$ for every integer m. Thus, the linear functional F, defined by

$$F(\Psi) := \lim_{k\to\infty} \langle \Psi, \phi_{n(k)} \rangle,$$

is well defined on $U := \operatorname{span}\{\phi_m\}$, and hence by continuity, on $\overline{U}$. Now let $P : X \to \overline{U}$ be the orthogonal projection operator and for arbitrary $\Psi \in X$ write $\Psi = P(\Psi) + (I - P)(\Psi)$. For arbitrary $\Psi \in X$, define $F(\Psi)$ by

$$\begin{aligned} F(\Psi) := \lim_{k\to\infty} \langle \Psi, \phi_{n(k)} \rangle &= \lim_{k\to\infty} \left[\langle P\Psi, \phi_{n(k)} \rangle + \langle (I-P)\Psi, \phi_{n(k)} \rangle \right] \\ &= \lim_{k\to\infty} \left[\langle P\Psi, \phi_{n(k)} \rangle + \overbrace{\langle \Psi, (I-P)\phi_{n(k)} \rangle}^{=0} \right] \\ &= \lim_{k\to\infty} \langle P\Psi, \phi_{n(k)} \rangle, \end{aligned}$$

where we have that $P = P^*$. Furthermore, $\|F\| < c$. Hence, the Riesz Representation Theorem says there exists a unique $\phi \in X$ such that

$$F(\Psi) = \langle \Psi, \phi \rangle, \qquad \text{for all } \Psi \in X.$$

Therefore, $\lim_{k\to\infty} \langle \Psi, \phi_{n(k)} \rangle = \langle \Psi, \phi \rangle$ for all $\Psi \in X$, i.e. $\phi_{n(k)}$ is weakly convergent to ϕ as $k \to \infty$. □

10 Spectral Theory

Let $X \neq \{0\}$ be a complex Banach space, and $T : X \to X$ a linear operator (not necessarily bounded). With T we associate the operator $T_\lambda = T - \lambda I$ where $\lambda \in \mathbb{C}$. If T_λ has an inverse, we denote it by $R_\lambda(T)$, i.e. $R_\lambda(T) = (T - \lambda I)^{-1}$. This operator, R_λ is called the resolvent of T (recall that an operator $T : X \to X$ has an inverse $T^{-1} : \mathscr{R}(T) \to X$ if $Tx_1 = Tx_2$ implies $x_1 = x_2$).

Definition 10.1. *A regular value λ of $T : X \to X$ is a complex number such that*

(a) *$R_\lambda(T)$ exists,*

(b) *$R_\lambda(T)$ is bounded on the range of T_λ,*

(c) *$R_\lambda(T)$ is defined on a dense subset of X.*

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

The spectrum is divided into three disjoint sets:

(1) The point spectrum: Denoted $\sigma_p(T)$ is the set such that $R_\lambda(T)$ doesn't exist. A point $\lambda \in \sigma_p(T)$ is called an eigenvalue of T.

(2) The continuous spectrum: Denoted $\sigma_c(T)$ is the set such that $R_\lambda(T)$ exists, is defined on a dense subset of X, but $R_\lambda(T)$ is unbounded.

(3) The residual spectrum: Denoted $\sigma_r(T)$ is the set such that $R_\lambda(T)$ exists (and may or may not be bounded), but the domain of $R_\lambda(T)$ is not dense in X.

Note: In the case when X has finite dimension, we have that $\sigma_c(T) = \sigma_r(T) = \emptyset$. Note also that $R_\lambda(T)$ exists if and only if $T_\lambda x = 0$ implies $x = 0$. Hence, if $T_\lambda x = 0$ for some $x \neq 0$, then $\lambda \in \sigma_p(T)$ is an eigenvalue. The corresponding vector x is called an eigenvector. The subspace consisting of zero and all eigenvectors of T corresponding to the eigenvalue λ of T is called the eigenspace corresponding to λ.

Example: Let $T : C[0,1] \to C[0,1]$ be defined by $\int_0^x f(t)\,dt$. Then $Tf \in C^1 \in [0,1]$ and $R_0 = \frac{d}{dx}$. Section 2.7-5, Kreyszigr, implies $\frac{d}{dx}$ is unbounded, and hence, $0 \in \sigma_c(T)$.

Example: Define $T : \ell^2 \to \ell^2$ by $(\xi_1, \xi_2, \ldots) \to (0, \xi_1, \xi_2, \ldots) \in \ell^2$. Note that $Tx = \lambda x$ implies $(0, \xi_1, \xi_2, \ldots) = (\lambda\xi_1, \lambda\xi_2, \ldots)$. Hence, $\lambda = 0$ which implies $x = 0$. If $\lambda \neq 0$, we have $x = 0$. Thus, T has not point spectrum. Consider the case when $\lambda = 0$. Then the operator $R_0(T)$ exists and equals $T^{-1} : T(X) \to X$, which is given by

$$T^{-1} : (0, \xi_1, \xi_2, \ldots) \to (\xi_1, \xi_2, \ldots).$$

But clearly $T(X)$ is not dense in X since $(1, 0, 0, \ldots)$ is orthogonal to $T(X)$ ($\overline{T(X)} \neq X$ because $(1, 0, 0, \ldots)$ is orthogonal to $T(X)$ by the continuity of the inner product). Hence, $\lambda = 0$ is in $\sigma_r(T)$.

Show that the operator in the first example has no point spectrum.

Theorem 10.2. *If X is a Banach space, $T \in B(X)$ and $\lambda \in \rho(T)$, then $R_\lambda(T)$ is defined on all of X and is bounded.*

Proof. Let $T_\lambda x_n = y_n \to y \in X$. Then since $R_\lambda(T)$ is bounded on the range of T_λ, we have that $R_\lambda y_n = x_n$ is a Cauchy sequence (since y_n is). Since X is complete, there exists $x \in X$ such that $x_n \to x$, and $T_\lambda \in B(X)$ implies $T_\lambda x_n \to T_\lambda x = y$, i.e. $T_\lambda(X)$ is closed (here, T_λ is bounded so its continuous). But $\lambda \in \rho(T)$ implies $R_\lambda(T)$ is defined on a dense subset of X, which implies $\mathscr{D}(R_\lambda) = \overline{\mathscr{D}(R_\lambda)} = X$. Thus, $R_\lambda(T)$ is bounded on all of X. ∎

Theorem 10.3. *Let X be a Banach space and $T \in B(X)$. If $\|T\| < 1$, then $(I-T)^{-1} \in B(X)$ and*

$$(I-T)^{-1} = \sum_{j=0}^{\infty} T^j. \tag{10.1}$$

Proof. Since $T_1, T_2 \in B(X)$ implies $\|T_1T_2\| \leq \|T_1\|\|T_2\|$ from the definition of the norm on $B(X)$, we have that $\|T^j\| \leq \|T\|^j$. Hence, the series (**??**) is convergent in $B(X)$ for $\|T\| < 1$. Since, X is complete, so is $B(X)$ by Theorem 6.1. Hence,

$$S = \sum_{j=0}^{\infty} T^j,$$

exists as an element of $B(X)$. We need to show that $S = (I-T)^{-1}$. To this end, we have $(I-T)(I+T+\cdots+T^n) = (I+T+\cdots+T^n)(I-T) = I - T^{n+1}$. Let $n \to \infty$ and noting that $\|T^{n+1}\| \leq \|T\|^{n+1} \to 0$, we have $(I-T)S = S(I-T) = I$. Hence, $S = (I-T)^{-1}$. □

Theorem 10.4. *Let X be a Banach space and $T \in B(X)$. Then $\rho(T)$ is an open set (and hence $\sigma(T)$ is closed).*

Proof. We will show shortly that $\rho(T) \neq 0$. For a fixed $\lambda_0 \in \rho(T)$ and $\lambda \in \mathbb{C}$, we have

$$T - \lambda I = T - \lambda_0 I - (\lambda - \lambda_0)I = (T - \lambda_0 I)\left[I - (\lambda-\lambda_0)(T-\lambda_0 I)^{-1}\right],$$

i.e. $T_\lambda = T_{\lambda_0}V$, where $V = I - (\lambda-\lambda_0)R_{\lambda_0}$. Since $\lambda_0 \in \rho(T)$ and T is bounded, then Theorem 10.2 implies $R_{\lambda_0} \in B(X)$. Furthermore, Theorem 10.3 implies $V^{-1} \in B(X)$ exists for all λ such that $\|(\lambda-\lambda_0)R_{\lambda_0}\| < 1$. In other words, $|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}$. Hence, for such λ,

$$R_\lambda = T_\lambda^{-1} = (T_{\lambda_0})^{-1} = V^{-1}R_{\lambda_0} \in B(X),$$

exists. This implies $\rho(T)$ is open, and $\sigma(T)$ is closed. □

Corollary 10.5. *Let X be a Banach space and $T \in B(X)$. Then for every $\lambda_0 \in \rho(T)$ and $|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}$ we have that*

$$R_\lambda = \sum_{j=0}^{\infty} (\lambda-\lambda_0)^j R_{\lambda_0}^{j+1},$$

is absolutely convergent. In particular, if $|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|} \subset \rho(T)$.

Proof. Follows from Theorem 10.3 and $R_\lambda = V^{-1}R_0$. □

Theorem 10.6. *Let X be a Banach space and $T \in B(X)$. Then $\sigma(T)$ is compact and lies in the disk $|\lambda| \leq \|T\|$. In particular, $\rho(T) \neq \emptyset$.*

Proof. Let $\lambda \neq 0$, then Theorem 10.3 implies

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

where the series converges for all λ such that $\left\|\frac{1}{\lambda}T\right\| < 1$, i.e. $|\lambda| > \|T\|$. Hence, $\lambda \in \sigma(T)$ implies $|\lambda| \leq \|T\|$. By Theorem 10.4 we have $\sigma(T)$ is closed which implies $\sigma(T)$ is compact. □

Theorem 10.7. *If $X \neq \{0\}$ is a complex Banach space and $T \in B(X)$ then $\sigma(T) \neq \emptyset$.*

Proof. If $T = 0$, then $\sigma(T) = \{0\} \neq \emptyset$. Let $T \neq 0$ (possible since $X \neq \{0\}$). Then $\|T\| \neq 0$ and from Theorem 10.6, we have $R_\lambda = -\frac{1}{\lambda}\sum_{j=0}^{\infty}\left(\frac{1}{\lambda}T\right)^j$, provided $|\lambda| > \|T\|$. Since this series converges for $\frac{1}{|\lambda|} < \frac{1}{\|T\|}$, it converges for $\frac{1}{|\lambda|} \leq \frac{1}{2\|T\|}$, i.e. for $|\lambda| \geq 2\|T\|$ and

$$\|R_\lambda\| \leq \frac{1}{|\lambda|}\sum_{j=0}^{\infty}\left\|\frac{1}{\lambda}T\right\|^{j} = \frac{1}{|\lambda| - \|T\|} \leq \frac{1}{\|T\|}, \qquad |\lambda| \geq 2\|T\|.$$

Now suppose $\sigma(T) = \emptyset$. Then $\rho(T) = \mathbb{C}$ implies from Corollary 10.5, that for each $f \in X'$ and a fixed $x \in X$, the function

$$h(\lambda) = f(R_\lambda x).$$

is an entire function of λ. In particular, h is continuous and hence bounded on $|\lambda| \leq 2\|T\|$. But for $|\lambda| \geq 2\|T\|$ we have

$$|h(\lambda)| = |f(R_\lambda x)| \leq \|f\|\|R_\lambda x\| \leq \|f\|\|R_\lambda\|\|x\| \leq \frac{\|f\|\|x\|}{\|T\|}.$$

Hence, $h(\lambda)$ is bounded on $\mathbb{C}$, which implies $h(\lambda)$ is constant, by Louiville's Theorem. Since x and f are arbitrary, h being constant implies $f\left(\frac{\partial}{\partial\lambda}R_\lambda x\right) = 0$, for all f. This implies $\frac{\partial}{\partial\lambda}R_\lambda x = 0$ for all x. Hence, $\frac{\partial}{\partial\lambda}R_\lambda = 0$, which means R_λ is independent of λ, which implies $R_\lambda^{-1} = T - \lambda I$ is independent of λ which is a contradiction. □

---

11 Compact Operators

Definition 11.1. *Let X be a Banach space and $T \in B(X)$. We denote the null space and range of T by $\mathscr{N}(T)$ and $\mathscr{R}(T)$, respectively. An operator T is called compact if it maps bounded sets into relatively compact sets (i.e. sets with compact closure). Equivalently, T is compact if whenever $\{x_n\}$ is a bounded sequence in X, the sequence $\{Tx_n\}$ has a strongly convergent subsequence (Theorem 8.1-3, Kreyszigr).*

Definition 11.2. *An operator $T \in B(X)$, where X is a Banach space is said to be of finite rank, if the range of T, $\mathscr{R}(T)$ is finite dimensional. It is easily seen that every finite operator of finite rank is compact (Theorem 8.1-4, Kreyszigr.)*

Theorem 11.3. *Let X be a Banach space. Then the set of compact operators on X is a closed two-sided ideal in the algebra of bounded operators on X with a norm topology (closed two sided ideal means AT and TA are compact).*

Proof. Suppose T_1 and T_2 are compact and $\{x_j\} \subset X$ is bounded. We can choose a subsequence $\{y_j\}$ of $\{x_j\}$ such that $\{T_1 y_j\}$ converges, by the definition of compact. Choose a subsequence $\{z_j\}$ of $\{y_j\}$ such that $\{T_2 z_j\}$ converges. It follows that $a_1T_1 + a_2T_2$ is compact for all $a_1, a_2 \in \mathbb{C}$. It is also clear that if T is compact and S is bounded, then TS and ST are compact (Lemma 8.3-2, Kreyszigr). Thus, the set of compact operators is a compact ideal. Suppose $\|T_m - T\| \to 0$ where each T_m is compact. Given a sequence $\{x_j\} \subset X$ with $\|x_j\} \leq c$ for all j, choose a subsequence $\{x_{1j}\}$ such that $\{T_1 x_{1j}\}$ converges. Proceeding inductively $m = 2, 3, \ldots$, choose a subsequence $\{x_{mj}\}$ of $\{x_{(m-1)j}\}$ such that $\{T_m x_{mj}\}$ converges. Setting $y_j = x_{jj}$, one easily sees that $\{T_m y_j\}$ converges for all m. But then

$$\|Ty_j - Ty_k\| \leq \|(T - T_m)y_j\| + \|T_m(y_j - y_k)\| + \|(T_m - T)y_k\| \leq 2c\|T - T_m\| + \|T_m y_j - T_m y_k\|.$$

Given $\epsilon > 0$ choose m large enough that $\|T - T_m\| \leq \frac{\epsilon}{4c}$, and then with this choice of m, we have that

$$\|T_m y_j - T_m y_k\| \leq \frac{\epsilon}{2},$$

for j and k sufficiently large. Hence, $\|Ty_j - Ty_k\| < \epsilon$. Thus, $\{Ty_j\}$ is convergent. Hence, T is compact. □

Note: $T \in B(X)$ since $B(x)$ is a Banach space.

Corollary 11.4. *If X is a Banach space and $T \in B(X)$ such that there is a sequence $\{T_m\}$ of operators of finite rank such that $\|T_m - T\| \to 0$, then T is compact.*

Theorem 11.5. *If H is a Hilbert space and T is a compact operator on H, then T is the norm limit of operators of finite rank.*

Proof. Suppose $\epsilon > 0$ and let B be the unit ball in H. Since $T(B)$ has compact closure (the closure of B is not compact though), it is totally bounded, i.e. there exists a finite set $y_1, \ldots, y_n \in T(B)$ such that every $y \in T(B)$ satisfies $\|y - y_j\| < \epsilon$ for some j (c.f. Lemma 8.2-2, Kreyszigr). Let P_ϵ be the orthogonal projection onto the closure of the space spanned by $y_1, \ldots, y_n$ and set $T_\epsilon = P_\epsilon T$. Then T_ϵ is of finite rank and also, by the Lemma 3.1, $T_\epsilon x$ is the element closest to Tx in $\mathscr{R}(P_\epsilon)$. This implies for $x \in B$, we have

$$\|Tx - T_\epsilon x\| < \min_{1 \leq j \leq n} \|Tx - y_j\| < \epsilon, \qquad \Rightarrow \qquad \|T - T_\epsilon\| < \epsilon,$$

which implies $T_\epsilon \to T$ as $\epsilon \to 0^+$. □

For the sake of simplicity, for the rest of this course, we shall only consider operators on a complex Hilbert space.

Theorem 11.6. *Let H be a Hilbert space and $T \in B(H)$, then if T^*T is compact, so is T.*

Proof. Let $\{x_j\} \in H$ with $\|x_j\| \leq c$. Then $\{T^*Tx_j\}$ has a strongly convergent subsequence which we again denote as $\{T^*Tx_j\}$. Then

$$\|Tx_m - Tx_n\|^2 = \langle T(x_m - x_n), T(x_m - x_n)\rangle = \langle T^*T(x_m - x_n), (x_m - x_n)\rangle \leq \|T^*T(x_m - x_n)\|\|x_m - x_n\| \to 0,$$

by Cauchy-Schwarz, as $m, n \to \infty$, since $\{T^*Tx_j\}$ is convergent and $\|x_m - x_n\| \leq 2c$. Hence, $\{Tx_j\}$ is a Cauchy sequence. Hence, $\{Tx_j\}$ is convergent so that T is compact. □

Corollary 11.7. *If T is a compact operator, then so is the adjoint, where $T \in B(H)$.*

Proof. We know T^* is bounded, and hence $TT^* = (T^*)^*T^*$ is compact. The corollary now follows from the theorem. □

Lemma 11.8. *Let M be a closed subspace of a Hilbert space H. Then $M^{\perp\perp} = M$.*

Proof. We have $H = M \oplus M^\perp$ and also $H = M^\perp \oplus M^{\perp\perp}$. Then for $x \in H$, we have $x = z_1 + w_1$, where $z_1 \in M$ and $w_1 \in M^\perp$ and $x = z_2 + w_2$, where $z_2 \in M^{\perp\perp}$ and $w_2 \in M^\perp$. Hence,

$$0 = z_1 + w_1 - z_2 - w_2 = (z_1 - z_2) + (w_1 - w_2) \qquad \Rightarrow \qquad z_1 - z_2 = w_2 - w_1 \in M^\perp.$$

But $M \subseteq M^{\perp\perp}$ implies $z_1 - z_2 \in M^{\perp\perp}$ which implies $z_1 = z_2$. Hence, $M = M^{\perp\perp}$. □

Theorem 11.9. *If H is a Hilbert space, $T \in B(H)$, then $\mathscr{R}(T)^\perp = \mathscr{N}(T^*)$ and $\mathscr{N}(T^*)^\perp = \overline{\mathscr{R}(T)}$.*

Proof. Let $y \in \mathscr{R}(T)^\perp$. This happens if $\langle Tx, y\rangle = 0$ for all $x \in H$. This is true if and only if $\langle x, T^*y\rangle = 0$ for all $x \in H$. This implies $T^*y = 0$. Hence, $y \in \mathscr{N}(T^*)$. On the other hand, by the above lemma, $\overline{\mathscr{R}(T)} = \overline{\mathscr{R}(T)}^{\perp\perp} = \mathscr{N}(T^*)^\perp$ (since $\mathscr{R}(T)^\perp = \overline{\mathscr{R}(T)}^\perp = \mathscr{N}(T^*)$). □

Definition 11.10. *Let H be a separable Hilbert space. Then $T \in B(H)$ is said to be a Hilbert-Schmidt operator if there exists a complete orthogonal sequence $\{\phi_n\}_{n=1}^\infty$ in H such that*

$$\sum_{n=1}^{\infty} \|T\phi_n\|^2 < \infty.$$

Theorem 11.11. *Hilbert-Schmidt operators are compact.*

Proof. Let T be a Hilbert-Schmidt operator and $\{\phi_n\}_{n=1}^\infty$ be a complete orthonormal sequence such that $\sum_{n=1}^\infty \|T\phi_n\|^2 < \infty$. Define $T_k \in B(H)$ by

$$T_k\left(\sum_{n=1}^\infty a_n\phi_n\right) = \sum_{n=1}^k a_n T\phi_n,$$

where $x = \sum a_n\phi_n$ is an arbitrary element of H. Then T_k is a finite rank operator and is compact. Since

$$(T - T_k)x = \sum_{k+1}^\infty a_k T\phi_n.$$

This implies

$$\|(T-T_k)x\| \le \sum_{k+1}^\infty |a_n|\|T\phi_n\| \le \left(\sum_{k+1}^\infty |a_n|^2\right)^{1/2} \left(\sum_{k+1}^\infty \|T\phi_n\|^2\right)^{1/2} = \|x\| \left(\sum_{k+1}^\infty \|T\phi_n\|^2\right)^{1/2}.$$

Hence, $\|T - T_k\| \le \left(\sum_{k+1}^\infty \|T\phi_n\|^2\right)^{1/2} \to 0$ as $k \to \infty$. Thus, $T_k \to T$ and T is compact. □

Corollary 11.12. *Let $k : [1,b] \times [a,b] \to \mathbb{C}$ be a continuous function of two variables. Then the integrand operator $k : L^2[a,b] \to L^2[a,b]$ with bound k is a Hilbert-Schmidt operator and hence compact.*

Proof. Let $\{\phi_n\}_{n=1}^\infty$ be a complete orthonormal sequence in $L^2[a,b]$. For $t \in [a,b]$ we have $(k\phi_n)(t) = \int_a^b k(t,s)\,\phi_n(s)\,ds = \langle k_t, \overline{\phi}_n\rangle$, where $k_t(s) = k(t,s)$, $a < s < b$. Thus,

$$\|k\phi_n\|^2 = \int_a^b |(k\phi_n)(t)|^2\,dt = \int_a^b |\langle k_t, \overline{\phi}_n\rangle|^2\,dt.$$

Now, $\{\overline{\phi}_n\}_{n=1}^\infty$ is also a complete orthonormal sequence in $L^2[a,b]$. This implies

$$\begin{aligned}\sum_{n=1}^\infty \|k\phi_n\|^2 &= \overbrace{\sum_{n=1}^\infty \int_a^b |\langle k_t, \overline{\phi}_n\rangle|^2\,dt = \int_a^b \sum_{n=1}^\infty |\langle k_t, \overline{\phi}_n\rangle|^2\,dt}^{\text{Monotone convergent sequence}} \\ &= \int_a^b \|k_t\|^2\,dt = \int_a^b\int_a^b |k(t,s)|^2\,ds\,dt < \infty,\end{aligned}$$

which follows from Parseval's Relation and since $k(t,s)$ is continuous. □

Remark: Note that the corollary remains valid under weaker conditions, i.e. $k : [a,b] \times [a,b] \to \mathbb{C}$ is piecewise continuous. In particular, the integral operator $k : L^2[a,b] \to L^2[a,b]$,

$$(K\phi)(t) \int_0^t k(t,s)\,\phi(s)\,ds = \int_0^1 \hat{k}(t,s)\,\phi(s)\,ds,$$

where

$$\hat{k}(t,s) = \begin{cases} 0 & t \le s \le 1, \\ k(t,s) & 0 \le s < t, \end{cases}$$

is compact if $k(s,t)$ is continuous for $0 \le s < t \le 1$.

---

12 The Spectral Theorem for Compact, Selft-Adjoint Operators

Theorem 12.1. *Let $T \neq 0$ be a compact self-adjoint operator on a Hilbert space H. Then either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof. By assumption, $\|T\| \neq 0$. From Theorem 8.7,

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

Hence, there exists a sequence $\{x_n\}$ in H, $\|x_n\| = 1$ such that

$$\|\langle Tx_n, x_n\rangle\| \to \|T\|, \qquad n \to \infty.$$

Since for every x in H, $\langle Tx, x\rangle = \overline{\langle x, Tx\rangle}$, we have that $\langle Tx, x\rangle$ is real. Hence, $\langle Tx_n, x_n\rangle$ is real and we can assume (replacing $\{x_n\}$ by a subsequence if necessary) that $\langle Tx_n, x_n\rangle \to \lambda$ where $\lambda \neq 0$ is either $\|T\|$ or $-\|T\|$, where we have taken away the absolute value. Since λ is real,

$$\begin{aligned}\|Tx_n - \lambda x_n\|^2 = \|Tx_n\|^2 - 2\lambda\langle Tx_n, x_n\rangle + \lambda^2\|x_n\|^2 &\leq (\|T\|\|x_n\|)^n - 2\lambda\langle Tx_n, x_n\rangle + \lambda^2 \\ &= 2\lambda^2 - 2\lambda\langle Tx_n, x_n\rangle \to 0, \qquad \text{as } n \to \infty.\end{aligned}$$

Hence $Tx_n - \lambda x_n \to 0$ as $n \to \infty$. Since T is compact, there exists a subsequence, $\{x_n'\} \subset \{x_n\}$ such that Tx_n' is convergent. Let y be its limit. Then $Tx_n' \to y$, $Tx_n' \to \lambda x_n'$ implies $\lambda x_n' \to y$. Hence, $\lambda T x_n' \to Ty$. But by the definition of y,

$$\lambda T x_n' \to \lambda y \qquad \Rightarrow \qquad Ty = \lambda y.$$

Since $\|\lambda x_n'\| = |\lambda| \neq 0$ and $\lambda x_n' \to y$ we can conclude, $y \neq 0$. □

Theorem 12.2. *Let $T \neq 0$ be a compact self-adjoint operator on a Hilbert space H. Then the set of eigenvalues of T is a set of real numbers which either is finite or consists of a countable sequence tending to zero.*

Proof. Since $\langle Tx, x\rangle$ is real, $Tx = \lambda x$ implies $\langle Tx, x\rangle = \lambda\|x\|^2$, which implies λ is real. We know there exists at least one eigenvalue. Suppose there exists infinitely many eigenvalues but there exists some subsequence which does not tend to zero. Then there exists some $\epsilon > 0$ and a sequence $\{\lambda_n\}_{n=1}^{\infty}$ of distinct eigenvalues of T such that $|\lambda_n| > \epsilon$ for all $n \in \mathbb{N}$. By Theorem 9.1-1 of Kreyszig, $\{\phi_n\}$ is an orthonormal sequence. Then for all n, m, $n \neq m$,

$$\|T\phi_m - T\phi_n\|^2 = \|\lambda_n\phi_n - \lambda_m\phi_m\|^2 = |\lambda_n|^2 + |\lambda_m|^2 > 2\epsilon^2.$$

This implies $\{T\phi_n\}$ does not have a convergent subsequence since $\|\phi_n\| = 1$. This contradicts the compactness of T. Hence, since $|\lambda| \leq \|T\|$ (Theorem 8.7), and by the Bolzano-Weierstrass Theorem, the set of eigenvalues satisfying $|\lambda| \geq \frac{1}{n}$ is finite. □

Lemma 12.3. *Let M be a subspace of a Hilbert space H and $T \in B(H)$ such that $T(M) \subseteq M$. Then $T^*(M^\perp) \subseteq M^\perp$.*

Proof. Let $x \in M^\perp$ and $y \in M$. Then if $Ty \in M$, then $\langle Ty, x\rangle = 0$ for all $y \in M$. This implies $T^*x \in M^\perp$ for all $x \in M^\perp$. □

Theorem 12.4. *(Hilbert-Schmidt Theorem) Let $T \neq 0$ be a compact self-adjoint operator on a Hilbert space H. Then there exists a finite or infinite orthonormal set $\{\phi_n\}$ of eigenvectors of T with corresponding real eigenvalues $\{\lambda_n\}$ such that for all $x \in H$*

$$Tx = \sum \lambda_n \langle x, \phi_n \rangle \phi_n.$$

Proof. We shall show the existence of the sequence $\{\phi_n\}$ by induction. We know there exists one eigenvalue $\lambda_1 = \pm\|T\|$ and we may pick a corresponding eigenvector ϕ_1 such that $\|\phi_1\| = 1$. Since the span of $\{\phi_1\}$ is invariant under T, $H_2 = \text{span}\{\phi_1\}^\perp$ is invariant under $T^* = T$ by the above lemma. Hence, $T(H_2) \subseteq H_2$ and thus we can define $T_2 := T\Big|_{H_2}$ (T restricted to H_2). Clearly, T_2 is compact and for all $x, y \in H_2$ we have

$$\langle T_2^* x, y\rangle = \langle x, T_2 y\rangle = \langle x, Ty\rangle = \langle T^* x, y\rangle = \langle Tx, y\rangle \qquad \Rightarrow \qquad Tx = T_2^* x,$$

for all $x \in H_2$. Hence, $T_2^* = T\Big|_{H_2} = T_2$. Hence, T_2 is a compact self-adjoint operator on H_2. That implies that there exists an eigenvalue $\lambda_2 = \pm\|T_2\|$ of T_2 and hence of $\|T\|$. Choose a unit eigenvector ϕ_2 of T_2 corresponding to λ_2 (Note, it is possible that $\lambda_2 = \lambda_1$). Now suppose inductively we have constructed unit eigenvectors $\phi_1, \phi_2, \ldots, \phi_n$ of T with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $|\lambda_j| = \|T_j\|$, for $1 \le j \le n$, where $T_j := T\Big|_{H_j}$, where $H_j := \text{span}\{\phi_1, \ldots, \phi_{j-1}\}^\perp$, for $2 \le j \le n+1$. Here, H_{n+1} is invariant under T and arguing as above, $T_{n+1} = T_n\Big|_{H_{n+1}}$ is a compact self-adjoint operator on H_{n+1}, which implies T_{n+1} has an eigenvalue $\lambda_{n+1} = \pm\|T_{n+1}\|$ with unit eigenvector $\phi_{n+1} \in H_{n+1}$. Then $\phi_1, \phi_2, \ldots, \phi_{n+1}$ are naturally orthogonal unit eigenvectors with eigenvalues λ_j such that $|\lambda_j| = \|T_j\|$ for $1 \le j \le n+1$ (note, this implies that infinite multiplicity contradicts the range of T being compact). The inductibe construction continues as long as $T_n \neq 0$. If we encounter an m such that $T_m = 0$, the construction stops. Since, $y = x - \sum_{j=1}^{m-1} \langle x, \phi_j\rangle \phi_j \in H_m$, for all $x \in H$, we have

$$0 = T_m y = Ty = Tx - \sum_{j=1}^{m-1} \langle x, \phi_j\rangle T\phi_j = Tx - \sum_{j=1}^{m-1} \langle x, \phi_j\rangle \lambda_j \phi_j,$$

and the theorem is proved if there exists such an m. Now suppose $T_n \neq 0$ for all $n \in \mathbb{N}$. Then for $x \in H$, define $y_n \in H_n$ by $y_n = x - \sum_{j=1}^{n-1} \langle x, \phi_j\rangle \phi_j$. Then

$$\|x\|^2 = \|y_n\|^2 + \sum_{j=1}^{m-1} |\langle x, \phi_j\rangle|^2 \qquad \Rightarrow \qquad \|y_n\| \le \|x\|.$$

But $\|Ty_n\| = \|T_n y_n\| \le \|T_n\|\|y_n\| \le |\lambda_n|\|x\|$, i.e.

$$\left\| Tx - \sum_{j=1}^{m-1} \lambda_j \langle x, \phi_j\rangle \phi_j \right\| = \|Ty_n\| \le |\lambda_n|\|x\|.$$

Since $\lambda_n \to 0$ as $n \to \infty$, the theorem follows. □

Remark: The orthonormal set $\{\phi_n\}$ given above may not be complete. Indeed if H is not separable then no countable orthonormal set is complete in H.

Corollary 12.5. *Let $T \neq 0$ be a compact self-adjoint operator on a separable Hilbert space H. Then there exists a complete orthonormal set for H consisting of eigenvectors of T.*

Proof. By the Hilbert-Schmidt Theorem, there exists a sequence $\{\phi_n\}$ such that

$$Tx = \sum \lambda_n \langle x, \phi_n \rangle \phi_n,$$

with $\lambda_n \neq 0$. Let $\{\Psi_n\}$ be a complete orthonormal set in the Hilbert space $\mathscr{N}(T)$ (recall that $\mathscr{N}(T)$ is closed). Then each Ψ_n is an eigenvector of T corresponding to the eigenvalue 0 and since $\lambda_n \neq 0$ we have that $\phi_n \perp \Psi_m$ for all n, m. That is

$$\langle \phi_n, \Psi_m \rangle = \frac{1}{\lambda_n} \langle T\phi_n, \Psi_m \rangle = \frac{1}{\lambda_n} \langle \phi_n, T\Psi_m \rangle = 0,$$

i.e. $\{\phi_n\} \cup \{\Psi_m\}$ is a countable orthonormal set in H. For every $x \in H$, $x - \sum \langle x, \phi_n \rangle \phi_n \in \mathscr{N}(T)$, implies

$$x - \sum \langle x, \phi_n \rangle \phi_n = \sum \langle x, \Psi_m \rangle \Psi_m,$$

which says that $\{\phi_n\} \cup \{\Psi_m\}$ is complete. □

13 Ill-Posed Problems

For problems in mathematical physics, Hadamard postulated three properties:

(1) Existence of a solution.

(2) Uniqueness of the solution.

(3) Continuous dependence of the solution on the data.

A problem satisfying all requirements is called well-posed. To be more precise, we make the following definition:

Definition 13.1. *Let $T : U \to V$ be an operator from a subset U of a normed space X into a subset V of a normed space Y. The equation $T\phi = f$ is called well posed if T is bijective and $T^{-1} : V \to U$ is continuous. Otherwise, $T\phi = f$ is called ill-posed or improperly posed.*

Example: Consider the initial-boundary valued problem $\frac{\partial}{\partial t}u = \frac{\partial^2}{\partial x^2}u$ in $[0, \pi] \times [0, T]$, where $u(0, t) = u(\pi, t) = 0$, $0 \le t \le T$ and $u(x, 0) = \phi(x)$, $0 \le x \le \pi$. then by separation of variables

$$u(x,t) = \sum_{n=1}^{\infty} a_n e^{-n^2 t} \sin nx, \qquad a_n = \frac{2}{\pi}\int_0^{\pi} \phi(y)\, \sin ny\, dy.$$

This problem is well-known to be well-posed.

Example: Now consider the inverse problem of determining ϕ from $f := U(\cdot, t)$. In this case $u(x, t) = \sum_{n=1}^{\infty} b_n e^{n^2(T-t)} \sin nx$, where $b_n = \frac{2}{\pi}\int_0^{\pi} f(y)\, \sin ny\, dy$, with

$$\|\phi\|_2^2 = \frac{\pi}{2}\sum_{n=1}^{\infty} b_n^2 e^{2n^2 T},$$

which is infinite unless the b_n decrease extremely rapidly. Even if this was the case, small purterbations of f (and hence b_n) result in the non-existence of a solution. Note that the inverse problem can be written as an integral equation of the first kind with smooth kernel:

$$\int_0^{\pi} k(x,y)\, \phi(y)\, dy = f(x), \qquad 0 \le x \le \pi,$$

where

$$k(x,y) = \frac{2}{\pi}\sum_{n=1}^{\infty} e^{-n^2 T} \sin nx\, \sin ny, \qquad 0 \le x, y \le \pi.$$

In particular, the above integral operator is compact in any reasonable function space, i.e. $L^2[0, \pi]$.

Theorem 13.3. *Let X and Y be Hilbert spaces and let $T : X \to Y$ be a compact operator. Then $T\phi = f$ is improperly posed if X is not of finite dimension.*

Proof. If T^{-1} exists and is continuous (bounded) then $I = T^{-1}T$ is compact. But if $\{\phi_n\}$ is a countable orthonormal set for X, then $\|\phi_n\| = 1$ for all n but $\|\phi_n - \phi_m\| = \sqrt{2}$ implies $\{\phi_n\}$ cannot have a convergent subsequence. Hence, I is not compact. This is a contradiction. □

From now on X and Y will always be infinite dimensional Hilbert spaces and $T : X \to Y$ will always be a compact operator. Note that T^*T is compact and self-adjoint. Hence, by the Hilbert-Schmidt Theorem, there exists at most a countable set of eigenvalues $\{\lambda_n\}$ of T^*T and $T^*T\phi_n = \lambda_n\phi_n$, then $\langle T^*T\phi_n, \phi_n\rangle = \lambda_n\|\phi_n\|^2$, i.e. $\|T\phi_n\|^2 = \lambda_n\|\phi_n\|^2$, which implies $\lambda_n \ge 0$. The non-negative square roots of the eigenvalues of $T^*T : X \to X$, are called the singular values of T.

Theorem 13.3. *Let $\{\mu_n\}$ be the sequence of non-zero singular values of the compact operator T ordered such that $\mu_1 \geq \mu_2 \geq \cdots$. Then there exists orthonormal sequence $\{\phi_n\}$ in X and $\{g_n\}$ in Y such that $T\phi_n = \mu_n g_n$, $T^*g_n = \mu_n\phi_n$. For every $\phi \in X$, we have the singular value decomposition*

$$\phi = \sum_{n=1}^{\infty} \langle \phi, \phi_n\rangle \phi_n + P\phi,$$

where $P : X \to \mathscr{N}(T)$ and $T\phi = \sum_{n=1}^{\infty} \mu_n \langle\phi, \phi_n\rangle g_n$. The system (μ_n, ϕ_n, g_n) is called a singular system of T

Proof. Let $\{\phi_n\}$ be an orthonormal eigenvectors of T^*T, i.e.

$$T^*T\phi_n = \mu_n^2\phi_n.$$

And define a second orthonormal sequence by

$$g_n := \frac{1}{\mu_n} T\phi_n.$$

Then $T\phi_n = \mu_n g_n$ and $T^*g_n\mu_n\phi_n$. The Hilbert-Schmidt Theorem implies

$$\phi = \sum_{n=1}^{\infty} \langle \phi, \phi_n\rangle \phi_n + P\phi,$$

where $P : X \to \mathscr{N}(T^*T)$ (if $\Psi \in \mathscr{N}(T^*T)$, then $\langle T\Psi, T\Psi\rangle = \langle \Psi, T^*T\Psi\rangle = 0$, which implies if $\Psi \in \mathscr{N}(T)$, then $\Psi \in \mathscr{N}(T^*T)$). Finally, applying T to the above expression, we have that

$$T\phi = \sum_{n=1}^{\infty} \mu_n \langle \phi, \phi_n\rangle \phi_n.$$

□

Theorem 13.4. *(Picard's Theorem) Let $T : X \to Y$ be a compact operator, not necessarily self-adjoint, with singular system (μ_n, ϕ_n, g_n). The equation $T\phi = f$ is solvable if and only if $f \in \mathscr{N}(T^*)^\perp = \overline{\mathscr{R}(T)}$ and, furthermore,*

$$\sum_{n=1}^{\infty} |\langle f, g_n\rangle|^2 < \infty. \tag{13.1}$$

In this case, a solution is given by

$$\phi = \sum_{n=1}^{\infty} \frac{1}{\mu_n} \langle f, g_n\rangle \phi_n.$$

Proof. The necessity of $f \in \mathscr{N}(T^*)^\perp = \overline{T(X)}$ (Theorem 11.10). If ϕ is a solution of $T\phi = f$, then

$$\mu_n\langle \phi, \phi_n\rangle = \langle \phi, T^*\phi_n\rangle = \langle T\phi, g_n\rangle = \langle f, g_n\rangle,$$

which implies the singular value decomposition,

$$\|\phi\|^2 = \sum_{n=1}^{\infty} |\langle \phi, \phi_n\rangle|^2 + \|P\phi\|^2.$$

Hence,

$$\sum_{n=1}^{\infty} \frac{1}{\mu_n^2} |\langle f, g_n \rangle|^2 = \sum_{n=1}^{\infty} \langle \phi, \phi_n \rangle|^2 \leq \|\phi\|^2,$$

which implies (**??**). Conversely, assume $f \in \mathscr{N}(T^*)^{\perp}$ and (**??**) is satisfied. Then

$$\sum_{n=1}^{\infty} \frac{1}{\mu_n} \langle f, g_n \rangle \phi_n,$$

converges on the Hilbert space X. Applying T to this series gives

$$T\phi = \sum_{n=1}^{\infty} \langle f, g_n \rangle g_n.$$

But since $f \in \mathscr{N}(T^*)^{\perp}$, this is the singular value decomposition of f corresponding to the operator T^*, which implies $T\phi = f$. □

For $f \in \mathscr{R}(T)$, Picard's Theorem suggests trying to regularize $T\phi = f$ by damping or filtering at the influence of the factor $\frac{1}{\mu_n}$ in the solution (if $f \in \mathscr{R}(T)$) given by

$$\phi = \sum_{n=1}^{\infty} \frac{1}{\mu_n} \langle f, g_n \rangle \phi_n.$$

The simplest regularization method is the spectral cut off method (truncate the series), where we define the "solution" of $T\phi = f$, $f \in Y$, by

$$R_m f := \sum_{\mu_n \leq \mu_m} \frac{1}{\mu_n} \langle f, g_n \rangle \phi_n, \qquad (\mu \to 0).$$

How should one choose m? The Morozov Discrepancy Principle says that the residual

$$\|TR_m f - f\|,$$

should not be smaller than the accuracy of the measurement, i.e. if f^{exact} is the "noise free" data and $\|f^{exact} - f\| \leq \delta$, then we want

$$\|TR_m f - f\| \geq \delta.$$

On the other hand, one wants the residual as small as possible, i.e. one should choose $m = m(\delta)$ such that $m(\delta)$ is the smallest of m such that $\|TR_m f - f\| \leq \delta$.

Theorem 13.5. *Let $T : X \to Y$ be an injective compact operator with dense range in Y. Let $f \in Y$ and $\delta > 0$. Then there exists a smallest integer m such that*

$$\|TR_m f - f\| \leq \delta.$$

Proof. Since $\overline{T(X)} = Y$, T^* is injective ($\mathscr{N}(T^*)^{\perp} = \overline{T(X)}$). Hence, the singular value decomposition with the singular system (μ_n, ϕ_n, g_n) for T^* implies for every $f \in Y$, we have that

$$f = \sum_{n=1}^{\infty} \langle f, g_n \rangle g_n. \tag{13.2}$$

Hence,

$$\|(TR_m - I)f\|^2 = \sum_{\mu_n < \mu_m} |\langle f, g_n \rangle|^2 \to 0 \qquad \text{as } m \to \infty \tag{13.3}$$

In particular, there exists a smallest integer $m = m(\delta)$ such that $\|TR_m f - f\| \leq \delta$. □

Note that (**??**) and (**??**) imply that $m = m(\delta)$ is the smallest integer m such that

$$\|TR_m f - f\|^2 = \|f\|^2 - \sum_{\mu_n \geq \mu_m} |\langle f, g_n \rangle|^2 \leq f.$$

In particular, for the backwards heat equation example, we have $g_n(x) = \sqrt{\frac{2}{\pi}} \sin nx$ and m, n such that

$$\|f\|^2 - \sum_{n=1}^{m} |b_n|^2 \leq \delta^2,$$

where b_n are the Fourier coefficients of f.

Note: An operator is compact if the kernel $k(s,t)$ is continuous or if the operator has finite rank. An operator is self-adjoint if the kernel $k(s,t)$ is symmetric.

14 Fredholm Alternative in a Hilbert Space

We are now interested in Fredholm integral equations of the second kind,

$$\lambda f(t) - \int_a^b k(t,s) f(s)\, ds = g(t),$$

where $\lambda \neq 0$ (If $\lambda = 0$ we have a first kind integral equation, Section 13). Assume $k(s,t)$ is continuous on $[a,b]\times[a,b]$ and $f, g \in L^2[a,b]$ or more generally operator equations of the form

$$(\lambda I - T)x = y, \tag{14.1}$$

where x and y are in a (infinite dimensional) Hilbert space H and $T : H \to H$ is compact.

Example: We want to find f such that

$$\lambda f(t) - \int_0^1 t f(s)\, ds = g(t), \tag{14.2}$$

if f exists, it is of the form $f(t) = c_1 t + \frac{1}{\lambda} g(t)$. This implies

$$\begin{aligned} & c_1 \lambda t + g(t) - \frac{c_1}{2} - \frac{t}{\lambda}\int_0^1 g(s)\, ds = g(t) \\ \Rightarrow \quad & c_1\left(\lambda - \frac{1}{2}\right) = \frac{1}{\lambda}\int_0^1 g(s)\, ds \\ \Rightarrow \quad & c_1 = \frac{1}{\lambda\left(\lambda - \frac{1}{2}\right)}\int_0^1 g(s)\, ds, \end{aligned}$$

provided that $\lambda \neq \frac{1}{2}$. Hence, we have a solution to (**??**) if $\lambda \neq \frac{1}{2}$, and this solution is unique, by construction. On the other hand, if $\lambda = \frac{1}{2}$, the solution to (**??**) exists, and equals $c_1 t + 2g(t)$ only if $\int_0^1 g(s)\, ds = 0$ and $f(t) = 1$ is a solution of the homogeneous adjoint equation

$$\frac{1}{2} f - \int_0^1 s f(s)\, ds = 0.$$

In particular, g must be orthogonal to solutions of the homogeneous adjoint equation.

Theorem 14.1. *Let $T : H \to H$ be compact and let $v_\lambda := \left\{ x \in H \,\middle|\, Tx = \lambda x \right\}$, and $w_\lambda := \left\{ x \in H \,\middle|\, T^* x = \lambda x \right\}$. Then*

(*a*) *The set $\lambda \in \mathbb{C}$ such that $v_\lambda \neq \{0\}$ is finite or countable and in the latter case its only accumulation point is zero. Moreover, the dimension of the Eigenspace, dim $v_\lambda < \infty$ for all $\lambda \neq 0$.*

(*b*) *If $\lambda \neq 0$, dim v_λ = dim $w_{\overline{\lambda}} < \infty$*

(*c*) *If $\lambda \neq 0$, the range of $\lambda I - T$ is closed.*

Proof. Part (a) is equivalent to the following statement: for all $\epsilon > 0$, the linear span of the spaces v_λ with $|\lambda| \geq \epsilon$ is finite dimensional (note that if $\lambda_1 \neq \lambda_2$, then $v_{\lambda_1} \cap v_{\lambda_2} = \{0\}$ since $T_x = \lambda_1 x$ and $Tx = \lambda_2 x$ implies $\lambda_1 = \lambda_2$ if $x \neq 0$). Suppose to the contrary that there exists $\epsilon > 0$ and an infinite sequence $(x_j) \subset H$ of linearly independent elements such that $Tx_j = \lambda_j x_j$ with $|\lambda_j| \geq \epsilon$ for all j (it may be that $\lambda_{j_0} = \lambda_{j_0+1} = \cdots$).

Since $|\lambda_j| \leq \|T\|$ by passing to a subsequence, we can assume that (λ_j) is a Cauchy Sequence. Let H_m be the linear span of $x_1, \ldots, x_m$ for all m choose $y_m \in H_m$ with $\|y_m\| = 1$ and $y_m \perp H_{m-1}$. Then

$$y_m = \sum_{j=1}^{m} c_{mj} x_j$$

for some constants c_{mj} so $\lambda_m^{-1} T y_m = c_{mm} x_m + \sum_{j=1}^{m-1} c_{mj} \lambda_j \lambda_m^{-1} x_j$ (here, we applied T and divided by λ_m). This equals

$$\lambda_m^{-1} T y_m = y_m + \sum_{j=1}^{m-1} c_{mj}(\lambda_j \lambda_m^{-1} - 1) x_j = y_m (\mathrm{mod} H_{m-1}).$$

If $n < m$, then

$$\lambda_m^{-1} T y_m - \lambda_n^{-1} T y_n = y_m (\mathrm{mod} H_{m-1}).$$

Therefore, since $y_m \perp H_{m-1}$ and $\|y_m\| = 1$, then by orthogonality, we have

$$\|\lambda_m^{-1} T y_m - \lambda_n^{-1} T y_n\| \geq 1.$$

But then

$$1 \leq \|\lambda_m^{-1}(T y_m - T y_n) + (\lambda_m^{-1} + \lambda_n^{-1}) T y_n\|$$
$$1 \leq |\lambda_m^{-1}| \|T y_m - T y_n\| + |\lambda_m^{-1} + \lambda_n^{-1}| \|T y_n\|.$$

This implies

$$\|T y_m - T y_n\| \geq |\lambda_m| - |1 - \lambda_m \lambda_n^{-1}| \|T y_n\|,$$

as $m, n \to \infty$, the second term on the right tends to zero, since $\|T y_n\| \leq \|T\|$ and $\lambda_m \lambda_n^{-1} \to 1$ (recall that (λ_j) is a Cauchy Sequence), and the first term on the right is bounded below by ϵ. Thus $(T y_m)$ has no convergent subsequence contradicting the fact that T is compact. Thus, (a) is proved. ■

Now consider part (b). Given $\lambda \neq 0$, from Theorem 11.5, we can write $T = T_0 + T_1$ where T_0 has finite rank and $\|T_1\| < \lambda$. Then the operator $\lambda I - T_1 = \lambda(I - \lambda^{-1} T_1)$ is invertible (the inverse being given by the Neumann Series $\sum_0^\infty \lambda^{-k-1} T_1^k$). Therefore,

$$(\lambda I - T_1)^{-1}(\lambda I - T) = (\lambda I - T_1)^{-1}(\lambda I - T_0 - T_1) = I - (\lambda I - T_1)^{-1} T_0. \tag{14.3}$$

Set $T_2 = (\lambda I - T_1)^{-1} T_0$, then (**??**) implies $x \in v_\lambda$ is equivalent to $x - T_2 x = 0$. Thus, T_2 conserves finite rank. On the other hand, taking the adjoint of both sides of (**??**) implies

$$(\lambda I - T^*)(\overline{\lambda} I - T_1^*)^{-1} = I - T_2^*,$$

so $y = (\overline{\lambda} I - T_1^*)^{-1} x$ is in w_λ which is equivalent to $x - T_2^* x = 0$. We must therefore show that the equations $x - T_2 x = 0$ and $x - T_2^* x = 0$ have the same finite number of independent solutions. Since T_0 has finite rank, so does T_2 ($T_2 = (\lambda I - T_1)^{-1} T_0$)). Hence, if $u_1, \ldots, u_n$ forms an orthonormal basis for the range of T_2, we can write:

$$T_2 x = \sum_{j=1}^{n} f_j(x) u_j,$$

and $\|T_2 x\|^2 = \sum_{j=1}^n |f_j(x)|^2$ implies $|f_j(x)| \leq \|T_2\| \|x\|$ which implies f_j is a bounded linear functional on H. By the Riesz Representation Theorem, there exists $v_1, \ldots, v_n \in H$ such that

$$T_2 x = \sum_{j=1}^{n} \langle x, v_j \rangle u_j.$$

Set $B_{jk} = \langle u_j, v_k\rangle$ and for $x \in H$, set $\alpha_j = \langle x, v_j\rangle$. If $x - T_2x = 0$, x must be a linear combination of $u_1, \dots, u_n$ and we see that

$$0 = \langle x - T_2x, v_k\rangle = \alpha_k - \sum_{j=1}^{n} B_{jk}\alpha_j, \tag{14.4}$$

for $k = 1, 2, \dots, n$. Conversely, if $\alpha_1, \dots, \alpha_n$ satisfy (**??**), then $x = \sum_{j=1}^n \alpha_j u_j$ satisfies $x - T_2x = \sum_{j=1}^n \alpha_j u_j - \sum_{j=1}^n \langle x, v_j\rangle u_j = \sum_{j=1}^n \alpha_j u_j - \sum_{j=1}^n \alpha_j u_j = 0$. On the other hand, one easily verifies that $T_2^* x = \sum_{j=1}^n \langle x, u_j\rangle v_j$, and so by the same reasoning, $x - T_2^*x = 0$ implies $x = \sum_{j=1}^n \alpha_j v_j$ where

$$\alpha_k - \sum_{j=1}^{n} \overline{B_{jk}}\alpha_j = 0., \tag{14.5}$$

for $k = 1, 2, \dots, n$. But the matrices $(\delta_{jk} - B_{jk})$ and $(\delta_{jk} - B_{jk})$ are adjoints of one another, and so have the same rank (dimension of range). Hence, they have the same nullity i.e. (**??**) and (**??**) have the same number of independent solutions. This proves part (b). ■

Finally, we prove (c). Suppose we have a sequence (y_j) which is contained in the range of $(\lambda I - T)$, which converges to some $y \in H$. We can write $y_j = (\lambda I - T)x_j$, for some $x_j \in H$. If we set $x_j = u_j + v_j$, where $u_j \in v_\lambda$ and $v_j \in v_\lambda^\perp$ (v_λ is closed because it is finite dimensional). We have $y_j = (\lambda I - T)v_j$. We claim that (v_j) is a bounded sequence. Otherwise, by passing to a subsequence we may assume $\|v_j\| \to \infty$. Set $w_j = \frac{v_j}{\|v_j\|}$. then, by passing to another subsequence, we may assume that (Tw_j) converges to a limit z (since T is compact). Since the y_j's are bounded and $\|v_j\| \to \infty$, $y_j = (\lambda I - T)v_j$ implies

$$\lambda w_j = Tw_j | \frac{y_j}{\|v_j\|} \to z, \qquad \text{as } j \to \infty. \tag{14.6}$$

Since $w_j \in v_\lambda^\perp$, we have that $z \in v_\lambda^\perp$. On the other hand,

$$(\lambda I - T)z = \lim_{j\to\infty} (\lambda I - T)\lambda w_j = \lim_{j\to\infty} \frac{\lambda y_j}{\|v_j\|} \to 0.$$

So $z \in v_\lambda$. Hence, $z = 0$, but from (**??**) we have $\|z\| = |\lambda|$, which is a contradiction since $\lambda \neq 0$. Hence, (v_j) is a bounded sequence. By passing to a subsequence and using the fact that T is compact, we can assume that (Tv_j) converges to a limit x. But then

$$y_j = (\lambda I - T)v_j \quad \Rightarrow \quad v_j = \lambda^{-1}(y_j + Tv_j) \to \lambda^{-1}(y + x).$$

Hence,

$$y = \lim_{j\to\infty} (\lambda I - T)v_j = (\lambda I - T)\lambda^{-1}(y + x),$$

which means y is in the range of $(\lambda I - T)$, which proves (c). □

Corollary 14.2. *(Fredholm Alternative) Suppose $\lambda \neq 0$ and $T : H \to H$ is compact. Then*

(a) $(\lambda I - T)x = y$ has a solution i.e. $y \perp w_{\overline{\lambda}}$.

(b) $(\lambda I - T)$ is surjective i.e. $(\lambda I - T)$ is injective.

Proof. Part (a) follows from (c) of Theorem 14.1 and the fact that the closure of the range of a bounded operator is the orthogonal complement of the kernel of its adjoint. Part (b) follows from part (a) of this corollary and part b of Theorem 14.1. □

Remark: The "alternative" of the Fredholm Alternative is either $\mathscr{N}(\lambda I - T) = 0$ and $(\lambda I - T)(H) = H$ or $\mathscr{N}(\lambda I - T) \neq 0$ and $(\lambda I - T)(H) = \left\{ y \in H \,\middle|\, y \in w_{\overline{\lambda}^\perp} \right\}$.

15 The Dirichlet Problem for the Laplace Equation

We consider the following Dirichlet problem

$$\begin{cases} \Delta u = 0 & \text{in} \quad \Omega \\ u = f & \text{on} \quad \partial\Omega \end{cases}.$$

Theorem 15.1. *The solution of the Dirichlet problem, if it exists, is unique.*

We do not prove the uniqueness here, but we can show existence. We will look for a solution in the form of a double layer potential, given by

$$u(x) = \frac{1}{2\pi} \int_{\partial\Omega} \Psi(y) \frac{\partial}{\partial n_y} \log \frac{1}{|x-y|}\, ds_y, \qquad \text{for } x \in \mathbb{R}^2 \setminus \partial\Omega,$$

where s is arclength, and $\Psi \in C(\partial\Omega)$ is known as the density. The following properties of the double layer potential can be shown

(a) The kernel is continuous for $x \in \partial\Omega$.

(b) If $\Omega^+ = \Omega$, $\Omega^- = \mathbb{R}^2 \setminus \overline{\Omega}$, and $x_0 \in \partial\Omega$, define

$$u^+(x_0) = \lim_{x \to x_0, x \in \Omega^+} u(x), \qquad u^-(x_0) = \lim_{x \to x_0, x \in \Omega^-} u(x),$$

then

$$u^+(x_0) - u(x_0) = -\frac{1}{2}\Psi(x_0), \qquad \text{and} \qquad u^-(x_0) - u(x_0) = \frac{1}{2}\Psi(x_0),$$

which describes a jump discontinuity across the boundary.

(c) Lastly,

$$\frac{\partial u^+}{\partial n}(x_0) = \frac{\partial u^-}{\partial n}(x_0).$$

We look for a solution of the Dirichlet problem in the form

$$u(x) = \frac{1}{2\pi} \int_{\partial\Omega} \Psi(y) \frac{\partial}{\partial n_y} \log \frac{1}{|x-y|}\, ds_y, \qquad \text{for } x \in \Omega. \tag{15.1}$$

Letting $x \to \partial\Omega$, and using property (b), implies

$$\Psi(x) - \frac{1}{\pi} \int_{\partial\Omega} \Psi(y) \frac{\partial}{\partial n_y} \log \frac{1}{|x-y|}\, ds_y = -2f(x), \qquad \text{for } x \in \partial\Omega. \tag{15.2}$$

We are done (by the Fredholm Alternative) if we can show that the only solution $\Psi \in L^2(\partial\Omega)$ of the homogeneous equation of (**??**) is $\Psi = 0$. Let Ψ satisfy

$$\Psi(x) - \frac{1}{\pi} \int_{\partial\Omega} \Psi(y) \frac{\partial}{\partial n_y} \log \frac{1}{|x-y|}\, ds_y = 0, \qquad \text{for } x \in \partial\Omega. \tag{15.3}$$

Then $\Psi \in C(\partial\Omega)$. Define u by (**??**) for this Ψ. By uniqueness of the solution to the Dirichlet problem, $u(x) = 0$ for $x \in \Omega^+$, and we have $\frac{\partial u^+}{\partial n}(x) = 0$ for $x \in \partial\Omega$. Now consider $u(x)$ for $x \in \mathbb{R}^2 \setminus \overline{\emptyset}$. Property ($c$) implies

$$\frac{\partial u^-}{\partial n}(x) = 0,$$

for $x \in \partial\Omega$. Furthermore, $u(x) = \mathcal{O}\left(\frac{1}{r}\right)$, $\frac{\partial}{\partial r} u(r) = \mathcal{O}\left(\frac{1}{r^2}\right)$, as $r = |x| \to \infty$. Hence, from Green's Formula

$$\iint_{\Omega} (u\Delta v - \nabla u \cdot \nabla v)\, dx = \int_{\partial\Omega} u \frac{\partial v}{\partial n}\, ds,$$

for any domain Ω, where n is the unit outward normal to Ω, we have $\Omega = \Omega_R \setminus D$, where $\Omega = \Omega_R \setminus D$ where $\Omega_R := \{x : |x| \leq R\}$, and $v = u$,

$$\iint_{\Omega_R \setminus D} |\nabla u|^2 \, dx = -\int_{\partial \Omega_R} \underbrace{u}_{\mathcal{O}(\frac{1}{r})} \underbrace{\frac{\partial u}{\partial n}}_{\mathcal{O}(\frac{1}{r^2})} \, ds \qquad \Rightarrow \qquad \iint_{\Omega_R \setminus D} |\nabla u|^2 \, dx = \mathcal{O}\left(\frac{1}{R^2}\right).$$

Letting $R \to \infty$, implies u is a constant in $\mathbb{R}^2 \setminus D$. But $u(x) \to 0$ as $R \to \infty$. This constant is zero, i.e. $u(x) = 0$ in $\mathbb{R}^2 \setminus \overline{D}$. By property (b) we have $\Psi(x) = 0$ for $x \in \partial\Omega$ and we are done.