ANDREW TULLOCH

# FUNCTIONAL ANALYSIS

SCHOOL OF MATHEMATICS AND STATISTICS
THE UNIVERSITY OF SYDNEY

# Contents

# 1

# Banach Spaces and Linear Operators

## 1.1 Banach Spaces

**Definition 1.1** (Norm)**.** Let $X$ be a vector space. A norm on $X$ is a function $\|\cdot\| : X \mapsto \mathbb{R}$ satisfying

- $\|x\| \geq 0$ with equality if and only if $x = 0$.
- $\|\alpha x\| = |\alpha|\|x\|$.
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

We call the pair $(X, \|\cdot\|)$ a **normed vector space.**

**Theorem 1.2** (Reverse triangle inequality)**.** *Let $X$ be a normed vector space. For any $x, y \in X$, we have*

$$|\|x\| - \|y\|| \leq \|x - y\|$$

**Definition 1.3** (Complete space)**.** Let $X$ be a normed vector space. Then $X$ is **complete** if every Cauchy sequence in $X$ converges to some $x \in X$.

**Definition 1.4** (Banach space)**.** A **Banach space** is a complete normed vector space.

**Proposition 1.5** (Convergence)**.** *Let $(V, \|\cdot\|)$ be a normed vector space. A sequence $(x_n)$ in $V$ converges to $x \in V$ if given $\epsilon > 0$, there exists $N$ such that $\|x - x_n\| < \epsilon$ whenever $n < N$.*

**Lemma 1.6.** *If $x_n \to x$, then $\|x_n\| \to \|x\| \in \mathbb{R}$.*

*Proof.* $|\|x_n\| - \|x\|| \leq \|x - x_n\| \to 0$. □

**Proposition 1.7.** *Every convergent sequence is Cauchy.*

**Definition 1.8** (Banach space)**.** A complete, normed, vector space is called a **Banach space**

**Proposition 1.9.** *$(\mathbb{K}, |\cdot|)$ is complete.*

**Proposition 1.10.** *$(\ell^p, \|\cdot\|_p)$ is a Banach space for all $1 \leq p \leq \infty$.*[1]

**Proposition 1.11.** *$(\ell([a,b]), \|\cdot\|_\infty)$ is a Banach space*

**Proposition 1.12.** *If $1 \leq p < \infty$, then $(\ell([a,b]), \|\cdot\|_p)$ is **not** a Banach space.*

*Proof.* Consider a sequence of functions that is equal to one on $[0, \frac{1}{2}]$, zero on $[\frac{1}{2} + \frac{1}{n}, 1]$, and linear between. This is a Cauchy sequence that does not converge to a continuous function. □

We've seen that $(\ell([a,b]), \|\cdot\|_p)$ is not complete for $1 \leq p < \infty$.

**Theorem 1.13** (Completion)**.** *Let $(V, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$. There exists a Banach space $(V_1, \|\cdot\|_1)$ such that $(V, \|\cdot\|)$ is isometrically isomorphic to a dense subspace of $(V_1, \|\cdot\|_1)$.*

*Furthermore, the space $(V_1, \|\cdot\|_1)$ is unique up to isometric isomorphisms.*[2]

**Definition 1.14.** $(V_1, \|\cdot\|_1)$ is called **the completion** of $(V, \|\cdot\|)$.

**Definition 1.15** (Dense)**.** If $X$ is a topological space and $Y \subseteq X$, then $Y$ is **dense** in $X$ if the closure of $Y$ in $X$ equals $X$, that is, $\overline{Y} = X$.

Alternatively, for each $x \in X$, there exists $(y_n)$ in $Y$ such that $y_n \to x$.

**Definition 1.16** (Isomorphism of vector spaces)**.** Two normed vector spaces $(X, \|\cdot\|X)$ and $(Y, \|\cdot\|Y)$ are **isometrically isomorphic** if there is a vector space isomorphism $\Psi : X \to Y$ such that

$$\|\Psi(x)\|_Y = \|x\|_X \quad \forall x \in X$$

**Example 1.17.** Let $\ell_0 = \{(x_i) \mid \#\{i, x_i \neq 0\} < \infty\}$. The completion of $\ell_0, \|\cdot\|_p$ is $(\ell^p, \|\cdot\|_p)$, because,

- $\ell_0$ is a subspace of $\ell^p$,
- It is dense, since we can easily construct a sequence in $\ell_0$ converging to arbitrary $x \in \ell^p$.

**Example 1.18** ( $L^p$ spaces)**.** Let $\mu$ be the Lebesgue measure on $\mathbb{R}$. Let

$$\mathcal{L}^p([a,b]) = \{\text{measurable } f : [a,b] \to \mathbb{K} \mid \int_a^b |f|^p \, d\mu < \infty\}$$

Let $\|f\|_p = \left(\int_a^b |f|^p \, d\mu\right)^{1/p}$. Since $\|f\|_p = 0 \iff f = 0 \, a.e$, we quotient out by the rule $f \equiv g \iff f - g = 0 \, a.e.$, and then our space of equivalence classes forms a normed vector space, denoted $L^p([a,b])$.

[1] A general proof outline follows.

- Use completeness of $\mathbb{R}$ to find a candidate for the limit.
- Show this limit function is in $V$.
- Show that $x_n \to x$ in $V$.

Let $x^{(n)}$ be a Cauchy sequence in $\ell^p$. Since $|x_j^{(n)} - x_j^{(n)}| \leq \|x^{(n)} - x^{(m)}\|$, we know that $x_j^{(n)}$ is a Cauchy sequence in $\mathbb{K}$. Hence, $\lim_{n\to\infty} x_j^{(n)} := x_j$ exists, and is our limit candidate.

We need only then show that $\sum_{j=1}^{\infty} |x_j|^p < \infty$.

[2] The proof of this fact is rather straightforward;

(i) construct Cauchy sequences,
(ii) append limits,
(iii) quotient out (as different sequences may converge to the same limit)

**Theorem 1.19** (Riesz-Fischer). *$(L^p([a,b]), \|\cdot\|_p)$ is the completion of $(\mathcal{C}[a,b], \|\cdot\|_p)$, and is a Banach space.*[3]

[3] This fact follows quite easily from the definition of the Lebesgue integral.

*Remark.*

- Let $X$ be any compact topological space, let $\mathcal{C}(X) = \{f : X \to \mathbb{K} \mid f \text{ is continuous}\}$, and let $\|f\|_\infty = \sup_{x \in X} \|f(x)|$. Then $\mathcal{C}(X, \|\cdot\|_\infty)$ is Banach.
- Let $X$ be any topological space. Then the set of all continuous and bounded functions with the supremum norm forms a Banach space.
- Let $(S, \mathcal{A}, \mu)$ be a measure space. Then we can define the $\mathcal{L}^p$ and $L^p$ analogously, and they are also Banach.

## 1.2 *Linear Operators*

**Definition 1.20** (Linear operators on normed vector spaces)**.** Let $X, Y$ be vector spaces over $\mathbb{K}$. A linear operator is a function $T : X \to Y$ such that

$$T(x+y) = T(x) + T(y)$$
$$T(\alpha x) = \alpha T(x)$$

for all $x, y \in X, \alpha \in \mathbb{K}$.

We write $\text{Hom}(X,Y) = \{T : X \to Y \mid T \text{ is linear}\}$

**Definition 1.21.** $T : X \to Y$ is continuous at $x \in X$ if for all $\epsilon > 0$, there exists $\delta > 0$ such that

$$\|x - y\|_X < \delta \Rightarrow \|Tx - Ty\|_Y < \epsilon$$

**Definition 1.22.**

$$\mathcal{L}(X,Y) = \{T : X \to Y \mid T \text{ is linear and continuous}\}$$

*Remark.* If $\dim(X) < \infty$ then $\text{Hom}(X,Y) = \mathcal{L}(X,Y)$. This is **not** true if $X$ has infinite dimension.

**Definition 1.23** (Bounded linear operator)**.** Let $T : X \to Y$ be linear, then $T$ is **bounded** if $T$ maps bounded sets in $X$ to bounded sets in $Y$. That is: for each $M > 0$ there exists $M' > 0$ such that

$$\|x\|_X \leq M \Rightarrow \|Tx\|_Y \leq M'$$

Consider the space $\mathcal{L}(X,Y)$, the set of all linear and continuous maps between two normed vector spaces $X$ and $Y$.

**Theorem 1.24** (Fundamental theorem of linear operators)**.** *Let $(X, \|\cdot\|_X)$ and $Y, \|\cdot\|_Y$ be normed vector spaces. Let $T \in Hom(X,Y)$, the set of all linear maps from $X$ to $Y$. Then the following are all equivalent.*

*1) T is uniformly continuous*

*2) T is continuous*

*3) T is continuous at 0*

*4) T is bounded*

*5) There exists a constant $c > 0$ such that*

$$\|Tx\|_Y \leq c\|x\|_X \quad \forall x \in X$$

*Proof.*

1) $\Rightarrow$ 2) $\Rightarrow$ 3) is clear.

3) $\Rightarrow$ 4). Since $T$ is continuous at 0, given $\epsilon = 1 > 0$, there exists $\delta$ such that

$$\|Tx - T0\| \leq 1 \quad \text{whenever} \quad \|X - 0\| \leq \delta,$$

i.e. that $\|x \leq \delta \Rightarrow \|Tx\| \leq 1$. Let $y \in X$. The $\|\frac{\delta y}{\|y\|}\| \leq \delta$, and so $\|T\left(\frac{\delta y}{\|y\|}\right)\| <\leq 1$. Hence,

$$\frac{\delta}{\|y\|}\|Ty\| \leq 1$$

and so

$$\|Ty\| \leq \frac{\|y\|}{\delta}$$

for all $y \in X$. Thus, for all $\|y\| \leq M$, we have $\|Ty\| \leq M'$, where $M' = \frac{M}{\delta}$, and so $T$ is bounded.

4) $\Rightarrow$ 5). If $T$ is bounded, given $M = 1 > 0$, there exists $c \geq 0$ such that $\|x\| \leq 1 \Rightarrow \|Tx\| \leq c$. Then

$$\left\|T\left(\frac{x}{\|x\|}\right)\right\| \leq c$$

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

5) $\Rightarrow$ 1). If 5) holds, then

$$\|Tx - Ty\| = \|T(x-y)\| \leq c\|x - y\|.$$

So if $\epsilon$ is given, taking $\delta = \frac{\epsilon}{c}$, we have

$$\|Tx - Ty\| \leq c\|x - y\| < c\frac{\epsilon}{c} = \epsilon. \qquad \square$$

**Corollary.** *If $T \in Hom(X, Y)$, then $T$ continuous $\iff$ $T$ bounded $\iff$ $\|Tx\| \leq c\|x\|$ for all $x \in X$.*

**Definition 1.25** (Operator norm)**.** The **operator norm** of $T \in \mathcal{L}(x, y)$, $\|T\|$ is defined by any one of the following equivalent expressions.

(a) $\|T\| = \inf\{c > 0 \mid \|Tx\| < c\|x\|\}$.

(b) $\|T\| = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|}$.

(c) $\|T\| = \sup_{\|x\|\leq 1} \|Tx\|$.

(d) $\|T\| = \sup_{\|x\|=1} \|Tx\|$.

**Proposition 1.26.** *The operator norm is a norm on $\mathcal{L}(x, y)$.*

*Proof.* The following are simple to verify.

(a) $\|T\| \geq 0$, with equality if and only if $T = 0$.

(b) $\|\alpha T\| = |\alpha| \|T\|$.

(c) $\|S + T\| \leq \|S\| + \|T\|$.

□

**Example 1.27** (Calculating $\|T\|$)**.** To calculate $\|T\|$, try the following.

1) Make sensible calculations to find $c$ such that

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

for all $x \in X$.

2) Find $x \in X$ such that $\|Tx\| = c\|x\|$.

**Definition 1.28** (Algebraic dual)**.** Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$. The **algebraic dual** of $X$ is

$$X^\star = \text{Hom}(X, \mathbb{K}) = \{\varphi : X \to \mathbb{K} \mid \varphi \text{ is linear}\}.$$

Elements of $X^\star$ are called linear functionals.

**Definition 1.29** (Continuous dual (just **dual**))**.** The **continuous dual** (just dual) of $X$ is

$$X' = \mathcal{L}(X, \mathbb{K}) = \{\varphi : X \to K \mid \varphi \text{ is linear and continuous}\}.$$

*Remark.* $X^\star \supsetneq X'$ if $\dim(X) = \infty$.

**Example 1.30.** Let $(\wp([a, b]), \|\cdot\|_\infty)$ be the normed vector space of polynomials $p : [a, b] \to \mathbb{K}$.

(a) The functional $D : \wp([0, 1]) \to \mathbb{K}$ given by $D(p) = p'(1)$ is linear, but **not** continuous.

(b) The functional $I : \wp([0, 1]) \to \mathbb{K}$ given by $I(p) = \int_0^1 p(t)\, dt$ is linear **and** continuous.

*Proof.* (a) Linearity is clear. Then consider the sequence of functions $p_n(t) = t^n$ for all $t \in [0, 1]$. Then $|D(p_n)| = n\|p_n\|_\infty$. So $D$ is not continuous, as continuity implies that there exists $c$ such that

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

(b) Exercise: Show $\|I\| = 1$.

$\square$

Describing the continuous dual space $X'$ is one of the first things to do when trying to understand a normed vector space. It is generally pretty difficult to describe $X'$.

**Proposition 1.31** (Dual of the $\ell^p$ space for $(1 < p < \infty)$)**.** *Let $1 < p < \infty$. Let $q$ be the "dual" of $p$, defined by $\frac{1}{q} + \frac{1}{p} = 1$. Then $(\ell^p)'$ is isometrically isomorphic to $\ell^q$.*

*Remark* (Observation before proof). Let $1 \leq p < \infty$. Let $e_i = (0, 0, \dots, 1, 0, \dots)$ where 1 is in the $i$-th place.

1) If $x = (x_i) \in \ell^p$, then

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

in the sense that the partial sums converge to $x$.

2) If $\varphi : \ell^p \to \mathbb{K}$ is linear and continuous, then

$$\varphi(x) = \sum_{i=1}^{\infty} x_i \varphi(e_i)$$

*Proof of observations.* Let $S_n = \sum_{i=1}^{n} x_i e_i$. Then

$$\begin{aligned} \|x - S_n\|_p^p &= \|(0, 0, \dots, x_{n+1}, x_{n+2}, \dots)\|_p^p \\ &= \sum_{i=n+1}^{\infty} |x_i|^p \\ &\to 0 \quad \text{as it is the tail of a convergent sum.} \end{aligned}$$

Write $\varphi(x)$ as

$$\begin{aligned} \varphi(x) &= \varphi(\lim_{n\to\infty} S_n) \quad \text{(continuity)} \\ &= \lim_{n\to\infty} (\varphi(S_n)) \\ &= \lim_{n\to\infty} \varphi\left(\sum_{i=1}^{n} x_i e_i\right) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} x_i \varphi(e_i) \quad \text{(linearity)} \\ &= \sum_{i=1}^{\infty} x_i \varphi(e_i) \square \end{aligned}$$

*Proof.* Define a map $\theta$ by

$$\begin{aligned} \theta : \ell^q &\to (\ell^p)' \\ y &\mapsto \varphi_y \end{aligned}$$

where $\varphi_y(x) = \sum x_i y_i$ for all $x \in \ell^p$.

(1) $\varphi_y$ is linear, as $\varphi_y(x+x') = \varphi_y(x) + \varphi_y(x')$ (valid as sums converge absolutely.)

(2) $\varphi_y$ is continuous, as

$$|\varphi_y(x)| = |\sum x_i y_i| \leq \sum |x_i y_i| \leq \|x\|_p \|y\|_q$$

by Hölder's inequality. From the fundamental theorem of linear operators, as $|\varphi_y(x)| \leq \|x\|_p \|y\|_q$, we have that $\varphi_y$ is continuous, and that

$$\|\varphi_y\| \leq \|y\|_q \qquad (\star)$$

(3) $\theta$ is linear.

(4) $\theta$ is injective, as

$$\theta(y) = \theta(y') \Rightarrow \varphi_y = \varphi_{y'} \Rightarrow \varphi_y(x) = \varphi_{y'}(x) \quad \forall x \in \ell^p$$

$$\Rightarrow \varphi_y(e_i) = \varphi_{y'}(e_i) \quad \forall i \in \mathbb{N} \Rightarrow y_i = y_i' \quad \forall i \in \mathbb{N} \Rightarrow y = y'$$

(5) $\theta$ is surjective. Let $\varphi \in (\ell^p)$. Let $y = (\varphi(e_1), \dots, \varphi(e_n), \dots) = (y_1, \dots, y_n, \dots)$. We now show $y \in \ell^q$.

Let $x^{(n)} \in \ell^q$ be defined by

$$x_i^{(n)} = \begin{cases} \frac{|y_i|^q}{y_i} & \text{if } i \leq n \text{ and } y_i \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\varphi(x^{(n)}) = \sum_{i=1}^{\infty} x_i^{(n)} \varphi(e_i) = \sum_{i=1}^{n} |y_i|^q \qquad (\dagger)$$

by Observation 2) above.

On the other hand, we know

$$\begin{aligned} \|\varphi(x^{(n)}) &\leq \|\varphi\| \|x^{(n)}\|_p \\ &= \|\varphi\| \left( \sum_{i=1}^{\infty} |x_i^{(n)}|^p \right)^{1/p} \\ &= \|\varphi\| \left( \sum_{i=1}^{n} |y_i|^{(q-1)p} \right)^{1/p} \\ &= \|\varphi\| \left( \sum_{i=1}^{n} |y_i|^q \right)^{1/p} \qquad \text{as } 1/p + 1/q = 1. \qquad (\star\star) \end{aligned}$$

Now, using $(\dagger)$ and $(\star\star)$, we have

$$\sum_{i=1}^{n} |y_i|^q \leq \|\varphi\| \left( \sum_{i=1}^{n} |y_i|^q \right)^{1/p}$$

and so we must have

$$\|y\|_q \leq \|\varphi\| \qquad (\star\star\star)$$

and so $y \in \ell^q$.

We also have, by $(\star\star)$,

$$\|y\|_q \leq \|\varphi_y\|$$

(6) Finally, we show that $\theta$ is an isometry. By $(\star)$ and $(\star\star\star)$, we have

$$\|\theta(y)\| = \|\varphi_y\| = \|y\|_q$$

as required. □

How big is $X'$? When is $X' \neq \{0\}$? Examples suggest that $X'$ is big with a rich structure. The following chapter gives a fundamental theorem regardig

# 2
# The Hahn-Banach Theorem

The Hahn-Banach theorem is a cornerstone of functional analysis. It is all about extending linear functionals defined on a subspace to linear functionals on the whole space, while preserving certain properties of the original functional.

**Definition 2.1** (Seminorm)**.** A let $X$ be a vector space over $\mathbb{K}$. A seminorm on $X$ is a function $p : X \to \mathbb{R}$ such that

(1) $p(x+y) \leq p(x) + p(y) \quad \forall x, y \in X$

(2) $p(\lambda x) = |\lambda| p(x) \quad \forall x \in X, \lambda \in \mathbb{K}$

**Theorem 2.2** (General Hahn-Banach)**.** *Let $X$ be a vector space over $\mathbb{K}$. Let $p : X \to \mathbb{R}$ be a seminorm on $X$. Let $Y \subseteq X$ be a subspace of $X$. If $f : Y \to \mathbb{K}$ is a linear functional such that*

$$|f(y)| \leq p(y) \quad \forall y \in Y$$

*then there is an extension $\tilde{f} : X \to \mathbb{K}$ such that*

- *$\tilde{f}$ is linear*
- $\tilde{f}(y) = f(y) \quad \forall y \in Y$
- $|f(x)| \leq p(x) \quad \forall x \in X$

*Remark.* This is great.

- $Y$ can be finite dimensional (and we know about linear functionals on finite dimensional spaces).
- If $p(x) = \|x\|$, then

  $$|\tilde{f}(x)| \leq \|x\| \quad \forall x \in X$$

  and so $\tilde{f} \in X'$

**Corollary.** *Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$. For each $y \in X$, with $y \neq 0$, there is $\varphi \in X'$ such that*

$$\varphi(y) = \|y\| \quad \text{and} \quad \|\varphi\| = 1$$

*Proof.* Fix $y \neq 0$ in $X$. Let $Y = \{\mathbb{K}y\} = \{\lambda y \,|\, \lambda \in \mathbb{K}\}$, a one-dimensional subspace.

Define $f : Y \to \mathbb{K}$, $f(\lambda y) = \lambda \|y\|$. This is linear. Set $p(x) = \|x\|$. Then

$$|f(\lambda y)| = p(\lambda y)$$

and so by Hahn-Banach, there exists $\tilde{f} : X \to \mathbb{K}$ such that

- $\tilde{f}$ is linear
- $\tilde{f}(\lambda y) = f(\lambda y) \quad \forall \lambda \in \mathbb{K}$
- $|\tilde{f}(x)| \leq \|x\| \quad \forall x \in X$

Then we have $\tilde{f} \in X'$ and $\|f\| = 1$ as required. □

## 2.1 *Zorn's Lemma*

**Theorem 2.3** (Axiom of Choice is equivalent to Zorn's Lemma)**.** *See handout for proof that*

$$A.C. \Rightarrow Z.L.$$

**Definition 2.4** (Partially ordered set)**.** A **partially ordered set** (poset) is a set $A$ with a relation $\leq$ such that

(1) $a \leq a$ for all $a \in A$,

(2) If $a \leq b$ and $b \leq a$then $a = b$,

(3) If $a \leq b$ and $b \leq c$, then $a \leq c$

**Definition 2.5** (Totally ordered set)**.** A **totally ordered set** is a poset $(A, \leq)$ such that if $a, b \in A$ then either $a \leq b$ or $b \leq a$.

**Definition 2.6** (Chain)**.** A **chain** in a poset $(A, \leq)$ is a totally ordered subset of $A$.

**Definition 2.7** (Upper bound)**.** Let $(A, \leq)$ be a poset. An **upper bound** for $B \subseteq A$ is an element $u \in A$ such that $b \leq u$ for all $b \in B$.

**Definition 2.8** (Maximal element)**.** A **maximal element** of a poset $(A, \leq)$ is an element $m \in A$ such that $m \leq x$ implies $x = m$, that is,

$$m \leq x \Rightarrow x = m$$

**Example 2.9.** Let $S$ be any set. Let $\mathcal{P}(S)$ be the power set of $S$ (the set of all subsets of $S$). Define $a \leq b \iff a \subseteq b$. Maximal element is $S$

**Theorem 2.10** (Zorn's Lemma)**.** *Let $(A, \leq)$ be a poset. Suppose that every chain in $A$ has an upper bound. Then $A$ has (at least one) maximal element.*

**Definition 2.11** (Linearly independent)**.** Let $X$ be a vector space over $\mathbb{F}$. We call $B \subseteq X$ **linearly independent** if

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = 0 \Rightarrow \lambda_1 = \cdots = \lambda_n = 0$$

for all finite $\{x_1, \ldots, x_n\} \subseteq B$.

**Definition 2.12** (Span)**.** We say $B \subseteq X$ **spans** $X$ if each $x \in X$ can be written as

$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n$$

for some $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and $\{x_1, \ldots, x_n\} \subseteq B$.

**Definition 2.13** (Hamel basis)**.** A Hamel basis is a linearly independent spanning set. Equivalently, $B \subseteq X$ is a Hamel basis if and only if each $x \in X$ can be written in exactly one way as a finite linear combination of elements of $B$.

**Theorem 2.14.** *Every vector space has a Hamel basis.*

*Proof.* Let $L = \{\text{linearly independent subsets}\}$, with subset ordering. Let $C$ be a chain in $L$. Let $u = \bigcup_{a \in C} a$. Then

(1) $u \in L$,

(2) $u$ is an upper bound for $C$.

So Zorn's Lemma says that $L$ has a maximal element $\mathbf{b}$.

Then $\mathbf{b}$ is a Hamel basis.

- $\mathbf{b}$ is linearly independent.

- If $\mathrm{Span}(\mathbf{b}) \neq X$, there exists $X \in X \backslash \mathrm{Span}(\mathbf{b})$, and $\mathbf{b}' = \mathbf{b} \bigcup \{x\} \in L$ is linearly independent, contradicting maximality of $\mathbf{b}$. □

**Theorem 2.15.** *If $(X, \|\cdot\|)$ is Banach, every Hamel basis is uncountable.*

**Theorem 2.16** (Hahn-Banach theorem over $\mathbb{R}$)**.** *Let $X$ be a real linear space and let $p(x)$ be a seminorm on $X$. Let $M$ be a real linear subspace of $X$ and $f_0$ a real-valued linear functional defined on $M$. Let $f_0$ satisfy $f_0(x) \leq p(x)$ on $M$. Then there exists a real valued linear functional $F$ defined on $X$ such that*

*(i) $F$ is an extension of $f_0$, that is, $F(x) = f_0(x)$ for all $x \in M$, and*

*(ii) $F(x) \leq p(x)$ on $X$.*

*Proof.* We first show that $f_0$ can be extended if $M$ has codimension one. Let $x_0 \in X \backslash M$ and assume that $\text{span}(M \cup \{x_0\}) = X$. As $x_0 \notin M$ be can write $x \in X$ uniquely in the form

$$x = m + \alpha x_0$$

for $\alpha \in \mathbb{R}$. Then for every $c \in \mathbb{R}$, the map $f_c \in \text{Hom}(X, \mathbb{R})$ given by $f_c(m + \alpha x) = f_0(m) + c\alpha$ is well defined, and $f_c(m) = f_0(m)$ for all $m \in M$. We now show that we can choose $c \in \mathbb{R}$ such that $f_c(x) \le p(x)$ for all $x \in X$. Equivalently, we must show

$$f_0(m) + c\alpha \le p(m + \alpha x_0)$$

for all $m \in M$ and $\alpha \in \mathbb{R}$. By positive homogeneity of $p$ and linearity of $f$ we have

$$\begin{aligned} f_0(m/\alpha) + c &\le p(x_0 + m/\alpha) \quad \alpha > 0 \\ f_0(-m/\alpha) - c &\le p(-x_0 - m/\alpha) \quad \alpha < 0 \end{aligned}$$

Hence we need to choose $c$ such that

$$\begin{aligned} c &\le p(x_0 + m) - f_0(m) \\ c &\ge -p(-x_0 + m) + f_0(m). \end{aligned}$$

This is possible if

$$-p(-x_0 + m_1) + f_0(m_1) \le p(x_0 + m_2) - f_0(m_2)$$

for all $m_1, m_2 \in M$. By subadditivity of $p$ we can verify this condition since

$$f_0(m_1 + m_2) \le p(m_1 m_2) = p(m_1 - x_0 + m_2 - x_0) \le p(m_1 - x_0) + p(m_2 + x_0)$$

for all $m_1, m_2 \in M$. Hence $c$ can be chosen as required.

Hence $D(F) = X$, and the theorem is proven. □

**Theorem 2.17** (Hahn-Banach over $\mathbb{C}$)**.** *Suppose that c is a seminorm on a complex vector space X and let M sub a subspace of X. If* $f_0 \in Hom(M, \mathbb{C})$ *is such that* $|f_0(x)| \le p(x)$ *for all* $x \in M$*, then there exists an extension* $f \in Hom(X, \mathbb{C})$ *such that* $f|_M = f_0$ *and* $|f(x)| \le p(x)$ *for all* $x \in X$*.*

*Proof.* Split $f_0$ into real and imaginary parts

$$f_0(x) = g_0(x) + ih_0(x).$$

By linearity of $f_0$ we have

$$\begin{aligned} 0 = if_0(x) - f_0(ix) &= ig_0(x) - h_0(x) - g_0(ix) - ih_0(ix) \\ &= -(g_0(ix) + h_0(x)) + i(g_0(x) - h_0(ix)) \end{aligned}$$

and so $h_0(x) = -g_0(ix)$. Therefore,

$$f_0(x) = g_0(x) - ig_0(ix)$$

for all $x \in M$. We now consider $X$ as a vector space over $\mathbb{R}$, $X_{\mathbb{R}}$. Now considering $M_{\mathbb{R}}$ as a subspace of $X_{\mathbb{R}}$. GSince $g_0 \in \text{Hom}(M_{\mathbb{R}}, \mathbb{R})$ and $g_0(x) \leq |f_0(x)| \leq p(x)$ and so by the real Hahn-Banach, there exists $g \in \text{Hom}(X_{\mathbb{R}}, \mathbb{R})$ such that $g|_{M_{\mathbb{R}}} = g_0$ and $g(x) \leq p(x)$ for all $x \in X_{\mathbb{R}}$. Now set $F(x) = g(x) - ig(ix)$ for all $x \in X_{\mathbb{R}}$. Then by showing $f(ix) = if(x)$, we have that $f$ is linear.

We now show $|f(x)| \leq p(x)$. For a fixed $x \in X$ choose $\lambda \in \mathbb{C}$ such that $\lambda f(x) = |f(x)|$. Then since $|f(x)| \in \mathbb{R}$ and by definition of $f$, we have

$$|f(x)| = \lambda f(x)| = f(\lambda x) = g(\lambda x) \leq p(\lambda x) = |\lambda p(x) = p(x)$$

as required. □

# 3
# An Introduction to Hilbert Spaces

## 3.1 Hilbert Spaces

**Definition 3.1** (Inner product)**.** Let $X$ be a vector space over $\mathbb{K}$. An **inner product** is a function

$$\langle\cdot,\cdot\rangle : X \times X \to \mathbb{K}$$

such that

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

(2) $\langle \alpha x, z\rangle = \alpha\langle x, z\rangle$

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

(4) $\langle x, x\rangle \geq 0$ with equality if and only if $x = 0$

We then have

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

and

$$\langle x, \alpha z\rangle = \overline{\alpha}\langle x, z\rangle$$

**Definition 3.2** (Inner product space)**.** Let $(X, \langle\cdot,\cdot\rangle)$ be an **inner product space**. Defining $\|x\| = \sqrt{\langle x, x\rangle}$ turns $X$ into a normed vector space. To prove the triangle inequality, we use the Cauchy-Swartz theorem.

**Theorem 3.3** (Cauchy-Schwarz)**.** *In an inner product space* $(X, \langle\cdot,\cdot\rangle)$*, we have*

$$|\langle x, y\rangle| \leq \|x\|\|y\| \quad \forall x, y \in X$$

*Proof.*

$$\begin{aligned}
0 &\leq \langle x-\lambda y, x-\lambda y\rangle \\
&= \langle x, x\rangle - \langle x, \lambda y\rangle - \langle \lambda y, x\rangle + \langle \lambda y, \lambda y\rangle \\
&= \|x\|^2 - \bar{\lambda}\langle x, y\rangle - \lambda\langle y, x\rangle + |\lambda|^2\|y\|^2 \\
&= \|x\|^2 - 2\mathrm{Re}(\lambda\langle y, x\rangle) + |\lambda|^2\|y\|^2
\end{aligned}$$

Set $\lambda = \frac{\langle x,y \rangle}{\|y\|^2}$. Then

$$\begin{aligned} 0 &\le \|x\|^2 - 2\text{Re}\left(\frac{|\langle x,y\rangle|^2}{\|y\|^2}\right) + \frac{|\langle x,y\rangle|^2}{\|y\|^2} \\ &= \|x\|^2 - \frac{|\langle x,y\rangle|^2}{\|y\|^2} \end{aligned}$$

as required. □

**Corollary.**

$$\|x+y\| \le \|x\| + \|y\|$$

**Definition 3.4** (Hilbert space)**.** If $(X, \langle \cdot, \cdot \rangle)$ is complete with respect to $\|\cdot\|$ then it is called a **Hilbert space.**

**Example 3.5.**

(a) $\ell^2$, where $\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$.

Cauchy-Schwarz then says

$$|\sum_{i=1}^{\infty} x_i \overline{y_i}| \le \sqrt{\sum_{i=1}^{\infty} |x_i|^2} \sqrt{\sum_{i=1}^{\infty} |y_i|^2}$$

(b) $L^2([a,b])$, where $\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\,dx$.

Cauchy-Swartz then says

$$|\int_a^b f(x)\overline{g(x)}\,dx \le .....$$

**Definition 3.6** (Orthogonality)**.** Let $(X, \langle \cdot, \cdot \rangle)$ be inner product spaces. Then $x, y \in X$ are orthogonal if $\langle x, y \rangle = 0$ where $x, y \neq 0$.

**Theorem 3.7.** *Let $x_i, \ldots, x_n$ be pairwise orthogonal elements in $(X, \langle \cdot, \cdot \rangle)$. Then*

$$\|\sum_{i=1}^{n} x_i\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

**Theorem 3.8** (Parallelogram identity)**.** *In $(X, \langle \cdot, \cdot \rangle)$ we have*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \qquad (\star)$$

*for all $x, y \in X$.*

*Remark.* If $(X, \|\cdot\|)$ is a normed vector space which satisfies parallelogram identity then $X$ is an inner product space with inner products defined by the polarisation equation

$$\langle x, y \rangle = \begin{cases} \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) & \mathbb{K} = \mathbb{R} \\ \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right) & \mathbb{K} = \mathbb{C} \end{cases}$$

## 3.2 Projections

**Definition 3.9** (Projection)**.** Let $X$ be a vector space over $\mathbb{K}$. A subset $M$ of $X$ is convex if for any $x, y \in M$, then

$$tx + (1-t)y \in M \quad \forall t \in [0,1]$$

**Theorem 3.10** (Projection)**.** *Let $(\mathcal{H}, \langle \cdot, \cdot, \rangle)$ be a Hilbert space. Let $M \subseteq \mathcal{H}$ be closed and convex. Let $x \in \mathcal{H}$. Then there exists a unique point $m_x \in M$ which is closest to $x$, i.e.*

$$\|x - m_x\| = \inf_{m \in M} \|x - m\| = d$$

*Proof.* For each $k \geq 1$ choose $m_k \in M$ such that

$$d^2 \leq \|x - m_k\|^2 \leq d^2 + \frac{1}{k}$$

Each $m_k$ exists as $d$ is defined as the infimum over all $m$.

Then

$$\begin{aligned} \|m_k - m_l\|^2 &= \|(m_k - x) - (m_k - x)\|^2 \\ &= 2\|m_k - x\|^2 + 2\|m_l - x\|^2 - \|m_k + m_l - 2x\|^2 \\ &\leq 2d^2 + \frac{2}{l} + 2d^2 + \frac{2}{k} - 4\|\frac{m_k + m_l}{2} - x\|^2 \end{aligned}$$

and as $m_k/2 + m_l/2 \in M$, we have $\|\frac{m_k+m_l}{2} - x\|^2 \geq d^2$. Then

$$\|m_k - m_l\|^2 \leq 2\left(\frac{1}{k} + \frac{1}{l}\right)$$

Thus $(m_k)$ is Cauchy. So $m_k \to m_x \in M$ as $\mathcal{H}$ is complete and $M$ is closed. We then have

$$\|x - m_x\| = d$$

and so now we show that $m_x$ is unique.

Suppose that there exists $m'_x \in M$ with $\|x - m'_x\| = d$. Then by the above inequality, we have

$$\|m_x - m'_x\|^2 = 2\|m_x - x\|^2 + 2\|m'_x - x\|^2 - 4\left\|\frac{m_x - m'_x}{2} - x\right\|^2 \leq 0$$

from above. □

**Definition 3.11** (Projection operator)**.** Let $(\mathcal{H}, \langle \cdot, \cdot, \rangle)$ be a Hilbert space. Let $M \subseteq \mathcal{H}$ be closed and convex. Define

$$P_M : \mathcal{H} \to \mathcal{H}$$

by $P_M(x) = m_x$ from above. This is the projection of $\mathcal{H}$ onto $M$.

**Definition 3.12** (Orthogonal decomposition)**.** If $S \subseteq \mathcal{H}$, let

$$S^{\perp} = \{x \in \mathcal{H} \,|\, \langle x, y \rangle = 0 \quad \forall y \in S.$$

We call $S^{\perp}$ the orthogonal component.

**Theorem 3.13** (From previous lecture)**.** *If $M \subseteq \mathcal{H}$, then the projection of $\mathcal{H}$ onto $M$ is*

$$\begin{aligned} P_m : \mathcal{H} &\to \mathcal{H} \\ x &\mapsto m_x \end{aligned}$$

*where $m_x \in M$ is the unique element with $\|x - m_x\| = \inf_{m \in M} \|x - m\|$.*

**Lemma 3.14.** *Let $M \subseteq \mathcal{H}$ be closed subspace. Then $x - P_M x \in M^{\perp}$ for all $x \in \mathcal{H}$.*

*Proof.* Let $m \in M$. We need to show $\langle x - P_M x, m \rangle = 0$. This is clear if $m = 0$. Without loss of generality, assuming $m \neq 0$, we can assume $\|m\| = 1$. Then write

$$x - P_M x = x - (P_M x + \langle x - P_M x, m \rangle m) + \langle x - P_M x, m \rangle m.$$

Let the bracketed term be $m'$. Then $x - m' \perp \langle x - P_M x, m \rangle m$ because

$$\begin{aligned} \langle x - m', \langle x - P_M x, m \rangle m \rangle &= \overline{\langle x - P_M x, m \rangle} \langle x - m', m \rangle \\ &= C \langle x - P_M x - \langle x - P_M x, m \rangle m, m \rangle \\ &= C(\langle x - P_M x, m \rangle - \langle x - P_M x, m \rangle \|m\|) \\ &= 0. \end{aligned}$$

So $\|x - P_M x\|^2 = \|x - m'\|^2 + |\langle x - P_M x, m \rangle|^2$. So $\|x - P_M x\|^2 \geq \|x - P_M x\|^2 + |\langle x - P_M x, m \rangle|^2$ by definition of $P_M x$. Thus,

$$\langle x - P_M x, m \rangle = 0$$

and thus $x - P_M x \in M^{\perp}$. □

**Theorem 3.15.** *The following theorem is the key fundamental result. Let $(\mathcal{H}, \langle \cdot, \cdot, \rangle)$ be a Hilbert space. Let $M$ be a closed subspace of $\mathcal{H}$. Then*

$$\mathcal{H} = M \oplus M^{\perp}.$$

*That is, each $x \in \mathcal{H}$ can be written in exactly one way as $x = m + m^{\perp}$ with $m \in M$, $m^{\perp} \in M^{\perp}$.*

*Proof.* **Existence** - Let $x = P_m x + (x - P_M x)$.

**Uniqueness** - Let $x = x_1 + x_1^{\perp}$, $x = x_2 + x_2^{\perp}$ with $x_1, x_2 \in M, x_1^{\perp}, x_2^{\perp} \in M^{\perp}$ . Then

$$x_1 - x_2 = x_2^{\perp} - x_1^{\perp} \in M^{\perp}$$

Then

$$\langle x_1 - x_2, x_1 - x_n \rangle = 0 \Rightarrow x_1 = x_2.$$

Thus $x_1^{\perp} = x_2^{\perp}$. □

**Corollary.** *Let $M \subseteq \mathcal{H}$ be a closed subspace. Then we have*

*(a)* $P_M \in \mathcal{L}(\mathcal{H}, \mathcal{H})$.

*(b)* $\|P_M\| \leq 1$.

*(c)* $ImP_m = M, \text{KER } P_M = M^{\perp}$.

*(d)* $P_M^2 = P_M$.

*(e)* $P_{M^{\perp}} = I - P_M$.

*Proof.* (c), (d), (e) exercises.

(a). Let $x, y \in H$. Write $x = x_1 + x_1^{\perp}$ and $y = y_1 + y_1^{\perp}$ with $x_1, y_1 \in M$ and $x_1^{\perp}, y_1^{\perp} \in M^{\perp}$. Then

$$x = y = (x_1 + y_1) + (x_1^{\perp} + y_1^{\perp})$$

and so

$$P_M(x + y) = x_1 + y_1$$

and similarly $P_M(\alpha x) = \alpha P_M x$. We also have

$$\begin{aligned} \|x\|^2 &= \|P_M x + (x - P_M x)\|^2 \\ &= \|P_M x\|^2 + \|x - P_M x\|^2 \\ &\geq \|P_M x\|^2 \end{aligned}$$

and so $\|P_M\| \leq 1$. □

If $y \in \mathcal{H}$ is fixed, then the map

$$\begin{aligned} \varphi_y : \mathcal{H} &\to \mathbb{K} \\ x &\mapsto \langle x, y \rangle \end{aligned}$$

is in $\mathcal{H}'$. Linearity is clear, and continuity is proven by Cauchy-Swartz,

$$|\varphi_y(x)| = |\langle x, y \rangle| \leq \|y\| \|x\|.$$

So $\|\varphi_y\| \leq \|y\|$. Since $|\varphi_y(y)| = \|y\|^2$, we then have

$$\|\varphi_y\| = \|y\|.$$

**Theorem 3.16** (Riesz Representation Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space. The map*

$$\begin{aligned} \theta : \mathcal{H} &\to \mathcal{H}' \\ y &\mapsto \varphi_y \end{aligned}$$

*is a conjugate linear bijection, and $\|\varphi_y\| = \|y\|$.*

*Proof.* Conjugate linearity is clear.

**Injectivity**

$$\varphi_y = \varphi_{y'} \Rightarrow \varphi_y(x) = \varphi_{y'}(x) \quad \forall x$$

so

$$\langle x, y = \langle x, y' \rangle = 0 \quad \Rightarrow \langle y - y', y - y' \rangle = 0$$

and so $y = y'$.

**Surjectivity** Let $\varphi \in H'$. We now find $y \in \mathcal{H}$ with $\varphi = \varphi_y$. If $\varphi = 0$, take $y = 0$. Suppose $\varphi \neq 0$. Then $\textsc{Ker}\ \varphi \neq \mathcal{H}$. But $\textsc{Ker}\ \varphi$ is a closed subspace of $\mathcal{H}$. So

$$H = (\textsc{Ker}\ \varphi) \oplus (\textsc{Ker}\ \varphi)^{\perp}.$$

Hence $(\textsc{Ker}\ \varphi)^{\perp} \neq \{0\}$. Pick $z \in (\textsc{Ker}\ \varphi)^{\perp}, z \neq 0$. For each $x \in \mathcal{H}$, the element

$$x - \frac{\varphi(x)}{\varphi(z)} z \in \textsc{Ker}\ \varphi$$

Note that $\varphi(z) \neq 0$ since $z \notin \textsc{Ker}\ \varphi$. Then

$$\begin{aligned} 0 &= \langle x - \frac{\varphi(x)}{\varphi(z)} z, z \rangle \\ &= \langle x, z \rangle - \frac{\varphi(x}{\varphi(z)} \|z\|^2 \end{aligned}$$

and so

$$\varphi(x) = \langle x, \frac{\overline{\varphi(z)}}{\|z\|^2} z \rangle \quad \forall x \in \mathcal{H},$$

and so letting $y = \frac{\overline{\varphi(z)}}{\|z\|^2} z$, we have $\varphi = \varphi_y$. □

**Example 3.17.** From Hahn-Banach given $y \in \mathcal{H}$ there exists $\varphi \in \mathcal{H}'$ such that

$$\|\varphi\| = 1$$

and $\varphi(y) = \|y\|$. We can be very constructive in the Hilbert case, and let

$$\varphi(x) = \langle x, \frac{y}{\|y\|} \rangle$$

**Example 3.18.** All continuous linear functionals on $L^2([a, b])$ are of the form

$$\varphi(f) = \int_a^b f(x) \overline{g(x)}\, dx$$

for some $g \in L^2([a, b])$.[1]

**Example 3.19** (Adjoint operators)**.** Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. Let $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. The **adjoint** of $T$ is $T^{\star} \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ given by

$$\langle Tx, y \rangle_2 = \langle x, T^{\star} y \rangle_1$$

for all $x \in \mathcal{H}_1, y \in \mathcal{H}_2$[2]

[1] As an exercise, show that this is true.

[2] As an exercise, show that for a matrix $T$, we have $T^{\star} = \overline{T^t}$ where $T^t$ is the matrix transpose.

## 3.3 Orthonormal Systems

**Definition 3.20** (Orthonormal system)**.** As subset $S \subseteq \mathcal{H}$ is an **orthonormal system** (orthonormal) if

$$\langle e, e' \rangle = \delta_{e,e'} \quad \forall e, e' \in S$$

**Definition 3.21** (Complete orthonormal system or Hilbert basis)**.** An orthonormal system $S$ is **complete** or a **Hilbert basis** [3] if

$$\overline{\text{span } S} = \mathcal{H}$$

[3] By Gram-Schmidt and Zorn's Lemma, every Hilbert space has a complete orthonormal system.

**Example 3.22.**

1. $\ell^2$. Then
$$S = \{e_i \mid i \geq 1\}$$
is orthonormal and is complete.

2. $L^2_{\mathbb{C}}([0, 2\pi])$. Then
$$S = \{\frac{1}{2\pi} e^{int} \mid n \in \mathbb{Z}\}$$
is orthonormal and is complete. Completeness follows from Stone-Weierstrass theorem.

3. $L^2_{\mathbb{R}}([0, 2\pi])$. Then
$$S = \{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos nt, \frac{1}{\sqrt{\pi}} \sin nt \mid n \geq 1\}$$
is orthonormal and is complete, again by Stone-Weierstrass.

We want to look at series $\sum_{e \in S}$ ..., which is tricky if $S$ is not countable.

**Lemma 3.23.** *If $\{e_k \mid k \geq 0\}$ is orthonormal, then*

$$\sum_{k=0}^{\infty} a_l e_k$$

*converges in $\mathcal{H}$ if and only if*

$$\sum_{k=0}^{\infty} |a_k|^2$$

*converges in $\mathbb{K}$.*

*If either series converges, then*

$$\left\| \sum_{k=0}^{\infty} a_k e_k \right\|^2 = \sum_{k=0}^{\infty} |a_k|^2$$

*Proof.* If $\sum_{k=0}^{\infty} a_k e_k$ converges to $x$, then[4]

$$\begin{aligned}\langle x, x\rangle &= \lim_{n\to\infty} \langle \sum_{k=0}^{n} a_k e_k, \sum_{k=0}^{n} a_k e_k\rangle \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} |a_k|^2\end{aligned}$$

[4] If $x_n \to x, y_n \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$

Conversely, if $\sum_{k=0}^{\infty} |a_k|^2$ converges, then writing $x_n = \sum_{k=0}^{n} a_k e_k$, we have

$$\begin{aligned}\|x_m - x_n\|^2 &= \|\sum_{k=n+1}^{m} a_k e_k\|^2 \\ &= \sum_{k=n+1}^{m} \|a_k e_k\|^2 \text{ by Pythagoras} \\ &= \sum_{k=n+1}^{m} |a_k|^2 \to 0\end{aligned}$$

and so $(x_n)$ is Cauchy, and hence converges by completeness of $\mathcal{H}$. □

**Lemma 3.24.** *Let $\{e_1, \ldots, e_n\}$ be orthonormal. Then*

$$\sum_{k=1}^{n} |\langle x, e_k\rangle|^2 \leq \|x\|^2$$

*for each $x \in \mathcal{H}$.*

*Proof.* Let $y = \sum_{k=1}^{n} \langle x, e_k\rangle e_k$. Let $z = x - y$. We claim that $z \perp y$. We have

$$\begin{aligned}\langle x, y\rangle &= \langle x - y, y\rangle \\ &= \langle x, y\rangle - \|y\|^2 \\ &= \sum_{k=1}^{n} \overline{\langle x, e_k\rangle} \langle x, e_k\rangle - \sum_{k=1}^{n} |\langle x, e_k\rangle|^2 \\ &= 0.\end{aligned}$$

So

$$\begin{aligned}\|x\|^2 &= \|y + z\|^2 \\ &= \|y\|^2 + \|z\|^2 \text{ Pythagoras} \\ &\geq \|y\|^2 = \sum_{k=1}^{n} |\langle x, e_k\rangle|^2\end{aligned}$$

□

We want to write expressions like $\sum_{e \in S} \langle x, e\rangle e$.

**Corollary.** *Let $x \in \mathcal{H}$ and $S$ orthonormal. Then*

$$\{e \in S \mid \langle x, e\rangle \neq 0\}$$

*is countable.*

*Proof.*

$$\{e \in S \mid \langle x, e \rangle \neq 0\} = \bigcup_{k \geq 1} \{e \in S \mid |\langle x, e \rangle| > \frac{1}{k}$$

From the lemma,

$$\#\{e \in S \mid |\langle x, e \rangle| > \frac{1}{k}\} \leq k^2 \|x^2\|$$

For if this number were greater than $k^2\|x\|^2$, then the LHS in Lemma is greater than $\frac{1}{k^2}k^2\|x\|^2$. □

Therefore:

**Corollary** (Bessel's Inequality)**.** *If $S$ is orthonormal, then*

$$\sum_{e \in S} |\langle x, e \rangle|^2 \leq \|x\|^2$$

*for all $x \in \mathcal{H}$*

*Proof.* $\sum_{e \in S} |\langle x, e \rangle|^2$ is a sum of countably many positive terms, and so order is not important. □

We want to write $\sum_{e \in S} \langle x, e \rangle e$. This sum is over a countable set, but is the order important?

**Theorem 3.25.** *Let $S$ be orthonormal. Let $M = \overline{\text{span } S}$. Then*

$$P_M x = \sum_{e \in S} \langle x, e \rangle e$$

*where the sum can be taken in any order.*

*Proof.* Fix $x \in H$. Choose an enumeration

$$\{e_k \mid k \geq 0\} = \{e \in S \mid \langle x, e \rangle \neq 0\}.$$

By Bessel's inequality, we have

$$\sum_{k=0}^{\infty} |\langle x, e_k \rangle|^2 \leq \|x\|^2$$

and so the LHS converges. By Lemma 3.23, we know

$$y = \sum_{k=0}^{\infty} \langle x, e_k \rangle e_k \in M$$

converges in $\mathcal{H}$.

Write $x = y + (x - y) = M + M^{\perp}$. We claim $(x - y) \in M^{\perp}$. Then $P_M x = y$ from characterisation of projection operator. Let $e \in S$. Then

$$\begin{aligned}\langle x - y, e \rangle &= \lim_{n \to \infty} \left\langle x - \sum_{k=0}^{n} \langle x, e_k \rangle e_k, e \right\rangle \\ &= \lim_{n \to \infty} (\langle x, e \rangle - \sum_{k=}^{n} \langle x, e_k \rangle \langle e_k, e \rangle) \\ &= \langle x, e \rangle - \sum_{k=0}^{\infty} \langle x, e_k \rangle \langle e_k, e \rangle .\end{aligned}$$

If $e \in \{e' \in S \mid \langle x, e' \rangle \neq 0\}$, then $e = e_j$ for some $j$, and so

$$\langle x - y, e \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0$$

If $\langle x, e \rangle = 0$, then $e \neq e_j$ for all $j$, and so $\langle e_j, e \rangle = 0$, and so

$$\langle x - y, e \rangle = 0 - 0 = 0.$$

Thus $x - y \in (\text{span } S)^{\perp}$.

**Exercise 3.26.** Show that

$$x - y \in \overline{(\text{span } S)}^{\perp} = M^{\perp}.$$

□

Recall that if $\{x_1 \ldots \}$ is a countable orthonormal system in a Hilbert space $\mathcal{H}$. Then

$$\sum_{k=1}^{\infty} a_k e_k < \infty \iff \sum_{k=1}^{\infty} |a_k|^2 < \infty$$

and

$$\|\sum_{k=1}^{\infty} a_k e_k\|^2 = \sum_{k=1}^{\infty} |a_k|^2 \qquad (\star)$$

We also had the following.

**Theorem 3.27.** *Let $S$ be orthonormal in $\mathcal{H}$. Let $M = \overline{\text{span } S}$. Then*

$$P_M x = \sum_{e \in S} \langle x, e \rangle \, e \quad \forall x \in \mathcal{H}$$

*where the sum has only countable many terms and convergence is unconditional.*

**Theorem 3.28.** *Let $S$ be orthonormal in $\mathcal{H}$. Then following are equivalent.*

*(a) $S$ is a complete orthonormal system ($\overline{\text{span } S} = \mathcal{H}$).*

*(b) $x = \sum_{e \in S} \langle x, e \rangle \, e$ for all $x$ (Fourier series).*

*(c) $\|x\|^2 = \sum_{e \in S} | \langle x, e \rangle |^2$ for all $x$ (Parseval's formula).*

*Proof.* (a) ⇒ (b). If $M = \overline{\text{span } S} = \mathcal{H}$, then

$$P_M x = x = \sum_{e \in S} \langle x, e \rangle \, e$$

by Theorem 3.27.

(b) ⇒ (c). By the infinite Pythagoras theorem ($\star$).

(c) ⇒ (a). Let $M = \overline{\text{span } S}$. Suppose that $z \in M^{\perp}$. Then $z = 0 + z \in M + M^{\perp}$. Hence

$$0 = \|P_M z\|^2 = \|\sum_{e \in S} \langle z, e \rangle \, e\|^2 = \sum_{e \in S} | \langle z, e \rangle |^2 = \|z\|^2$$

which implies $z = 0$, so $M = \mathcal{H}$, and so $S$ is complete. □

*Remark.* Consider $L^2([0, 2\pi])$, and let $S = \{e_n \mid n \in Z\}$. Then we can write

$$f = \sum_{n \in \mathbb{Z}} c_n e_n$$

where $c_n = \langle f, e_n \rangle = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(t) e^{-int} \, dt$.

We do not claim that convergence is pointwise, what we have proven is convergence is in $L^2$,

$$\|f - \sum_{|n| \le N} c_n e_n\|_2 \to 0$$

as $N \to \infty$. This is not the same as pointwise or uniform convergence ($\|\cdot\|_\infty$).

## 3.4 The Stone-Weierstrass Theorem

This is a useful tool to show an orthonormal system is complete. In fact, this theorem is about uniformly approximating elements of $\mathcal{C}(X)$, where $X$ is a compact Hausdorff space. it is a generalisation of the Weierstrass approximation theorem.

**Theorem 3.29** (Weierstrass approximation theorem)**.** *Let $f \in \mathcal{C}([a, b])$ and let $\epsilon > 0$ be given. Then there exists a polynomial $p(x)$ such that*

$$|f(x) - p(x)| < \infty \quad \forall x \in [a, b],$$

*that is, $\|f - p\|_\infty < \epsilon$.*

**Corollary.** *This implies the following important results:*

- *Continuous functions can be uniformly approximated by polynomials.*
- *$\mathcal{P}([a, b])$, the space of polynomials on $[a, b]$, is dense in $\mathcal{C}([a, b])$.*
- *$\overline{\mathcal{P}([a, b])} = \mathcal{C}([a, b])$.*

We now prove Stone's 1930's generalisation.

**First some setup:** Let $X$ be a compact Hausdorff space throughout. We then know that $\mathcal{C}(X)$ is a vector space. It also has sensible vector multiplication,

$$(fg)(x) = f(x)g(x).$$

Thus $\mathcal{C}(X)$ is a unital, commutative, associative ring. As we have

$$f(\lambda g) = \lambda(fg)$$

then $\mathcal{C}(X)$ is a unital, commutative, associative algebra over $\mathbb{K}$.

**Definition 3.30** (Subalgebra)**.** A subalgebra of $\mathcal{C}(X)$ is a subset $\mathcal{A}$ which is closed under scalar multiplication, vector addition, and vector multiplication. $\mathcal{A}$ is unital if it contains the constant function $f(x) = 1$.

**Example 3.31.** $\mathcal{P}([a,b])$ is a subalgebra of $\mathcal{C}([a,b])$.

When is $\mathcal{A}$ dense in $\mathcal{C}(X)$?

**Theorem 3.32** (Stone-Weierstrass theorem)**.** *Let X be a compact Hausdorff space, and let $\mathcal{A}$ be a subalgebra of $\mathcal{C}(X)$. If*

*(1) $\mathcal{A}$ is unital,*

*(2) $f \in \mathcal{A} \Rightarrow f^\star \in \mathcal{A}$, where $f^\star(x) = \overline{f(x)}$,*

*(3) $\mathcal{A}$ separates points of X.*

*Then $\overline{\mathcal{A}} = \mathcal{C}(X)$.*

**Definition 3.33.** $\mathcal{A}$ separates points of $X$ if, given $x \neq y$, there is a function $f \in \mathcal{A}$ with $f(x) \neq f(y)$.

**Corollary.** *(a) $\mathcal{P}([a,b])$ is dense in $\mathcal{C}([a,b])$, as $f(x) = x$ separates points.*

*(b) Trigonometric polynomials are dense in*

$$\{f \in \mathcal{C}([0,2\pi]) \mid f(0) = f(2\pi)\}.$$

*(c) Trigonometric polynomials are dense in $L^2([0,2\pi])$, and*

$$S = \{e_n \mid n \in \mathbb{Z}\}$$

*is complete.*

**Setup**

**Lemma 3.34.** *The function $f(t) = |t|$ can be uniformly approximated by polynomials on $[-1,1]$*

*Proof.* The binomial theorem says

$$(1+x)^{1/2} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} x^n \quad \forall x \in [-1,1]$$

We then have

$$|t| = \sqrt{t^2} = \sqrt{1+(t^2-1)} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (t^2-1)^n \quad t \in [-\sqrt{2}, \sqrt{2}]$$

Now let $p_N(t) = sum_{n=0}^{N} (\frac{\frac{1}{2}}{n})(t^2-1)^n$, and

$$||t| - p_N(t)| = |\sum_{n=N+1}^{\infty} \binom{\frac{1}{2}}{n} (t^2-1)^n| \leq \sum_{n=N+1}^{\infty} |\binom{\frac{1}{2}}{n}|$$

and so $\||t| - p_n\|_\infty \to 0$ as $N \to \infty$ on $[-1,1]$. □

**Theorem 3.35** (Stone-Weierstrass theorem). *Let $X$ be a compact Hausdorff space, and let $\mathcal{A}$ be a subalgebra of $\mathcal{C}(X)$. If*

*(1)* *$\mathcal{A}$ is unital,*

*(2)* *$f \in \mathcal{A} \Rightarrow f^\star \in \mathcal{A}$, where $f^\star(x) = \overline{f(x)}$,*

*(3)* *$\mathcal{A}$ separates points of $X$.*

*Then $\overline{\mathcal{A}} = \mathcal{C}(X)$.*

*Proof.* We first prove for $\mathcal{C}_{\mathbb{R}}(X)$.

**Lemma 3.36.** *Let $\mathcal{A}$ be a unital subalgebra of $\mathcal{C}_{\mathbb{R}}(X)$ . Then*

*(a)* $|f| \in \overline{\mathcal{A}}$,

*(b)* $\min(f_1, \ldots, f_n), \max(f_1, \ldots, f_n) \in \overline{\mathcal{A}}$

*for all $f, f_1, \ldots, f_n \in \mathcal{A} \subseteq \mathcal{C}_{\mathbb{R}}(X)$.*

*Proof.* (a) Replace $f$ by $\frac{f}{\|f\|_\infty}$ so we can assume that $\|f\|_\infty = 1$. From the previous lemma, we know for each $n \geq 1$ there is a polynomial $p_n : [-1, 1] \to \mathbb{R}$ such that $||t| - p_n(t)| < \frac{1}{n}$ for all $t \in [-1, 1]$.

Since $|f(x)| \leq \|f\|_\infty = 1$ for all $x \in X$, we have

$$\||f| - p_n(f)\| \leq \frac{1}{n}$$

But $p_n(f)$ is a finite linear combination of $1, f, f^2, f^3, \ldots$ and so in in $\mathcal{A}$, as $\mathcal{A}$ is unital. Thus $|f| \in \overline{\mathcal{A}}$.

(b) Use the formulas

$$\max(f, g) = \frac{f + g - |f - g|}{2}, \quad \min(f, g) = \frac{f + g - |f - g|}{2} \in \overline{\mathcal{A}}$$

and induction.

□

*Proof of Stone-Weierstrass for $\mathcal{C}_{\mathbb{R}}(X)$.* Let $f \in \mathcal{C}_{\mathbb{R}}(X)$ and let $\epsilon > 0$ be given. We need to find $g \in \mathcal{A}$ such that

$$|f(z) - g(z)| < \epsilon \quad \forall z \in X$$

**Step 0.** We can assume that $\mathcal{A}$ is closed.

**Exercise 3.37.** Why?

**Step 1.** Let $x, y \in X$ be fixed.

**Proposition 3.38.** *There exists $f_{xy} \in \mathcal{A}$ with*

$$f_{xy}(x) = f(x), \quad f_{xy}(x) = f(y)$$

*Proof.* If $x = y$ then trivial (take $f_{xy}(z) = f(x)\mathbf{1}(z)$).

If $x \neq y$, since $A$ separates points, there is $h \in \mathcal{A}$ with $h(x) \neq h(y)$. Then take

$$f_{xy} = ah + b1 \in \mathcal{A}$$

we can invert the coefficient matrix to find our coefficients $a$ and $b$. □

**Step 2.** Let $x \in X$ be fixed.

**Proposition 3.39.** *There exists $f_x \in \mathcal{A}$ such that*

- $f_x(x) = f(x)$.
- $f_x(z) < f(z) + \epsilon$

*Proof.* For each $y \in X$,let

$$O_y = \{z \in X \mid f_{xy}(z) < f(z) + \epsilon\}$$

where $f_{xy}$ is the function from Step 1. These are all open sets (why?) and thus

$$X = \bigcup_{y \in X} O_y$$

since $y \in O_y$.

By compactness of $X$, we have

$$X = \bigcup_{i=1}^{m} O_{y_i}$$

Letting $f_x = \min(f_{xy_1}, \ldots, f_{xy_n})$. Then

- Since $f_{xy_i}(x) = f(x)$ for all i,

$$f_x(x) = f(x)$$

- If $z \in X$, then $z \in O_{y_i}$ for some $i$, and so

$$f_x(z) \leq f_{xy_i}(z) < f(z) + \epsilon$$

as required. □

**Step 3.**

**Proposition 3.40.** *There exists a function $g \in \mathcal{A}$ such that*

$$|f(z) - g(z)| < \epsilon$$

*for all $z \in X$.*

*Proof.* For each $x \in X$, let

$$U_x = \{z \in X \mid f_x(z) > f(x) - \epsilon\}$$

where $f_x$ is from Step 2. These sets $U_i$ are open and since $x \in U_x$, for an open cover, we can write

$$X = \bigcup_{x \in X} U_x = \bigcup_{j=1}^{n} U_{x_j}.$$

Define $g = \max(f_{x_1}, \dots, f_{x_n})$. If $z \in X$,

- $g(z) = f_{x_i}(z)$ for some $i$, which is less than $f(z) + \epsilon$ from Step 2.
- If $z \in U_{x_j}$ for some $j = 1, \dots, n$, then

$$g(z) \geq f_{x_j}(z) > f(x) - \epsilon.$$

□

□

**Exercise 3.41.** Where did we use the Hausdorff property?

We now prove for $\mathcal{C}_{\mathbb{C}}(X)$.

Let

$$\mathcal{A}_{\mathbb{R}} = \{f \in \mathcal{A} \mid f \text{ is real valued}\}.$$

Then $\mathcal{A}_{\mathbb{R}}$ is an $\mathbb{R}$-subalgebra of $\mathcal{C}_{\mathbb{R}}(X)$. It is unital, as $1 \in \mathcal{A}$ and it is real valued.

We now show $\mathcal{A}_{\mathbb{R}}$ separates points. If $x \neq y$, there is $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. Write $f = u + iv$ with $u, v$ real valued. Either $u(x) \neq u(y)$ or $v(x) \neq v(y)$, and so $\mathcal{A}_{\mathbb{R}}$ separates points.

Hence $\mathcal{A}_{\mathbb{R}}$ is dense in $\mathcal{C}_{\mathbb{R}}$.

Now, let $f \in \mathcal{C}_{\mathbb{C}}(X)$. Then write $f = u + iv$. Then $u, v \in \mathcal{C}_{\mathbb{R}}(X)$. Then given $\epsilon > 0$, there exists $u_1, v_1 \in \mathcal{A}_{\mathbb{R}}$ such that

$$\|u - u_1\|_\infty \leq \frac{\epsilon}{2}, \quad \|v - v_1\|_\infty \leq \frac{\epsilon}{2}$$

Writing $f_1 = u_1 + iv_1 \in \mathcal{A}$, we have

$$\|f - f_1\|_\infty \leq \|(u - u_1) + i(v - v_1)\|_\infty \leq \|u - u_1\|_\infty + \|v - v_1\|_\infty < \epsilon$$

and thus $\mathcal{A}$ is dense in $\mathcal{C}_{\mathbb{C}}(X)$. □

**Corollary.** *Polynomials are dense in $\mathcal{C}([a, b])$.*

*Proof.* $\mathcal{A} = \mathcal{P}([a, b])$ is an algebra, is unital, is closed under complex conjugation, and separates points. Thus, $\mathcal{A}$ is dense in $\mathcal{C}([a, b])$. □

**Definition 3.42** (Trigonometric polynomials)**.** A **trigonometric polynomail** is an expression

$$\sum_{n\in\mathbb{Z}} c_n e^{int}$$

with finitely many $c_n \neq 0$. So these are polynomials in $s = e^{it}$ and $s^{-1} = \bar{s} = e^{-it}$.

**Corollary.** *The space $\mathcal{A}$ of all trigonometric polynomials is dense in $\mathcal{C}(\Pi)$, where $\Pi = \{z \in \mathbb{C} \mid |z| = 1\}$*

*Proof.* $\mathcal{A}$ is a sub-algebra of $\mathcal{C}(\Pi)$, it is unital, closed under complex conjugation,

$$\overline{\sum_{n\in\mathbb{Z}} c_n e^{int}} = \sum_{n\in\mathbb{Z}} \overline{c_{-n}} e^{int}$$

and separates points. $T$ is a compact Hausdorff space, and thus Stone-Weierstrass states that $\mathcal{A}$ is dense in $\mathcal{C}(\Pi)$. ☐

**Corollary.** *The orthonomal system*

$$S = \{\frac{1}{\sqrt{2\pi}} e^{int} \mid n \in \mathbb{Z}\}$$

*is complete in $L^2([0,2\pi])$.*

*Proof.* span $S = \mathcal{A}$ is a space of trigonometric polynomials, which is dense in $\mathcal{C}(\Pi)$. Define

$$\begin{aligned} \Phi :\ \mathcal{C}_p([0,2\pi]))] &\to \mathcal{C}(\Pi) \\ f &\mapsto \tilde{f} \end{aligned}$$

where $\mathcal{C}_p([0,2\pi]) = \{f \in \mathcal{C}([0,2\pi]) \mid f(0 = f(2\pi))\}$. Then $\Phi$ is an isometric isomorphism, and therefore functions of the form $f(t) = \sum c_n e^{int}$ is dense in $\mathcal{C}_p([0,2\pi])$.

By the construction of the Lebesgue integral, **simple** functions

$$\sum_{i=1}^{n} a_i \mathbf{1}_{A_i}$$

are dense in $L^2([0,2\pi])$.

**Exercise 3.43.** Given $f \in L^2([0,2\pi])$ and $\epsilon > 0$, there exists $g \in \mathcal{C}_p([0,2\pi])$ such that $\|f - g\|_2 < \epsilon$.

Thus $\mathcal{A}$ is dense in $L^2([0,2\pi])$. ☐

**Corollary.** *The following are separable (have a countable dense subset):*

*(a)* $\mathcal{C}([a,b])$,

*(b)* $L^p([a,b])$ , $1 \leq p < \infty$

*Proof.* (a) We have $\mathcal{P}([a,b])$ is dense in $\mathcal{C}([a,b])$ and set $\mathcal{P}_{\mathbb{Q}}([a,b])$ with rational coefficients is dense in $\mathcal{P}([a,b])$. Clearly, $\mathcal{P}_{\mathbb{Q}}([a,b])$ is countable, and thus is dense in $\mathcal{C}([a,b])$.

(b) Use the fact that $\mathcal{C}([a,b])$ is dense in $L^p([a,b])$.

□

**Corollary.** *Let $X$ be a compact metric space. Then $\mathcal{C}(X)$ is separable.*

*Proof.* As $X$ is a compact metric space, then $X$ is separable. [5] Let $\{x_n \mid n \geq 1\}$ be a countable dense subset of $X$. For each $n \geq 1$ and $m \geq 1$ define

$$f_{n,m} : X \to \mathbb{K}$$

by

$$f_{n,m}(x) = \inf_{z \notin B(x_n, \frac{1}{m})} d(x,z)$$

We then claim $f_{n,m}$ is continuous. Now, let $\mathcal{A}$ be the space of all $\mathbb{K}$-linear combinations of

$$f^{k_1}_{n_1,m_1}, \ldots, f^{k_l}_{n_l,m_l}, k_1, \ldots, k_l \in \mathcal{N}. \qquad (\star)$$

This is a sub-algebra of $\mathcal{C}(X)$, as $\mathcal{A}$ is unital, closed under conjugation, and separates points - if $z_1, z_2 \in X$ with $z_1 \neq z_2$, Choose $n, m$ such that $z_1 \in B(x_n, \frac{1}{m})$, $z_n \notin B(x_n, \frac{1}{m})$. Thus the sub-algebra $\mathcal{A}$ is dense by Stone-Weierstrass.

The subset of $\mathbb{Q}$-linear combinations of $(\star)$ is countable and dense.

□

[5] As an exercise, show that if $X$ is a compact metric space, then $X$ is separable.

**Lemma 3.44.** *If $X$ is compact metric space then $X$ is separable.*

*Proof.* For each $m \geq 1$,

$$X = \bigcup_{x \in X} B(x; \frac{1}{m})$$

has a finite subcover

$$X = \bigcup_{n=1}^{N_m} B(x_{m,n} \frac{1}{m})$$

and thus the subset of all $\{x_{m,n}\}$ is a countably dense subset. □

**Corollary.**

$$\frac{pi^2}{6} = \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

*Proof.* $S = \{\frac{1}{\sqrt{2\pi}} e^{int} \mid n \in \mathbb{Z}\}$ is complete, and so Parseval's formula holds,

$$\|f\|_2^2 = \sum_{n \in \mathbb{Z}} | \langle f, e_n \rangle |^2.$$

Apply to $f(x) = x$. □

A common strategy is to prove for polynomials, and then Stone-Weierstrass proves it for continuous functions.

**Corollary.** *If* $f \in \mathcal{C}([a,b] \times [c,d])$ *then*

$$\int_a^b \int_c^d f(x,y)\, dydx = \int_c^d \int_a^b f(x,y)\, dxdy$$

*Proof.* By direct calculation, the result is true for two-variable polynomials. Let $f \in \mathcal{C}([a,b] \times [c,d])$ and $\epsilon > 0$ be given. By Stone-Weierstrass, the space of polynomials in 2 variables is dense in $\mathcal{C}([a,b] \times [c,d])$ and so there exists a polynomial $p(x,y)$ with

$$|f(x,y) - p(x,y)| < \frac{\epsilon}{(b-a)(d-c)}.$$

The result then follows by direct calculation. □

# 4
# Uniform Boundedness and the Open Mapping Theorem

The following sections introduce two of the cornerstone theorems of functional analysis - the uniform boundedness principle and the open mapping theorem.

## 4.1 The Principle of Uniform Boundedness

**Theorem 4.1** (Baire's theorem). *Let $X$ be a complete metric space. If $U_1, U_2, \dots$ are open dense subsets of $X$, then*

$$U = \bigcap_{n=1}^{\infty} U_n$$

*is dense in $X$.*

*Proof.* Let $x \in X$ and $\epsilon > 0$ be given. We need to show that

$$B(x, \epsilon) \cap U \neq \emptyset.$$

**Lemma 4.2.** *There exists sequences $(x_n)$ in $X$ and $(\epsilon_n)$ in $\mathbb{R}^+$ with the property that*

*(a)* $x_1 = x$, $\epsilon_1 = \epsilon$.

*(b)* $\epsilon_n \downarrow 0$

*(c)* $\overline{B(x_{n+1}, \epsilon_{n+1})} \subseteq B(x_n, \epsilon_n) \cap U_n$ *for all* $n \geq 1$.

*Proof.* Let $x_1, \dots, x_n$ and $\epsilon_1, \dots, \epsilon_n$ be chosen. By density of $U_n$,

$$B(x_n, \epsilon_n) \cap U_n \neq \emptyset.$$

Choose $x_{n+1} \in B(x_n, \epsilon_n) \cap U_n$. Choose $\epsilon'_{n+1} > 0$ such that $B(x_{n+1}, \epsilon'_{n+1}) \subseteq B(x_n, \epsilon_n) \cap U_n$ (openness). We have $\epsilon'_{n+1} \leq \epsilon_n$. Choose $0 \leq \epsilon_{n+1} \leq \min(\frac{\epsilon'_{n+1}}{2}, \frac{1}{n+1})$, then we have

$$\begin{aligned} \overline{B}(x_{n+1}, \epsilon_{n+1}) &\subseteq B(x_{n+1}, \epsilon'_{n+1}) \\ &\subseteq B(x_n, \epsilon_n) \cap U_n \end{aligned}$$

and $\epsilon_{n+1} < \epsilon_n$ with $\epsilon_{n+1} < \frac{1}{n+1}$. □

Given the lemma, the theorem follows. If $m \geq n$, then by (c),

$$B_(x_m, \epsilon_m) \subseteq B(x_n, \epsilon_n) \cap U_n \quad (\star)$$

In particular, $x_m \in B(x_n, \epsilon_n)$. Thus, $d(x_n, x_m) < \epsilon_n$ for all $m \geq n$. Thus $(x_n)$ is Cauchy, and so $x_n \to \zeta$ in $X$ by completeness. By $(\star)$, we then have $d(x_n, \zeta) \leq \epsilon_n$ for all $n \geq 1$. So $\zeta \in \overline{B(x_n, \epsilon_n)}$. So by (c), $\zeta \in \overline{B(x_{n+1}, \epsilon_{n+1})} \subseteq B(x_n, \epsilon_n) \cap U_n$.

Thus $\zeta \in B(x, \epsilon)$ and thus $\zeta \in U = \bigcap_{n=1}^{\infty} U_n$. □

The following corollary is often used

**Corollary.** *Let $X$ be a complete metric space. If $C_1, C_2, \ldots$ are closed with $X = \bigcup_{n=1}^{\infty}$ then $Int(C_n) \neq \emptyset$ for some $n$.*

*Proof.* If $\text{Int}(C_n) = \emptyset$ for all $n$ then $U_n = X \backslash C_n$ are open and dense. So by Baire's theorem, $\cap_{n=1}^{\infty} U_n$ is sense, and in particular, $\bigcap n = 1^{\infty} U_n \neq \emptyset$. We have

$$\begin{aligned} X = \bigcup_{n=1}^{\infty} C_n &= \bigcup_{n=1}^{\infty} (X \backslash U_n) \\ &= X \backslash (\bigcap_{n=1}^{\infty} U_n) \\ &\subsetneq X, \end{aligned}$$

a contradiction. □

There are three cornerstone theorems.

- Hahn-Banach,
- Uniform Boundedness,
- Open Mapping.

**Theorem 4.3** (Uniform boundedness)**.** *Let $X, Y$ be Banach spaces. Let $T_\alpha$, $\alpha \in A$, a family of continuous linear operators $T_\alpha : X \to Y$. Then if*

$$\sup_{\alpha \in A} \|T_\alpha x\| < \infty$$

*for each fixed $x \in X$, then*[1]

$$\sup_{\alpha \in A} \|T_\alpha\| < \infty$$

*Proof.* For each $n \geq 1$, let

$$X_n = \{x \in X \mid \|T_\alpha x\| \leq n \, \forall \alpha \in A\}$$

[1] This is an amazing result — we obtain a global bound from pointwise bounds.

These are **closed** ($T_\alpha$ is continuous) and

$$X = \bigcup_{n=1}^{\infty} X_n$$

by the hypothesis.

By the corollary to Baire's theorem, we know there exists $n_0 \geq 1$ with $\mathrm{Int}(X_{n_0}) \neq \emptyset$. Choose $x_0 \in \mathrm{Int}(X_{n_0})$, and let $r > 0$ such that

$$B(x_0, r) \subseteq \mathrm{Int}(X_{n_0}).$$

If $\|z\| \leq 1$ then $x_0 + rz \in \overline{B}(x_0, r)$. So $x_0 + rz \in X_{n_0}$, and

$$\|T_\alpha(x_0 + rz)\| \leq n_0 \, \forall \alpha \in A,$$

but $|\|a\| - \|b\|| \leq \|a + b\|$, so

$$\|T_\alpha(rz)\| - \|T_\alpha(x_0)\| \leq \|T_\alpha(x_0 + rz)\| \leq n_0.$$

So $r\|T_\alpha z\| \ \leq n_0 + n_0$, and

$$\|T_\alpha z\| \leq \frac{2n_0}{r} \, \forall \|z\| \leq 1, \forall \alpha \in A$$

For a general $x \in X$,

$$\|T_\alpha x\| = \|T_\alpha(\frac{x}{\|x\|})\|x\| \leq \frac{2n_0}{r}\|x\|$$

and thus $\|T_\alpha\| \leq \frac{2n_0}{r}$, which implies

$$\sup_{\alpha \in A} \|T_\alpha\| < \infty$$ □

Recall, the Fourier series of $f \in L^2([-\pi, \pi])$ is

$$\sum_{k \in \mathbb{Z}} \langle f, e_k \rangle e_k$$

where $e_k(t) = \frac{e^{ikt}}{\sqrt{2\pi}}$. This converges to $f$ in the $L^2$ norm.[2]

There are explicit (complicated) examples, but the easiest existence is using the uniform boundedness principle.

[2] It is useful here to attempt to prove or disprove this statement.

**Exercise 4.4.** If $f$ is $2\pi$-periodic and continuous, does the Fourier series converge pointwise?

**Proposition 4.5.** *There is a $2\pi$ periodic continuous function whose Fourier series does not converge at 0.*

*Proof.* Let $\mathcal{C}_p([-\pi, \pi]) = \{f \in \mathcal{C}([-\pi, \pi]) \mid f(-\pi) = f(\pi)\}$. This is a Banach space with $\|\cdot\|_\infty$. If $f \in \mathcal{C}_p$, let

$$f_n = \sum_{|k| \leq n} \langle f, e_k \rangle e_k.$$

*Remark.* We can now define, for each $n \geq 1$, a linear operator $T_n : \mathcal{C}_p \to \mathbb{K}$ by

$$T_n(f) = f_n(0).$$

If $f_n(0)$ converges (as $n \to \infty$) for each $f \in \mathcal{C}_p$, then

$$\sup_{n\geq 1} |T_n f| = \sup_{n\geq 1} |f_n(0)| < \infty$$

for all $f \in \mathcal{C}_p$, which by uniform boundedness implies

$$\sup_{n\geq 1} \|T_n\| \leq \infty. \qquad (\star)$$

We now show that $(\star)$ is false.

We have

$$\begin{aligned} f_n(x) &= \sum_{|k\leq n} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt}\, dt e^{ikx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left( \sum_{|k|\leq n} e^{-ik(x-t)} \right) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_n(x-t)\, dt \end{aligned}$$

where $D_n(t) = \sum_{|k|\leq n} e^{ikt}$ is the **Dirichlet Kernel**. The Dirichlet kernel is real, and even, with

$$D_n(t) = \frac{\sin(n+\frac{1}{2})t}{\sin\frac{t}{2}}.$$

*Note.* $T_n$ is continuous, with norm $\|T_n\| = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\, dt$.

*Proof.*

$$\begin{aligned} |T_n(f)| &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)||D_n(t)|\, dt \\ &\leq \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)|\, dt \right) \|f\|_\infty \end{aligned}$$

and so $\|T_n\| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\, dt$.

Going the other way, let

$$s_t = \begin{cases} 1 & D_n(t) \geq 0 \\ -1 & D_n(t) < 0 \end{cases}.$$

We have seen hat set functions can be approximated in $L^1$-norm by continuous (periodic) functions. So if $\epsilon > 0$ is given, there is a $g \in \mathcal{C}_p$ such that

$$\left| \frac{1}{2\pi} \int_{-\pi}^{\pi} (g(t) - s(t)) D_n(t)\, dt \right| < \epsilon.$$

$g$ can be chosen with $\|g\|_\infty = 1$.

So

$$\left| T_n(g) - \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)|\, dt \right| < \epsilon.$$

Thus

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)|\, dt - |T_n(g)| < \epsilon.$$

So

$$|T_n(g)| \geq \frac{\|g\|_\infty}{2\pi} \int_{-\pi}^{\pi} |D_n(t)|\, dt - \epsilon.$$

Since $\epsilon > 0$ was arbitrary,

$$\|T_n\| \geq \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)|\, dt. \qquad \square$$

All that remains is to show that

$$\|T_n\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)|\, dt \to \infty$$

We have

$$\begin{aligned}
\|T_n\| &= \frac{1}{\pi} \int_0^{\pi} |D_n(t)|\, dt \\
&= \frac{1}{\pi} \int_0^{\pi} \frac{|\sin(n+\frac{1}{2})t|}{|\sin\frac{t}{2}|}\, dt \\
&\geq \frac{2}{\pi} \int_0^{\pi} \frac{|\sin(n+\frac{1}{2})t|}{t}\, dt \\
&= \frac{2}{\pi} \int_0^{(n+\frac{1}{2})\pi} \frac{\sin v}{v}\, dv \\
&\geq \frac{2}{\pi} \int_0^{n\pi} \frac{\sin v}{v}\, dv \\
&= \frac{2}{\pi} \sum_{k=1}^{n} \int_{(k-1)\pi}^{k\pi} \frac{|\sin v|}{v}\, dv \\
&\geq \sum_{k=1}^{n} \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin v|\, dv \\
&= \frac{4}{\pi^2} \sum_{k=1}^{n} \frac{1}{k} \to \infty
\end{aligned}$$

as $n \to \infty$.

Thus there exists $f \in \mathcal{C}_p$ such that the Fourier series of $f$ diverges at $x = 0$. $\square$

## 4.2 *The Open Mapping Theorem*

THIS THEOREM IS TAILOR-MADE to deal with inverse operators.

**Definition 4.6** (Open mapping)**.** Let $X, Y$ be metric spaces. A function $f : X \to Y$ is **open** if open sets in $X$ are mapped to open sets in $Y$.

**Theorem 4.7** (Open mapping theorem)**.** *Let $X, Y$ be Banach spaces. If $T \in \mathcal{L}(X, Y)$ is surjective then $T$ is open.*

**Corollary** (Bounded inverse theorem)**.** *Let $X, Y$ be Banach spaces. If $T \in \mathcal{L}(X, Y)$ is bijective, then*

$$T^{-1} \in \mathcal{L}(Y, X).$$

*Proof.* Let $O \subseteq X$ be open. Then $(T^{-1})^{-1}(O) = T(O)$ is open (by the open mapping theorem). Thus $T^{-1}$ is continuous. □

**Corollary.** *Let $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ be Banach spaces. If*

$$\|x\|_1 \leq C\|x\|_2 \quad \forall x \in X$$

*then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

*Proof.*

$$\begin{aligned} i : \ (X, \|\cdot\|_2) &\to (X, \|\cdot\|_1) \\ x &\mapsto x \end{aligned}$$

is linear, surjective and injective, and also continuous, as

$$\|i(x)\| = \|x\|_1 \leq C\|x\|_2.$$

So the bounded inverse theorem gives

$$i^{-1} : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$$

is continuous. Thus there exists $A > 0$ such that $\|i^{-1}(x)\|_2 \leq A\|x\|_1$, which implies $\|x\|_2 \leq A\|x\|_1$. So

$$\frac{1}{A}\|x\|_2 \leq \|x\|_1 \quad \forall x \in X$$ □

More generally, if $T \in \mathcal{L}(X, Y)$ is bijective, then by the bounded inverse theorem,

$$c\|x\| \leq \|Tx\| \leq C\|x\|$$

where $c = \frac{1}{\|T^{-1}\|}$, $C = \|T\|$.

**Lemma 4.8.** *Let $X$ be a Banach space and $Y$ a normed space. Then for $T \in \mathcal{L}(X, Y)$, the following are equivalent.*

*(a) $T$ is open*

*(b) There exists $r > 0$ such that $B(0, r) \subseteq T(\overline{B(0, 1)})$*

*(c) There exists $r > 0$ such that $B(0, r) \subseteq \overline{T(\overline{B(0, 1)})}$.*

*Proof.* $(a) \Rightarrow (b),(c)$. As $B(0,1)$ is open, the set $T(B(0,1))$ is open in $Y$. Since $0 \in T(B(0,1))$ there exists $> 0$ such that the set

$$B(0,r) \subseteq T(B(0,1)) \subseteq T(\overline{B(0,1)}) \subseteq \overline{T(\overline{B(0,1)})}.$$

$(c) \Rightarrow (b)$. Assume that there exists $r > 0$ such that

$$B(0,r) \subseteq \overline{T(\overline{B(0,1)})}.$$

We now show that $B(0, \frac{r}{2}) \subseteq T(\overline{B(0,1)})$ which proves $(b)$. Let $y \in B(0, \frac{r}{2})$. Then $2y \in B(0,r)$ and since $B(0,r) \subseteq \overline{T(\overline{B(0,1)})}$ there exists $x_1 \in \overline{B(0,1)}$ such that

$$\|2y - Tx_1\| \leq \frac{r}{2}$$

Hence $4y - 2Tx_1 \in B(0,r)$ and by the same argument as before there exists $x_2 \in \overline{B(0,1)}$ such that

$$\|4y - 2Tx_1 - Tx_2\| \leq \frac{r}{2}$$

Continuing this way we construct a sequence $(x_n) \in \overline{B(0,1)}$ such that

$$\|2^n y - 2^{n-1}Tx_1 - \cdots - 2Tx_{n-1} - Tx_n\| \leq \frac{r}{2}$$

for all $n$. Dividing by $2^n$ we obtain

$$\|y - \sum_{k=1}^{n} 2^{-k}Tx_k\| \leq \frac{r}{2^{n+1}}$$

Hence $y = \sum_{k=1}^{\infty} 2^{-k}Tx_k$. Since $\|x_k\| \leq 1$ for all $k \in \mathbb{N}$ we have that

$$\sum_{k=1}^{\infty} 2^{-k}\|x_k\| \leq \sum_{k=1}^{\infty} 2^{-k} = 1$$

and so the series

$$x = \sum_{k=1}^{\infty} 2^{-k}x_k$$

converges absolutely in $X$ as $X$ is Banach and hence complete. We have also that $\|x\| \leq 1$ and so $x \in \overline{B(0,1)}$. Because $T$ is continuous we have

$$Tx = \lim_{n\to\infty} \sum_{k=1}^{n} 2^{-k}Tx_k = y$$

by construction of $x$. Hence $y \in T(\overline{B(0,1)})$ and $(b)$ follows.

$(b) \Rightarrow (a)$. By $(b)$ and the linearity of $T$ we have

$$T(\overline{B(0,\epsilon)}) = \epsilon T(\overline{B(0,1)})$$

for all $\epsilon > 0$. Since the map $x \mapsto \epsilon x$ is a homeomorphism on $Y$ the set $T(\overline{B(0,\epsilon)})$ is a neighbourhood of zero for all $\epsilon > 0$. Now let $U \subseteq X$ be open and $y \in T(U)$. As $U$ is open there exists $\epsilon > 0$ such that

$$\overline{B(x,\epsilon)} = x + \overline{B(0,\epsilon)} \subseteq U$$

where $y = Tx$. Since $z \mapsto x + z$ is a homeomorphism and $T$ is linear we have

$$T(\overline{B(x,\epsilon)}) = Tx + T(\overline{B(0,\epsilon)}) = y + T(\overline{B(0,\epsilon)}) \subseteq T(U).$$

Hence $T(\overline{B(x,\epsilon)})$ is a neighbourhood of $y$ in $T(U)$. As $y$ was arbitrary in $T(U)$ it follows that $T(U)$ is open. □

**Lemma 4.9.** *Let $X$ be a normed vector space and $S \subseteq X$ convex with $S = -S$. If $\overline{S}$ has a non-empty interior, then $\overline{S}$ is a neighbourhood of zero.*

*Proof.* First note that $\overline{S}$ is convex. If $x, y \in S$ and $x_n, y_n \in S$ with $x_n, y_n \to x, y$ then $tx_n + (1 - ty_n) \in S$ for all $n$ and $t \in [0,1]$. Letting $n \to \infty$ we get $tx + (1-t)y \in \overline{S}$ for all $t \in [0,1]$ and so $\overline{S}$ is convex. We also easily have $\overline{S} = -\overline{S}$. If $\overline{S}$ has a non-empty interior, there exists $z \in \overline{S}$ and $\epsilon > 0$ such that $B(z,\epsilon) \subseteq \overline{S}$. Therefore $z \pm h \in \overline{S}$ whenever $\|h\| < \epsilon$ and since $\overline{S} = -\overline{S}$ we also have $-(z \pm h) \in \overline{S}$. By the convexity of $\overline{S}$ we have

$$y = \frac{1}{2}((x+h) + (-x+h)) \in \overline{S}$$

whenever $\|h\| < \epsilon$. Hence $B(0,\epsilon) \subseteq \overline{S}$, and so $\overline{S}$ is a neighbourhood of zero. □

**Theorem 4.10** (Open mapping theorem)**.** *Suppose that $X$ and $Y$ are Banach spaces. If $T \in \mathcal{L}(X,Y)$ is surjective, then $T$ is open.*

*Proof.* As $T$ is surjective we have

$$Y = \bigcup_{n \in \mathbb{N}} \overline{T(\overline{B(0,n)})}$$

with $[T(\overline{B(0,n)})]$ closed for all $n \in \mathbb{N}$. Since $Y$ is complete, by a corollary to Baire's theorem, there exists $n \in \mathbb{N}$ such that $\overline{T(\overline{B(0,n)})}$ has non-empty interior. Since the map $x \mapsto nx$ is a homeomorphism and $T$ is linear, the set $\overline{T(\overline{B(0,1)})}$ has non-empty interior as well. Now $\overline{B(0,1)}$ is convex and $\overline{B(0,1)} = -\overline{B(0,1)}$. By linearity of $T$ we have that

$$T(\overline{B(0,1)}) = -T(\overline{B(0,1)})$$

is convex as well. Since we know that $\overline{T(\overline{B(0,1)})}$ has non-empty interior, the previous lemma implies that $\overline{T(\overline{B(0,1)})}$ is a neighbourhood of zero, and thus there exists $r > 0$ such that

$$B(0,r) \subseteq \overline{T(\overline{B(0,1)})}$$

and since $X$ is Banach the previous lemma shows that $T$ is open. □

**Theorem 4.11** (Closed Graph theorem)**.** *Let $X, Y$ be Banach spaces, and $T \in Hom(X,Y)$. Then $T \in \mathcal{L}(X,Y)$ if and only if $\Gamma(T)$ is closed in $X \times Y$.*[3]

[3] Prove the following statement.

**Exercise 4.12.** If $X, Y$ are vector spaces, and if $T : X \to Y$ is linear, then $\Gamma(T)$ is a subspace of $X \times Y$. Moreover, if $X, Y$ are normed vectors paces, with

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

*Proof.* Suppose $T \in \mathcal{L}(X, Y)$. If $x_n \to x$ in $X$, then

$$(x_n, Tx_n) \to (x, Tx)$$

by continuity of $T$, and so $\Gamma(T)$ is closed.

Conversely, suppose that $\Gamma(T)$ is closed in $X \times Y$. Define a norm $\|\cdot\|_\Gamma$ on $X$ by $\|x\|_\Gamma = \|x\| + \|Tx\|$. Since $\Gamma(T)$ is closed, and since $(X, \|\cdot\|)$ is Banach, then $(X, \|\cdot\|_\Gamma)$ is also a Banach space (exercise). Note that $\|x\| \leq \|x\|_\Gamma$. So by a corollary to the Open Mapping theorem, $\|\cdot\|$ and $\|\cdot\|_\Gamma$ are equivalent. So there is $c > 0$ with

$$\|x\|_\Gamma \leq c\|x\| \quad \forall x \in X.$$

So $\|x\| + \|Tx\| \leq c\|x\|$, and so $\|Tx\| \leq (c-1)\|x\|$, and so $T$ is continuous. □

# 5
# *Spectral Theory*

RECALL THAT THE eigenvalues of an $n \times n$ matrix $T$ over $\mathbb{C}$ are the $\lambda \in \mathbb{C}$ with

$$\det(\lambda I - T) = 0$$

that is, $\lambda I - T$ is not invertible.[1,2]

[1] Recall that showing the existence of eigenvalues is equivalent to the fundamental theorem of algebra.

[2] We need our base field to be $\mathbb{C}$ to get reasonable spectral theory.

**Definition 5.1.** Write $\mathcal{L}(X) = \mathcal{L}(X, X)$.

**Definition 5.2.** Let $X$ be a Banach space over $\mathbb{K}$, and let $T \in \mathcal{L}(X)$. Then the spectrum of $T$ is[3]

$$\sigma(T) = \{\lambda \in \mathbb{K} \mid \lambda I - T \text{ is not invertible}\}.$$

[3] $\lambda I - T$ is non invertible if either $\lambda I - T$ is not injective, or $\lambda I - T$ is not surjective.

**Definition 5.3** (Eigenvalue)**.** $\lambda \in \mathbb{K}$ is an eigenvalue of $T \in \mathcal{L}(X)$ if there is $x \neq 0$ with $Tx = \lambda x$, i.e. $\lambda$ is an eigenvalue if and only if $\lambda I - T$ is not injective.[4]

[4] If $\dim(X) < \infty$, then $X \backslash \textsc{Ker}\,(T) \simeq \operatorname{im}(T)$, and so $T$ is injective if and only if $T$ is surjective. This fails in the infinite dimensional case - consider the left and right shift operators on $\ell^2$.

**Theorem 5.4.** *Let $X \neq \{0\}$ be a Banach space over $\mathbb{C}$, and let $T \in \mathcal{L}(X)$. Then $\sigma(T)$ is a non-empty, compact (closed and bounded) subset of*

$$\{\lambda \in \mathbb{C} \mid |\lambda| \leq \|T\|\}$$

**Example 5.5.** Let $L, R : \ell^2 \to \ell^2$ be the left and right shift operators.

Then $\|L\| = 1$, and so $\sigma(L) \subseteq \overline{D}(0,1)$. If $|\lambda| < 1$, then

$$L(\lambda, \lambda^2, \lambda^3, \dots) = (\lambda^2, \lambda^3, \lambda^4, \dots) = \lambda(\lambda, \lambda^2, \lambda^3, \dots)$$

and so $\lambda$ is an eigenvalue. Thus $D(0,1) \subseteq \sigma(L) \subseteq \overline{D}(0,1)$. But $\sigma(L)$ is closed, and so $\sigma(L) = \overline{D}(0,1)$. Are the $\lambda$ with $|\lambda| = 1$ eigenvalues? No - suppose $|\lambda| = 1$ and $x \neq 0$ with $Lx = \lambda x$.

Then

$$L^n(x) = \lambda^n x.$$

Thus, $x_{n+1} = \lambda^n x_1$. Then $x = (x_1, \lambda, x_1, \lambda^2 x_1, \dots)$ which is not in $\ell^2$.

Then $\|R\| = 1$, and so $\sigma(R) \subseteq \overline{D}(0,1)$.[5]

[5] $LR(x) = L(0, x_1, \dots) = (x_1, x_2, \dots)$, so

$$LR = I \qquad (\star)$$

*Remark.* Unlike $\dim(X) < \infty$, $(\star)$ does NOT say that $R$ is invertible ($RL = I$).

Consider the operator $L(\lambda I - R) = \lambda L - I = -\lambda(\lambda^{-1}I - L)$. If $0 < |\lambda| < 1$, then we know that $\lambda^{-1}I - L$ is invertible (as $\lambda^{-1} \notin \sigma(L)$). So if $\lambda I - R$ were invertible, then $L$ is invertible, which is false. Thus $\lambda \in \sigma(R)$. Hence

$$D(0,1)\backslash\{0\} \subseteq \sigma(R) \subseteq \overline{D}(0,1).$$

Since $\sigma(R)$ is closed, $\sigma(R) = \overline{D}(0,1)$.

**Theorem 5.6.** *Let $X \neq \{0\}$ be a Banach space over $\mathbb{C}$. Let $T \in \mathcal{L}(X)$. Then $\sigma(T)$ is a nonempty, compact subset of*

$$\{\lambda \in \mathbb{C} \mid |\lambda| \leq \|T\|\}.$$

**Lemma 5.7.** *With above assumptions $\sigma(T) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq \|T\|\}$.*

*Proof.* We need to show that if $|\lambda| > \|T\|$ then $\lambda I - T$ is invertible.

**Technique: Geometric series**. We guess

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

We now verify this guess. Since

$$\sum_{k=0}^{\infty} \frac{\|T^k\|}{|\lambda|^{k+1}} \leq \sum_{k=0}^{\infty} \frac{\|T\|^k}{|\lambda|^{k+1}} < \infty,$$

the series $S = \sum_{k=0}^{\infty} \frac{T^k}{\lambda^{k+1}}$ converges in X.

We now show that $S$ is the inverse of $\lambda I - T$. As we are working in infinite dimensions, we ned to check left and right inverses. Let $S_n = \sum_{k=1}^{n-1} \frac{T^k}{\lambda^{k+1}}$. Then

$$\begin{aligned} S_n(\lambda I - T) &= \left(\sum_{k=0}^{n-1} \frac{T^k}{\lambda^{k+1}}\right)(\lambda I - T) \\ &= I - \frac{T^n}{\lambda^n} \to I \\ (\lambda I - T)S_n &= I - \frac{T^n}{\lambda^n} \to I \end{aligned}$$

and so $S(\lambda I - T) = (\lambda I - T)S$ and so $\lambda I - T$ is invertible. □

**Exercise 5.8.** Show that if $\|I - T\| < 1$ then $T$ is invertible with inverse $\sum_{k=0}^{\infty}(I - T)^k$ *Hint: Consider*

$$\frac{1}{T} = \frac{1}{I - (I - T)}.$$

In particular, the ball $B(I, 1)$ in $\mathcal{L}(X)$ consists of invertible elements.

The following is used to show $\sigma(T)$ is closed and nonempty, it is also interesting in its own right.

**Proposition 5.9.** *Let $X$ be Banach over $\mathbb{K}$. Let $GL(X) = \{T \in \mathcal{L}(X) \mid T$ invertible. Then*

*(a) $GL(X)$ is a group under composition of operators.*

*(b) $GL(X)$ is open in $\mathcal{L}(X)$.*

*(c) The map*

$$\begin{aligned} \varphi : GL(X) &\to GL(X) \\ T &\mapsto T^{-1} \end{aligned}$$

*is continuous.*

*Proof.* (a) The open mapping theorem tells us that if $T \in \mathrm{GL}(X)$ then $T^{-1} \in \mathcal{L}(X)$, and so $T^{-1} \in \mathrm{GL}(X)$. The rest is clear.

(b) Let $T_0 \in \mathrm{GL}(X)$. We claim

$$B\left(T_0, \frac{1}{\|T_0^{-1}}\|\right) \subseteq \text{GL(X)}.$$

We have

$$\begin{aligned} \|I - T_0^{-1}T\| &= \|T_0^{-1}(T_0 - T)\| \\ &\leq \|T_0^{-1}\|\|T_0 - T\| \\ &< 1 \quad \text{as } T \in B\left(T_0, \frac{1}{\|T_0^{-1}\|}\right) \end{aligned}$$

(c) We have

$$\begin{aligned} \|T_0^{-1} - T^{-1}\| &= \|T^{-1}(T - T_0)T_0^{-1}\ | \\ &\leq \|T^{-1}\|T - T_0\|\|T_0^{-1}\| \qquad (\star) \end{aligned}$$

If $\|T - T_0\| \leq \frac{1}{2\|T_0^{-1}\|}$, then

$$\begin{aligned} \|I - TT_0^{-1}\| &= \|(T_0 - T)T_0^{-1}\| \\ &\leq \|T_0 - T\|\|T_0^{-1}\| \\ &\leq \frac{1}{2}. \end{aligned}$$

We then have

$$\begin{aligned} \|T_0T^{-1}\| &= \|(TT_0^{-1})^{-1}\| \\ &= \|\sum_{k=0}^{\infty}(I - TT_0^{-1})^k\| \\ &\leq \sum_{k=0}^{\infty}\|I - TT_0^{-1}\|^k \\ &\leq 2 \end{aligned}$$

Hence $\|T^{-1}\| = \|T_0^{-1}(T_0T^{-1})\| \le \|T_0^{-1}\|\|T_0T^{-1}\| \le 2\|T_0^{-1}\|$, and from $(\star)$, we have

$$\|T_0^{-1} - T^{-1}\| \le 2\|T_0^{-1}\|^2\|T - T_0\|$$

and so $T \mapsto T^{-1}$ is continuous.

□

**Corollary.** *$\sigma(T)$ is closed.*

*Proof.* Let

$$\begin{aligned} f : \mathbb{C} &\to \mathcal{L}(X) \\ \lambda &\mapsto \lambda I - T \end{aligned}$$

This is continuous, as

$$\begin{aligned} \|f(\lambda) - f(\lambda_0)\| &= \|(\lambda - \lambda_0)I\| \\ &= |\lambda - \lambda_0| \end{aligned}$$

and

$$\sigma(T) = f^{-1}\left(\mathcal{L}(X)\backslash \mathrm{GL}(X)\right)$$

which is the inverse image of a closed set, and hence is closed. □

So $\sigma(T)$ is a compact subset of $\{\lambda \in \mathbb{C} \,|\, |\lambda| \le \|T\|\}$. Write $\rho(T) = \mathbb{C}\backslash\sigma(T)$ (the **resolvent set**), and let $R_T = R : \rho(T) \to \mathcal{L}(X)$ with $R_T(\lambda) = (\lambda I - T)^{-1}$.

**Theorem 5.10.** *Let $\mathbb{K} = \mathbb{C}$ and $X \ne \{0\}$ and $T \in \mathcal{L}(X)$. Then $\sigma(T) \ne \emptyset$.*

*Proof.* We use Lioville's theorem - a bounded entire function must be constant.

Let $\varphi = \mathcal{L}(X)'$ (hence $\varphi : \mathcal{L}(X) \to C$.) Let

$$\begin{aligned} f_\varphi : \rho(T) &\to \mathbb{C} \\ \lambda &\mapsto \varphi(R(\lambda)) \end{aligned}$$

**Lemma 5.11.** *$f_\varphi$ is analytic on $\rho(T)$.*

*Proof.* We show $f_\varphi$ is differentiable. Consider

$$\begin{aligned} \frac{f_\varphi(\lambda) - f_\varphi(\lambda_0)}{\lambda - \lambda_0} &= \varphi\left(\frac{R(\lambda) - R(\lambda_0)}{\lambda - \lambda_0}\right) \\ &= \varphi\left(\frac{(\lambda I - T)^{-1} - (\lambda_0 I - T)^{-1}}{\lambda - \lambda_0}\right) \\ &= \varphi\left(\frac{(\lambda_0 I - T)^{-1}((\lambda_0 - \lambda)I)(\lambda I - T)^{-1}}{\lambda - \lambda_0}\right) \\ &= -\varphi\left((\lambda_0 I - T)^{-1}(\lambda I - T)^{-1}\right) \\ &\to -\varphi\left((\lambda_0 I - T)^{-2}\right) \end{aligned}$$

as $\lambda \to \lambda_0$, where we use the fact that $\varphi$ is continuous and $T \to T^{-1}$ is continuous. So $f_\varphi$ is analytic on $\rho(T)$ for all $\varphi \in \mathcal{L}(X)'$. $\square$

Now suppose that $\sigma(T) = \emptyset$. Then $f_\varphi : \mathbb{C} \to \mathbb{C}$ is analytic.

**Lemma 5.12.** *$f_\varphi$ is bounded.*

*Proof.* If $|\lambda > \|T\|$, then

$$\begin{aligned} f_\varphi(\lambda) &= \left|\varphi\left((\lambda I - T)^{-1}\right)\right| \\ &= \left|\varphi\left(\sum_{k=0}^{\infty} \frac{T^k}{\lambda^{k+1}}\right)\right| \\ &\leq \|\varphi\| \left\|\sum_{k=0}^{\infty} \frac{T^k}{\lambda^{k+1}}\right\| \\ &\leq \|\varphi\| \sum_{k=0}^{\infty} \frac{\|T\|^k}{|\lambda|^{k+1}} \\ &= \frac{\|\varphi\|}{|\lambda| - \|T\|} \to 0 \end{aligned}$$

as $|\lambda| \to \infty$. So $f_\varphi$ is bounded, entire, and thus $f_\varphi = c$ by Lioville's theorem. By the above, $f_\varphi(\lambda) = 0$ for all $\lambda$. Hence $\varphi(R(\lambda)) = 0$ for all $\lambda, \varphi$.

Thus from Hahn-Banach, $R(\lambda) = 0$ for all $\lambda$ which is a contradiction, as the zero operator is not invertible if $X \neq \{0\}$. $\square$

$\square$

**Theorem 5.13** (Spectral mapping theorem (polynomials))**.** *Let $T$ be an $n \times n$ matrix over $\mathbb{C}$. If we know all the eigenvalues of $T$, then we know the eigenvalues of every polynomial $p(T) = a_0 + a_1 T + \cdots + a_n T^n$. Specifically,*

$$\{\text{eigenvalues of } p(T)\} = \{p(\lambda) \mid \lambda \text{ is an eigenvalue of } T\}$$

*Therefore*

$$\sigma(p(T)) = p(\sigma(T)).$$

*This is called the **spectral mapping theorem** (for matrices/polynomials).*

*This also holds for $X$ Banach over $\mathbb{C}$, and $T \in \mathcal{L}(X)$.*

**Lemma 5.14.** *Let $\mathbb{C}[t]$ be the algebra of polynomials in $t$ with complex coefficients. Multiplication is defined as usual.*

**Lemma 5.15.** *Let $X$ be Banach over $\mathbb{C}$. Let $T \in \mathcal{L}(X)$. Then*

$$\begin{aligned} \varphi : \mathbb{C}[t] &\to \mathcal{L}(X) \\ p &\mapsto p(T) \end{aligned}$$

*is an algebra homomorphism (multiplication corresponds to composition in $\mathcal{L}(X)$.)*

*Proof.* Simply check

$$\begin{aligned}\varphi(p_1 + p_2) &= \varphi(p_1) + \varphi(p_2)\\ \varphi(p_1 p_2) &= \varphi(p_1)\varphi(p_2)\\ \varphi(\alpha p) &= \alpha\varphi(p)\end{aligned}$$

for all $p_1, p_2, p \in \mathbb{C}[t], \alpha \in \mathbb{C}$. □

**Theorem 5.16.** *Let $X$ be Banach over $\mathbb{C}$, and let $T \in \mathcal{L}(X)$. Then*

$$\sigma(p(T)) = p(\sigma(T)).$$

*Proof.* If $p = c$ is constant, then $p(T) = cI$ has spectrum

$$\sigma(p(T)) = \sigma(cI) = \{c\}$$

On the other hand,

$$p(\sigma(T)) = \{c\}$$

Now, suppose that $p$ is non constant. Let $\mu \in \mathbb{C}$ fixed. By the fundamental theorem of algebra, we can factorise $\mu - p(t)$ as

$$\alpha(t - \lambda_1)^{m_1} \dots (t - \lambda_n)^{m_n}$$

where $\lambda_1, \dots, \lambda_n$ are the distinct roots of $\mu - p(t)$. Note that $\mu = p(\lambda_i)$ for each $i$. Applying $\psi : \mathbb{C}[t] \to \mathcal{L}(X)$ from above, we have

$$\mu I - p(T) = \alpha(T - \lambda_1 I)^{m_1} \dots (T - \lambda_n I)^{m_n}$$

We know

$$\begin{aligned}\mu \in \sigma(p(T)) &\iff \mu - p(T) \text{ is not invertible}\\ &\iff T - \lambda I \text{ non invertible for some } i\\ &\iff \lambda \in \sigma(T) \text{ for some } i\\ &\iff \mu = p(\lambda_i) \in p(\sigma(T))\end{aligned}$$

and so

$$\sigma(p(T)) = p(\sigma(T))$$

□

**Exercise 5.17.** If $T_1, \dots, T_n \in \mathcal{L}(X)$ which commute with each other, then $T_1 \dots T_n$ is invertible if and only if the individual elements are invertible.

**Definition 5.18** (Spectral radius)**.** Let $X \neq \{0\}$ be a Banach space over $\mathbb{C}$. The **spectral radius** of $T \in \mathcal{L}(X)$ is

$$\begin{aligned}r(T) &= \sup\{|\lambda| \mid \lambda \in \sigma(T)\}\\ &= \max\{|\lambda| : \lambda \in \sigma(T)\}\end{aligned}$$

*Note.*

$$r(T) \leq \|T\|$$

since $\sigma(T) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq \|T\|\}$. Strict inequality can (and often does) occur.

**Example 5.19.** Let

$$T = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then consider $T : \mathbb{C}^2 \to \mathbb{C}^2$ where $\|(x,y)\|_2 = \sqrt{|x|^2 + |y|^2}$. Then

$$\begin{aligned} \|T\| &= \sup\{\|Tx\|_2 \mid x \in \mathbb{C}^2\} \\ &= \sqrt{\lambda_{max}(T^*T)} \end{aligned}$$

where

$$T^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

is conjugate transpose. Then $\|T\| = 1$. But $\sigma(T) = \{0\}$, and so $r(T) = 0 < 1 = \|T\|$.

**Theorem 5.20** (Gelfand, 1941)**.** *Let $X \neq \{0\}$ be Banach over $\mathbb{C}$, and let $T \in \mathcal{L}(X)$. Then*

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

*In particular, the limit exists.*

*Proof.* By the spectral mapping theorem,

$$\sigma(T^n) = \{\sigma(T)\}^n = \{\lambda^n \mid |\lambda \in \sigma(T)\}.$$

So

$$\begin{aligned} r(T) &= r(T^n)^{1/n} \\ &\leq \|T^n\|^{1/n}. \end{aligned}$$

So

$$r(T) \leq \liminf_{n\to\infty} \|T^n\|^{1/n}$$

Now, we must show that

$$\limsup_{n\to\infty} \|T^n\|^{1/n} \leq r(T).$$

Let $\varphi \in \mathcal{L}(X)$ and let

$$\begin{aligned} f_\varphi : \rho(T) &\to \mathbb{C} \\ \lambda &\mapsto \varphi((\lambda I - T)^{-1}) \end{aligned}$$

We saw that $f_\varphi$ is analytic on $\rho(T)$. We also have

$$f_\varphi(\lambda) = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} \varphi(T^n) \qquad (\star)$$

if $|\lambda| > \|T\|$. By general theory of Laurent series, $(\star)$ actually holds for all $\lambda \in \rho(T)$. In particular, it holds if $|\lambda| > r(T)$.

Thus,

$$\lim_{n\to\infty} \frac{1}{\lambda^{n+1}} \varphi(T^n) = 0 \quad \boxed{|\lambda| > r(T)}$$

Sp for each $\varphi \in \mathcal{L}(X)'$, and each $|\lambda| > r(T)$, there is $C_{\lambda,\varphi}$ such that

$$\left|\varphi\left(\frac{1}{\lambda^{n+1}} T^n\right)\right| \le C_{\lambda,\varphi} \quad \forall n \ge 0$$

Then by the principle of uniform boundedness, there exists a constant $C_\lambda$ such that

$$\left\|\frac{1}{\lambda^{n+1}} T^n\right\| \le C_\lambda \quad \forall n \ge 0$$

So $\|T^n\|^{1/n} \le |\lambda|(C_\lambda|\lambda|)^{1/n}$, which gives

$$\limsup_{n\to\infty} \|T^n\|^{1/n} \le \lambda$$

for all $|\lambda| > r(T)$. So

$$\limsup_{n\to\infty} \|T^n\|^{1/n} \le r(T)$$

We used the following lemma.

**Lemma 5.21.** *Let $X$ be a normed vector space, $A \subseteq X$ a subset. We say that*

*(1) $A$ is **bounded** if there exists $C > 0$ with $\|x\| \le C$, for all $x \in A$.*

*(2) $A$ is **weakly bounded** if for each $\varphi \in X'$, there exists $C_\varphi > 0$ such that*

$$|\varphi(x)| \le C_\varphi$$

*for all $x \in A$.*

*Then we have*

$$A \subseteq X \text{ is bounded} \iff \text{weakly bounded}$$

*Proof.* $A$ bounded $\Rightarrow \|x\| \le C$ for all $x \in A \Rightarrow |\varphi(x)| \le \|\varphi\|\|x\| \le \|\varphi\|C$. So $A$ is weakly bounded.

Now, suppose $A$ is weakly bounded. For each $x \in X$, let $\hat{x} \in X''$ with

$$\hat{x}(\varphi) = \varphi(x).$$

So $|\hat{x}(\varphi)| \le C_\varphi$ for all $x \in A$. By the principle of uniform boundedness,

$$\|\hat{x}\| \le C$$

for all $x \in A$, and since $\|\hat{x}\| = \|x\|$. Thus $A$ is bounded. □

□

# 6
# Compact Operators

We now turn to compact operators. In general, calculating $\sigma(T)$ is difficult, but for compact operators on a complex Banach space, we have a fairly explicit theory.

**Theorem 6.1.** *Let $X$ be a complex Banach space, with $\dim(X) = \infty$. Let $T : X \to X$ be a compact operator. Then*

*(1)* $0 \in \sigma(T)$.

*(2)* $\sigma(T)\backslash\{0\} = \sigma_p(T)\backslash\{0\}$*, that is, each* $\lambda \in \sigma(T)\backslash\{0\}$ *is an eigenvalue of* $T$ *(*$0$ *may or may not be an eigenvalue.)*

*(3) We are in exactly one of the cases:*

- $\sigma(T) = \{0\}$.
- $\sigma(T)\backslash\{0\}$ *is finite (nonempty).*
- $\sigma(T)\backslash\{0\}$ *is a sequence of points converging to* $0$.

*(4) Each* $\lambda \in \sigma(T)\backslash\{0\}$ *is isolated, and the eigenspace* $\text{Ker}\,(\lambda I - T)$ *is finite dimensional.*

*where* $\sigma_p(T)$ *is the* ***point spectrum of*** $T$*, where*

$$\begin{aligned}\sigma_p(T) &= \{\lambda \in \mathbb{K} \mid \lambda I - T \text{ is not injective}\}\\ &= \{\lambda \in \mathbb{K} \mid \text{there exits nonzero vector } x \text{ with } (\lambda I - T)x = 0\}\\ &= \{\text{eigenvalues of } T\}\end{aligned}$$

*Proof.* We shall prove these results next week. □

**Definition 6.2.** Let $X, Y$ be normed vector spaces. An operator $T : X \to Y$ is **compact** if $T$ is linear, and if $B \subseteq X$ is bounded then $T(B)$ is relatively compact (a set is relatively compact if its closure is compact.) Symbolically,

$$B \subseteq X \text{ bounded} \Rightarrow \overline{T(B)} \text{ compact}$$

**Lemma 6.3.** *If $T$ is compact, then $T$ is continuous.*

*Proof.* The closed ball $B = \{x \in X \mid \|x\| \leq 1\}$ is bounded, and so if $T$ is a compact operator, then $\overline{T(B)}$ is compact, and hence bounded. Hence $\|Tx\| \leq M$ for all $\|x\| \leq 1$, so $T$ is continuous, with $\|T\| \leq M$. □

We now recall definitions of compactness

**Theorem 6.4** (Characterisations of compactness)**.** *Let $X$ be a metric space. The following are equivalent.*

*(1) $X$ is **compact** (every open cover has a finite subcover).*

*(2) $X$ is **sequentially compact** (every sequence in $X$ has a convergent subsequence)*

**Lemma 6.5.** *Let $X$ be a compact set. Let $Y \subseteq X$. If $Y \subseteq X$ is closed, then $Y$ is compact.*

**Lemma 6.6.** *Let $V$ be a finite dimensional vector space. If $X \subseteq V$ is closed and bounded, then $X$ is compact.*

**Theorem 6.7** (Characterisations of compact operators)**.** *Let $X, Y$ be normed vector spaces over $\mathbb{K}$. Let $T \in \mathcal{L}(X, Y)$. Then the following are equivalent.*

*(a) $T$ is compact.*

*(b) $\overline{T(B)}$ is compact, where $B = \{x \in X \mid \|x\| \leq 1\}$.*

*(c) If $(x_n)_{n\geq 1}$ is bounded in $X$, then $(Tx_n)_{n\geq 1}$ has a convergent subsequence (**sequentially compact**).*

*Proof.* $(a) \Rightarrow (b)$ by definition.

$(b) \Rightarrow (a)$. Suppose $(b)$ holds. Let $B_1 \subseteq X$ be bounded. Then $B_1 \subseteq \alpha B$ for some $\alpha > 0$. So

$$\overline{T(B)} \subseteq \overline{T(\alpha B)} = \alpha \overline{T(B)}$$

which is a closed subset of a compact set, and hence compact.

$(a) \Rightarrow (c)$. Suppose $T$ is compact. Let $(x_n)_{n\geq 1}$ be bounded sequence in $X$. Then $T(B) = \{Tx_n \mid n \geq 1\}$ is relatively compact. So $\overline{T(B)}$ is compact, and hence is sequentially compact, and so has a convergence subsequence.

$(c) \Rightarrow (a)$. Let $B \subseteq X$ be bounded. Let $(y_n)_{n\geq 1}$ be a sequence in $T(B)$. Then there is $x_n \in B$ with $Tx_n = y_n$. So $(x_n)_{n\geq 1}$ is a bonded sequence. By assumption $(Tx_n)_{n\geq 1}$ has a convergent subsequence. So $\overline{T(B)}$ is sequentially compact, and hence compact. □

**Corollary.** *The set $\{$compact operators $T : X \to Y\}$ is a vector space. That is, if $T_1, T_2$ are compact, then $T_1 + T_2$ and $\alpha T_1$ are compact.*[1]

[1] As an exercise, prove this statement. Use $(c)$ from Theorem 6.7.

**Corollary.**

$$\mathcal{K}(X,Y) \subseteq \mathcal{L}(X,Y) \subseteq Hom(X,Y)$$

*where $\mathcal{K}(X,Y)$ is the set of compact operators $T : X \to Y$.*

**Example 6.8** (Finite rank operators)**.** Let $X, Y$ be normed vector spaces, and let $T \in \mathcal{L}(X,Y)$. If $\dim(\text{Im } T) < \infty$, then $T$ is said to have **finite rank.** Then if $T$ has finite rank, then $T$ is compact.

*Proof.* Let $(x_n)$ be a bounded sequence in $X$. Then $\|Tx_n\| \leq \|T\|\|x_n\|$ so $(Tx_n)$ is a bounded sequence in $\text{Im } T$. But $\text{Im } T$ is finite dimensional, and so $\overline{\{Tx_n \mid n \geq 1\}}$ is compact (closed and bounded), and so $(Tx_n)_{n\geq 1}$ has a convergent subsequence. By $(c)$ in Theorem 6.7, $T$ is compact. □

**Lemma 6.9.** *Let $X, Y$ be normed vector spaces. If $T \in \mathcal{L}(X,Y)$ has finite rank, then there exists $y_1, \ldots, y_n \in \text{Im } T$ and $\varphi_1, \ldots, \varphi_n \in X'$ with $Tx = \sum_{j=1}^{n} \varphi_j(x) y_j$ for all $x \in X$, with $n = \dim(\text{Im } T)$.*

*Proof.* Choose a basis $y_1, \ldots, y_n$ of $\text{Im } T$. For each $j = 1, \ldots, n$, define $\alpha_j \in (\text{Im } T)'$ by

$$\alpha_j(a_1 y_1 + \cdots + a_n y_n) = a_j$$

i.e. coordinate projection. By Hahn-Banach, we can extend $a_j$ to a continuous linear functional $\tilde{a}_j \in Y'$. Let $\varphi_j = \tilde{a}_j \circ T : X \to \mathbb{K}$. So $\varphi_j \in X'$. Since

$$y = \sum_{j=1}^{n} \tilde{a}_j(y) y_j \quad \forall y \in \text{Im } T$$

we have

$$\begin{aligned} Tx &= \sum_{j=1}^{n} \tilde{a}_j(Tx) y_j \\ &= \sum_{j=1}^{n} (\alpha_j \circ T)(x) y_j \\ &= \sum_{j=1}^{n} \varphi_j(x) y_j \quad \forall x \in X. \end{aligned}$$

□

Recall that the closed unit ball in $X$ is compact if and only if $\dim(X) < \infty$. Then it follows that the identity map $I : X \to X$ is compact if and only if $\dim(X) < \infty$. Hence,

$$\mathcal{K}(X) \subsetneq \mathcal{L}(X) \subsetneq \text{Hom}(X,X)$$

when $\dim(X) = \infty$.

Consider a sequence of compact operators $T_n$. If $T_n$ is compact and $T_n \to T$, then $T$ is compact.

**Lemma 6.10** (Riesz's Lemma). *Let $X$ be a normed vector space. Let $Y \subsetneq X$ be a proper **closed** subspace. Let $\theta \in (0,1)$ be given. Then there exists $x$ with $\|x\| = 1$ such that $\|x - y\| \geq \theta$ for all $y \in Y$.*

*Proof.* Pick any $z \in X \backslash Y$. Let $\alpha = \inf_{y \in Y} \|z - y\| > 0$ since $Y$ is closed. Then by the definition of the infimum, there is $y_0 \in Y$ with $\alpha \leq \|z - y_0\| \leq \frac{\alpha}{\theta}$. Now let $x = \frac{x - y_0}{\|z - y_0\|}$. Then $\|x\| = 1$.

Now,

$$\begin{aligned}
\|x - y\| &= \left\| \frac{z - y_0}{\|z - y_0\|} - y \right\| \\
&= \frac{1}{\|z - y_0\|} \|z - y_- + \|z - y_0\| y\| \\
&\geq \frac{\theta}{\alpha} \alpha = \theta
\end{aligned}$$

□

**Corollary.** *Let $X$ be a normed vector space. The closed unit ball $\overline{B}(0,1)$ is compact if and only if $\dim(X) < \infty$.*

*Proof.* If $\dim(X) < \infty$ then $\overline{B}(0,1)$ is compact (since closed and bounded if and only if compact in finite dimensions). Now suppose $\dim(X) = \infty$. Build a sequence $(x_n)$ with $\|x_n\| = 1$ with no convergent subsequence. Choose finite dimensional subspaces

$$\{0\} = X_0 \subsetneq X_1 \subsetneq X_2 \subsetneq \dots$$

These are all closed (finite dimensional spaces are complete, and hence closed). Use the lemma to choose $x_k \in X_k$ with $\|x_k\| = 1$, $\|x_k - x\| \geq \frac{1}{2}$ for al $x \in X_{k-1}$. So $x_k - x\| \geq \frac{1}{2}$ for all $x \in X_j (j \leq k-1)$. So $\|x_n - x_m\| \geq \frac{1}{2}$ for all $m, n \geq 1$. So $(x_n)$ has no convergent subsequence, and so $\overline{B}(0,1)$ is not compact. □

**Corollary.** *$I : X \to X$ is compact if and only if $\dim(X) < \infty$.*

*Proof.* Recall $T$ is compact if and only if $T(\overline{B}(0,1))$ is relatively compact. □

One way to show that an operator is compact is to apply the following.

**Proposition 6.11.** *Let $X$ be a normed vector space, and let $Y$ be Banach. Suppose that $T_n \in \mathcal{K}(X,Y)$ for each $n \geq 1$. If $T_n \to T$ (in operator norm, $\|T_N - T\| \to 0$) then $T$ is compact.*

*Proof.* Let $(x_n)$ be a bounded sequence in $X$. We now construct a subsequence $(x'_n)$ for which $(Tx'_n)$ converges.

- Since $T_1$ is compact, $(x_n)$ has a subsequence $x_n^{(1)}$ such that $(T_1 x_n^{(1)})$ converges.

- Since $T_2$ is compact and $x_n^{(1)}$ is bounded, there is a subsequence $x_n^{(2)}$ such that $T_2 x_n^{(2)}$ converges.
- Continuing, we can form a subsequence $x_n^{(k)}$ such that $T_k x_n^k$ converges.

Let $x'_n = x_n^{(n)}$. Then $(x'_n)$ is a subsequence of $(x_n^{(1)})$, and $(x'_n)_{n\geq 2}$ is a subsequence of $(x_n^{(2)})$, etc. So for each fixed $k \geq 1$, $(T_k x'_n)$ converges.

We now show $Tx'_n$ is Cauchy, and hence converges. We have

$$\|Tx'_m - Tx'_n\| \leq \|Tx'_m - T_k x'_m\| + \|T_K x'_m - T_k x'_n\| + T_k x'_n - Tx'_n\|$$

where $k$ is to be chosen. Suppose $\|x_n\| \leq M$ for all $n \geq 1$. Then

$$\|Tx'_m - Tx'_n\| \leq 2M\|T - T_k\| + \|T_k x' + m - T_k x'_n\|$$

Let $\epsilon > 0$ be given. Since $\|T - T_k\| \to 0$ as $k \to \infty$, fix a $k$ for which $\|T - T_k\| \leq \frac{\epsilon}{3M}$. For this fixed $k$, we know $(T_k x'_n)$ converges, and so is Cauchy. So there exists $N < 0$ such that $\|T_k x' m - T_k x'_n\| < frac\epsilon 3$ for all $m, n < N$. Hence $\|Tx'_m - Tx'_n\| \leq \frac{2M}{\epsilon} 3M + \frac{\epsilon}{3} = \epsilon$ for all $m, n > N$, so is Cauchy, and so converges. □

**Example 6.12.** Let $K(x, y) \in L^2(\mathbb{R}^2)$. Define $T : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ by

$$Tf(x) = \int_{\mathbb{R}} K(x, y) f(y)\, dy$$

(Hilbert-Schimidt Integral operator)

**Proposition 6.13.** *$T$ is compact.*

*Proof.* Note that $\|Tf\|_2 \leq \|K\|_2 \|f\|_2$ for all $f \in L^2(\mathbb{R})$, where $\|K\|_2 = \left(\iint_{\mathbb{R}^2} |K(x, y)|^2\, dx\, dy\right)^{1/2}$. So $T$ is continuous, with $\|T\| \leq \|K\|_2$. We now exhibit $T$ as a limit of finite rank (hence compact) operators, with $T_n : L^2(\mathbb{R}) \to L^2(\mathbb{R})$. Once can see that there is a sequence $K_n \in L^2(\mathbb{R}^2)$ of the form

$$K_n(x, y) = \sum_{k=1}^{N_n} \alpha_k^{(n)}(x) \beta_k^{(n)}(y)$$

with $K_n \to K$ in $L^2(\mathbb{R}^2)$. Then $\|T_n - T\| \leq \|K_n - K\|_2 \to 0$, and so $T_n \to T$. Hence

$$T_n f(x) \sum_{k=1}^{N_n} \int_{\mathbb{R}} \alpha_k^{(n)}(x) \beta_k^{(n)}(y) f(y)\, dy$$

$$= \sum_{k=1}^{N_n} \left\langle f, \overline{\beta_k^{(n)}} \right\rangle \alpha_k^{(n)}(x)$$

and so $T_n f = \sum_{k=1}^{N_n} \langle f \rangle, \overline{\beta_k^{(n)}} \alpha_k^{(n)}$ from which we use that $T_n$ has finite rank. □

**Theorem 6.14.** *Let $X$ be a complex Banach space, with $\dim(X) = \infty$. Let $T : X \to X$ be a compact operator. Then*

*(1)* $0 \in \sigma(T)$.

*(2)* $\sigma(T)\backslash\{0\} = \sigma_p(T)\backslash\{0\}$*, that is, each* $\lambda \in \sigma(T)\backslash\{0\}$ *is an eigenvalue of $T$ (0 may or may not be an eigenvalue.)*

*(3) We are in exactly one of the cases:*

- $\sigma(T) = \{0\}$.
- $\sigma(T)\backslash\{0\}$ *is finite (nonempty).*
- $\sigma(T)\backslash\{0\}$ *is a sequence of points converging to* $0$.

*(4) Each* $\lambda \in \sigma(T)\backslash\{0\}$ *is isolated, and the eigenspace* $\mathrm{Ker}\ (\lambda I - T)$ *is finite dimensional.*

*where $\sigma_p(T)$ is the* ***point spectrum of*** *$T$, where*

$$\begin{aligned}\sigma_p(T) &= \{\lambda \in \mathbb{K} \,|\, \lambda I - T \text{ is not injective} \\ &= \{\lambda \in \mathbb{K} \,|\, \text{there exits nonzero vector } x \text{ with } (\lambda I - T)x = 0 \\ &= \{\text{eigenvalues of } T\}\end{aligned}$$

Compact operators are very well behaved with respect to composition.

**Proposition 6.15.** *Let $X, Y, Z$ be normed vector spaces.*

*(a) If $T \in \mathcal{K}(X,Y)$ and $S \in \mathcal{L}(Y,Z)$, then $ST \in \mathcal{K}(X,Z)$.*

*(b) If $S \in \mathcal{L}(X,Y)$ and $T \in \mathcal{K}(Y,Z)$, then $TS \in \mathcal{K}(X,Z)$.*

*Proof.* (a) Let $(x_n)$ be a bounded sequence in $X$. Since $T$ is compact, $Tx_n$ has a convergent subsequence, say $T_{x_{n_k}} \to y \in Y$. Then $(STx_n)$ has a convergent subsequence, namely $STx_{n_k} = S(T_{n_k}) \to Sy$ by continuity of $S$. So $ST$ is compact.

(b) Let $B \subseteq X$ be bounded. Then $S(B)$ is bounded in $Y$, as $S$ is continuous. So $TS(B) = T(S(B))$ is relatively compact since $T$ is compact. Hence $TS$ is compact.

□

**Corollary** (Part (1) of theorem)**.** *If $X$ is infinite dimensional Banach space, then $0 \in \sigma(T)$.*

*Proof.* If $0 \notin \sigma(T)$ then $T$ is invertible. By bounded inverse theorem $T^{-1}$ is continuous, and then $I = TT^{-1}$ is compact, which is a contradiction. □

**Theorem 6.16** (Part (3) of theorem). *Let $X$ be a normed vector space. Let $T \in \mathcal{K}(X)$. Then $T$ has at most countably many eigenvalues. If $T$ has infinitely many eigenvalues, then they can be arranged in a sequence converging to zero.*

*Proof.* We show that for each $N > 0$, we have

$$\#\{\lambda \in \sigma_p(T) \mid |\lambda| \geq N\} < \infty \qquad (\star)$$

Suppose that there is $N > 0$ such that $(\star)$ fails. So $\lambda_1, \lambda_2, \dots$ are distinct eigenvalues with $|\lambda_n| \geq N$ for $n = 1, 2, \dots$. Let $x_n \neq 0$ be an eigenvector. $Tx_n = \lambda_n x_n$, $n = 1, 2, \dots$. Let $X_n = \text{span}\,\{x_1, \dots, x_n\}$. Since $\{x_n \mid \geq 1\}$ are linearly independent, we have

$$X_1 \subsetneq X_2 \subsetneq \dots$$

and each $X_n$ is closed (finite dimensional).

By Reisz's Lemma from previous lecture, choose $y_n \in X_n$ such that $\|y_n\| = 1$ and $\|y_n - x\| \geq \frac{1}{2}$ for all $x \in X_{n-1}$. So $(y_n)$ is bounded in $X$. We show that $Ty_n$ has no convergence subsequence, contradicting compactness of $T$.

Let $m > n$. Then

$$\begin{aligned}\|Ty_m - Ty_n\| &= \|\lambda_m y_m - (\lambda_m y_m - Ty_m + Ty_n)| \\ &= |\lambda_m| \, \|y_m - (\text{something in } X_{m-1})\| \\ &\geq \frac{1}{2}|\lambda_m| \geq \frac{1}{2}N\end{aligned}$$

as required.

Note that $y_m = a_1 x_1 + \dots + a_m x_m$. Then

$$\begin{aligned}\lambda_m y_m - Ty_m &= \lambda_m a_1 x_1 + \dots + \lambda_m a_m x_m - (a_1 \lambda_1 x_1 + \dots + a_m \lambda_m x_m) \\ &= a_1(\lambda_m - \lambda_1)x_1 + \dots + a_{m-1}(\lambda_m - \lambda_{m-1})x_{m-1} \in X_{m-1}\end{aligned}$$

and $Ty_n \in X_{m-1}$ since $n < m$.

□

**Definition 6.17** (Projection operator). Let $X$ be a vector space. A linear operator $P : X \to X$ is called a projection if $P^2 = P$.

**Proposition 6.18.** *If $P : X \to X$ is a projection then $I - P$ is a projection, and*

$$\text{Im}\ I - P = \text{Ker}\ P, \quad \text{Ker}\ I - P = \text{Im}\ P$$

*Proof.* If $P^2 = P$ then $(I - P)^2 = I - 2P + P^2 = I - P$ and so $I - P$ is a projection. Let $x \in \text{Im}\ I - P$ Then $x = (I - P)y$ for some $y \in X$. So $Px = P(I - P)y = (P - P^2)y = 0$. So $x \in \text{Ker}\ P$ and $\text{Im}\ I - P \subseteq \text{Ker}\ P$. If $x \in \text{Ker}\ P$ the $Px = 0$. So $(I - P)x = x$, and $x \in \text{Im}\ (I - P)$. □

**Definition 6.19** (Direct sum)**.** Let $X$ be a vector space, and let $X_1, X_2$ be subspaces. Then $X = X_1 \oplus X_2$ (direct sum) if

$$X = X_1 + X_2$$

and $X_1 \cap X_2 = \{0\}$. Equivalently, $X = X_1 \oplus X_2$ if and only if each $x \in X$ can be written in exactly one way as $x = x_1 + x_2$ with $x_1 \in X_1, x_2 \in X_2$.

**Theorem 6.20** (Equivalence of direct sums and projections)**.** *Let $X$ be a vector space.*

*(a) If $P : X \to X$ is a projection, then*

$$X = (\text{Im } P) \oplus (\text{Ker } P)$$

*(b) If $X = X_1 \oplus X_2$, there exists a unique projection with*

$$\text{Im } P = X_1, \quad \text{Ker } P = X_2.$$

*Specifically, $Px = x_1$ if $x = x_1 + x_2$.*

*Proof.* (a) Let $P : X \to X$ be a projection. Then we show $X = (\text{Im } P) \oplus (\text{Im } I - P)$, $x = Px + (I - P)x$. This shows that $X = \text{Im } P + \text{Im } I - P$. If $x \in \text{Im } P \cap \text{Ker } P$ then $x = Py$ and $Px = 0$. Hence, $Px = P^2 y = P^y = 0$ and so $x = 0$.

(b) Exercise.

□

**Proposition 6.21.** *Let $X$ be Banach. Let $X = X_1 \oplus X_2$. Let $P : X \to X$ be the corresponding projection operator. Then*

$$P \in \mathcal{L}(X) \iff X_1, X_2 \text{ closed}$$

*Proof.* ($\Rightarrow$). Suppose $P$ is continuous. Then $X_1 = \text{Im } P = \text{Ker } I - P$ and $X_2 = \text{Ker } P$ are both closed. For example, if $x_n \in \text{Ker } P$ and $x_n \to x$, then $0 = Px_n \to Px$ and so $x \in \text{Ker } P$.

($\Leftarrow$). Suppose that $X_1, X_2$ are closed. Since $X = X_1 \oplus X_2$, we can define a new norm $\| \cdot \|'$ by $\|x\|' = \|x_1\| + \|x_2\|$ where $x = x_1 + x_2$.

**Exercise 6.22.**

(a) Show that $\| \cdot \|'$ is a norm.

(b) Show that $(X, \| \cdot \|')$ is Banach. This relies on the fact that $(X, \| \cdot \|)$ is Banach and $X_1, X_2$ are closed.

Note that $\|x\| = \|x_1 + x_2\| \leq \|x_1\| + \|x_2\| = \|x\|'$, and so by a corollary to the open mapping theorem, there is a $c > 0$ with $\|x\|' \leq c\|x\|$ for all $x \in X$, and so

$$\|Px\| = \|x_1\| \leq \|x_1\| + \|x_2\| = \|x\|' \leq c\|x\|$$

and hence $P$ is continuous. □

**Corollary.** *Let $X$ be Banach, and let $M$ be a finite dimensional subspace. Then there exists a* ***closed*** *$N$ with*

$$X = M \oplus N.$$

*Proof.* Let $v_1, \dots, v_n$ be a basis of $M$. Define, for each $j = 1, \dots, n$, $\varphi_j \in M'$ by $\varphi_j(a_1 v_1 + \cdots + a_n v_n) = a_j$. Then using Hahn-Banach to extend $\tilde{\varphi}_j \in X'$. Let $P : X \in X$ be defined by

$$Px = \sum_{j=1}^{n} \tilde{\varphi}_j(x) v_j.$$

Then we need only check that $P$ is linear and continuous, $\textsc{Im}\ P = M$, and $P^2 = P$. Now take $N = \textsc{Ker}\ P$ and then $X = M \oplus N$. □

We are now ready to prove the following theorem.

**Theorem 6.23.** *Let $X$ be Banach, and let $T \in \mathcal{K}(X)$, and let $\lambda \in \mathbb{K}\backslash\{0\}$. For all $k \in \mathbb{N}$, we have*

*(a)* $\underbrace{\textsc{Ker}\ (\lambda I - T)^k}_{\text{generalised eigenspace}}$ *is finite dimensional.*

*(b)* $\textsc{Im}\ (\lambda I - T)^k$ *is closed.*

*Proof.* **Reductions**. Since $\textsc{Ker}\ (\lambda I - T)^k = \textsc{Ker}\ (I - \lambda^{-1} T)^k$, and similarly for the image, by replacing $T \in \mathcal{K}(X)$ by $\lambda T \in \mathcal{K}(X)$, we can assume that $\lambda = 1$.

Also, we have

$$\begin{aligned}
(I - T)^k &= \sum_{n=0}^{k} \binom{k}{n} (-1)^n T^n \\
&= I - T \underbrace{\sum_{n=1}^{k} \binom{k}{n} (-1)^{n-1} T^{n-1}}_{\text{continuous}} \\
&= I - \tilde{T}.
\end{aligned}$$

where $\tilde{T}$ is the composition of compact and continuos operators, and so is compact. So we can take $\lambda = 1, k = 1$.

(a) The closed unit ball in $\textsc{Ker}\ I - T$ is

$$\begin{aligned}
\{x \in \textsc{Ker}\ I - T \mid \|x\| \leq 1\} &= \{Tx \mid x \in \textsc{Ker}\ I - T,\ |x\| \leq 1\} \\
&\subseteq \overline{T(\overline{B}(0,1))}
\end{aligned}$$

which is compact as $T$ is compact. Hence, the closed unit ball in $\textsc{Ker}\ I - T$ is compact, and thus $\textsc{Ker}\ I - T$ is finite dimensional.

(b) Let $S = I - T$. We then need to show that $\textsc{Im}\ S$ is closed. Since $\textsc{Ker}\ S$ is finite dimensional from above, there is a **closed** subspace $N$ with

$$X = (\textsc{Ker}\ S) \oplus N$$

Note that $\textsc{Im}\ S = S(X) = S(N)$, and that $S|_N : N \to X$ is injective.

Suppose that $S(N)$ is not closed. So there is a sequence $(x_n)$ in $N$ such that $Sx_n \to y \in X \backslash S(N)$. Then there are two cases

*Case* 1 ($\|x_n\| \to \infty$). Let $y_n = \frac{1}{\|x_n\|} x_n$. Then $Sy_n = \frac{1}{\|x_n\|} Sx_n \to 0$. But $(y_n)_{n \geq 1}$ is bounded in $X$, and so there exists a subsequence $y_{n_k}$ such that $Ty_{n_k} \to z$ (as $T$ is compact). Hence $y_{n_k} = Sy_{n_k} + Ty_{n_k} \to 0 + z$. Thus $z \in N$ (as $y_{n_k} \in N$, and $N$ is closed), and $\|z\| = 1$.

So $Sy_{n_k} \to 0$, but $Sy_{n_k} \to Sz$ with $z \in N \backslash \{0\}$, by the continuity of $S$. This contradicts the injectivity of $S|_N$.

*Case* 2 ($\|x_n\|$ does not tend to infinity). So $(x_n)$ has a bounded subsequence $(x_{n_k})$. Since $T$ is compact, $(x_{n_k})$ has a subsequence such that $(Tx_{n_{k_l}})$ converges, to $z_1$ say. By replacing $x_n$ by this subsequence we can assume that $Sx_n \to y$, and that $Tx_n \to z$. A before, we can write

$$x_n = Sx_n + Tx_n \to y + z.$$

So $x_n$ converges to $x \in N$. So $Sx_n \to Sx \in S(N)$ by continuity, but we assume that $Sx_n \to y \in X \backslash S(N)$, which achieves our contradiction.

□

Let $T : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator. Then in the simplest case, $T$ has $n$ distinct eigenvalues, and the corresponding eigenvectors are linearly independent, forming a basis for $\mathbb{C}^n$.

Hence, $\mathbb{C}^n = \mathbb{C}x_1 \oplus \cdots \oplus \mathbb{C}x_n$ and the matrix of $T$ relative to this basis is simply diagonal with $\lambda_1, \ldots, \lambda_n$.

This is not always possible, because there is not always a basis of eigenvectors. Instead look at the generalised eigenspace,

$$\{x \in \mathbb{C}^n \mid (\lambda I - T)^k x = 0 \text{ for some } k \geq 1.$$

But $\{0\} \subseteq \textsc{Ker}\ (\lambda I - T)^1 \subseteq \textsc{Ker}\ (\lambda I - T)^2 \subseteq \ldots$ and since $\dim(\mathbb{C}^n) < \infty$ this must stabilise. Let $r \geq 1$ be the fist time that $\textsc{Ker}\ (\lambda I - T)^r = \textsc{Ker}\ (\lambda I - T)^{r+1}$. Then the generalised $\lambda$-eigenspace is just $\textsc{Ker}\ (\lambda I - T)^r$. There **is** a basis of $\mathbb{C}^n$ consisting of generalised eigenvectors, and the matrix of $T$ relative to this basis is in block form.

**Definition 6.24** (Complete reduction)**.** Let $T : X \to X$ be linear. If $X = X_1 \oplus X_2$ be can write

$$Tx = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

where we identify $x_1 + x_2 \iff (x_1, x_2)$. Here,

$$T_{11} : X_1 \to X_1$$
$$T_{12} : X_2 \to X_1$$
$$T_{21} : X_2 \to X_2$$
$$T_{22} : X_2 \to X_2$$

we say that $X = X_1 \oplus X_2$ **completely reduces** $T$ (well adapted to $T$) if

$$Tx = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

We write $T = T_1 \oplus T_2$.

**Exercise 6.25.** If $X = X_1 \oplus X_2$ completely reduces $T = T_1 \oplus T_2$, then

(a) $\textsc{Ker}\ T = \textsc{Ker}\ T_1 \oplus \textsc{Ker}\ T_2$

(b) $\textsc{Im}\ T = \textsc{Im}\ T_1 \oplus \textsc{Im}\ T_2$

(c) $T$ is injective if and only if $T_1, T_2$ are injective

(d) $T$ is surjective if and only if $T_1, T_2$ are surjective

(e) If $T$ is bijective, then $X = X_1 \oplus X_2$ completely reduces $T^{-1} = T_1^{-1} \oplus T_2^{-1}$.

**Corollary.** *Let $X = X_1 \oplus X_2$ be Banach, with $X_1, X_2$ closed subspaces. If $X = X_1 \oplus X_2$ completely reduces $T = T_1 \oplus T_2 \in \mathcal{L}(X)$, then*

*(a)* $T_1 \in \mathcal{L}(X_1), T_2 \in \mathcal{L}(X_2)$

*(b)* $\sigma(T) = \sigma(T_1) \cup \sigma(T_2)$

*(c)* $\sigma_p(T) = \sigma_p(T_1) \cup \sigma_p(T_2)$

*Proof.* Exercise. □

Consider the following chains

$$\{0\} \subseteq \textsc{Ker}\ S^1 \subseteq \textsc{Ker}\ S^2 \subseteq \dots$$
$$X \supseteq \textsc{Im}\ S^1 \supseteq \textsc{Im}\ S^2 \supseteq \dots$$

where $X$ is a vector space and $S \in \mathrm{Hom}(X, X)$. It is easy to see that if $\textsc{Ker}\ S^r = \textsc{Ker}\ S^{r+1}$ then $\textsc{Ker}\ S^r = \textsc{Ker}\ S^{r+k}$. Similarly for images (p. 109 in Daners.)

There is no reason that these should stabilise in general.

**Theorem 6.26.** *Let $X$ be Banach, $T \in \mathcal{K}(X), \lambda \neq 0$. Then both chains (with $S = \lambda I - T$) stabilise.*

*Proof.* Without loss of generality, assume $\lambda = 1$, so we can write $S = I - T$. Suppose that the kernel chain does not stabilise. Since we assume

$$\text{Ker } S^1 \subsetneq \text{Ker } S^2 \subsetneq \text{Ker } S^3 \subsetneq$$

We know that these are closed (being finite dimensional) subspaces. So Reisz's Lemma gives $x_n \in \text{Ker } S^n$ with $\|x_n\| = 1$, $\|x_n - x\| \geq \frac{1}{2}$ for all $x \in \text{Ker } S^{n+1}$. This is a bounded sequence. We claim that $Tx_n$ has no convergent subsequence.

Let $m > n$. Then

$$\begin{aligned} \|Tx_m - Tx_n\| &= \|(I-T)x_n - (I-T)x_m + x_m - x_n\| \\ &= \|Sx_n - Sx_m - x_m - x_n\| \\ &= \|x_m - \underbrace{(Sx_m - Sx_n + x_n)}_{\text{in Ker } S^{m-1}}\| \\ &\geq \frac{1}{2} \end{aligned}$$

The image argument is similar - using the fact that the images are closed - proved in the previous lecture. □

**Theorem 6.27.** *Let $X$ be a vector space, $S \in Hom(X, X)$. Suppose that*

$$\begin{aligned} \alpha(S) &= \inf\{r \geq 1 \mid \text{Ker } S^r = \text{Ker } S^{r+1}\} \\ \delta(S) &= \inf\{r \geq 1 \mid \text{Im } S^r = \text{Im } S^{r+1}\}, \end{aligned}$$

*the **ascent** and **descent** of $S$ respectively, are both finite.*

*Then*

*(a)* $\alpha(S) = \delta(S) = r$, *say*

*(b)* $X = \text{Ker } S^r \oplus \text{Im } S^r$

*(c) The direct sum in (b) completely reduces $S$.*

*Proof.* Daner's notes, p. 109. □

**Corollary.** *Let $X$ be Banach, $T \in \mathcal{K}(X)$, $\lambda \neq 0$. Let $r = \alpha(\lambda I - T) = \delta(\lambda I - T)$. Then $X = \text{Ker } (\lambda I - T)^r \oplus \text{Im } (\lambda I - T)^r$ and this completely reduces $\mu I - T, \mu \in \mathbb{K}$.*

**Corollary.** *If $X$ is Banach, $T \in \mathcal{K}(X), \lambda \neq 0$ then $\lambda I - T$ is injective if and only if $\lambda I - T$ is surjective.*

*Proof.*

$$\begin{aligned} & \lambda I - T \text{ injective} \\ \Rightarrow\ & 0 \in \textsc{Ker}\ (\lambda I - T)^1 = \textsc{Ker}\ (\lambda I - T)^2 \\ \Rightarrow\ & \alpha(\lambda I - T) = 1 \\ \Rightarrow\ & \delta(\lambda I - T) = 1 \\ \Rightarrow\ & X = \underbrace{\textsc{Ker}\ (\lambda I - T)}_{=\{0\}} \oplus \textsc{Im}\ (\lambda I - T) \\ \Rightarrow\ & X = \textsc{Im}\ (\lambda I - T) \\ \Rightarrow\ & X \text{ is surjective} \end{aligned}$$

The other direction is similar. □

**Corollary.** *Let $X$ be Banach, $T \in \mathcal{K}(X)$. Thus each $\lambda \in \sigma(T) \backslash \{0\}$ is an eigenvalue.*

*Proof.* Immediate from the previous corollary. □

# 7
# The Hilbert Space Decomposition

RECALL THAT WE HAD the following theorem, a corollary in the previous chapter.

**Theorem 7.1.** *Let $X$ be Banach, $T \in \mathcal{K}(X)$, $\lambda \neq 0$. Let $r = \alpha(\lambda I - T) = \delta(\lambda I - T)$. Then $X = \text{Ker}\,(\lambda I - T)^r \oplus \text{Im}\,(\lambda I - T)^r$ and this completely reduces $\mu I - T, \mu \in \mathbb{K}$.*[1]

Also note that $\text{Im Ker}\,(\lambda I - T)^r$ is closed, and $\text{Ker}\,(\lambda I - T)^r$ is finite dimensional.

In Hilbert spaces we can say even more. Recall that the adjoint of $T \in \mathcal{L}(\mathcal{H})$ is defined by

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \quad \forall x, y \in \mathcal{H}$$

Then $T^* \in \mathcal{L}(\mathcal{H})$.

**Definition 7.3** (Self-adjoint)**.** $T \in \mathcal{L}(\mathcal{H})$ is[2]

(a) **Hermitian (self-adjoint)** if $T^* = T$.

(b) **Unitary** if $T^* T = T T^* = I$.

(c) **Normal** if $T^* T = T T^*$.

**Proposition 7.4.** *Let $\mathcal{H}$ be Hilbert over $\mathbb{C}$. If $T \in \mathcal{L}(\mathcal{H})$ is normal, then $r(T) = \|T\|$.*

*Proof.* For Hermitian operators it is easy. We have

$$\|T\|^2 = \|T^* T\| = \|T^2\|.$$

By induction ,we then have $\|T\|^{2^n} = \|T^{2^n}\|$. So

$$\begin{aligned} r(T) &= \lim_{n \to \infty} \|T^n\|^{1/n} \\ &= \lim_{n \to \infty} \|T^{2^n}\|^{1/2^n} \\ &= \|T\|. \end{aligned}$$

[1] Indeed, consider the following exercise, which shows us how we can diagonalise an arbitrary operator.

**Exercise 7.2.** Let $\lambda_1, \ldots, \lambda_n \in \sigma(T) \backslash \{0\}$. Let $N_j = \text{Ker}\,(\lambda_j I - T)^{r_j}$ be the generalised $\lambda_j$-eigenspace. Show that there exists closed subspaces $M$ with

$$X = N_1 \oplus N_2 \oplus \cdots \oplus M$$

with $T = T_1 \oplus T_2 \oplus \cdots \oplus T_M$.

[2] Recall that for a matrix $A$, we can restate this as follows,

(a) Hermitian if and only if $\overline{A^T} = A$.

(b) Unitary if and only if the columns of $A$ are orthonormal.

(c) Hermitian and unitary operators are normal.

For normal operators, we have

$$\begin{aligned}\|T^2\|^2 &= \|(T^2)^*T^2\| \\ &= \|T^*(T^*T)T\| \\ &= \|T^*TT^*T\| \quad \text{normal} \\ &= \|(T^*T)^*(T^*T)\| \\ &= \|T^*T\|^2 \\ &= \|T^4\|\end{aligned}$$

and then we have $\|T^2\| = \|T\|^2$ and the proof follows by induction. □

**Corollary.** *Let $\mathcal{H}$ be a Hilbert space over $\mathbb{C}$.*

*(a) If $T \in \mathcal{L}(\mathcal{H})$ is unitary, then*

$$\sigma(T) \subseteq \mathbb{T} = \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$$

*(b) If $T \in \mathcal{L}(\mathcal{H})$ is Hermitian, then*

$$\sigma(T) \subseteq \mathbb{R}.$$

*Proof.*

(a) On practice sheet. Use the fact that $\sigma(T^*) = \overline{\sigma(T)}$.

(b) Let $\lambda = a + ib \in \sigma(T)$. So $\lambda I - T$ is not invertible. Hence, $(\lambda + it)I - (T + itI)$ is not invertible for all $t \in \mathbb{R}$. Then

$$\begin{aligned}\|\lambda + it\|^2 &\leq r(T + itI)^2 \\ &\leq \|T + itI\|^2 \\ &= \|(T + itI)^*(T + itI)\| \\ &= \|(T - itI)(T + itI)\| \\ &= \|T^2 + t^2 I\| \\ &\leq \|T^2 + t^2\|\end{aligned}$$

However, the left hand side is equal to

$$a^2 + b^2 + 2bt + t^2,$$

and so we obtain

$$a^2 + b^2 + 2bt \leq \|T\|^2 \quad \forall t \in \mathbb{R}$$

and so $b = 0$.

□

**Lemma 7.5.** *Let $\mathcal{H}$ be Hilbert over $\mathbb{C}$. Let $T \in \mathcal{L}(\mathcal{H})$, and let*

$$M_\lambda = \{x \in \mathcal{H} \mid Tx = \lambda x\} = \operatorname{Ker} \lambda I - T$$

*be the $\lambda$-eigenspace of $T$. Then*

*(a)* $M_\lambda \perp M_\mu$ *if* $\lambda \neq \mu$.

*(b) If $T$ is normal, each $M_\lambda$ is $T$ and $T^*$ invariant. That is,*

$$T(M_\lambda) \subseteq M_\lambda, \quad T^*(M_\lambda) \subseteq M_\lambda.$$

*Proof.*

(a) Let $u \in M_\lambda, v \in M_\mu$. Then

$$\begin{aligned}(\lambda - \mu)\langle u, v\rangle &= \langle \lambda u, v\rangle - \langle u, \overline{\mu} v\rangle \\ &= \langle Tu, v\rangle - \langle u, T^* v\rangle \\ &= \langle Tu, v\rangle - \langle Tu, v\rangle \\ &= 0\end{aligned}$$

and so $\langle u, v\rangle = 0$.

(b) If $T$ is normal, then $\operatorname{Ker} T = \operatorname{Ker} T^*$ as

$$\begin{aligned}\|Tx\|^2 &= \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle \\ &= \langle x, TT^*x\rangle = \langle T * x, T^*x\rangle \\ &= \|T^*x\|^2.\end{aligned}$$

Similarly, if $T$ is normal then $\lambda I - T$ is normal. Then

$$\begin{aligned}M_\lambda &= \operatorname{Ker} \lambda I - T \quad (T \text{ invariant}) \\ &= \operatorname{Ker} \overline{\lambda} I - T^* \quad (T^* \text{ invariant}).\end{aligned}$$

□

THE SPECTRAL THEORY FOR compact normal operators in a Hilbert space is particularly nice, as the following theorem demonstrates.

**Theorem 7.6.** *Let $T \in \mathcal{L}(\mathcal{H})$ be compact and normal. Then*

$$\mathcal{H} = \overline{\bigoplus_{\lambda \in \sigma(T)} M_\lambda},$$

*the closure of the span of the eigenspaces, and $\mathcal{H}$ has an orthonormal basis consisting of eigenvectors. Moreover, $T$ acts diagonally with respect to this basis.*

*Proof.* Let

$$M = \overline{\bigoplus_{\lambda \in \sigma(T)} M_\lambda},$$

a closed subspace. Hence $H = M \oplus M^\perp$, where

$$M^\perp = \{x \in \mathcal{H} \mid \langle x, m \rangle = 0 \, \forall m \in M\}.$$

We must show that $M^\perp = \{0\}$. Assume the contrary. Then consider $\tilde{T} = M^\perp \to \mathcal{H}$ be the restriction of $T$ to $M^\perp$. Then we have

$$\tilde{T} : M^\perp \to M^\perp$$

is compact and normal.[3] Then

(a) $\sigma(\tilde{T}) = \{0\}$. Then $r(\tilde{T}) = 0$, and so $\|\tilde{T}\| = 0$, and so $\tilde{T} = 0$. Then each $x \in M^\perp \backslash \{0\}$ satisfies $\tilde{T}x = 0 = 0x$, and so $x \in M_0$ with $M^\perp \subseteq M_0 \subseteq M$, a contradiction (from direct sum decomposition). Hence $M = \{0\}$.

(b) $\sigma(\tilde{T}) \neq \{0\}$. So there is an eigenvalue $\lambda \in \sigma(T) \backslash \{0\}$. So there is $x \in M^\perp \backslash \{0\}$ with $\tilde{T}x = \lambda x$. o $Tx = \lambda x$, and so $x \in (M_\lambda \cap M^\perp) \backslash \{0\}$, a contradiction. Hence $M^\perp = \{0\}$.

Choose an orthonormal basis for each $M_\lambda$, and combine to get an orthonormal basis of $\mathcal{H}$, using $M_\lambda \perp M_\mu$. □

[3] As an exercise, prove this statement.

## 7.1 *The Fredholm Alternative*

Recall that for matrices, we have the following result, known as the Fredholm alternative.

**Theorem 7.7** (Fredholm alternative (Finite dimensional spaces))**.** *Let $A : \mathbb{C}^n \to \mathbb{C}^n$ be linear. Then exactly one of the following two things occur:*

*(1) $Ax = 0$ has only the trivial solution $x = 0$, in which case $Ax = b$ has a unique solution for each $b \in \mathbb{C}^n$.*

*(2) $Ax = 0$ has a non-trivial solution, in which case $Ax = b$ has either no solutions, or infinitely many solutions.*

**Definition 7.8** (Hilbert-Schmidt integral operators)**.**

$$\begin{aligned} T : \; & L^2([a,b]) \to L^2([a,b]) \\ & (Tf)(x) \mapsto \int_a^b K(x,y) f(y) \, dy \end{aligned}$$

where $\|K\|_2$ is finite. These are compact operators.

Consider equations of the following form

$$\lambda f(x) - \int_a^b K(x,y) f(y)\, dy = g(x),$$

where $\lambda \neq 0$ and $g \in L^2$ are given. This can be rewritten in the form

$$(\lambda I - T) f = g.$$

Then we have the following theorem, due to Fredholm.

**Theorem 7.9** (Fredholm alternative (Hilbert spaces))**.** *Let $\mathcal{H}$ be Hilbert over $\mathbb{C}$, and let $T \in \mathcal{K}(\mathcal{H})$. Then exactly one of the following occurs.*

*(a)* $(\lambda I - T) = 0$ *has only the trivial solution, in which case* $(\lambda I - T)x = b$ *has a unique solution for each* $b \in \mathcal{H}$.

*(b)* $(\lambda I - T)x = 0$ *has a non trivial solution, in which case* $(\lambda I - T)x = b$ *has a solution if and only if* $b \perp y$ *for every solution* $y$ *of the equation*

$$(\overline{\lambda} I - T^*) y = 0$$

*This is finite dimensional, as it is the kernel of* $(\lambda I - T)^*$.

*Proof.*

(a) If $(\lambda I - T)x = 0$ has only the trivial solution, then $\textsc{Ker}\ \lambda I - T = \{0\}$ and so it is injective. Hence $\lambda$ is not an eigenvalue, and so $\lambda$ is not a spectral value. So $\lambda I - T$ is invertible, and so $(\lambda I - T)x = b$ has a unique solution $x = (\lambda I - T)^{-1} b$, which can be expanded into a series expression if $|\lambda| > r(T)$.

(b) Suppose $(\lambda I - T)x = 0$ has a non-trivial solution. Then

$$\begin{aligned} &(\lambda I - T)x = b \text{ has a solution} \\ \iff\ & b \in \textsc{Im}\ \lambda I - T \quad \text{which is closed} \\ \iff\ & b \in ((\textsc{Im}\ \lambda I - T)^{\perp})^{\perp} \\ \iff\ & b \in (\textsc{Ker}\ \overline{\lambda} - T^*)^{\perp} \\ \iff\ & b \perp y \quad \forall y \in \textsc{Ker}\ \overline{\lambda} I - T^*. \square \end{aligned}$$

**Proposition 7.10** (Miscelaneous)**.**

*(a) If $M$ is a closed subspace of $\mathcal{H}$, then $M = M^{\perp\perp}$.*

*(b) If $S : \mathcal{H} \to \mathcal{H}$ and $S \in \mathcal{L}(\mathcal{H})$, then $(\textsc{Im}\ S)^{\perp} = \textsc{Ker}\ S^*$.*

*Proof.*

(a) Let $m \in M$, then $\langle m, x \rangle = 0$ for all $x \in M^{\perp}$, and so $m \in (M^{\perp})^{\perp} = M^{\perp\perp}$, and so $M \subseteq M^{\perp\perp}$.

Let $x \in M^{\perp\perp}$. Since $M$ is closed, $\mathcal{H} = M \oplus M^{\perp}$, and so $x = m + m^{\perp}$. So $x - m \in M^{\perp\perp} + M \subseteq M^{\perp\perp}$, and so $x - m = m^{\perp} \in M^{\perp\perp}$. But $M^{\perp}$ is closed, and so $\mathcal{H} = M^{\perp} \oplus M^{\perp\perp}$. So $x - m - 0$, and $x = m \in M$.

(b)

$$\begin{aligned}
(\textsc{Im}\ S)^{\perp} &= \{x \in \mathcal{H} \mid \langle x, sy \rangle = 0 \quad \forall y \in \mathcal{H}\} \\
&= \{x \in \mathcal{H} \mid \langle S^{*}x, y \rangle = 0 \quad \forall y \in \mathcal{H}\} \\
&= \{x \in H \,|\, S^{*}x = 0\} \\
&= \textsc{Ker}\ S^{*}
\end{aligned}$$

□