

Linear Analysis

Lecturer:

András Zsák

a.zsak@dpmms.cam.ac.uk

Notes by:

Robert Frederik Uy

rfu20@cam.ac.uk

Course schedule

Normed and Banach spaces. Linear mappings, continuity, boundedness, and norms. Finite-dimensional normed spaces. [4]

The Baire category theorem. The principle of uniform boundedness, the closed graph theorem and the inversion theorem; other applications. [5]

The normality of compact Hausdorff spaces. Urysohn's lemma and Tietze's extension theorem. Spaces of continuous functions. The Stone–Weierstrass theorem and applications. Equicontinuity: the Ascoli–Arzelà theorem. [5]

Inner product spaces and Hilbert spaces; examples and elementary properties. Orthonormal systems, and the orthogonalization process. Bessel's inequality, the Parseval equation, and the Riesz–Fischer theorem. Duality; the self duality of Hilbert space. [5]

Bounded linear operations, invariant subspaces, eigenvectors; the spectrum and resolvent set. Compact operators on Hilbert space; discreteness of spectrum. Spectral theorem for compact Hermitian operators. [5]

Recommended books

B. Bollobas *Linear Analysis*. Cambridge University Press 1999.

G.J.O. Jameson *Topology and Normed Spaces*. Chapman and Hall 1974.

G. Allan *Introduction to Banach Spaces and Algebras*. Oxford University Press 2010.

W. Rudin *Real and Complex Analysis*. McGraw–Hill International Edition 1987.

Contents

1 Normed spaces and bounded linear maps	3
1.1 Definitions and examples	3
1.2 Inequalities of Minkowski and Hölder	5
1.3 More examples: function spaces	7
1.4 More on the normed topology	7
1.5 Bounded linear maps	8
2 Dual spaces	12
2.1 Basics	12
2.2 Dual space of ℓ_p	12
2.3 Bidual	14
2.4 Dual operators	14
3 Finite-dimensional normed spaces	16
4 The Baire category theorem and its applications	19
4.1 Baire category theorem	19
4.2 Consequences for Banach spaces	21
4.3 Applications	24
5 $C(K)$ spaces	26
5.1 Normal spaces: Urysohn and Tietze	27
5.2 Stone-Weierstrass theorem	29
5.3 Application to Fourier analysis	32
5.4 Arzelà-Ascoli theorem	35
5.5 Application to compact operators	37
6 Hilbert spaces	40
6.1 Basics	40
6.2 Riesz representation theorem	42
6.3 Orthonormal bases	44
6.4 Operators on Hilbert spaces	46
7 Spectral theory	50
7.1 Operators on complex Banach spaces	51
7.2 Operators on complex Hilbert spaces	56

1 Normed spaces and bounded linear maps

1.1 Definitions and examples

Let X be a vector space over \mathbb{R} or \mathbb{C} . For ease of notation and discussion, we will sometimes just take our scalars to be in \mathbb{R} , although the statement may be easily generalised to \mathbb{C} -vector spaces.

Definition Norm

A norm on X is a function $\|\cdot\|: X \rightarrow \mathbb{R}$ such that

- (i) $\|x\| \geq 0$ for all $x \in X$, with $\|x\| = 0$ iff $x = 0$
- (ii) $\|\lambda x\| = |\lambda| \|x\|$ for all $x \in X$ and any scalar λ
- (iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$

Definition Normed space

A normed space is a pair $(X, \|\cdot\|)$ where X is a vector space and $\|\cdot\|$ is a norm on X .

Example Some finite-dimensional normed spaces

- (1) $\ell_2^n = (\mathbb{R}^n, \|\cdot\|_2)$ or $(\mathbb{C}^n, \|\cdot\|_2)$, where the norm is given by

$$\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

This is called the ℓ_2 -norm or euclidean norm.

(i),(ii) are easy to check, whereas (iii) follows from Cauchy-Schwarz.

$$(2) \ell_1^n = (\mathbb{R}^n, \|\cdot\|_1) \text{ where } \|x\|_1 = \sum_{i=1}^n |x_i| \text{ (called the } \ell_1\text{-norm)}$$

$$(3) \ell_\infty^n = (\mathbb{R}^n, \|\cdot\|_\infty) \text{ where } \|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \text{ (called the } \ell_\infty\text{-norm or the sup-norm)}$$

Given a normed space X , its norm $\|\cdot\|$ induces a metric on X :

$$d(x, y) = \|x - y\|$$

Indeed, d is a metric:

- $d(x, y) \geq 0$ for all $x, y \in X$, with $d(x, y) = 0 \iff x - y = 0 \iff x = y$
- $d(x, y) = \|x - y\| = \|y - x\| = d(y, x)$
- $d(x, z) = \|x - z\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$

This metric, in turn, induces a topology on X , called the *norm topology*. This allows us talk about open/closed sets, convergence, and continuity, as we illustrate in the following examples.

Example

The algebraic operations are continuous:

- if $x_n \rightarrow x$ and $y_n \rightarrow y$ in X , then $x_n + y_n \rightarrow x + y$
- if $x_n \rightarrow x$ in X and $\lambda_n \rightarrow \lambda$ in \mathbb{R} , then $\lambda_n x_n \rightarrow \lambda x$

Example

The norm $\|\cdot\|: X \rightarrow \mathbb{R}$ is continuous: by the triangle inequality, we have

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

so $\|\cdot\|$ is, in fact, Lipschitz.

Definition Banach space

A Banach space is a complete normed space, i.e., a normed space that is complete in its norm topology.

Example

$\ell_2^n, \ell_1^n, \ell_\infty^n$ are complete: for any of these spaces,

- $x^{(k)} \rightarrow x \iff x_i^{(k)} \rightarrow x_i$ for all $1 \leq i \leq n$
- $(x_i^{(k)})_{k \in \mathbb{N}}$ is Cauchy $\iff (x_i^{(k)})_{k \in \mathbb{N}}$ is Cauchy for all $1 \leq i \leq n$

In a normed space, a useful object is the *unit ball*

$$B_X := \{x \in X : \|x\| \leq 1\}$$

Remarks

- B_X defines a norm on X :

$$\|x\| = \inf\{t \geq 0 : x \in tB_X\}$$

- B_X is symmetric ($x \in B_X \implies -x \in B_X$), convex, and closed
- If $B \subset \mathbb{R}^n$ is a closed, convex, symmetric, bounded neighbourhood of 0, then B is the unit ball of $(\mathbb{R}^n, \|\cdot\|)$ for some norm $\|\cdot\|$
- ‘Geometry of Banach spaces’

Previously, we gave $\ell_2, \ell_1, \ell_\infty$ as examples of finite-dimensional normed spaces. More generally, we have the following family of examples

Example

(4) $\ell_p^n = (\mathbb{R}^n, \|\cdot\|_p)$ for $1 \leq p < \infty$, where $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ (called the ℓ_p -norm)

Again, (i) and (ii) are easy to check, whereas (iii) is not obvious.¹

Now, let S denote the set of all scalar sequences. This is a vector spaces under the coordinate operations $(x_n) + (y_n) = (x_n + y_n)$ and $\lambda(x_n) = (\lambda x_n)$.

Example Sequence spaces

$$(5) \quad \ell_1 = \left\{ (x_n) \in S : \sum_{n=1}^{\infty} |x_n| < \infty \right\}, \quad \|(x_n)\|_1 = \sum_{n=1}^{\infty} |x_n| \quad (\ell_1\text{-norm})$$

(i) and (ii): easy to check.

(iii): Given $(x_n), (y_n) \in \ell_1$, we have $|x_n + y_n| \leq |x_n| + |y_n|$ for all $n \in \mathbb{N}$. Summing over all $n \in \mathbb{N}$, we deduce that $(x_n) + (y_n) \in \ell_1$ and $\|(x_n) + (y_n)\|_1 \leq \|(x_n)\|_1 + \|(y_n)\|_1$.

Hence, ℓ_1 is a subspace of S and $\|\cdot\|_1$ is a norm on ℓ_1 .

$$(6) \quad \ell_2 = \left\{ (x_n) \in S : \sum_{n=1}^{\infty} |x_n|^2 < \infty \right\}, \quad \|(x_n)\|_2 = \left(\sum_{n=1}^{\infty} |x_n|^2 \right)^{1/2} \quad (\ell_2\text{-norm})$$

(i) and (ii): easy to check.

(iii): Given $(x_n), (y_n) \in \ell_2$, the triangle inequality in ℓ_2^N gives us

$$\left(\sum_{k=1}^N |x_k + y_k|^2 \right)^{1/2} \leq \left(\sum_{k=1}^N |x_k|^2 \right)^{1/2} + \left(\sum_{k=1}^N |y_k|^2 \right)^{1/2}.$$

Taking $N \rightarrow \infty$, we get $(x_n) + (y_n) \in \ell_2$ and $\|(x_n) + (y_n)\|_2 \leq \|(x_n)\|_2 + \|(y_n)\|_2$

¹We will return to this later in the next subsection.

More generally, for $1 \leq p < \infty$, the set

$$\ell_p = \left\{ (x_n) \in S : \sum_{n=1}^{\infty} |x_n|^p < \infty \right\}$$

is a subspace of S , and

$$\| (x_n) \|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} \quad (\ell_p\text{-norm})$$

is a norm on ℓ_p . [(iii) follows from the triangle inequality on ℓ_p^n , which we will see later.]

Example More sequence spaces

$$(7) \quad \ell_\infty = \{ (x_n) \in S : \exists M \geq 0 \ \forall n \in \mathbb{N} \ |x_n| \leq M \}, \quad \| (x_n) \|_\infty = \sup_{n \in \mathbb{N}} |x_n| \quad (\ell_\infty\text{-norm})$$

(i) and (ii): easy to check.

(iii): Given $x = (x_n), y = (y_n) \in \ell_\infty$,

$$|x_n + y_n| \leq |x_n| + |y_n| \leq \|x\|_\infty + \|y\|_\infty \quad \forall n \in \mathbb{N}$$

so $x + y \in \ell_\infty$ and $\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty$.

$$(8) \quad c_0 = \{ (x_n) \in S : x_n \rightarrow 0 \text{ as } n \rightarrow \infty \}$$

$$c = \{ (x_n) \in S : \lim_{n \rightarrow \infty} x_n \text{ exists} \}$$

Both c_0 and c are subspaces of ℓ_∞ and are hence normed spaces in the ℓ_∞ -norm.

1.2 Inequalities of Minkowski and Hölder

Recall that a function $f: (0, \infty) \rightarrow \mathbb{R}$ is *convex* if

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) \quad \forall x, y \in (0, \infty) \ \forall t \in [0, 1]$$

and *concave* if the above holds with \leq replaced by \geq .

Lemma 1.1

Let $1 \leq p < \infty$. Then the map

$$\begin{aligned} (0, \infty) &\rightarrow \mathbb{R} \\ x &\mapsto x^p \end{aligned}$$

is convex.

Proof. Fix $y > 0, t \in [0, 1]$, and define

$$g(x) = [(1-t)x + ty]^p - [(1-t)x^p + ty^p], \quad x > 0.$$

Differentiating, we get

$$g'(x) = p(1-t)[(1-t)x + ty]^{p-1} - p(1-t)x^{p-1}.$$

Observe that $0 < x < y \implies g'(x) \geq 0$ and that $x > y \implies g'(x) \leq 0$. By the MVT, we deduce that $g(x) \leq g(y) = 0$ for all $x \in (0, \infty)$. \blacksquare

Theorem 1.2 Minkowski's inequality

Let $1 \leq p < \infty, n \in \mathbb{N}$. For $x, y \in \mathbb{R}^n$,

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Remark. This shows that ℓ_p^n and ℓ_p are normed spaces.

Exercise. Show that $\ell_p, 1 \leq p \leq \infty$, is complete.²

Proof of Theorem 1.2. Let $B = \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$. We first show that B is convex. Let $x, y \in B$ and $t \in [0, 1]$. For $1 \leq k \leq n$,

$$|(1-t)x_k + ty_k|^p \leq ((1-t)|x_k| + t|y_k|)^p \leq (1-t)|x_k|^p + t|y_k|^p$$

by Lemma 1.1 for $x_k \neq 0, y_k \neq 0$; the inequality holds trivially if $x_k = 0$ or $y_k = 0$. Summing over k , we then get

$$\|(1-t)x + ty\|_p^p \leq (1-t)\|x\|_p^p + t\|y\|_p^p \leq 1,$$

so $(1-t)x + ty \in B$.

We then complete the proof as follows. Let $x, y \in \mathbb{R}^n$. WLOG, $x, y, x+y$ are nonzero. By convexity of B , we have

$$\frac{x+y}{\|x\|_p + \|y\|_p} = \frac{\|x\|_p}{\|x\|_p + \|y\|_p} \cdot \underbrace{\frac{x}{\|x\|_p}}_{\in B} + \frac{\|y\|_p}{\|x\|_p + \|y\|_p} \cdot \underbrace{\frac{y}{\|y\|_p}}_{\in B} \in B.$$

Thus, it follows that

$$\left\| \frac{x+y}{\|x\|_p + \|y\|_p} \right\| \leq 1 \implies \|x+y\|_p \leq \|x\|_p + \|y\|_p,$$

as required. ■

Let $x = (x_n) \in \ell_1$ and $y = (y_n) \in \ell_\infty$. We then write $x \cdot y = (x_n y_n)$. Note that, for all $n \in \mathbb{N}$, $|x_n y_n| = |x_n| |y_n| \leq |x_n| \|y\|_\infty$. Thus, $x \cdot y \in \ell_1$ and $\|x \cdot y\|_1 \leq \|x\|_1 \|y\|_\infty$.

Definition Conjugate index

Let $p \in (1, \infty)$. The conjugate index of p is the unique $q \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$.

Lemma 1.3

Let $1 < p, q < \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then, for $a, b \geq 0$, we have

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

Proof. The inequality holds trivially if $a = 0$ or $b = 0$, so it remains to consider the case $a, b > 0$. A proof similar to that of Lemma 1.1 shows that $\log: (0, \infty) \rightarrow \mathbb{R}$ is concave. Hence,

$$\log\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \geq \frac{1}{p} \log(a^p) + \frac{1}{q} \log(b^q) = \log(ab).$$

We then apply exp to get the required result. ■

Theorem 1.4 Hölder's inequality

Let $1 < p, q < \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$. If $x \in \ell_p$ and $y \in \ell_q$, then $x \cdot y \in \ell_1$ and

$$\|x \cdot y\|_1 \leq \|x\|_p \|y\|_q.$$

Remark. As discussed above, $p = 1, q = \infty$ also works. Moreover, setting $p = q = 2$, we recover Cauchy-Schwarz.

Exercise. Deduce Minkowski's inequality from Hölder's inequality.

²A slick proof of this will be provided later.

Proof of Theorem 1.4. WLOG, $x \neq 0$ and $y \neq 0$. By homogeneity, we may also take $\|x\|_p = \|y\|_q = 1$ WLOG. Now, by Lemma 1.3, we have $|x_n y_n| \leq |x_n|^p/p + |y_n|^q/q$ for all $n \in \mathbb{N}$. Summing over n , we have

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \frac{\|x\|_p^p}{p} + \frac{\|y\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1 = \|x\|_p \|y\|_q,$$

as required. ■

1.3 More examples: function spaces

Example

$$(9) C[0, 1] = \{f: [0, 1] \rightarrow \mathbb{R} \mid f \text{ cts}\}, \quad \|f\|_{\infty} = \sup_{[0,1]} |f| \quad (\text{sup norm or uniform norm})$$

By the uniform limit theorem, this is a Banach space.

$$(10) \text{ More generally, given a compact, Hausdorff topological space } K,$$

$$C(K) = \{f: K \rightarrow \mathbb{R} \mid f \text{ cts}\}$$

is a Banach space in the sup norm $\|f\|_{\infty} = \sup_K |f|$.

$$(11) (C[0, 1], \|\cdot\|_1), \quad \|\cdot\|_1 = \int_0^1 |f(t)| dt \quad (L_1\text{-norm})$$

This is an *incomplete* normed space — see Example Sheet 1.

More generally, $C[0, 1]$ is incomplete in the L_p -norm, $1 \leq p < \infty$, given by

$$\|f\|_p = \left(\int_0^1 |f(t)|^p dt \right)^{1/p}.$$

In II Probability and Measure, you will encounter the completion of $(C[0, 1], \|\cdot\|_p)$, which is the Lebesgue space $L_p[0, 1]$.

$$(12) C^1[0, 1] = \{f \in C[0, 1] \mid f \text{ continuously differentiable}\} \text{ is a subspace of } C[0, 1], \text{ so it is a normed space in } \|\cdot\|_{\infty} \text{ but incomplete, i.e. not closed in } C[0, 1]. \text{ However, it is complete in the norm } \|f\| = \|f\|_{\infty} + \|f'\|_{\infty} \text{ — see Example Sheet 1.}$$

$$(13) \text{ Let } \Delta = \{z \in \mathbb{C} \mid |z| \leq 1\}. \text{ The set}$$

$$A(\Delta) = \{f \in C(\Delta) \mid f \text{ analytic on } \text{int } \Delta\}$$

is a subspace of $C(\Delta)$. In fact, it is closed in $C(\Delta)$ and hence a Banach space in $\|\cdot\|_{\infty}$.

1.4 More on the normed topology

Let X be a normed space and $A \subset C$. Recall that the *closure* of A in X is

$$\overline{A} = \{x \in X \mid \exists (a_n) \text{ in } X \text{ s.t. } a_n \rightarrow x \text{ as } n \rightarrow \infty\}.$$

We then say that A is *dense* in X if $\overline{A} = X$. Moreover, A is *separable* if it has a countable dense subset.

If $Y \subset X$ is a subspace, then so is \overline{Y} : if $x, y \in \overline{Y}$, then there exists $(x_n), (y_n)$ in Y such that $x_n \rightarrow x$ and $y_n \rightarrow y$. So $\lambda x_n + \mu y_n \rightarrow \lambda x + \mu y \in \overline{Y}$. Similarly, if $A \subset X$ is convex, then so is \overline{A} .

For a subset $A \subset X$, the *closed linear span* of A , denoted by $\overline{\text{span}} A$, is the closure of $\text{span} A$.

Remarks

- If A is countable, then $\overline{\text{span}} A$ is separable.
- The set of all rational linear combinations of elements of A is countable and dense in $\overline{\text{span}} A$.

Example

- $\overline{\mathbb{Q}} = \mathbb{R}$, so \mathbb{R} is separable.
- $\ell_p, 1 \leq p < \infty$, is separable.

Let $e_n = (0, \dots, 0, \underset{n}{1}, 0, \dots)$, $n \in \mathbb{N}$ (unit vector basis)

Let $c_{00} = \text{span}\{e_n : n \in \mathbb{N}\} = \{(x_n) \in S : \exists N \in \mathbb{N} \forall n > N x_n = 0\}$

We then show that $\ell_p = \overline{\text{span}}\{e_n : n \in \mathbb{N}\}$: if $x = (x_n) \in \ell_p$, then

$$\left\| x - \sum_{i=1}^N x_i e_i \right\|_p = \left(\sum_{i>n} |x_i|^p \right)^{1/p} \rightarrow 0 \text{ as } N \rightarrow \infty$$

- Similarly, in ℓ_∞ , we have $\overline{\text{span}}\{e_n : n \in \mathbb{N}\} = c_0$. Moreover, c is separable, whereas ℓ_∞ is not.

Exercise. Prove the claims in the last example above.

1.5 Bounded linear maps**Theorem 1.5**

Let X, Y be normed spaces and $T: X \rightarrow Y$ be a linear map. The following are equivalent:

- (i) T is continuous at 0
- (ii) T is continuous
- (iii) T is Lipschitz
- (iv) T is bounded, i.e., $\exists C \geq 0 \forall x \in X \|Tx\| \leq C\|x\|$.

Proof. (iv) \implies (iii): Observe that

$$d(Tx, Ty) = \|Tx - Ty\| = \|T(x - y)\| \leq C\|x - y\| = Cd(x, y)$$

(iii) \implies (ii): Given $\varepsilon > 0$ take $\delta = \varepsilon/(C + 1)$.

(ii) \implies (i): Trivial.

(i) \implies (iv): $\exists \delta > 0 \forall x \in X d(x, 0) = \|x\| \leq \delta \implies d(Tx, T0) = \|Tx\| \leq 1$. For $x \neq 0$, $\|\delta x/\|x\|\| = \delta$, so $\|T(\delta x/\|x\|\)\| \leq 1$. Hence, $\|Tx\| \leq \delta^{-1}\|x\|$. \blacksquare

For normed spaces X, Y , let $\mathcal{B}(X, Y) = \{T: X \rightarrow Y \mid T \text{ linear and bounded}\}$. For $T \in \mathcal{B}(X, Y)$, its *operator norm* is

$$\|T\| = \sup\{\|Tx\| : x \in B_X\}.$$

Remark. Since $T \in \mathcal{B}(X, Y)$, we have $C \geq 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in X$. So if $\|x\| \leq 1$, then $\|Tx\| \leq C$. Thus, by definition, $\|T\| \leq C$. Conversely, for all $x \in B_X$, we have $\|Tx\| \leq \|T\|$, so by homogeneity, $\|Tx\| \leq \|T\|\|x\|$. Hence, $\|T\|$ is the least C such that (iv) in Theorem 1.5 above holds.

The operator norm is a norm on $\mathcal{B}(X, Y)$: given $S, T \in \mathcal{B}(X, Y)$, we have, for all $x \in X$,

$$\|(S + T)x\| = \|Sx + Tx\| \leq \|Sx\| + \|Tx\| \leq \|S\|\|x\| + \|T\|\|x\| \leq (\|S\| + \|T\|)\|x\|,$$

from which it follows that $S + T \in \mathcal{B}(X, Y)$ and $\|S + T\| \leq \|S\| + \|T\|$.

Notation. We write $\mathcal{B}(X)$ for $\mathcal{B}(X, X)$.

Proposition 1.6

Let X, Y, Z be normed spaces, $S \in \mathcal{B}(X, Y)$, $T \in \mathcal{B}(Y, Z)$. Then $TS \in \mathcal{B}(X, Z)$ and $\|TS\| \leq \|T\| \|S\|$.

Proof. For all $x \in X$, we have $\|TSx\| \leq \|T\| \|Sx\| \leq \|T\| \|S\| \|x\|$. ■

Example

- (1) $T: \ell_2^n \rightarrow \ell_2^n, (x_1, \dots, x_n) \mapsto (x_1, \dots, x_r, 0, \dots, 0)$

$$\|Tx\|_2 = \left(\sum_{i=1}^r |x_i|^2 \right)^{1/2} \leq \|x\|_2 \implies \|T\| \leq 1$$

But $Te_1 = e_1$ so $\|T\| = 1$.

More generally, if T is represented by a matrix A wrt the standard basis, then Cauchy-Schwarz gives us

$$\|T\| \leq \left(\sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2}$$

- (2) Let $1 \leq p < \infty$; $R: \ell_p \rightarrow \ell_p, (x_1, x_2, x_3, \dots) \mapsto (0, x_1, x_2, \dots)$ (right shift)

For all $x \in \ell_p$, $\|Rx\|_p = \|x\|_p$, so R is isometric and $\|R\| = 1$. Note that R is injective but not surjective.

- (3) Let $1 \leq p < \infty$; $L: \ell_p \rightarrow \ell_p, (x_1, x_2, x_3, \dots) \mapsto (x_2, x_3, x_4, \dots)$ (left shift)

For all $x \in \ell_p$, $\|Lx\|_p \leq \|x\|_p$, so $L \in \mathcal{B}(\ell_p)$ with $\|L\| \leq 1$. Since $Le_2 = e_1$ and $\|e_1\|_p = \|e_2\|_p = 1$, we in fact have $\|L\| = 1$. Note that L is surjective but not injective.

- (4) $T: \ell_1 \rightarrow \ell_2, x \mapsto x$

► **Claim.** $\ell_1 \subset \ell_2$, and $\forall x \in \ell_2 \quad \|x\|_2 \leq \|x\|_1$

Proof. WLOG assume $\|x\|_1 = 1$ by homogeneity. Since $\sum_{n=1}^{\infty} |x_i| = 1$, we have $|x_i| \leq 1$ for all i . Thus,

$$|x_i|^2 \leq |x_i| \quad \forall i \implies \|x\|_2^2 \leq \|x\|_1 = 1 \implies \|x\|_2 = 1 = \|x\|_1$$

as claimed. ■

Using the above claim, we have $T \in \mathcal{B}(\ell_1, \ell_2)$ and $\|T\| = 1$.

- (5) $T: \ell_2 \rightarrow \ell_1, (x_n) \mapsto (x_n/n)$

By Cauchy-Schwarz,

$$\sum_{n=1}^{\infty} \left| \frac{x_i}{n} \right| \leq \left(\sum_{n=1}^{\infty} x_i^2 \right)^{1/2} \left(\sum_{n=1}^{\infty} \frac{1}{n^2} \right)^{1/2}$$

so $T \in \mathcal{B}(\ell_2, \ell_1)$ with $\|T\| \leq \left(\sum_{n=1}^{\infty} \frac{1}{n^2} \right)^{1/2}$. In fact, we can replace \leq with $=$.

- (6) $D: (C^1[0, 1], \|\cdot\|) \rightarrow (C[0, 1], \|\cdot\|_{\infty}), f \mapsto f'$

Note that $\|Df\|_{\infty} = \|f'\|_{\infty} \leq \|f\|_{\infty} + \|f'\|_{\infty} = \|f\|$, so $\|D\| \leq 1$. But taking $f(x) = \sin(n\pi x)$, we have

$$\|Df\|_{\infty} = n\pi, \quad \|f\| = n\pi + 1,$$

so in fact $\|D\| = 1$. Note also that, for $f \neq 0$, $\|Df\|_\infty < \|f\|$, so $\|D\|$ is not attained.

- (7) On a normed space X , the identity $x \mapsto x$ is denoted by Id , I , Id_X or I_X . This map is isometric, i.e., $\|\text{Id}(x)\| = \|x\| \forall x \in X$.

- (8) For normed spaces X, Y , we let

$$X \oplus Y = \{(x, y) : x \in X, y \in Y\}$$

with norm $\|(x, y)\|_1 = \|x\| + \|y\|$. The corresponding norm topology is the product topology.

Define $P: X \oplus Y \rightarrow X, (x, y) \mapsto x$ (projection onto X). Note that $P \in \mathcal{B}(X \oplus Y, X)$ with $\|P\| = 1$.

Let X, Y be normed spaces. We introduce some terminology:

- An *isomorphism* $X \rightarrow Y$ is a linear homeomorphism $T: X \rightarrow Y$, i.e., T is a linear bijection such that T and T^{-1} are bounded. Equivalently, T is a linear bijection³ such that

$$\exists a, b > 0 \forall x \in X a\|x\| \leq \|Tx\| \leq b\|x\|$$

If such T exists, we say that X and Y are *isomorphic*, and we write $X \sim Y$.

- An *isometric isomorphism* is a linear bijection $T: X \rightarrow Y$ such that

$$\forall x \in X \|Tx\| = \|x\|$$

If such T exists, we say that X and Y are *isometrically isomorphic*, and we write $X \cong Y$.

The Banach-Mazur distance is defined as

$$d(X, Y) = \begin{cases} \infty, & \text{if } X \not\sim Y \\ \inf\{\|T\| \|T^{-1}\| \mid T: X \rightarrow Y \text{ is an isomorphism}\}, & \text{otherwise} \end{cases}$$

Note that $\|T\| \|T^{-1}\| \geq \|TT^{-1}\| = 1$. If $X \cong Y$, then $d(X, Y) = 1$. Does the converse hold?

- An *isomorphic embedding* $X \rightarrow Y$ is a linear map $T: X \rightarrow Y$ such that $T: X \rightarrow TX = \text{im } T$ is an isomorphism. If such T exists, we say that X (*isomorphically*) *embeds into* Y , and we write $X \hookrightarrow Y$.

Definition Equivalent norms

Let X be a normed space. Two norms $\|\cdot\|, \|\cdot\|'$ are equivalent if

$$\begin{aligned} & \text{Id}: (X, \|\cdot\|) \rightarrow (X, \|\cdot\|') \text{ is an isomorphism} \\ \iff & \|\cdot\|, \|\cdot\|' \text{ induce the same norm topology on } X \\ \iff & \exists a, b > 0 \forall x \in X a\|x\| \leq \|x\|' \leq b\|x\| \\ \iff & \exists a, b > 0 aB'_X \subset B_X \subset bB'_X \end{aligned}$$

Remarks

- If $X \sim Y$, then X is complete iff Y is complete.

If $\|\cdot\|, \|\cdot\|'$ are equivalent norms on a vector space X , then $(X, \|\cdot\|)$ is complete iff $(X, \|\cdot\|')$ is complete.

³We can actually replace ‘bijection’ with ‘surjection’.

- Let X and Y be normed spaces. On $X \oplus Y$, the norm $\|(x, y)\|_1 = \|x\| + \|y\|$ is equivalent to $\|(x, y)\|_p = (\|x\|^p + \|y\|^p)^{1/p}$ for all $1 \leq p < \infty$ and to $\|(x, y)\|_\infty = \max\{\|x\|, \|y\|\}$.
- $(C[0, 1], \|\cdot\|_\infty)$ is complete whereas $(C[0, 1], \|\cdot\|_1)$ is incomplete. Thus, we can use the first remark above to deduce that $\|\cdot\|_\infty \not\sim \|\cdot\|_1$ (but this can easily be proven directly as well). However, $\|f\|_1 = \int_0^1 |f(t)| dt \leq \|f\|_\infty$, so

$$\text{Id}: (C[0, 1], \|\cdot\|_\infty) \rightarrow (C[0, 1], \|\cdot\|_1)$$

is a continuous linear bijection but its inverse is not continuous.

- On c_{00} , $\|\cdot\|_1 \not\sim \|\cdot\|_2$. To see why, consider $x = (\underbrace{1, \dots, 1}_n, 0, 0, \dots)$ and note that $\|x\|_1 = n$, $\|x\|_2 = \sqrt{n}$.

Finally, we discuss convergence and completeness. Let X, Y be normed spaces. In $\mathcal{B}(X, Y)$, convergence implies pointwise convergence, i.e., if $T_n \rightarrow T$ in $\mathcal{B}(X, Y)$, then, for all $x \in X$, $T_n x \rightarrow T x$ in Y . To see why, note that, for fixed $x \in X$, we have $\|T_n x - T x\| \leq \|T_n - T\| \|x\| \rightarrow 0$. However, the converse is false in general, e.g., $T_n: \ell_1 \rightarrow \mathbb{R}$, $x \mapsto x_n$. We have $T_n \rightarrow 0$ pointwise, but $\|T_n\| = 1$ for all $n \in \mathbb{N}$.

Theorem 1.7

Let X, Y be normed spaces. If Y is complete, then $\mathcal{B}(X, Y)$ is complete.

Proof. Let (T_n) be a Cauchy sequence in $\mathcal{B}(X, Y)$. Fix $x \in X$. Then

$$\|T_m x - T_n x\| \leq \|T_m - T_n\| \|x\| \rightarrow 0 \text{ as } m, n \rightarrow \infty$$

So $(T_n x)$ is Cauchy in Y and thus convergent. Now, define $T: X \rightarrow Y$ by $x \mapsto \lim_{n \rightarrow \infty} T_n x$. Observe that

- T is linear

$$T(\lambda x + \mu y) = \lim_{n \rightarrow \infty} T_n(\lambda x + \mu y) = \lim_{n \rightarrow \infty} [\lambda T_n x + \mu T_n y] = \lambda T x + \mu T y$$

- T is bounded

(T_n) is Cauchy implies (T_n) is bounded, i.e., there exists $M \geq 0$ such that $\|T_n\| \leq M$ for all $n \in \mathbb{N}$. Fix $x \in X$. Then, for all $n \in \mathbb{N}$, we have $\|T_n x\| \leq \|T_n\| \|x\| \leq M \|x\|$. Letting $n \rightarrow \infty$, we obtain $\|T x\| \leq M \|x\|$.

- $T_n \rightarrow T$ in $\mathcal{B}(X, Y)$

Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $\|T_m - T_n\| \leq \varepsilon$ for all $m, n \geq N$. Fix $x \in X$. Note that, for all $m, n \geq N$, we have

$$\|T_m x - T_n x\| \leq \|T_m - T_n\| \|x\| \leq \varepsilon \|x\|$$

Letting $n \rightarrow \infty$ with $m \geq N$ fixed yields $\|T_m x - T x\| \leq \varepsilon \|x\|$. Hence, $\|T_m - T\| \leq \varepsilon$ for all $m \geq N$. ■

2 Dual spaces

2.1 Basics

Let X be a normed space. A *functional* on X is a map $X \rightarrow \mathbb{R}$. The *dual space* X^* of X is the space of all bounded linear functionals on X , i.e., $X^* = \mathcal{B}(X, \mathbb{R})$ equipped with the operator norm. Since \mathbb{R} is complete, Theorem 1.7 gives us the following result.

Theorem 2.1

For any normed space X , its dual X^* is a Banach space.

Notation. For $x \in X$ and $f \in X^*$, we let $\langle x, f \rangle = f(x)$.

Now, we know that $0 \in X^*$. Are there other elements?

Theorem 2.2 Hahn-Banach theorem

Let X be a normed space, $Y \subset X$ be a subspace and $g \in Y^*$. Then $f \in X^*$ such that $f|_Y = g$ and $\|f\| = \|g\|$.

Proof. See II Analysis of Functions. ■

Corollary 2.3

Let X be a normed space, $x_0 \in X \setminus \{0\}$. Then there exists $f \in S_{X^*} = \{f \in X^* : \|f\| = 1\}$ such that $f(x_0) = \|x_0\|$.

Remarks

- For any $g \in B_{X^*}$, $|g(x_0)| \leq \|g\| \|x_0\| \leq \|x_0\|$. Corollary 2.3 says that there exists $f \in B_{X^*}$ such that $f(x_0) = \|x_0\|$, so

$$\|x_0\| = \sup\{g(x_0) : g \in B_{X^*}\} = \max\{g(x_0) : g \in B_{X^*}\}.$$

We call f a *norming functional* at x_0 .

- Given $x \neq y$ in X , we can set $x_0 = x - y$ and Corollary 2.3 implies that there exists $f \in X^*$ such that $f(x) \neq f(y)$. Thus, X^* separates the points of X .

Proof of Corollary 2.3. Set $Y = \text{span}\{x_0\}$ and define $g(\lambda x_0) = \lambda \|x_0\|$. Then $g \in S_{Y^*}$ with $g(x_0) = \|x_0\|$. Finally, apply Theorem 2.2. ■

2.2 Dual space of ℓ_p

Motivation: Recall that, for $1 \leq p < \infty$, we have $\ell_p = \overline{\text{span}}\{e_n : n \in \mathbb{N}\} = \overline{c_0}$. Given $\varphi \in \ell_p^*$ and $x = (x_n) \in \ell_p$,

$$\varphi(x) = \varphi \left(\lim_{n \rightarrow \infty} \sum_{k=1}^n x_k e_k \right) = \sum_{k=1}^{\infty} x_k \varphi(e_k)$$

so φ corresponds to the sequence $y = (\varphi(e_n))_{n \in \mathbb{N}}$. We may then ask: is $\ell_p^* \cong \ell_q$ for some q ?

Fix $1 < p < \infty$, and let q be the conjugate index of p . Fix $y = (y_n) \in \ell_q$. Define

$$\begin{aligned} \varphi_y : \ell_p &\rightarrow \mathbb{R} \\ x &\mapsto \sum_{n=1}^{\infty} x_n y_n \end{aligned}$$

By Holder's inequality (Theorem 1.4), this is well-defined and $|\varphi_y(x)| \leq \|x\|_p \|y\|_q$. So φ_y is linear and bounded: $\|\varphi_y\| \leq \|y\|_q$. Thus, $\varphi_y \in \ell_p^*$, which means that we have a map

$$\begin{aligned}\varphi: \ell_q &\rightarrow \ell_p^* \\ y &\mapsto \varphi_y\end{aligned}$$

Note that φ is linear and bounded with $\|\varphi\| \leq 1$.

Theorem 2.4

Let p, q, φ be as above. Then φ is an isometric isomorphism $\ell_q \rightarrow \ell_p^*$.

Proof. It remains to check that φ is isometric and surjective:

- φ is isometric

Fix $y \in \ell_q$. WLOG $y \neq 0$. Define

$$x_n = \begin{cases} \frac{|y_n|^q}{y_n}, & y_n \neq 0 \\ 0, & y_n = 0 \end{cases}$$

Observe that $\sum_{n=1}^{\infty} |x_n|^p = \sum_{n=1}^{\infty} |y_n|^{(q-1)p} = \sum_{n=1}^{\infty} |y_n|^q = \|y\|_q^q < \infty$, so $x \in \ell_p$ with $\|x\|_p^p = \|y\|_q^q$.

Since $y \neq 0$, we have $x \neq 0$, so $x/\|x\|_p \in B_{\ell_p}$. Note that

$$\|\varphi_y\| \geq \varphi_y \left(\frac{x}{\|x\|_p} \right) = \frac{1}{\|x\|_p} \sum_{n=1}^{\infty} x_n y_n = \frac{\|y\|_q^q}{\|y\|_q^{q/p}} = \|y\|_q.$$

Hence, $\|\varphi_y\| = \|y\|_q$.

- φ is surjective

Fix $f \in \ell_p^*$. Define $y_n = f(e_n)$, $n \in \mathbb{N}$. Let $y = (y_n)$. For some fixed $N \in \mathbb{N}$, set

$$x_n = \begin{cases} \frac{|y_n|^q}{y_n}, & y_n \neq 0, n \leq N \\ 0, & \text{otherwise} \end{cases}$$

Then $x = (x_n) \in \ell_p$, so

$$\begin{aligned}f(x) &= \sum_{n=1}^N x_n f(e_n) = \sum_{n=1}^N x_n y_n = \sum_{n=1}^N |y_n|^q \leq \|f\| \|x\|_p \\ \|x\|_p &= \left(\sum_{n=1}^N |x_n|^p \right)^{1/p} = \left(\sum_{n=1}^N |y_n|^{(q-1)p} \right)^{1/p} = \left(\sum_{n=1}^N |y_n|^q \right)^{1/p}\end{aligned}$$

Hence, $\sum_{n=1}^N |y_n|^q \leq \|f\| \left(\sum_{n=1}^N |y_n|^q \right)^{1/p}$, i.e.

$$\left(\sum_{n=1}^N |y_n|^q \right)^{1/q} \leq \|f\|$$

Let $N \rightarrow \infty$ to deduce that $y \in \ell_q$. Finally, observe that

$$\begin{aligned}f(e_n) &= y_n = \varphi_y(e_n) \quad \forall n \in \mathbb{N} \\ \implies f(x) &= \varphi_y(x) \quad \forall x \in \text{span}\{e_n : n \in \mathbb{N}\} = c_{00} \quad \text{by linearity}\end{aligned}$$

$$\implies f(x) = \varphi_y(x) \quad \forall x \in \overline{\text{span}}\{e_n : n \in \mathbb{N}\} = \ell_p \quad \text{by continuity}$$

Thus, $f = \varphi_y$, so φ is surjective. ■

Remarks

- We also have $\ell_1^* \cong \ell_\infty$ and $c_0^* \cong \ell_1$. The proof also shows that $\ell_1 \hookrightarrow \ell_\infty^*$ isometrically. However, the proof of surjectivity breaks down since $\overline{\text{span}}\{e_n : n \in \mathbb{N}\}$ in ℓ_∞ is $c_0 \subsetneq \ell_\infty$.
- From the proof, we can show Corollary 2.3 holds for ℓ_p .
- We've shown that $\ell_p, 1 \leq p \leq \infty$, is complete as they are dual spaces. For c_0 , one simply has to show that c_0 is closed in ℓ_∞ .

2.3 Bidual

Let X be a normed space. Then $X^{**} = (X^*)^* = \mathcal{B}(X^*, \mathbb{R})$ is the *bidual* or *second dual* of X .

For each $x \in X$, define the map

$$\begin{aligned} \hat{x} : X^* &\rightarrow \mathbb{R} \\ f &\mapsto f(x) \end{aligned}$$

Note that \hat{x} is linear and bounded: $|\hat{x}(f)| = |f(x)| \leq \|f\| \|x\|$. So $\hat{x} \in X^{**}$ with $\|\hat{x}\| \leq \|x\|$. Thus, we have

$$\begin{aligned} \hat{\cdot} : X &\rightarrow X^{**} \\ x &\mapsto \hat{x} \end{aligned}$$

This is linear: $\widehat{\lambda x + \mu y}(f) = f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = (\lambda \hat{x} + \mu \hat{y})(f)$.

For $x \neq 0$, let $f \in X^*$ be a norming functional at x . Then

$$\hat{x}(f) = f(x) = \|x\| \implies \|\hat{x}\| = \|x\|$$

so the canonical map $X \rightarrow X^{**}, x \mapsto \hat{x}$ is an isometric embedding into X^{**} . If f is surjective, we say that X is *reflexive*.

2.4 Dual operators

Let X, Y be normed spaces and $T \in \mathcal{B}(X, Y)$. The *dual operator* T^* of T is the map

$$\begin{aligned} T^* : Y^* &\rightarrow X^* \\ g &\mapsto g \circ T \end{aligned}$$

By Proposition 1.6, $T^*(g) = g \circ T \in X^*$ and $\|T^*(g)\| \leq \|g\| \|T\|$, so T^* is well-defined. Moreover, it is clearly linear and bounded with $\|T^*\| \leq \|T\|$.

Remark. Note that $\langle \cdot, \cdot \rangle : X \times X^* \rightarrow \mathbb{R}$ is bilinear. Moreover, for $x \in X$ and $g \in Y^*$, we have $\langle x, T^*(g) \rangle = \langle T(x), g \rangle$.

It turns out that $\|T^*\| = \|T\|$:

$$\|T^*\| = \sup_{g \in B_{Y^*}} \|T^* g\| = \sup_{g \in B_{Y^*}} \sup_{x \in B_X} |\langle x, T^*(g) \rangle| = \sup_{x \in B_X} \sup_{g \in B_{Y^*}} |\langle T x, g \rangle| = \sup_{x \in B_X} \|T x\| = \|T\|,$$

where the penultimate equality follows from Corollary 2.3.

Example

Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Consider the right-shift map $R: \ell_p \rightarrow \ell_p$. What is $R^*: \ell_p^* \rightarrow \ell_p^*$? Recall that $\ell_p^* \cong \ell_q$. Thought of as a map $\ell_q \rightarrow \ell_q$, it turns out that $R^* = L$, the left-shift map.

Now, let's note some properties of dual operators:

- (1) $(\text{Id}_X)^* = \text{Id}_{X^*}$
- (2) $(\lambda S + \mu T)^* = \lambda S^* + \mu T^*$ for all $S, T \in \mathcal{B}(X, Y)$ and all scalars λ, μ

Indeed, for $g \in Y^*$, $x \in X$,

$$\begin{aligned} \langle x, (\lambda S + \mu T)^* g \rangle &= \langle (\lambda S + \mu T)x, g \rangle \\ &= \langle \lambda Sx + \mu Tx, g \rangle \\ &= \lambda \langle Sx, g \rangle + \mu \langle Tx, g \rangle \\ &= \lambda \langle x, S^* g \rangle + \mu \langle x, T^* g \rangle \\ &= \langle x, (\lambda S^* + \mu T^*)g \rangle \end{aligned}$$

Since x is arbitrary, $(\lambda S + \mu T)^* g = (\lambda S^* + \mu T^*)g$ for all $g \in Y^*$, and we are done.

- (3) $(ST)^* = T^*S^*$ for all $T \in \mathcal{B}(X, Y)$ and all $S \in \mathcal{B}(Y, Z)$

$$\langle x, (ST)^* g \rangle = \langle STx, g \rangle = \langle S(Tx), g \rangle = \langle Tx, S^* g \rangle = \langle x, T^*S^* g \rangle$$

- (4) Let $T \in \mathcal{B}(X, Y)$. We have $T^* \in \mathcal{B}(Y^*, X^*)$ and $T^{**} \in \mathcal{B}(X^{**}, Y^{**})$. The diagram

$$\begin{array}{ccc} X & \xrightarrow{T} & Y \\ \downarrow \wedge & & \downarrow \\ X^{**} & \xrightarrow{T^{**}} & Y^{**} \end{array}$$

commutes, i.e., $\hat{T}x = T^{**}\hat{x}$ for all $x \in X$. For $x \in X, g \in Y^*$,

$$\langle g, T^{**}\hat{x} \rangle = \langle T^*g, \hat{x} \rangle = \langle x, T^*g \rangle = \langle Tx, g \rangle = \langle g, \hat{Tx} \rangle$$

Remark. Properties (1) and (3) imply that $X \sim Y \implies X^* \sim Y^*$.

3 Finite-dimensional normed spaces

Recall that norms $\|\cdot\|$ and $\|\cdot\|'$ on a vector space X are equivalent if $\text{Id}: (X, \|\cdot\|) \rightarrow (X, \|\cdot\|')$ is an isomorphism or, equivalently, if $\exists a, b > 0 \forall x \in X a\|x\| \leq \|x\|' \leq b\|x\|$.

Example

On \mathbb{R}^n , the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent. We've already seen that $\|x\|_2 \leq \|x\|_1$ for all $x \in \mathbb{R}^n$. Moreover, by Cauchy-Schwarz, we have

$$\|x\|_1 = \sum_{i=1}^n |x_i| \leq \sqrt{n} \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} = \sqrt{n}\|x\|_2.$$

Theorem 3.1

Any two norms on a finite-dimensional vector space are equivalent.

Proof. Let X be a f.d. vector space. Fix a basis (e_1, \dots, e_n) of X . For $x = \sum_{k=1}^n x_k e_k \in X$, define $\|x\|_1 = \sum_{k=1}^n |x_k|$. Let $\|\cdot\|$ be an arbitrary norm on X .

We show that $\|\cdot\|$ is equivalent to $\|\cdot\|_1$. For $x = \sum_{k=1}^m x_k e_k \in X$, we have

$$\|x\| \leq \sum_{k=1}^n |x_k| \|e_k\| \leq M \|x\|_1$$

where $M = \max_{1 \leq k \leq n} \|e_k\|$.

Now, let $S = \{x \in X : \|x\|_1 = 1\}$, the unit sphere of $(X, \|\cdot\|_1)$. We have the following result:

► **Claim.** S is compact.

Proof. Let $(x^{(r)})_{r \in \mathbb{N}}$ be a sequence in S . Write $x^{(r)} = \sum_{k=1}^n x_k^{(r)} e_k$. For each $1 \leq k \leq n$, $|x_k^{(r)}| \leq \|x^{(r)}\| = 1$ for all $r \in \mathbb{N}$. By repeated application of Bolzano-Weierstrass, there exists $r_1 < r_2 < r_3 < \dots$ in \mathbb{N} such that $(x_k^{(r_\ell)})_{\ell \in \mathbb{N}}$ is convergent for each $1 \leq k \leq n$. Let $x_k = \lim_{\ell \rightarrow \infty} x_k^{(r_\ell)}$ and $x = \sum_{k=1}^n x_k e_k$. Note that

$$\|x\|_1 = \sum_{k=1}^n |x_k| = \lim_{\ell \rightarrow \infty} \sum_{k=1}^n |x_k^{(r_\ell)}| = 1$$

so $x \in S$. Moreover,

$$\|x^{(r_\ell)} - x\|_1 = \sum_{k=1}^n |x_k^{(r_\ell)} - x_k| \rightarrow 0 \quad \text{as } \ell \rightarrow \infty$$

so $x^{(r_\ell)} \rightarrow x$ in S . Thus, S is sequentially compact and hence compact. ■

For any $x, y \in S$, $\| \|x\| - \|y\| \| \leq \|x - y\| \leq M \|x - y\|_1$. So $\|\cdot\|$ is continuous on S with respect to $\|\cdot\|_1$. So $c = \inf\{\|x\| : x \in S\}$ is achieved: $\exists x \in S \ \|x\| = c$. Since $0 \notin S$ and $c > 0$, we have $\|y\| \geq c = c\|y\|_1$ for all $y \in S$. By homogeneity, $\|y\| \geq c\|y\|_1$ for all $y \in X$. ■

Corollary 3.2

Let $T: X \rightarrow Y$ be a linear map between two normed spaces. If X is f.d., then T is continuous.

Proof. Let $\|\cdot\|$ denote the norm on X and Y . Define $\|x\|' = \|Tx\| + \|x\|$ for all $x \in X$. It is easy to check that this is a norm on X . By Theorem 3.1, there exists $b > 0$ such that, for all $x \in X$, $\|x\|' \leq b\|x\|$. In particular, $\|Tx\| \leq b\|x\|$ for all $x \in X$. ■

Corollary 3.3

If $\dim X = \dim Y < \infty$, then $X \sim Y$.

Proof. We have a linear bijection $T: X \rightarrow Y$. By Corollary 3.2, T and T^{-1} are continuous. ■

Remark. Corollary 3.3 does *not* imply that the theory of f.d. normed spaces is uninteresting.

Recall that, for X a metric space and $Y \subset X$, we have

- Y complete $\implies Y$ is closed in X
- Y closed in X and X complete $\implies Y$ complete

Corollary 3.4

- (i) A f.d. normed space X is complete.
- (ii) A f.d. subspace X of a normed space Y is closed in Y .

Proof. (i) Let $n = \dim X$. By Corollary 3.3, $X \sim \ell_2^n$ which is complete. (ii) follows from above properties of metric spaces. ■

Corollary 3.5

Let X be a f.d. normed space and $A \subset X$. Then

$$A \text{ is compact} \iff A \text{ is closed and bounded}$$

Proof. If $X = \ell_2^n$, then this is simply Heine-Borel. For general X , the result follows by invoking Corollary 3.3 to deduce that $X \sim \ell_2^n$ and noting isomorphisms map compact subsets to compact subsets (ditto for closed and bounded subsets). ■

In particular, $B_X = \{x \in X : \|x\| = 1\}$ is compact. How about if $\dim X = \infty$? Note that, in ℓ_p , $1 \leq p < \infty$, $\|e_n\|_p = 1$ for all n and $\|e_m - e_n\| = 2^{1/p}$ for all $m \neq n$, so (e_n) has no convergent subsequence. Hence, B_{ℓ_p} is not compact.

A similar obstruction does, in fact, hold for any infinite-dimensional normed space. To show this, we need the following lemma:

Proposition 3.6 Riesz's lemma

Let Y be a proper, closed subspace of a normed space X . Then

$$\forall \varepsilon > 0 \exists x \in B_X \quad d(x, Y) = \inf\{\|x - y\| : y \in Y\} > 1 - \varepsilon.$$

Proof. WLOG, $0 < \varepsilon < 1$. Fix $z \in X \setminus Y$. Since Y is closed, $d = d(z, Y) > 0$. Pick $y \in Y$ such that $d \leq \|z - y\| < d/(1 - \varepsilon)$. Set $x = \frac{(z-y)}{\|z-y\|}$. Note that $d(x, Y) > 1 - \varepsilon$: for $y' \in Y$,

$$\|x - y'\| = \left\| \frac{z - y - \|z - y\|y'}{\|z - y\|} \right\| \geq \frac{d}{\|z - y\|} > 1 - \varepsilon$$

so $d(x, Y) \geq d/\|z - y\| > 1 - \varepsilon$. ■

Theorem 3.7

Let X be a normed space. Then B_X is compact if and only if $\dim X < \infty$.

Proof. (\Leftarrow) Corollary 3.5

(\Rightarrow) Similar to the ℓ_p case, we construct (x_n) in B_X such that $\|x_m - x_n\| > 1/2$ for all $m \neq n$. As before, we then deduce that (x_n) has no convergent subsequence and so B_X is not compact.

Pick any $x_1 \in B_X$. Suppose we have already picked x_1, \dots, x_n for some $n \in \mathbb{N}$. We then set $Y = \text{span}\{x_1, \dots, x_n\}$. Then Y is a proper ($\dim X = \infty$) and closed (Corollary 3.4) subspace of X . By Proposition 3.6, we can then pick $x_{n+1} \in B_X$ such that $d(x_{n+1}, Y) > 1/2$. In particular, $\|x_{n+1} - x_m\| > 1/2$ for $1 \leq m \leq n$. \blacksquare

4 The Baire category theorem and its applications

Let (X, d) be a metric space. In this course, we will denote closed and open balls as

$$\begin{aligned} B_r(x) &= \{y \in X : d(x, y) \leq r\} \\ D_r(x) &= \{y \in X : d(x, y) < r\} \end{aligned}$$

Recall that, for $A \subset X$, the *closure of A in X* is

$$\begin{aligned} \overline{A} &:= \bigcap_{\substack{F \text{ closed in } X \\ A \subset F}} F \\ &= \{x \in X : \forall r > 0 \ D_r(x) \cap A \neq \emptyset\} \\ &= \{x \in X : \exists (a_n) \text{ in } A \text{ s.t. } a_n \rightarrow x\} \end{aligned}$$

Note that $\overline{D_r(x)} \subset B_r(x)$. In general, this inclusion can be strict. But normed spaces are nice:

Exercise. Show that, in a normed space, $\overline{D_r(x)} = B_r(x)$.

Recall also that, for $A \subset X$, we say that A is *dense in X* if

$$\begin{aligned} \overline{A} &= X \\ \iff \forall x \in X \ \forall r > 0 \ D_r(x) \cap A &\neq \emptyset \\ \iff \forall \text{ non-empty open } U \subset X \ U \cap A &\neq \emptyset \end{aligned}$$

Example

\mathbb{Q} is dense in \mathbb{R} and so is $\sqrt{2} + \mathbb{Q}$. But $\mathbb{Q} \cap (\sqrt{2} + \mathbb{Q}) = \emptyset$.

4.1 Baire category theorem

Theorem 4.1 Baire category theorem

Let (X, d) be a complete metric space and $U_n \subset X$ be open and dense in X for each $n \in \mathbb{N}$. Then $\bigcap_{n \in \mathbb{N}} U_n$ is dense in X .

Proof. Fix $x_0 \in X$ and $r_0 > 0$. Since U_1 is dense, $U_1 \cap D_{r_0}(x_0) \neq \emptyset$. Then we can pick $x_1 \in U_1 \cap D_{r_0}(x_0)$. Since $U_1 \cap D_{r_0}(x_0)$ is open, there exists $r_1 > 0$ such that $B_{r_1}(x_1) \subset U_1 \cap D_{r_0}(x_0)$. WLOG, we can pick $r_1 < 1$. We then continue inductively. At the n^{th} stage, density of U_n implies that $U_n \cap D_{r_{n-1}}(x_{n-1}) \neq \emptyset$, so we can pick $x_n \in U_n \cap D_{r_{n-1}}(x_{n-1})$. Since $U_n \cap D_{r_{n-1}}(x_{n-1})$ is open, there exists $r_n > 0$ such that $B_{r_n}(x_n) \subset U_n \cap D_{r_{n-1}}(x_{n-1})$. WLOG, $r_n < 1/n$.

We end up with $(x_n)_{n=0}^\infty$ in X and $(r_n)_{n=0}^\infty$ with $0 < r_n < 1/n$ for all $n \in \mathbb{N}$ and, for all $n > N \geq 0$,

$$\begin{aligned} B_{r_n}(x_n) &\subset U_n \cap D_{r_{n-1}}(x_{n-1}) \\ &\subset U_n \cap U_{n-1} \cap D_{r_{n-2}}(x_{n-2}) \\ &\vdots \\ &\subset U_n \cap U_{n-1} \cap \cdots \cap U_{N+1} \cap D_{r_N}(x_N) \end{aligned}$$

so, for all $m, n \geq N$, we have $d(x_m, x_n) \leq 2r_N < 2/N$. Thus, $(x_n)_{n=0}^\infty$ is Cauchy and thus convergent in X . Write $x = \lim_{n \rightarrow \infty} x_n$. Note that, for $n \geq m$, $x_n \in B_{r_m}(x_m)$ so $x \in B_{r_m}(x_m)$.

By fixing $N = 0$ above and taking $n \rightarrow \infty$, we get

$$x \in \left(\bigcap_{n \in \mathbb{N}} U_n \right) \cap D_{r_0}(x_0)$$

as required. ■

Remark. A countable intersection of open sets is called a \mathcal{G}_δ -set. Theorem 4.1 then says that a countable intersection of open dense sets in a complete metric space is a dense \mathcal{G}_δ -set.

Application Uncountability of \mathbb{R}

Suppose, on the contrary, that \mathbb{R} is countable, so we can write $\mathbb{R} = \{r_1, r_2, r_3, \dots\}$. Let $U_n = \mathbb{R} \setminus \{r_n\}$. Then U_n is open and dense in \mathbb{R} . Since \mathbb{R} is complete, Theorem 4.1 tells us that $\bigcap_{n \in \mathbb{N}} U_n$ is dense in \mathbb{R} . But $\bigcap_{n \in \mathbb{N}} U_n = \emptyset$ — a contradiction!

Observe that, if $U \subset X$ is open and dense in X , then $F = X \setminus U$ is closed in X and $\text{int } F = \emptyset$.

Definition Nowhere dense

Let (X, d) be a topological space. We say that $A \subset X$ is nowhere dense in X if $\text{int } \overline{A} = \emptyset$.

Remarks

- For $A \subset Y \subset X$, it is possible that A is nowhere dense in X but not in Y (e.g. take $A = Y \neq \emptyset$)
- A is nowhere dense in X if and only if $U \not\subset \overline{U \cap A}$ for any nonempty open $U \subset X$.
 A is dense in X if and only if $U \subset \overline{U \cap A}$ for every open $U \subset X$.

Example

- In \mathbb{R} , any finite set and the Cantor set are nowhere dense.
- Write $\mathbb{Q} = \{q_1, q_2, q_3, \dots\}$ and let $(\delta_n)_{n \in \mathbb{N}}$ in $(0, 1)$. Then $U = \bigcup_{n \in \mathbb{N}} (q_n - \delta_n, q_n + \delta_n)$ is open and dense in \mathbb{R} . So $\mathbb{R} \setminus U$ is closed and nowhere dense in \mathbb{R} .

Theorem 4.1'

Let (X, d) be a non-empty complete metric space. Suppose $X = \bigcup_{n \in \mathbb{N}} A_n$ for some $A_n \subset X$. Then there exists $N \in \mathbb{N}$ such that $\text{int } \overline{A_n} \neq \emptyset$.

Proof. Suppose, on the contrary, that $\text{int } \overline{A_n} = \emptyset$ for all $n \in \mathbb{N}$. Then $\forall x \in X \ \forall r > 0 \ D_r(x) \not\subset \overline{A_n}$ and thus $D_r(x) \cap U_n = \emptyset$. Thus, $U_n = X \setminus \overline{A_n}$ is open and dense in X . Hence, by Theorem 4.1, $\bigcap_{n \in \mathbb{N}} U_n$ is dense in X . But note that $\bigcap_{n \in \mathbb{N}} U_n = (\bigcup_{n \in \mathbb{N}} \overline{A_n})^c = \emptyset$ — a contradiction! ■

Exercise. Deduce Theorem 4.1 from Theorem 4.1'.

Definition First and second category

Let X be a topological space and $A \subset X$.

- We say that A is meagre in X or is of first category in X if $A = \bigcup_{n \in \mathbb{N}} A_n$ where A_n is nowhere dense in X for all $n \in \mathbb{N}$.
- We say that A is of second category in X if A is not of first category.

Remarks

- Intuition: Think of meagre sets as ‘small’.
- Typical Baire argument: Theorem 4.1' is useful as, to find some element $x \in X$ (in a non-empty complete metric space) with some property P , we just have to show that $A = \{x \in X : x \text{ fails } P\}$ is meagre.

Application Existence of a nowhere differentiable function in $C[0, 1]$

Note that $(C[0, 1], \|\cdot\|_\infty)$ is a nonempty complete metric space. Let

$$A = \{f \in C[0, 1] : \exists x \in [0, 1] \text{ s.t. } f \text{ differentiable at } x\}$$

Observe that, if $f'(x)$ exists, i.e. $[f(y) - f(x)]/(y - x) \rightarrow f'(x)$ as $y \rightarrow x$, then there exists $N \in \mathbb{N}$ such that, for all $y \in X$,

$$|y - x| < \frac{1}{N} \implies \left| \frac{f(y) - f(x)}{y - x} \right| \leq N$$

Thus, for $n \in \mathbb{N}$, consider the set

$$A_n = \left\{ f \in C[0, 1] : \exists x \in [0, 1] \forall y \in [0, 1] |y - x| < \frac{1}{n} \implies |f(y) - f(x)| \leq n|y - x| \right\}$$

and note that $A \subset \bigcup_{n \in \mathbb{N}} A_n$.

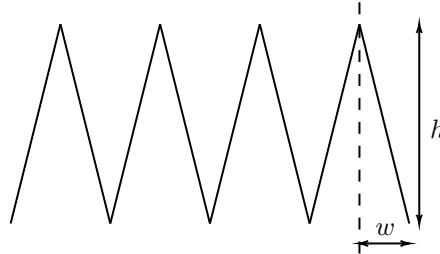
It then remains to show that, for all $n \in \mathbb{N}$, A_n is closed and $\text{int } A_n = \emptyset$.

- A_n is closed: Consider $(f_k)_{k \in \mathbb{N}}$ in A_n with $f_k \rightarrow f$ in $C[0, 1]$. For each $k \in \mathbb{N}$, we can pick $x_k \in [0, 1]$ such that, for all $y \in [0, 1]$, $|y - x_k| < 1/n \implies |f_k(y) - f_k(x_k)| \leq n|y - x_k|$. Passing to a subsequence if necessary, $x_k \rightarrow x$ in $[0, 1]$ WLOG. By IB Analysis and Topology Example Sheet 1 Q5 (2024), $f_k(x_k) \rightarrow f(x)$ and hence

$$\forall y \in [0, 1] |y - x| < \frac{1}{n} \implies |f(y) - f(x)| \leq n|y - x|$$

as required.

- Fix $f \in A_n$ and $r > 0$. To get $D_r(f) \not\subset A_n$, the idea is to consider a small but rapidly oscillating perturbation of f . Let $0 < \varepsilon < r/4$. Pick $\delta > 0$ such that $|y - x| < \delta \implies |f(y) - f(x)| < \varepsilon$. Choose h, w such that $4\varepsilon < h < r$ and $w = \min\{\varepsilon/n, \delta\}$. Set g to be the function



We can check that $f + g \in D_r(f) \setminus A_n$.

Direct proof: Take g_n similar to above with height h_n and width w_n , where $h_n \searrow 0$ fast and $h_n/w_n \rightarrow \infty$ fast. Then $\sum g_n$ is nowhere differentiable.

4.2 Consequences for Banach spaces

Theorem 4.2 Principle of uniform boundedness⁴

Let X be a Banach space, Y a normed space and $\mathcal{T} \subset \mathcal{B}(X, Y)$. If T is pointwise bounded (i.e., $\forall x \in X \sup_{T \in \mathcal{T}} \|Tx\| < \infty$), then T is uniformly bounded (i.e., $\sup_{T \in \mathcal{T}} \|T\| < \infty$).

Proof. Let $A_n = \{x \in X : \sup_{T \in \mathcal{T}} \|Tx\| \leq n\}$. By hypothesis, $X = \bigcup_{n \in \mathbb{N}} A_n$. By Theorem 4.1', there exists $n \in \mathbb{N}$ such that $\text{int } \overline{A_n} \neq \emptyset$. Note that $A_n = \bigcap_{T \in \mathcal{T}} \{x \in X : \|Tx\| \leq n\}$ is closed as

⁴This result is also known as the *Banach-Steinhaus theorem*.

the map $x \mapsto \|Tx\|$ is continuous. Thus, there exists $r > 0$ and $x \in A_n$ such that $B_r(x) \subset A_n$. Given $y \in B_X$, $T \in \mathcal{T}$, we have $x + ry, x - ry \in B_r(x)$ and thus

$$\|Ty\| = \left\| \frac{T(x+ry) - T(x-ry)}{2r} \right\| \leq \frac{2n}{2r} = \frac{n}{r}$$

Hence, $\|T\| \leq n/r$ for all $T \in \mathcal{T}$. ■

Corollary 4.3

Let X be a Banach space, Y a normed space, and $(T_n)_{n \in \mathbb{N}}$ a sequence in $\mathcal{B}(X, Y)$ that pointwise converges to T . Then T is linear and bounded. Moreover, $\sup_n \|T_n\| < \infty$.

Proof. For all $x \in X$, $(T_n x)_{n \in \mathbb{N}}$ is convergent and thus bounded. So $\{T_n : n \in \mathbb{N}\}$ is pointwise bounded. Hence, by Theorem 4.2, there exists $M \geq 0$ such that, for all $n \in \mathbb{N}$, we have $\|T_n\| \leq M$.

- T linear: $T(\lambda x + \mu y) = \lim_{n \rightarrow \infty} T_n(\lambda x + \mu y) = \lim_{n \rightarrow \infty} [\lambda T_n(x) + \mu T_n(y)] = \lambda T(x) + \mu T(y)$
- T bounded: $\forall x \in B_X \ \forall n \in \mathbb{N} \|T_n x\| \leq \|T_n\| \leq M$, so $\|Tx\| = \lim_{n \rightarrow \infty} \|T_n x\| \leq M$ for all $x \in B_X$. Hence, T is bounded with $\|T\| \leq M$. ■

Exercise. Show that $\|T\| \leq \liminf_{n \rightarrow \infty} \|T_n\|$.

Definition δ -dense

Let A, B be subsets of a metric space (X, d) and $\delta > 0$. We say that A is δ -dense in B if $\forall b \in B \ \exists a \in A \ d(a, b) \leq \delta$.

Remark. If $\overline{A} \supset B$, then A is δ -dense in B for all $\delta > 0$.

Lemma 4.4 Open mapping lemma

Let X be a Banach space, Y a normed space, $T \in \mathcal{B}(X, Y)$. Suppose that $T(MB_X)$ is δ -dense in B_Y for some $M \geq 0$ and $0 \leq \delta < 1$. Then $T(\frac{M}{1-\delta}B_X) \supset B_Y$.

Remarks

- Another way to think of the open mapping lemma is as follows.
Condition: for all $y \in B_Y$, $y = Tx$ has a δ -approximate solution in MB_X
Conclusion: for all $y \in B_Y$, $y = Tx$ has an exact solution in $\frac{M}{1-\delta}B_X$
- For any $M \geq 0$, $T(MB_X)$ is 1-dense in B_Y since $0 = T(0) \in T(MB_X)$ and $\forall y \in Y \ \|y - T(0)\| = \|y\| \leq 1$
- Lemma 4.4 implies that T is surjective
- Lemma 4.4 shows that $\overline{T(B_X)} \supset B_Y \implies T(D_X) \supset D_Y$.

Proof of Lemma 4.4. The strategy is to use ‘successive approximations’. Fix $y \in B_Y$. Pick $x_1 \in MB_X$ such that $\|y - Tx_1\| \leq \delta$. Note that $\frac{y-Tx_1}{\delta} \in B_X$, so we can pick $x_2 \in MB_X$ such that $\|\frac{y-Tx_1}{\delta} - Tx_2\| \leq \delta$. Note that $\frac{y-T(x_1+\delta x_2)}{\delta^2} \in B_Y$, so we can pick $x_3 \in MB_X$ such that $\|\frac{y-T(x_1+\delta x_2)}{\delta^2} - Tx_3\| \leq \delta$. Continuing inductively, we obtain a sequence $(x_n)_{n \in \mathbb{N}}$ in MB_X such that

$$\left\| y - T \left(\sum_{k=1}^n \delta^{k-1} x_k \right) \right\| \leq \delta^n \quad \forall n \in \mathbb{N}$$

Note that $\|\delta^{n-1} x_n\| \leq M\delta^{n-1}$ for all $n \in \mathbb{N}$, so the series $\sum_{n=1}^{\infty} \delta^{n-1} x_n$ converges absolutely and hence converges since X is complete (cf. Example Sheet 1 Q8). Let $x = \sum_{n=1}^{\infty} \delta^{n-1} x_n$. Then $\|x\| \leq \sum_{n=1}^{\infty} \delta^{n-1} M = \frac{M}{1-\delta}$, so $x \in \frac{M}{1-\delta}B_X$. Since T is continuous, $Tx = \sum_{n=1}^{\infty} T(\delta^{n-1} x_n) = \lim_{n \rightarrow \infty} T(\sum_{k=1}^n \delta^{k-1} x_k) = y$. ■

Theorem 4.5 Open mapping theorem

Let X, Y be Banach spaces and $T \in \mathcal{B}(X, Y)$. If T is surjective, then T is open.

Remark. In particular, we have $T(B_X) \supset T(D_X) \supset rB_Y$ for some $r > 0$. Equivalently, $T(MB_X) \supset B_Y$ for some $M > 0$. So the conclusion of Theorem 4.5 is that, for all $y \in Y$, $y = Tx$ has a solution such that $\|x\| \leq M\|y\|$.

Proof of Theorem 4.5. Observe that it suffices to show that $T(MB_X) \supset B_Y$ for some $M > 0$. Indeed, it would mean that, given open $U \subset X$ and $y \in T(U)$, we can pick $x \in U$ such that $y = Tx$. As U is open, $B_r(x) \subset U$ for some $r > 0$. Then $T(U) \supset T(B_r(x)) = T(x + rB_X) = T(x + \frac{r}{M}MB_X) \supset y + \frac{r}{M}B_Y = B_{r/M}(y)$.

Note that $Y = T(X) = T(\bigcup_{n \in \mathbb{N}} nB_X) = \bigcup_{n \in \mathbb{N}} T(nB_X)$. As Y is non-empty and complete, Theorem 4.1 implies that there exists $n \in \mathbb{N}$ such that $\text{int } \overline{T(nB_X)} \neq \emptyset$. So there exists $y \in Y$ and $r > 0$ such that $B_r(y) \subset \overline{T(nB_X)}$. Moreover, since B_X is convex and symmetric, so is $\overline{T(nB_X)}$. Thus, given $z \in B_Y$, we have $y \pm rz \in B_r(y) \subset \overline{T(nB_X)}$ and also $-y \pm rz \in \overline{T(nB_X)}$. Now, note that $rz = \frac{1}{2}(y+rz) + \frac{1}{2}(-y+rz) \in \overline{T(nB_X)}$. Thus, $rB_Y \subset \overline{T(nB_X)}$ or $B_Y \subset \overline{T(\frac{n}{r}B_X)}$. So $T(\frac{n}{r}B_X)$ is $\frac{1}{2}$ -dense in B_Y . By Lemma 4.4, $T(\frac{2n}{r}B_X) \supset B_Y$. ■

Theorem 4.6 Inversion theorem

Let X, Y be Banach spaces and $T: X \rightarrow Y$ a continuous linear bijection. Then $T^{-1}: Y \rightarrow X$ is also continuous.

Proof 1. T is surjective so it is open by Theorem 4.5. So for open $U \subset X$, $(T^{-1})^{-1}(U) = T(U)$ is open in Y . So T^{-1} is continuous. ■

Proof 2. By Theorem 4.5, we have $T(MB_X) \supset B_Y$ for some $M > 0$. Given $y \in Y$, there exists $x \in X$ such that $y = Tx$ and $\|x\| \leq M\|y\|$, i.e., $\|T^{-1}y\| \leq M\|y\|$. So T^{-1} is bounded. ■

Corollary 4.7

Let $\|\cdot\|$ and $\|\cdot\|'$ be complete norms on a vector space X . If there exists $b > 0$ such that $\|x\|' \leq b\|x\|$ for all $x \in X$, then $\|\cdot\|$ and $\|\cdot\|'$ are equivalent.

Proof. $\text{Id}: (X, \|\cdot\|) \rightarrow (X, \|\cdot\|')$ is a linear bijection and bounded by hypothesis. ■

Remark. This gives us another proof that the L_1 -norm on $C[0, 1]$ is complete.

Recall that, for a function $f: X \rightarrow Y$ between sets, the *graph of f* is the set

$$\Gamma(f) = \{(x, y) \in X \times Y : y = f(x)\}.$$

If X, Y are topological spaces with Y Hausdorff and f continuous, then $\Gamma(f)$ is closed.⁵

Theorem 4.8 Closed graph theorem

Let X, Y be Banach spaces and $T: X \rightarrow Y$ a linear map. If $\Gamma(T)$ is closed, then T is continuous.

Remark. Note that continuity of T means:

$$x_n \rightarrow x \text{ in } X \implies (Tx_n) \text{ converges in } Y \text{ and } \lim_{n \rightarrow \infty} Tx_n = Tx$$

But by Theorem 4.8, to prove continuity of T , it suffices to show that

$$x_n \rightarrow x \text{ in } X \text{ and } Tx_n \rightarrow y \text{ in } Y \implies y = Tx$$

⁵You may recognise this from Tripos 2025 Paper 2 Section II Question 10G

Proof of Theorem 4.8. Consider the map $S: X \rightarrow \Gamma(T)$, $x \mapsto (x, Tx)$. This is a linear bijection with $S^{-1}: \Gamma(T) \rightarrow X$, $(x, y) \mapsto x$. So $S^{-1} = P_X|_{\Gamma(T)}$ is continuous. Since $\Gamma(T)$ is a closed subspace of the Banach space $X \oplus Y$, $\Gamma(T)$ is complete. By Theorem 4.6, $S = (S^{-1})^{-1}$ is continuous. Hence, $T = P_Y \circ S$ is continuous. \blacksquare

Remark. Note that the condition ‘ $x_n \rightarrow x$ in X and $Tx_n \rightarrow y$ in $Y \implies y = Tx$ ’ is equivalent to ‘ $x_n \rightarrow 0$ in X and $Tx_n \rightarrow y$ in $Y \implies y = 0$ ’. The latter version will be quite useful for applications.

Exercise. Deduce Theorem 4.6 from Theorem 4.8.

4.3 Applications

Application

Let X be a closed subspace of ℓ_2 . We further assume that $X \subset \ell_1$.

► **Claim.** There exists $C > 0$ such that, for all $x \in X$, $\|x\|_1 \leq C\|x\|_2$.

Remark. We know that $\ell_1 \subset \ell_2$ with $\|x\|_2 \leq \|x\|_1$ for all $x \in \ell_1$. We might want to use Corollary 4.7, but we don't know if $\|\cdot\|_1$ is complete on X . Similarly, $\text{Id}: (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|_2)$ is a continuous linear bijection. However, we cannot use Theorem 4.6 to deduce that Id^{-1} is continuous since we don't know if $(X, \|\cdot\|_1)$ is complete.

Proof of Claim. Consider $T: (X, \|\cdot\|_1) \rightarrow \ell_1$, $x \mapsto x$. Since X is closed in ℓ_2 , $(X, \|\cdot\|_2)$ is complete and so is ℓ_1 . Moreover, T is linear, so by Theorem 4.8, it suffices to show that $\Gamma(T)$ is closed. To do this, suppose $x^n \rightarrow 0$ in $(X, \|\cdot\|_2)$ and $Tx^n = x^n \rightarrow y \in \ell_1$. Write $x^n = (x_k^n)_{k \in \mathbb{N}}$ and $y = (y_k)_{k \in \mathbb{N}}$. For $k \in \mathbb{N}$, $|x_k^n| \leq \|x^n\|_2 \rightarrow 0$ and $|x_k^n - y_k| \leq \|x^n - y\|_1 \rightarrow 0$ as $n \rightarrow \infty$. So $y = 0$ which means that $\Gamma(T)$ is closed. \blacksquare

Application

Let X be a normed space and $(x_n)_{n \in \mathbb{N}}$ a sequence in X such that $\forall f \in X^* \sum_{n \in \mathbb{N}} |f(x_n)| < \infty$.

► **Claim.** There exists $C \geq 0$ such that, for all $f \in X^*$, $\sum_{n \in \mathbb{N}} |f(x_n)| \leq C\|f\|$.

Proof of Claim. Define $T: X^* \rightarrow \ell_1$, $f \mapsto f(x_n)$. This is well-defined by hypothesis and also linear. Moreover, X^* and ℓ_1 are complete, so we can again use Theorem 4.8. Suppose $f_n \rightarrow 0$ in X^* and $Tf_n \rightarrow y = (y_k)_{k \in \mathbb{N}}$ in ℓ_1 . As before, for $k \in \mathbb{N}$, $f_n(x^k) \rightarrow y(x^k)$ as $n \rightarrow \infty$. But $|f_n(x_k)| \leq \|f_n\| \|x^k\| \rightarrow 0$ as $n \rightarrow \infty$. Hence, $y = 0$ as required. \blacksquare

Exercise. Prove the above claims using Theorem 4.2 and Corollary 4.3.

Application

► **Claim.** An infinite-dimensional Banach space X has no countable (algebraic) basis.

Proof. Suppose $B = \{e_1, e_2, e_3, \dots\}$ is a countable basis of X . Note that B is countably infinite since $\dim X = \infty$. For each $n \in \mathbb{N}$, let $F_n = \text{span}\{e_1, e_2, \dots, e_n\}$. Since $\dim F_n < \infty$, we know from Corollary 3.2 that F_n is closed in X . Let $x \in F_n$, $r > 0$. Then

$$x + \frac{r}{\|e_{n+1}\| + 1} e_{n+1} \in D_r(x) \setminus F_n$$

So $D_r(x) \not\subseteq F_n$ and thus $\text{int } F_n \neq \emptyset$, i.e., F_n is nowhere dense. Since $X = \text{span}B$, we have $X = \bigcup_{n \in \mathbb{N}} F_n$ — a contradiction to Theorem 4.1'. \blacksquare

Application

Let X be a vector space. Suppose Y and Z are subspaces of X such that $X = Y + Z$ and $Y \cap Z = \{0\}$, i.e., X is an algebraic direct sum of Y and Z . Then $T: Y \times Z \rightarrow X, (y, z) \mapsto y + z$ is a linear bijection. Note that $P: X \rightarrow Y, x = y + z \mapsto y$ is linear with $\text{im } P = Y$ and $\ker P = Z$. This is called the *projection of X onto Y along Z* . On the other hand, $I - P$ is the projection of X onto Z along Y .

Now, suppose X is a normed space. Recall that $Y \times Z$ is a normed space when equipped with the norm $\|(y, z)\| = \|y\| + \|z\|$. We denote this normed space by $Y \oplus Z$. The norm topology on $Y \oplus Z$ is the product topology. Note that $T: Y \oplus Z \rightarrow X$ is a continuous linear bijection:

$$\|T(y, z)\| = \|y + z\| \leq \|y\| + \|z\| = \|(y, z)\|$$

If T^{-1} is also continuous (i.e., $\|T^{-1}(y + z)\| = \|(y, z)\| = \|y\| + \|z\| \leq C\|y + z\|$ for some $C > 0$), then $X \sim Y \oplus Z$ and we write $X = Y \oplus Z$ and say that X is the (*topological*) *direct sum of Y and Z* .

► **Claim.** Suppose X is a Banach space, with subspaces Y and Z such that $X = Y + Z$ and $Y \cap Z = \{0\}$. The following are equivalent:

- (i) Y and Z are closed subspaces
- (ii) $X = Y \oplus Z$
- (iii) P is continuous

Proof. (i) \implies (ii): Y and Z are Banach spaces and hence so is $Y \oplus Z$. Since $T: Y \oplus Z \rightarrow X$ is a continuous linear bijection, T^{-1} is continuous by Theorem 4.6.

(ii) \implies (iii): $\|P(y + z)\| = \|y\| \leq \|y\| + \|z\| \leq C\|y + z\|$

(iii) \implies (i): $Z = \ker P = P^{-1}(\{0\})$ and $Y = \ker(I - P) = (I - P)^{-1}(\{0\})$ ■

Exercise. Show that (i) \implies (iii) directly using the closed graph theorem.

5 $C(K)$ spaces

Let K be a set, and let $\ell_\infty(K) := \{f: K \rightarrow \mathbb{R} \mid f \text{ bdd}\}$. This is a Banach space in the sup-norm $\|f\|_\infty = \sup_{x \in K} |f(x)|$.

Example. For $K = \mathbb{N}$, $\ell_\infty(K) = \ell_\infty$.

Let K be a topological space. Let $C_b(K) := \{f \in \ell_\infty(K) \mid f \text{ cts}\}$. By the uniform limit theorem, this is a closed subspace of $\ell_\infty(K)$ and hence a Banach space. In the case that K is a *compact* topological space, we also consider $C(K) = \{f: K \rightarrow \mathbb{R} \mid f \text{ cts}\} = C_b(K)$.

Remark. We can replace \mathbb{R} by \mathbb{C} in all of the above.

Interlude: Completions

Let (M, d) be a metric space. A *completion* of (M, d) is a complete metric space (\tilde{M}, \tilde{d}) with an isometric map $j: M \rightarrow \tilde{M}$ such that $j(M)$ is dense in \tilde{M} .

Theorem 5.1

Every metric space (M, d) has a completion.

Proof. WLOG, $M \neq \emptyset$ (result is trivial in that case). For $x \in M$, define $f_x: M \rightarrow \mathbb{R}$, $y \mapsto d(x, y)$. Fix $x_0 \in M$. Define $j: M \rightarrow \ell_\infty(M)$, $x \mapsto f_x - f_{x_0}$. Then, for every $y \in M$,

$$|f_x(y) - f_{x_0}(y)| = |d(x, y) - d(x_0, y)| \leq d(x, x_0)$$

so $j(x) \in \ell_\infty(M)$ for all $x \in M$. Moreover, for any $x, z \in M$, we have

$$\|j(x) - j(z)\|_\infty = \sup_{y \in M} |d(x, y) - d(z, y)| \leq d(x, z)$$

with equality attained when $y = z$, so j is an isometric map. Let $\tilde{M} = \overline{j(M)}$ with $\tilde{d}(f, g) = \|f - g\|_\infty$. Since \tilde{M} is a closed subset of the complete space $\ell_\infty(M)$, it is complete. ■

Remark. Completions are unique. Suppose $(\tilde{M}_1, \tilde{d}_1)$ and $(\tilde{M}_2, \tilde{d}_2)$ are completions of (M, d) with isometric maps $j_1: M \rightarrow \tilde{M}_1$ and $j_2: M \rightarrow \tilde{M}_2$. Then there exists an isometry $\theta: \tilde{M}_1 \rightarrow \tilde{M}_2$ such that $\theta \circ j_1 = j_2$. Indeed, given $x \in \tilde{M}_1$, pick a sequence $(x_n)_{n \in \mathbb{N}}$ in M such that $j_1(x_n) \rightarrow x$ and then define $\theta(x) = \lim_{n \rightarrow \infty} j_2(x_n)$.

Theorem 5.2

The completion (\tilde{X}, \tilde{d}) of a normed space X is a normed space. Moreover, the isometric map $j: X \rightarrow \tilde{X}$ is linear.

Proof. Given $x, y \in \tilde{X}$ and scalar λ , pick sequence $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ in X such that $j(x_n) \rightarrow x$ and $j(y_n) \rightarrow y$ in \tilde{X} . Then define $x + y = \lim_{n \rightarrow \infty} j(x_n + y_n)$, $\lambda x = \lim_{n \rightarrow \infty} j(\lambda x_n)$ and $\|x\| = \lim_{n \rightarrow \infty} \|x_n\|$. It is routine to verify that these are well-defined, that \tilde{X} becomes a normed space with \tilde{d} the metric induced by the norm, and that j is linear. ■

Note that we have the canonical embedding $X \rightarrow X^{**}$, $x \mapsto \hat{x}$, which is isometric and linear. We can thus take $\tilde{X} = \overline{\{\hat{x}: x \in X\}}$.

5.1 Normal spaces: Urysohn and Tietze

Definition Normal space

Let X be a topological space. We say that X is normal if, for every $E, F \subset X$ such that E, F are closed and $E \cap F = \emptyset$, there exist open sets $U, V \subset X$ such that $E \subset U$, $F \subset V$ and $U \cap V = \emptyset$.

Proposition 5.3

- (i) Metric spaces are normal.
- (ii) Compact Hausdorff spaces are normal.

Proof. (i): Let (M, d) be a metric space. For $x, y \in M$, $|d(x, A) - d(y, A)| \leq d(x, y)$. For $x \in M$, $d(x, A) = 0 \iff x \in \overline{A}$. Let E, F be disjoint closed subsets of M , and let

$$\begin{aligned} U &= \{x \in X : d(x, E) < d(x, F)\} \\ V &= \{x \in X : d(x, E) > d(x, F)\} \end{aligned}$$

Note that $U \cap V = \emptyset$ and that U, V are open since $x \mapsto d(x, E) - d(x, F)$ is continuous. Finally, for $x \in E$, we have $x \notin F = \overline{F}$ so $d(x, E) = 0 < d(x, F)$. Thus $E \subset U$ and, similarly, $F \subset V$.

(ii): Let K be a compact Hausdorff topological space. Let E, F be disjoint closed subsets of K .

STEP 1: We first prove that we can separate $x \in E$ and F .

For all $y \in F$, there exist disjoint open sets $U_y, V_y \subset K$ such that $x \in U_y$ and $y \in V_y$.

Note that $\{V_y\}_{y \in F}$ is an open cover for F . As a closed subset of a compact space, F is compact, which implies that there exists $y_1, \dots, y_n \in F$ such that $F \subset \bigcup_{i=1}^n V_{y_i}$.

Now, let $\tilde{U} = \bigcap_{i=1}^n U_{y_i}$ and $\tilde{V} = \bigcup_{i=1}^n V_{y_i}$. Note that \tilde{U}, \tilde{V} are open and disjoint, with $x \in \tilde{U}$ and $F \subset \tilde{V}$.

STEP 2: Separate E and F

Using STEP 1, for each $x \in E$, we can pick disjoint open sets $U_x, V_x \subset K$ such that $x \in U_x$ and $F \subset V_x$. Note that $\{U_x\}_{x \in E}$ is an open cover of E . As a closed subset of a compact space, E is compact, so we can pick $x_1, \dots, x_m \in E$ such that $E \subset \bigcup_{i=1}^m U_{x_i}$.

Let $U = \bigcup_{i=1}^m U_{x_i}$ and $V = \bigcap_{i=1}^m V_{x_i}$. Notably, U and V are open and disjoint, with $E \subset U$ and $F \subset V$. ■

Theorem 5.4 Urysohn's lemma

Let K be a normal topological space. Then, for disjoint closed subsets $E, F \subset K$, there exists a continuous function $f: K \rightarrow [0, 1]$ such that $f|_E = 0$ and $f|_F = 1$.

Remark. In a metric space, we can take

$$f(x) = \frac{d(x, E)}{d(x, E) + d(x, F)}$$

Of course, for a general topological space, we need a different construction. In that case, we use the following lemma.

Lemma 5.5

Let K be a topological space and $F_t, t \in \mathbb{Q}^+$ be open subsets of K such that

- (i) $K = \bigcup_{t \in \mathbb{Q}^+} F_t$
- (ii) for every $s < t$ in \mathbb{Q}^+ , $\overline{F}_s \subset F_t$

Then $f: K \rightarrow [0, \infty)$, $x \mapsto \inf\{t \in \mathbb{Q}^+ : x \in F_t\}$ is continuous.

Proof. Note that (i) implies that f is well-defined. It is easy to check that, for $t \in \mathbb{R}$,

$$f(x) < t \iff \exists s \in \mathbb{Q}^+ \text{ with } s < t \text{ s.t. } x \in F_s$$

Moreover, observe that, for $t \in \mathbb{R}$,

$$f(x) \leq t \iff (\forall s \in \mathbb{Q}^+, t < s \implies x \in \overline{F}_s)$$

Indeed, (\implies) follows from the definition of inf:

$$t < s \text{ in } \mathbb{Q}^+ \implies \exists u \in \mathbb{Q}^+ \cap (t, s) \text{ s.t. } x \in F_u \subset F_s \subset \overline{F}_s$$

To see (\iff) , note that, if $t < s$, we can pick $u \in \mathbb{Q}^+$ such that $u \in (t, s)$. Then $x \in \overline{F}_u \subset F_s$, so $f(x) \leq s$. This is true for all $s \in \mathbb{Q}^+$ with $s > t$, so we have $f(x) \leq t$.

Putting everything together, we have that, for all $a < b$ in \mathbb{R} ,

$$\{x \in K : f(x) \in (a, b)\} = \left(\bigcup_{\substack{t \in \mathbb{Q}^+ \\ t > b}} F_t \right) \setminus \left(\bigcap_{\substack{t \in \mathbb{Q}^+ \\ t > a}} \overline{F}_t \right)$$

is open, so f is continuous. ■

Remark. Suppose $E \subset W$ in some normal space, with E closed in K and W open in K . Observe that there exists open $U \subset K$ such that $E \subset U \subset \overline{U} \subset W$. Indeed, E and $K \setminus W$ are disjoint closed sets, so there exist disjoint open sets U, V such that $E \subset U$ and $K \setminus W \subset V$. Then $U \subset K \setminus V$, so $\overline{U} \subset K \setminus V$ and thus $E \subset U \subset \overline{U} \subset K \setminus V \subset W$.

With the above lemma established, we can now prove Urysohn's lemma.

Proof of Theorem 5.4. Enumerate $\mathbb{Q} \cap [0, 1]$ as $q_0 = 0, q_1 = 1, q_2, q_3, \dots$. Let $F_0 = E$ and $F_1 = K \setminus E$, $F_t = K$ for all $t \in \mathbb{Q}_{>1}$. We have $\overline{F}_0 = F_0 \subset F_1$ and $\overline{F}_1 \subset F_t$ for all $t \in \mathbb{Q}_{>1}$. We then construct open sets F_{q_n} , $n \in \mathbb{N}$, inductively as follows: Suppose we have $F_{q_0}, F_{q_1}, \dots, F_{q_n}$ for some $n \geq 2$ such that, letting π be a permutation of $\{0, 1, \dots, n\}$ such that $q_{\pi(0)} < q_{\pi(1)} < \dots < q_{\pi(n)}$, we have $\overline{F}_{q_{\pi(i-1)}} \subset F_{q_{\pi(i)}}$ for all $1 \leq i \leq n$. To define $F_{q_{n+1}}$, pick $1 \leq i \leq n$ such that $q_{\pi(i-1)} < q_{n+1} < q_{\pi(i)}$. Since $\overline{F}_{q_{\pi(i-1)}} \subset F_{q_{\pi(i)}}$, the Remark above tells us that we can pick an open set $F_{q_{n+1}}$ such that $\overline{F}_{q_{\pi(i-1)}} \subset F_{q_{n+1}} \subset \overline{F}_{q_{n+1}} \subset F_{q_{\pi(i)}}$, completing the inductive construction.

By Lemma 5.5, $f(x) = \inf\{t \in \mathbb{Q}^+ : x \in F_t\}$ is a continuous function on K . On K , $f \geq 0$ and, for $x \in K$, $x \in F_t$ for $t \in \mathbb{Q}_{>1}$ and so $f(x) \leq 1$. Thus, f is a continuous function $K \rightarrow [0, 1]$. On E , $x \in F_t \forall t \in \mathbb{Q}^+$, so $f(x) = 0$. On F , $x \notin F_1$, so $x \in F_t \iff t \in \mathbb{Q}_{>1}$. Thus, $f(x) = 1$ for $x \in F$. ■

Remark. If K is compact Hausdorff, then $C(K)$ separates points of K : given $x \neq y$ in K , there exists $f \in C(K)$ such that $f(x) \neq f(y)$.

Theorem 5.6 Tietze's extension theorem

Let K be a normal topological space and L a closed subset of K . Then, for all $g \in C_b(L)$, there exists $f \in C_b(K)$ such that $f|_L = g$ and $\|f\|_\infty = \|g\|_\infty$.

Proof. Let $X = C_b(K)$ and $Y = C_b(L)$. Recall that these are Banach spaces in $\|\cdot\|_\infty$. Define $R: X \rightarrow Y$, $f \mapsto f|_L$. Then R is a bounded linear map with $\|R\| \leq 1$. By homogeneity, it suffices to show that $R(B_X) \supset B_Y$.

Let $G \subset B_Y$. So $g: L \rightarrow [-1, 1]$. Let $E = g^{-1}([-1, -\frac{1}{3}])$ and $F = g^{-1}([\frac{1}{3}, 1])$. Then E, F are disjoint and closed in K . By Theorem 5.4, there exists a continuous function $h: K \rightarrow [0, 1]$ with

$h|_E = 0$ and $h|_F = 1$. Let $f = \frac{2}{3}(h - \frac{1}{2})$. Then $f: K \rightarrow [-\frac{1}{3}, \frac{1}{3}]$ is continuous and $f|_E = -\frac{1}{3}$ and $f|_F = \frac{1}{3}$. Now, it is easy to check that, for all $x \in L$, $|f(x) - g(x)| \leq \frac{2}{3}$. Then $f \in \frac{1}{3}B_X$ and $\|R(f) - g\| \leq \frac{2}{3}$, i.e., $R(\frac{1}{3}B_X)$ is $\frac{2}{3}$ -dense in B_Y . By 4.4, we conclude that

$$R\left(\frac{\frac{1}{3}}{1-\frac{2}{3}}B_X\right) = R(B_X) \supset B_Y.$$

as required. ■

Remark. The above proof is for the real case.

Exercise. Deduce the complex case from Theorem 5.6.

5.2 Stone-Weierstrass theorem

In this subsection, we take K to be a compact topological space. We know that $C(K)$ is a Banach space in $\|\cdot\|_\infty$. As well as being a vector space, $C(K)$ is an algebra: there exists a ‘multiplication’ binary operation $(f, g) \mapsto fg$ such that, for all $f, g, h \in C(K)$ and scalars λ ,

- (i) $(fg)h = f(gh)$
- (ii) $f(g+h) = fg + fh$ and $(f+g)h = fh + gh$
- (iii) $f(\lambda g) = (\lambda f)g = \lambda(fg)$

In fact, it is a commutative unital algebra. Moreover, for all $f, g \in C(K)$, $\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$ since for all $x \in K$, $|(fg)(x)| = |f(x)||g(x)| \leq \|f\|_\infty \|g\|_\infty$.

An algebra A which is also a normed space such that $\|ab\| \leq \|a\| \|b\|$ for all $a, b \in A$ is called a *normed algebra*. If A is complete, then it is a *Banach algebra*. For instance, $C(K)$ is a commutative unital Banach algebra.

Notation. In $C(K)$, $|f|$, $f \geq g$, $\max\{f, g\}$ are understood pointwise.

Definition Separates points

We say that $A \subset K$ separates points of K if, for every $x \neq y$ in K , there exists $f \in A$ such that $f(x) \neq f(y)$.

Remarks

- If $C(K)$ separates points of K , then K is Hausdorff.
- By Urysohn’s lemma, if K is compact Hausdorff, then $C(K)$ separates points of K .

Definition Strongly separates points

We say that $A \subset K$ strongly separates points of K if A separates points of K and, for all $x \in K$, there exists $f \in A$ such that $f(x) \neq 0$.

Definition Subalgebra

We say that $A \subset C(K)$ is a subalgebra if it is a subspace such that

$$\forall f, g \in A \quad f \cdot g \in A$$

If, in addition, $1 \in A$, then we say that A is a unital subalgebra.

Remark. If $A \subset C(K)$ is a subalgebra, then so is \overline{A} . This is because multiplication is continuous.

Thus far, $C(K)$ denotes both the real and complex case. In what follows, we will differentiate between these by letting

$$C^{\mathbb{R}}(K) = \{f: K \rightarrow \mathbb{R} \mid f \text{ cts}\}, \quad C^{\mathbb{C}}(K) = \{f: K \rightarrow \mathbb{C} \mid f \text{ cts}\}$$

Theorem 5.7 Stone-Weierstrass theorem

If A is a subalgebra of $C^{\mathbb{R}}(K)$ that strongly separates points of K , then $\overline{A} = C^{\mathbb{R}}(K)$.

Remark. The condition in the Stone-Weierstrass theorem implies that K is Hausdorff.

Lemma 5.8

For every $\varepsilon > 0$, there exists a real polynomial p such that $p(0) = 0$ and

$$\forall t \in [-1, 1] \quad |p(t) - |t|| \leq \varepsilon$$

To get an idea for how to prove this, note that $|t| = (t^2)^{1/2}$. On \mathbb{C} , $z \mapsto z^{1/2}$ has a holomorphic branch, so it has Taylor expansions on sufficiently small neighbourhoods about $z_0 \in (0, \infty)$. We can achieve this by perturbing $t^2 \mapsto t^2 + \varepsilon$.

Proof. Fix $0 < \varepsilon < 1$. Consider $f(t) = (t^2 + \varepsilon^2)^{1/2}$ for $t \in [-1, 1]$. For every $t \in [-1, 1]$, we have

$$|t| \leq \sqrt{t^2 + \varepsilon^2} \leq |t| + \varepsilon$$

so $|f(t) - |t|| \leq \varepsilon$. On \mathbb{C} ,

$$\forall t \in [-1, 1] \quad t^2 + \varepsilon^2 \in [\varepsilon^2, 1 + \varepsilon^2] \subset D_1(1) = \{z \in \mathbb{C} : |z - 1| < 1\}$$

Let $z^{1/2}$ be the holomorphic branch on $\mathbb{C} \setminus (-\infty, 0]$ such that, for $x \geq 0$, $x^{1/2} \geq 0$. This has Taylor expansion $\sum_{n=0}^{\infty} a_n(z - 1)^n$ on $D_1(1)$ (with $a_n \in \mathbb{R}$ for all $n \in \mathbb{N}$) which converges uniformly on any compact subset of $D_1(1)$, e.g. on $[\varepsilon^2, 1 + \varepsilon^2]$.

Now, choose $N \in \mathbb{N}$ such that, for all $s \in [\varepsilon^2, 1 + \varepsilon^2]$, we have

$$\left| s^{1/2} - \sum_{n=0}^N a_n(s - 1)^n \right| < \varepsilon$$

Then, for all $t \in [-1, 1]$,

$$\left| f(t) - \sum_{n=0}^N a_n(t^2 + \varepsilon^2 - 1)^n \right| < \varepsilon$$

so q is a real polynomial such that, for all $t \in [-1, 1]$, $||t| - q(t)|| \leq 2\varepsilon$. Setting $t = 0$, we see that $|q(t)| \leq 2\varepsilon$. Then $p(t) = q(t) - q(0)$ is a real polynomial such that $p(0) = 0$ and $||t| - p(t)|| \leq 4\varepsilon$. ■

Corollary 5.9

If A is a closed subalgebra of $C^{\mathbb{R}}(K)$, then A is a lattice:

$$\forall f, g \in A \quad \max\{f, g\} \in A, \min\{f, g\} \in A$$

Proof. Note that

$$\max\{f, g\} = \frac{1}{2}(f + g) + \frac{1}{2}|f - g|, \quad \min\{f, g\} = \frac{1}{2}(f + g) - \frac{1}{2}|f - g|$$

It suffices to show that $f \in A \implies |f| \in A$. Fix $f \in A$. WLOG, we may assume that $\|f\|_{\infty} \leq 1$. Let $\varepsilon > 0$ and p a real polynomial as in Lemma 5.8. Write $p = \sum_{k=1}^n a_k t^k$. Then $p(f) = \sum_{k=1}^n a_k f^k \in A$. But for every $x \in K$, $f(x) \in [-1, 1]$, so

$$\left| |f(x)| - \sum_{k=1}^n a_k f(x)^k \right| \leq \varepsilon,$$

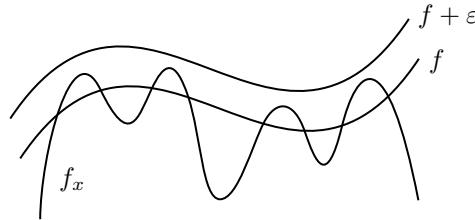
i.e., $\| |f| - p(f) \|_{\infty} \leq \varepsilon$. Thus, $|f| \in \overline{A} = A$. ■

Proof of Theorem 5.7. We observe that, for any $x \neq y$ in K and for any $a, b \in \mathbb{R}$, there exists $f \in A$ such that $f(x) = a$ and $f(y) = b$. To see this, consider the linear map $T: A \rightarrow \mathbb{R}^2, f \mapsto (f(x), f(y))$. Choose $f \in A$ such that $f(x) \neq f(y)$. If $f(x) \neq 0$ and $f(y) \neq 0$, then Tf and Tf^2 are linearly independent (note that f^2 means squaring pointwise). If WLOG $f(x) = 0$, then choose $g \in A$ such that $g(x) \neq 0$. Then Tf and Tg are linearly independent. In either case, we get that T is surjective as claimed.

Fix $f \in C^\mathbb{R}(K)$ and $\varepsilon > 0$. Fix $x \in K$. For any $y \in K$, pick $f_{x,y} \in A$ such that $f_{x,y}(x) = f(x)$ and $f_{x,y}(y) = f(y)$.

STEP 1: Approximate f within ε above

Let $U_y = \{z \in K : f_{x,y}(z) < f(z) + \varepsilon\}$. This is open and $y \in U_y$, so $\{U_y\}_{y \in K}$ is an open cover for K . By compactness, there exists $y_1, \dots, y_n \in K$ such that $K = \bigcup_{j=1}^n U_{y_j}$. Set $f_x = \min_{1 \leq j \leq n} f_{x,y_j}$. Then $f_x \in \overline{A}$ (by Corollary 5.9), with $f_x(x) = f(x)$. Given $z \in K$, there exists $1 \leq j \leq n$ so that $z \in U_{y_j}$. Thus, $f_x(z) \leq f_{x,y_j}(z) \leq f(z) + \varepsilon$.



STEP 2: Approximate f within ε below

Let $V_x = \{z \in K : f_x(z) > f(z) - \varepsilon\}$. Again this is an open cover with a finite subcover $\{V_{x_i}\}_{i=1}^m$. Let $g = \max_{1 \leq i \leq m} f_{x_i}$. Then $g \in \overline{A}$ with

$$\forall z \in K \quad g(z) = \max_{1 \leq i \leq m} f_{x_i}(z) < f(z) + \varepsilon$$

Given $z \in K$, there exists $1 \leq i \leq m$ such that $z \in V_{x_i}$. So $g(z) \geq f_{x_i}(z) > f(z) - \varepsilon$. Thus, $\|f - g\|_\infty < \infty$ and hence $f \in \overline{\overline{A}} = \overline{A}$. ■

Remark. The above result does not hold for the complex case. Consider $C^\mathbb{C}(\Delta)$ and $A = A(\Delta) = \{f \in C^\mathbb{C}(\Delta) : f \text{ holomorphic on } \text{int } \Delta\}$. A is a subalgebra, strongly separates points of Δ ($1, z \in A$). However, $\overline{A} = A \neq C^\mathbb{C}(\Delta)$ since $z \mapsto \bar{z}$ is in $C^\mathbb{C}(\Delta) \setminus A$.

Theorem 5.10 Complex Stone-Weierstrass theorem

Let A be a subalgebra of $C^\mathbb{C}(K)$ that strongly separates points of K and closed under complex conjugation. Then $\overline{A} = C^\mathbb{C}(K)$.

Proof. Let $A^\mathbb{R} = A \cap C^\mathbb{R}(K)$. This is a subalgebra of $C^\mathbb{R}(K)$. $A^\mathbb{R}$ strongly separates points of K . A is also closed under complex conjugation, so $\text{Re } f \in A^\mathbb{R}$ and $\text{Im } f \in A^\mathbb{R}$. By Theorem 5.7, $\overline{A^\mathbb{R}} = C^\mathbb{R}(K)$. Noting that $A \supset A^\mathbb{R} + iA^\mathbb{R}$, we have

$$\overline{A} \supset \overline{A^\mathbb{R} + iA^\mathbb{R}} = \overline{A^\mathbb{R}} + i\overline{A^\mathbb{R}} = C^\mathbb{R}(K) + iC^\mathbb{R}(K) = C^\mathbb{C}(K)$$

as required. ■

Now, let us explore some applications of the Stone-Weierstrass theorem.

Example. The polynomials are dense in $C[a, b]$ for any $a < b$ in \mathbb{R}

Example. Let K be a compact space of \mathbb{R}^d ($d \in \mathbb{N}$). Then the polynomials are dense in $C(K)$.

Example. Let $\mathbb{T} = \{z \in \mathbb{C}: |z| = 1\}$. Then the trigonometric polynomials, i.e., maps of the form

$$z \mapsto \sum_{k=-n}^n a_k z^k, \quad z \in \mathbb{T}, \quad a_{-n}, \dots, a_n \in \mathbb{C}$$

are dense in $C(\mathbb{T})$. (A key observation in the proof is that $\bar{z} = z^{-1}$ in \mathbb{T} .)

Example. Let K and L be compact Hausdorff. Then the continuous functions on $K \times L$ of the form

$$(x, y) \mapsto \sum_{i=1}^n f_i(x) g_i(y)$$

where $f_i \in C(K)$ and $g_i \in C(L)$ are dense in $C(K \times L)$.

Example. If K is a compact metric space, then $C(K)$ is separable.

5.3 Application to Fourier analysis

Let $\mathbb{T} = \{z \in \mathbb{C}: |z| = 1\}$. As usual, we also identify \mathbb{T} with $\mathbb{R}/2\pi\mathbb{Z}$. Recall that $C(\mathbb{T}) = \{f: \mathbb{T} \rightarrow \mathbb{C} \mid f \text{ cts}\}$ is a Banach space in $\|\cdot\|_\infty$ — in fact, its a Banach algebra. We saw above that the trigonometric polynomials are dense in $C(\mathbb{T})$, which you may recall from IB Methods is exactly what partial sums of Fourier series are.

For $f \in C(\mathbb{T})$ and $n \in \mathbb{Z}$, we call

$$\hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta$$

the n^{th} Fourier coefficient of f . The series $\sum_{n \in \mathbb{Z}} \hat{f}_n e^{in\theta}$ is called the Fourier series of f . Immediately, it is natural to ask: Does it converge? If so, in what sense?

Defines the *partial sums* by

$$S_N(f)(z) = \sum_{n=-N}^N \hat{f}_n z^n = \sum_{n=-N}^N \hat{f}_n e^{int}, \quad (z = e^{it})$$

Note that

$$S_N(f)(e^{-it}) = \sum_{n=-N}^N \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta e^{int} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) \sum_{n=-N}^N e^{in(t-\theta)} d\theta$$

Now, define the N^{th} Dirichlet kernel

$$D_N(t) = \sum_{n=-N}^N e^{int}$$

Observe that $S_N(f) = f * D_N$ and that

$$\begin{aligned} D_N(t) &= e^{-iNt} (1 + e^{it} + e^{2it} + \dots + e^{2iNt}) \\ &= e^{-iNt} \cdot \frac{e^{i(2N+1)t} - 1}{e^{it} - 1} \cdot \frac{e^{-it/2}}{e^{-it/2}} \\ &= \frac{e^{i(N+\frac{1}{2})t} - e^{-i(N+\frac{1}{2})t}}{e^{it/2} - e^{-it/2}} \\ &= \frac{\sin[(N + \frac{1}{2})t]}{\sin(\frac{1}{2}t)} \end{aligned}$$

Moreover, $D_N(0) = 2N + 1$ and $\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t) dt = 1$.

Now, consider the linear functional

$$\begin{aligned} T_N : C(\mathbb{T}) &\rightarrow \mathbb{C} \\ f &\mapsto (f * D_N)(0) \end{aligned}$$

Note that, by symmetry of D_N , we can write

$$T_N(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) D_N(\theta) d\theta$$

Lemma 5.11

$T_N \in C(\mathbb{T})^*$ and $\|T_N\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)| dt \rightarrow \infty$ as $N \rightarrow \infty$.

Before we prove this lemma, let us quickly observe a corollary that immediately follows.

Corollary 5.12

There exists $f \in C(\mathbb{T})$ such that $(T_N(f))_{N=0}^{\infty}$ is not convergent. In particular, $S_N(f)(0) \not\rightarrow f(0)$.

Proof. If $T_N(f)$ is convergent for all $f \in C(\mathbb{T})$, then $(T_N(f))_{N=0}^{\infty}$ is pointwise bounded and hence uniformly bounded by the principle of uniform boundedness (Theorem 4.2). ■

Proof of Lemma 5.11. For $f \in C(\mathbb{T})$,

$$|T_N(f)| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\theta)| |D_N(\theta)| d\theta \leq \|f\|_{\theta} \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)| d\theta$$

so T_N is bounded with $\|T_N\| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)| d\theta$.

Consider a dissection $-\pi = a_0 < a_1 < \dots < a_k = \pi$ of $[-\pi, \pi]$ such that $D_N(\theta)$ has constant sign on (a_{i-1}, a_i) for all i . Fix $\delta > 0$. Let $f(\theta) = \text{sign}(D_N(\theta))$ on $[a_{i-1} + \delta, a_i - \delta]$, $f(\pi) = f(-\pi) = 0$, and extend linearly. Then $\|f\|_{\infty} = 1$, so

$$\begin{aligned} \|T_N\| \geq T_N(f) &\geq \sum_{i=1}^k \frac{1}{2\pi} \int_{a_{i-1} + \delta}^{a_i - \delta} |D_N(\theta)| d\theta - \frac{1}{2\pi} \int_{[-\pi, \pi] \setminus \cup_{i=1}^k [a_{i-1} + \delta, a_i - \delta]} |D_N(\theta)| |f(\theta)| d\theta \\ &\geq \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)| d\theta - \frac{1}{\pi} \int_{[-\pi, \pi] \setminus \cup_{i=1}^k [a_{i-1} + \delta, a_i - \delta]} |D_N(\theta)| d\theta \\ &\geq \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)| d\theta - \frac{1}{\pi} (k+1)(2\delta)(2N+1) \end{aligned}$$

Taking $\delta \rightarrow 0$, we get $\|T_N\| \geq \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)| d\theta$.

Proving that $\|T_N\| \rightarrow \infty$ as $N \rightarrow \infty$ is left as an exercise. ■

Despite this, it turns out that, thanks to Fejér, we can remedy this by taking averages. Define

$$\sigma_N(f) = \frac{1}{N+1} \sum_{n=0}^{N+1} S_N(f) = f * \left(\frac{1}{N+1} \sum_{n=0}^N D_n \right)$$

We call

$$K_N = \frac{1}{N+1} \sum_{n=0}^N D_n$$

the N^{th} Fejér kernel. Note that

$$K_N(t) = \frac{1}{N+1} \sum_{n=0}^N e^{-int} \frac{e^{i(2n+1)t} - 1}{e^{it} - 1}$$

$$\begin{aligned}
 &= \frac{1}{N+1} \frac{1}{e^{it}-1} \sum_{n=0}^N [e^{i(n+1)t} - e^{-int}] \\
 &= \frac{1}{N+1} \frac{1}{e^{it}-1} \left(e^{it} \cdot \frac{e^{i(N+1)t} - 1}{e^{it}-1} - e^{-iNt} \frac{e^{i(N+1)t} - 1}{e^{it}-1} \right) \\
 &= \frac{1}{N+1} \frac{1}{(e^{it}-1)^2} \left[e^{i(N+2)t} - 2e^{it} + e^{-iNt} \right] \cdot \frac{e^{-it}}{e^{-it}} \\
 &= \frac{1}{N+1} \frac{e^{i(N+1)t} - 2 + e^{-i(N+1)t}}{(e^{it/2} - e^{-it/2})^2} \\
 &= \frac{1}{N+1} \left[\frac{\sin\left(\frac{N+1}{2}t\right)}{\sin\left(\frac{1}{2}t\right)} \right]^2
 \end{aligned}$$

Some other properties of K_N are

- (1) $K_N \geq 0$ on $[-\pi, \pi]$
- (2) $\frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(t) dt = 1$
- (3) for every $\delta > 0$, $K_N \rightarrow 0$ uniformly on $[-\pi, \pi] \setminus (-\delta, \delta)$

Theorem 5.13

For every $f \in C(\mathbb{T})$, $f * K_N \rightarrow f$ uniformly on \mathbb{T}

Proof. Let $f \in C(\mathbb{T})$ and $\varepsilon > 0$. Since f is uniformly continuous, we can pick $\delta > 0$ such that

$$\forall s, t \in \mathbb{T} \quad |s - t| < \delta \implies |f(s) - f(t)| < \varepsilon$$

By (3) above, we can pick $N_0 \in \mathbb{N}$ such that

$$\forall N \geq N_0 \quad \forall s \in [-\pi, \pi] \setminus (-\delta, \delta) \quad K_N(s) < \varepsilon \quad (\dagger)$$

For $N \geq N_0$ and $t \in \mathbb{T}$,

$$\begin{aligned}
 |(f * K_N)(t) - f(t)| &= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(s) f(t-s) ds - f(t) \right| \\
 &\stackrel{(2)}{=} \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(s) [f(t-s) - f(t)] ds \right| \\
 &\leq \left| \frac{1}{2\pi} \int_{-\delta}^{\delta} K_N(s) [f(t-s) - f(t)] ds \right| \\
 &\quad + \left| \frac{1}{2\pi} \int_{[-\pi, \pi] \setminus (-\delta, \delta)} K_N(s) [f(t-s) - f(t)] ds \right| \\
 &\stackrel{(1), (\dagger)}{\leq} \frac{1}{2\pi} \int_{-\delta}^{\delta} K_N(s) ds + \frac{1}{2\pi} \int_{[-\pi, \pi] \setminus (-\delta, \delta)} \varepsilon \cdot 2||f||_{\infty} ds \\
 &\leq (1 + 2||f||_{\infty}) \varepsilon
 \end{aligned}$$

as required. ■

Remark. This gives us another proof (a direct one) that the trigonometric polynomials are dense in $C(\mathbb{T})$.

Now, recall that $\hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt$, so $|\hat{f}_n| \leq ||f||_{\infty}$ for all $n \in \mathbb{Z}$. Thus, $(\hat{f}_n)_{n \in \mathbb{Z}} \in \ell_{\infty}(\mathbb{Z})$.

Proposition 5.14 Riemann-Lebesgue lemma

For every $f \in C(\mathbb{T})$, $\hat{f}_n \rightarrow 0$ as $|n| \rightarrow \infty$.

Proof. Let $f \in C(\mathbb{T})$ and $\varepsilon > 0$. By density, we can choose a trigonometric polynomial g such that $\|f - g\|_\infty < \varepsilon$. Write $g(z) = \sum_{n=-N}^N a_n z^n$. Note that

$$\hat{g}_n = \begin{cases} a_n & |n| \leq N \\ 0 & |n| > N \end{cases}$$

This implies that, for $|n| > N$, we have $|\hat{f}_n| \leq |\hat{f}_n - \hat{g}_n| = |(\widehat{f-g})_n| \leq \|f - g\|_\infty < \varepsilon$. ■

Thus, $(\hat{f}_n)_{n \in \mathbb{Z}} \in c_0(\mathbb{Z}) = \{(a_n)_{n \in \mathbb{Z}} \in \mathbb{C}^\mathbb{Z} : a_n \rightarrow 0 \text{ as } |n| \rightarrow \infty\}$, which is a closed subspace of $\ell_\infty(\mathbb{Z})$.

Next, define the *Fourier transform* as

$$\begin{aligned} \mathcal{F} : C(\mathbb{T}) &\rightarrow c_0(\mathbb{Z}) \\ f &\mapsto (\hat{f}_n)_{n \in \mathbb{Z}} \end{aligned}$$

This is a linear and bounded ($\|\mathcal{F}(f)\|_\infty = \sup_{n \in \mathbb{Z}} |\hat{f}_n| \leq \|f\|_\infty$). Moreover, $\text{im } \mathcal{F}$ contains all finite sequences and hence dense in $c_0(\mathbb{Z})$.

Theorem 5.15

The Fourier transform \mathcal{F} is injective but not surjective.

Proof. We first prove injectivity. Let $f \in \ker \mathcal{F}$. Then for all $n \in \mathbb{Z}$, $\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt = 0$, so $\int_{-\pi}^{\pi} f(t) g(t) dt = 0$ for all trigonometric polynomials g . Given $h \in C(\mathbb{T})$ and $\varepsilon > 0$, pick a trigonometric polynomial g such that $\|g - h\|_\infty < \varepsilon$. Then

$$\begin{aligned} \left| \int_{-\pi}^{\pi} f(t) h(t) dt \right| &\leq \left| \int_{-\pi}^{\pi} f(t) (h(t) - g(t)) dt \right| \\ &\leq 2\pi \|f\|_\infty \|h - g\|_\infty \\ &< 2\pi \|f\|_\infty \varepsilon \end{aligned}$$

So $\int_{-\pi}^{\pi} f(t) h(t) dt = 0$ for all $h \in C(\mathbb{T})$. Take $h = \tilde{f}$ to get $\int_{-\pi}^{\pi} |f|^2 dt = 0$, which implies that $f = 0$.

Next, suppose, on the contrary, that \mathcal{F} is surjective. Then by Theorem 4.5, \mathcal{F}^{-1} is continuous, i.e., there exists $\delta > 0$ such that $\|\mathcal{F}(f)\|_\infty \geq \delta \|f\|_\infty$ for all $f \in C(\mathbb{T})$. Note that

$$D_N(t) = \sum_{n=-N}^N e^{int}, \quad \mathcal{F}(D_N) = \mathbf{1}_{\{-N, -N+1, \dots, N-1, N\}}$$

so $1 = \|\mathcal{F}(D_N)\|_\infty \geq \delta \|D_N\|_\infty \geq \delta D_N(0) = \delta(2N + 1)$ for all $N \in \mathbb{N}$ — a contradiction! ■

For those doing II Probability and Measure, you may wonder whether we can get surjectivity by extending the domain to $L^1(\mathbb{T})$. Since $C(\mathbb{T})$ is dense in $L^1(\mathbb{T})$, the trigonometric polynomials are still dense in $L^1(\mathbb{T})$. However, \mathcal{F} is still not surjective since we still have $\|D_N\|_1 \rightarrow \infty$.

5.4 Arzelà-Ascoli theorem

The aim of this subsection is to characterise compact subsets of $C(K)$ for a compact topological space K .

Before proceeding, let us recall some relevant concepts from IB Analysis and Topology. Let (M, d) be a metric space.

- The *diameter* of a nonempty subset $A \subset M$ is

$$\text{diam } A = \sup\{d(x, y) : x, y \in A\}$$

For instance, $\text{diam } B_r(x) \leq 2r$. If $\text{diam } A \leq r$, then for any $x \in A$, $A \subset B_r(x)$. Moreover, $\text{diam } \bar{A} = \text{diam } A$.

- For $\varepsilon > 0$ and $F \subset M$, we say that F is an ε -net for M if

$$\forall x \in M \exists y \in F \quad d(x, y) \leq \varepsilon$$

Equivalently, $M = \bigcup_{y \in F} B_\varepsilon(y)$.

- We say that M is *totally bounded* if

$$\forall \varepsilon > 0 \exists \text{ finite } \varepsilon\text{-net for } M$$

$$\iff \forall \varepsilon > 0 \exists \text{ nonempty subsets } A_1, \dots, A_n \text{ of } M \text{ s.t. } M = \bigcup_{i=1}^n A_i \text{ and } \text{diam } A_i \leq \varepsilon \forall i$$

$$\iff \forall \varepsilon > 0 \exists \text{ closed subsets } A_1, \dots, A_n \text{ of } M \text{ s.t. } M = \bigcup_{i=1}^n A_i \text{ and } \text{diam } A_i \leq \varepsilon \forall i$$

Example. $(0, 1)$ is totally bounded but not compact.

Lemma 5.16

Let (M, d) be a metric space and $N \subset M$.

- (i) If M is totally bounded, then so is N .
- (ii) If N is totally bounded, then so is \bar{N} .

Proof. (i): Given $\varepsilon > 0$, $M = \bigcup_{i=1}^n A_i$ with $\text{diam } A_i \leq \varepsilon$ for all i . Then $N = \bigcup_{i=1}^n (A_i \cap N)$. Discard $A_i \cap N$ if empty; otherwise, $\text{diam } (A_i \cap N) \leq \varepsilon$.

(ii): Given $\varepsilon > 0$, $N = \bigcup_{i=1}^n A_i$ with $\text{diam } A_i \leq \varepsilon$ for all i . Then $N \subset \bigcup_{i=1}^n \bar{A}_i$ (closure in M) with $\text{diam } \bar{A}_i \leq \varepsilon$ for all i . Finite union of closed sets is closed, so $\bar{N} \subset \bigcup_{i=1}^n \bar{A}_i$. ■

Theorem 5.17

For a metric space (M, d) , the following are equivalent:

- (i) M is compact
- (ii) M is sequentially compact
- (iii) M is totally bounded and complete

Proof. Recall from IB Analysis and Topology. ■

Now, we introduce some new topological concepts.

Definition Relatively compact

We say that $N \subset M$ is relatively compact in M if \bar{N} is compact.

Corollary 5.18

Let (M, d) be a complete metric space and $N \subset M$. The following are equivalent:

- (i) N is relatively compact in M
- (ii) Every sequence in N has a subsequence that converges in M
- (iii) N is totally bounded

Proof. (i) \implies (ii): \bar{N} is compact and hence sequentially compact.

(ii) \implies (iii): Given a sequence $(x_n)_{n \in \mathbb{N}}$ in \bar{N} , for each $n \in \mathbb{N}$, we can choose $y_n \in \mathbb{N}$ such that $d(x_n, y_n) < \frac{1}{n}$. By (ii), there exists $n_1 < n_2 < \dots$, there exists $z \in M$ such that $y_{n_k} \rightarrow z$. Then $x_{n_k} \rightarrow z \in \bar{N}$. Thus, \bar{N} is sequentially compact and hence totally bounded.

(iii) \implies (i): Note that \bar{N} is totally bounded and complete, so it is compact. \blacksquare

Definition Equicontinuity and uniform boundedness

Let K be a compact topological space and $\mathcal{F} \subset C(K)$.

- We say that \mathcal{F} is equicontinuous if, for every $x \in K$ and every $\varepsilon > 0$, there exists an open neighbourhood U of x such that, for every $f \in \mathcal{F}$ and $y \in U$, we have $|f(y) - f(x)| < \varepsilon$.
- We say that \mathcal{F} is uniformly bounded if \mathcal{F} is bounded with respect to $\|\cdot\|_\infty$.

Theorem 5.19 Arzelà-Ascoli theorem

\mathcal{F} is relatively compact if and only if \mathcal{F} is uniformly bounded and equicontinuous.

Remark. \implies is trivial for finite \mathcal{F} , and compact is ‘almost’ finite.

Proof. By Corollary 5.18, \mathcal{F} is relatively compact iff \mathcal{F} is totally bounded.

(\implies) \mathcal{F} equicontinuous: Fix $x \in K$ and $\varepsilon > 0$. Let f_1, \dots, f_n be an ε -net for K . For $1 \leq i \leq n$, f_i is continuous (at x), so there exists an open neighbourhood U_i of x such that $y \in U_i \implies |f_i(y) - f_i(x)| < \varepsilon$. Then $U = \bigcap_{i=1}^n U_i$ is an open neighbourhood of x . Given $f \in \mathcal{F}$ and $y \in U$, pick $1 \leq i \leq n$ such that $\|f - f_i\| < \varepsilon$. Then

$$|f(y) - f(x)| \leq |f(y) - f_i(y)| + |f_i(y) - f_i(x)| + |f_i(x) - f(x)| < 3\varepsilon$$

\mathcal{F} uniformly bounded: Let f_1, \dots, f_n be an 1-net for K . Let $M = 1 + \max_{1 \leq i \leq n} \|f_i\|_\infty$.

For every $f \in \mathcal{F}$, choose $1 \leq i \leq n$ such that $\|f - f_i\|_\infty < 1$. Then we have

$$\|f\|_\infty \leq \|f - f_i\|_\infty + \|f_i\|_\infty \leq M$$

(\Leftarrow) Fix $\varepsilon > 0$. For each $x \in K$, pick a neighbourhood U_x of x such that, for every $f \in \mathcal{F}$ and every $y \in U_x$, $|f(y) - f(x)| < \varepsilon$. Note that $\{U_x : x \in K\}$ is an open cover for K , so we can pick $x_1, \dots, x_m \in K$ such that $K = \bigcup_{i=1}^m U_{x_i}$. For $f \in \mathcal{F}$, let $Tf = (f(x_i))_{i=1}^m$ and $S = \{Tf : f \in \mathcal{F}\} \subset \ell_\infty$ (norms of finite-dimensional spaces are equivalent). \mathcal{F} is uniformly bounded, so there exists $M \geq 0$ such that, for every $f \in \mathcal{F}$, $\|f\|_\infty \leq M$. So $\|Tf\|_\infty \leq M$ for all $f \in \mathcal{F}$. Thus, S is bounded in ℓ_∞^m and hence totally bounded. This implies that we can pick f_1, \dots, f_n in \mathcal{F} such that $\{Tf_1, \dots, Tf_n\}$ is an ε -net for S . Finally, we show that f_1, \dots, f_n is a 3ε -net for \mathcal{F} . Fix $f \in \mathcal{F}$. Pick $1 \leq j \leq n$ such that $\|Tf - Tf_j\| < \varepsilon$. Given $y \in K$, choose $1 \leq i \leq m$ such that $y \in U_{x_i}$. Then

$$|f(y) - f_j(y)| \leq |f(y) - f(x_i)| + |f(x_i) - f_j(x_i)| + |f_j(x_i) - f_j(y)| < 3\varepsilon$$

as required. \blacksquare

5.5 Application to compact operators

Let X, Y be Banach spaces, and $T: X \rightarrow Y$ be a linear map. We say that T is *compact* if $\overline{T(B_X)}$ is compact. Let $\mathcal{K}(X, Y) = \{T: X \rightarrow Y \mid T \text{ linear and compact}\}$.

Remarks

- $\mathcal{K}(X, Y) \subset \mathcal{B}(X, Y)$, i.e., compact \implies bounded
- For linear $T: X \rightarrow Y$, the following are equivalent:

- T is compact
- $T(B_X)$ is totally bounded
- For every bounded sequence (x_n) in X , Tx_n has a convergent subsequence in Y
- $\mathcal{K}(X, Y)$ is a closed subspace of $\mathcal{B}(X, Y)$.

Indeed, we have

$$- S, T \in \mathcal{K}(X, Y) \implies S + T \in \mathcal{K}(X, Y)$$

Given a bounded sequence (x_n) in X , compactness of S implies that we can pick infinite subset $L \subset \mathbb{N}$ such that $(Sx_n)_{n \in L}$ is convergent in Y . Since T is compact, we can pick an infinite subset $M \subset L$ such that $(Tx_n)_{n \in M}$ is convergent in Y . Hence, $((S+T)x_n)_{n \in M} = (Sx_n + Tx_n)_{n \in M}$ is convergent in Y , implying that $S+T \in \mathcal{K}(X, Y)$.

$$- \text{For any sequence } (T_n)_{n \in \mathbb{N}} \text{ in } \mathcal{K}(X, Y) \text{ with } T_n \rightarrow T \text{ in } \mathcal{B}(X, Y), \text{ we have } T \in \mathcal{K}(X, Y)$$

Fix $\varepsilon > 0$. Pick $n \in \mathbb{N}$ such that $\|T_n - T\| \leq \varepsilon$. Since T_n is compact, $T_n(B_X)$ is totally bounded. Thus, we can pick $x_1, \dots, x_m \in B_X$ such that $\{T_n x_1, \dots, T_n x_m\}$ is an ε -net for $T_n(B_X)$.

Given $x \in B_X$, pick $1 \leq i \leq m$ such that $\|T_n x - T_n x_i\| \leq \varepsilon$. Then $\|Tx - T x_i\| \leq \|Tx - T_n x\| + \|T_n x - T_n x_i\| + \|T_n x_i - Tx_i\| \leq 3\varepsilon$. Thus, $\{Tx_1, \dots, Tx_m\}$ is a 3ε -net for $T(B_X)$. Thus, $T \in \mathcal{K}(X, Y)$.

- $\mathcal{K}(X, Y)$ has the *ideal property*: given bounded linear maps

$$W \xrightarrow{B} X \xrightarrow{T} Y \xrightarrow{A} Z$$

between Banach spaces, $T \in \mathcal{K}(X, Y)$ implies that $ATB \in \mathcal{K}(W, Z)$. Indeed, given a bounded sequence (x_n) in W , (Bx_n) is bounded in X , so there exists an infinite subset $M \subset \mathbb{N}$ such that $(Tx_n)_{n \in M}$ is convergent in Y and thus $(ATBx_n)_{n \in M}$ is convergent in Z .

Example

- (1) Every finite-rank operator is compact.

Recall that $T \in \mathcal{B}(X, Y)$ is *of finite rank* if $\text{rk}(T) = \dim T(X) < \infty$. Let $\mathcal{F}(X, Y) = \{T \in \mathcal{B}(X, Y) : \text{rk}(T) < \infty\}$. If $T \in \mathcal{F}(X, Y)$, then $T(B_X)$ lies in the finite-dimensional $T(X)$, so $\overline{T(B_X)}$ is compact (Heine-Borel). Thus, $\mathcal{F}(X, Y) \subset \mathcal{K}(X, Y)$. But $\mathcal{K}(X, Y)$ is closed, so $\overline{\mathcal{F}(X, Y)} \subset \mathcal{K}(X, Y)$.

- (2) $T: \ell_2 \rightarrow \ell_2, (x_n) \mapsto (x_n/n)$ is compact.

Consider $T_n: \ell_2 \rightarrow \ell_2, (x_n) \mapsto \sum_{i=1}^n \frac{1}{n} x_i e_i$. Then $\|T_n - T\| \leq \frac{1}{n+1}$, so $T_n \rightarrow T$ with $\text{rk}(T_n) = n$. Hence, T is compact.

- (3) If $\dim X = \infty$, then Id_X is not compact: $\overline{\text{Id}(B_X)} = B_X$, which is not compact. More generally, if $T \in \mathcal{B}(X, Y)$ is surjective and $\dim Y = \infty$, then T is not compact. Indeed, by Theorem 4.5, there exists $\delta > 0$ such that $T(B_X) \supset \delta B_Y$.

- (4) Let $K: [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ be continuous. Define

$$\begin{aligned} T: C[0, 1] &\rightarrow C[0, 1] \\ f &\mapsto \left(x \mapsto \int_0^1 f(y) K(x, y) dy \right) \end{aligned}$$

We first show that this is well-defined. Since $[0, 1] \times [0, 1]$ is compact, K is uniformly continuous. Given $\varepsilon > 0$, there exists $\delta > 0$ such that $\|(x_1, y_1) - (x_2, y_2)\|_2 < \delta \implies$

$|K(x_1, y_1) - K(x_2, y_2)| < \varepsilon$. Then, for $|x - z| < \delta$, we have

$$(Tf)(x) - (Tf)(z) = \left| \int_0^1 f(y)[K(x, y) - K(z, y)] dy \right| \leq \int_0^1 |f(y)|\varepsilon dy = \varepsilon \|f\|_\infty$$

Next, we note that T is manifestly linear. Moreover, for any $x \in [0, 1]$,

$$|(Tf)(x)| \leq \int_0^1 |f(y)||K(x, y)| dy \leq \|f\|_\infty \|K\|_\infty$$

so $\|Tf\|_\infty \leq \|K\|_\infty \|f\|_\infty$. Hence, T is bounded with $\|T\|_\infty \leq \|K\|_\infty$.

Theorem 5.20

Let X, Y be Banach spaces and $T \in \mathcal{B}(X, Y)$. Then T is compact if and only if T^* is compact.

Remark. Note that $T^*(B_{Y^*}) = \{T^*g : g \in Y^*, \|g\| \leq 1\}$. For $g \in B_{Y^*}$, we have

$$\|T^*g\| = \sup_{x \in B_X} |(T^*g)(x)| = \sup_{x \in B_X} |g(Tx)| = \sup_{y \in T(B_X)} |g(y)| = \sup_{y \in \overline{T(B_X)}} |g(y)|$$

Thus, $T^*(B_{Y^*}) \hookrightarrow C(\overline{T(B_X)})$, which motivates the use of Arzelà-Ascoli in the proof of the Theorem.

Proof of Theorem 5.20.

(\Rightarrow) Let $K = \overline{T(B_X)}$, which is a compact metric space with $K \subset Y$. Define $\theta : T^*(B_{Y^*}) \rightarrow C(K), T^*g \mapsto g|_K$. By the computation in the Remark above, $\|T^*g\| = \|g|_K\|_\infty$. So θ is well-defined and isometric. Thus, $T^*(B_{Y^*})$ is totally bounded iff $\text{im } \theta$ is totally bounded. Now, we show that $\text{im } \theta$ is uniformly bounded and equicontinuous:

- As K is compact, it is bounded, so we can pick $M \geq 0$ such that $\|y\| \leq M$ for all $y \in K$. For $g \in B_{Y^*}$ and $y \in K$, $|g(y)| \leq \|g\|\|y\| \leq M$, so $\|g|_K\| \leq M$. So $\text{im } \theta$ is uniformly bounded.
- For $g \in B_{Y^*}$ and $y, z \in K$,

$$|g(y) - g(z)| \leq \|g\|\|y - z\| \leq \|y - z\|$$

Given $y \in K$ and $\varepsilon > 0$, we have $|g(y) - g(z)| < \varepsilon$ for every $z \in K$ with $\|y - z\| < \varepsilon$ and every $g \in B_{Y^*}$. Thus, $\text{im } \theta$ is equicontinuous.

By Arzelà-Ascoli, $\text{im } \theta$ is totally bounded and hence so is $T^*(B_{Y^*})$.

(\Leftarrow) Suppose T^* is compact, By (\Rightarrow), T^{**} is compact. But recall that we have the commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{T} & Y \\ \downarrow & & \downarrow \\ X^{**} & \xrightarrow[T^{**}]{} & Y^{**} \end{array}$$

Suitably interpreted via identifications, $T = T^{**}|_X$. Previously, we used Hahn-Banach to deduce that the canonical map $\hat{\cdot} : X \rightarrow X^{**}$ is an isometric embedding, which completes the proof. ■

6 Hilbert spaces

6.1 Basics

Definition Inner product space

Let X be a vector space over \mathbb{R} (or \mathbb{C}). An inner product on X is a function $\langle \cdot, \cdot \rangle : X \rightarrow X \rightarrow \mathbb{R}$ (resp. \mathbb{C}) such that

- (i) $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$ (linearity in the first variable)
- (ii) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ (conjugate symmetry)
- (iii) $\langle x, x \rangle \geq 0$ with equality iff $x = 0$ (positive-definiteness)

The pair $(X, \langle \cdot, \cdot \rangle)$ is called an inner product space.

Remark. Of course (i) and (ii) above imply conjugate linearity in the second variable.

Example

The following are inner product spaces:

- ℓ_2^n with $\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i$
- ℓ_2 with $\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \bar{y}_i$
- $C[0, 1]$ with $\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt$
- $C(\mathbb{T})$ with $\langle f, g \rangle = \frac{1}{2\pi} \int_0^1 f(t) \overline{g(t)} dt$
- $L^2(\Omega)$ with $\langle f, g \rangle = \int_{\Omega} f \overline{g} d\mu$ where $(\Omega, \mathcal{F}, \mu)$ is a measure space

Definition Norm

Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space. The norm of $x \in X$ is $\|x\| = \sqrt{\langle x, x \rangle}$.

Theorem 6.1

Let X be an inner product space and $x, y \in X$.

- (i) $|\langle x, y \rangle| \leq \|x\| \|y\|$ (Cauchy-Schwarz)
- (ii) $\|x + y\| \leq \|x\| + \|y\|$ (triangle)

Remarks

- The above result tells us that the object we called ‘norm’ is indeed a norm.
- $\langle \cdot, \cdot \rangle$ is continuous: if $x_n \rightarrow x$ and $y_n \rightarrow y$ in X , then $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$

Proof of Theorem 6.1. (i): For $t \in \mathbb{R}$,

$$0 \leq \langle x + ty, x + ty \rangle = \|x\|^2 + t^2 \|y\|^2 + 2t \operatorname{Re}\langle x, y \rangle$$

Choose θ such that $e^{i\theta} \langle x, y \rangle = |\langle x, y \rangle|$. Apply above to $x, e^{-i\theta} y$:

$$0 \leq \|x\|^2 + t^2 \|y\|^2 + 2t |\langle x, y \rangle|$$

We must have discriminant ≤ 0 , which gives us the result.

(ii): We have

$$\|x + y\|^2 \leq \|x\|^2 + 2 \operatorname{Re}\langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2$$

as required. ■

Definition Hilbert space

A Hilbert space is a complete inner product space.

Example

ℓ_2^n , ℓ_2 and L^2 are Hilbert spaces, whereas $C[0, 1]$ and $C(\mathbb{T})$ are not.

Proposition 6.2 Polarisation identity

Let X be an inner product space and $x, y \in X$. Then

$$\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) \quad (\mathbb{R})$$

$$\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 + i\|x + iy\|^2 - \|x - y\|^2 - i\|x - iy\|^2) \quad (\mathbb{C})$$

Proof. Expand the RHS. ■

Remark. The polarisation identity implies that, if a norm $\|\cdot\|$ arises from an inner product, then that is the unique inner product that gives rise to $\|\cdot\|$. This, for instance, allows us to talk about *the* inner product space $(C[0, 1], \|\cdot\|_2)$.

Theorem 6.3

The completion \tilde{X} of an inner product space X is a Hilbert space.

Proof. Given $x, y \in \tilde{X}$, choose sequences $(x_n), (y_n)$ in X such that $x_n \rightarrow x$ and $y_n \rightarrow y$. Define the inner product $\langle x, y \rangle = \lim_{n \rightarrow \infty} \langle x_n, y_n \rangle$. It is routine to verify that this is well-defined and defines an inner product on \tilde{X} that extends the one on X . Given $x \in \tilde{X}$, choose (x_n) in X such that $x_n \rightarrow x$. Then

$$\|x\| = \lim_{n \rightarrow \infty} \|x_n\| = \lim_{n \rightarrow \infty} \sqrt{\langle x_n, x_n \rangle} = \sqrt{\langle x, x \rangle}$$

as required. ■

Proposition 6.4 Parallelogram law

Let X be an inner product space and $x, y \in X$. Then

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

Proof. $\|x + y\|^2 + \|x - y\|^2 = \langle x + y, x + y \rangle + \langle x - y, x - y \rangle = 2\|x\|^2 + 2\|y\|^2$ ■

Remarks

- There is a converse: if a normed space X satisfies the parallelogram law, then X is an inner product space. This gives another proof of Theorem 6.3.
- We also have the *generalised parallelogram law*: Given x_1, \dots, x_n in an inner product space X , we have

$$\frac{1}{2^n} \sum_{\varepsilon \in \{\pm 1\}^n} \left\| \sum_{i=1}^n \varepsilon_i x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2$$

Definition Orthogonal

Let X be an inner product space and $x, y \in X$. We say that x and y are orthogonal and write $x \perp y$ if $\langle x, y \rangle = 0$.

Proposition 6.5 Pythagoras' theorem

$$x \perp y \implies \|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Proof. $\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\langle x, y \rangle = \|x\|^2 + \|y\|^2$ ■

Remark. We can generalise the above result. For $x_1, \dots, x_n \in X$ pairwise orthogonal,

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

Note that, for $S = \{(1 + \frac{1}{n})e_n : n \in \mathbb{N}\} \subset \ell_2$, $d(0, S) = \inf_{s \in S} \|0 - s\| = 1$ is not attained by any $s \in S$.

Theorem 6.6 Closed point theorem

Let Y be a closed subspace of a Hilbert space \mathcal{H} , and let $x \in \mathcal{H}$. Then there is a unique $y \in Y$ for which $\|x - y\| = d(x, Y)$.

Proof. Pick a sequence (y_n) in Y such that $\|x - y_n\| \rightarrow d(x, Y)$. Applying Proposition 6.4 to $x - y_m$ and $x - y_n$, we get

$$\|2x - y_m - y_n\|^2 + \|y_m - y_n\|^2 = 2\|x - y_m\|^2 + 2\|x - y_n\|^2$$

Rearranging, we get

$$\begin{aligned} \|y_m - y_n\|^2 &= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4\left\|x - \underbrace{\frac{y_m + y_n}{2}}_{\in Y}\right\|^2 \\ &\leq 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4d(x, Y)^2 \end{aligned}$$

which implies that (y_n) is Cauchy. Since Y is a closed subspace of a complete metric space, it is complete and thus (y_n) converges to some $y \in Y$. Finally, note that $\|x - y\| = \lim_{n \rightarrow \infty} \|x - y_n\| = d(x, Y)$.

Now, suppose $z \in Y$ also satisfies $\|x - z\| = d(x, Y)$. By Proposition 6.4, we have

$$\begin{aligned} \|2x - y - z\|^2 + \|y - z\|^2 &= 2\|x - y\|^2 + 2\|x - z\|^2 \\ \implies \|y - z\|^2 &= 4d(x, Y)^2 - 4\left\|x - \frac{y+z}{2}\right\|^2 \leq 0 \\ \implies y &= z \end{aligned}$$

as required. ■

6.2 Riesz representation theorem

In this subsection, we are concerned with identifying \mathcal{H}^* given a Hilbert space \mathcal{H} .

Let X be an inner product space and $y \in X$. Define $\theta_y: X \rightarrow \{\text{scalars}\}, x \mapsto \langle x, y \rangle$. Note that θ_y is linear, and by Cauchy-Schwarz, $|\theta_y(x)| \leq \|x\| \|y\|$, so $\theta_y \in X^*$ and $\|\theta_y\| \leq \|y\|$. In fact, taking $x = y$ shows that $\|\theta_y\| = \|y\|$.

Definition

Let X be an inner product space. For $y \in X$, let

$$y^\perp = \{x \in X : \langle x, y \rangle = 0\} = \ker \theta_y$$

For $S \subset X$, let

$$S^\perp = \{x \in X : \forall y \in S \quad \langle x, y \rangle = 0\} = \bigcap_{s \in S} s^\perp$$

Remarks

- We read the above as ‘ y perp’ and ‘ S perp’.
- Both are closed subspaces of X .

Theorem 6.7

Let Y be a closed subspace of a Hilbert space \mathcal{H} . Then $\mathcal{H} = Y \oplus Y^\perp$.

Remark. We say that \mathcal{H} is the *orthogonal direct sum* of Y and Y^\perp and that Y^\perp is the *orthogonal complement* of Y .

Proof of Theorem 6.7. Since Y and Y^\perp are closed subspaces and \mathcal{H} is complete, it suffices (by one of our applications of the Baire category theorem) to show that \mathcal{H} is the algebraic direct sum of Y and Y^\perp :

- $Y \cap Y^\perp = \{0\}$
- $x \in Y \cap Y^\perp \implies \|x\|^2 = \langle x, x \rangle = 0 \implies x = 0$
- $Y + Y^\perp = \mathcal{H}$

Fix $x \in \mathcal{H}$. Let $y \in Y$ be the unique closest point of Y to x . Let $z = x - y$. It remains to show that $z \in Y^\perp$.

Suppose, on the contrary, that $z \notin Y^\perp$. Then we can pick $w \in Y$ such that $\langle z, w \rangle \neq 0$. By rotating $w \mapsto e^{i\theta}w$ if necessary, we may assume WLOG that $\langle z, w \rangle \in \mathbb{R}_{>0}$. Observe that

$$\begin{aligned} \|x - y - tw\|^2 &= \|z - tw\|^2 \\ &= \langle z - tw, z - tw \rangle \\ &= \|z\|^2 + t^2\|w\|^2 - 2t\langle z, w \rangle \\ &= \|z\|^2 - t(2\langle z, w \rangle - t\|w\|^2) \\ &< \|z\|^2 \end{aligned}$$

for small $t > 0$, contradicting the fact that y is the unique closest point. ■

Remark. The projection onto Y along Y^\perp

$$\begin{aligned} P: \mathcal{H} &\rightarrow \mathcal{H} \\ y + z &\mapsto y \quad (y \in Y, z \in Y^\perp) \end{aligned}$$

is called the *orthogonal projection onto Y* . For $x \in \mathcal{H}$, we have

$$x = \underbrace{Px}_{\in Y} + \underbrace{(x - Px)}_{\in Y^\perp}$$

so Px is the unique closest point of Y to x .

Theorem 6.8 Riesz representation theorem

Let \mathcal{H} be a Hilbert space and $f \in \mathcal{H}^*$. Then there exists $y \in \mathcal{H}$ such that $f = \theta_y$.

Proof. The result is trivial when $f = 0$, so it remains to carefully consider the case $f \neq 0$. Let $Y = \ker f$, which is a closed subspace of \mathcal{H} . By Theorem 6.7, $\mathcal{H} = Y \oplus Y^\perp$ and $Y^\perp \neq \{0\}$ as $f \neq 0$. For $z, w \in Y^\perp$, $f(w)z - f(z)w \in Y$, so $\dim Y^\perp = 1$. Fix $y \in Y^\perp$ such that $f(y) = \|y\|^2$. Given $x \in \mathcal{H}$, $x = u + \lambda y$ for some $u \in Y$ and scalar λ . Then $f(x) = f(u) + \lambda f(y) = \lambda\|y\|^2$ and $\theta_y(x) = \langle x, y \rangle = \lambda\|y\|^2 = f(x)$. ■

Corollary 6.9

For a Hilbert space \mathcal{H} , the map

$$\begin{aligned} \theta: \mathcal{H} &\rightarrow \mathcal{H}^* \\ y &\mapsto \theta_y \end{aligned}$$

is an isometric, conjugate-linear isomorphism.

Proof. From Theorem 6.8 and our computation at the start of this subsection, it remains to show that θ is conjugate-linear: $\theta_{y+\lambda z}(x) = \langle x, y + \lambda z \rangle = \langle x, y \rangle + \bar{\lambda} \langle x, z \rangle = (\theta_y + \bar{\lambda} \theta_z)(x)$. ■

Remark. In Chapter 2, we had $\varphi: \ell_2 \rightarrow \ell_2^*$ is an isometric isomorphism, given by $\varphi_y(x) = \sum x_i y_i$. On the other hand, $\theta_y(x) = \langle x, y \rangle = \sum x_i \bar{y}_i$.

6.3 Orthonormal bases

Definition Orthonormal sequence and basis

Let X be an inner product space. A sequence (x_n) in X is called an orthonormal sequence if $\langle x_m, x_n \rangle = 0$ for all $m \neq n$ and $\|x_n\| = 1$ for all n . If, in addition, $\overline{\text{span}}\{x_n\} = X$, then (x_n) is called an orthonormal basis.

Example

- In ℓ_2 , (e_n) is an orthonormal basis
- In ℓ_2 , (e_{2n}) is an orthonormal sequence but not an orthonormal basis
- In $C(\mathbb{T})$, define $e_n(z) = z^n$ ($n \in \mathbb{Z}, z \in \mathbb{T}$). Note that (e_n) is an orthonormal sequence:

$$\langle e_m, e_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e_m(t) \overline{e_n(t)} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(m-n)t} dt = \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases}$$

Moreover, $\text{span}\{e_n : n \in \mathbb{Z}\}$ is the space of trig polynomials. Given $f \in C(\mathbb{T})$, there exists a trigonometric polynomial g such that $\|f - g\|_\infty < \varepsilon$. Then

$$\|f - g\|_2 = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t) - g(t)|^2 dt \right)^{1/2} \leq \varepsilon$$

Hence, $(e_n)_{n \in \mathbb{Z}}$ is an orthonormal basis.

Remarks

- If (x_n) is an orthonormal sequence, then it is linearly independent. If $\sum_{i=1}^n a_i x_i = 0$, then $a_j = \langle \sum_{i=1}^n a_i x_i, x_j \rangle = 0$ for all $1 \leq j \leq n$.
- If $\dim X < \infty$, then any orthonormal basis of X is an algebraic basis. If $\dim X = \infty$, then an orthonormal basis need not be an algebraic basis.
- If X has an orthonormal basis, then X is separable — the finite rational combinations of the basis vectors is countable and dense in X .

Theorem 6.10 Gram-Schmidt process

Let x_1, x_2, x_3, \dots (possibly finite) be a linearly independent sequence in X . Then there exists an orthonormal sequence (e_n) in X such that $\text{span}\{x_1, \dots, x_n\} = \text{span}\{e_1, \dots, e_n\}$.

Proof. Set $e_1 = x_1/\|x_1\|$. Suppose we already have constructed e_1, \dots, e_n . Let $e'_{n+1} = x_{n+1} - \sum_{i=1}^n \langle x_{n+1}, e_i \rangle e_i$. Then for $1 \leq i \leq n$, $\langle e'_{n+1}, e_i \rangle = 0$. Let $e_{n+1} = e'_{n+1}/\|e'_{n+1}\|$ (note that $e'_{n+1} \neq 0$; otherwise, $x_{n+1} \in \text{span}\{e_1, \dots, e_n\} = \text{span}\{x_1, \dots, x_n\}$ contradicting linear independence). Finally, we have

$$\text{span}\{x_1, \dots, x_n, x_{n+1}\} = \text{span}\{e_1, \dots, e_n, x_{n+1}\} = \text{span}\{e_1, \dots, e_n, e_{n+1}\}$$

as required. ■

Corollary 6.11

If X is a separable inner product space, then X has an orthonormal basis.

Proof. Let $\{x_1, x_2, x_3, \dots\}$ be a countable dense set. Pass to a subsequence y_1, y_2, \dots such that, for all n , $y_n \notin \text{span}\{x_i : i < n\}$. Note that $\{y_1, y_2, \dots\}$ is a linearly independent set whose span contains all the x_n 's, so $X = \overline{\text{span}}\{y_1, y_2, \dots\}$. By Theorem 6.10, we can pick an orthonormal sequence e_1, e_2, e_3, \dots in X such that $\text{span}\{e_1, \dots, e_n\} = \text{span}\{y_1, \dots, y_n\}$ for all n . Hence, (e_n) is an orthonormal basis of X . \blacksquare

Corollary 6.12

Let X be an inner product space with $\dim X = n$. Then $X \cong \ell_2^n$.

Proof. Let $\{e_1, \dots, e_n\}$ be an orthonormal basis of X . Then $T: \ell_2^n \rightarrow X, (a_i)_{i=1}^n \mapsto \sum_{i=1}^n a_i e_i$ is a linear bijection. By Pythagoras,

$$\forall a = (a_i)_{i=1}^n \in \ell_2^n \quad \|Ta\|^2 = \sum_{i=1}^n \|a_i e_i\|^2 = \sum_{i=1}^n |a_i|^2 = \|a\|_2^2$$

so T is isometric. \blacksquare

Theorem 6.13

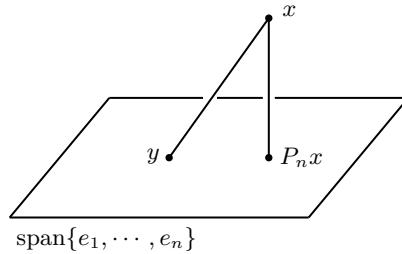
Let $(e_n)_{n \in \mathbb{N}}$ be an orthonormal basis of an inner product space X . Then for every $x \in X$, $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$.

Proof. Fix $x \in X$. For $n \in \mathbb{N}$, let $P_n x = \sum_{i=1}^n \langle x, e_i \rangle e_i$. So $x - P_n x \perp e_j$ for all $1 \leq j \leq n$ and hence $x - P_n x \perp z$ for all $z \in \text{span}\{e_1, \dots, e_n\}$. Since $X = \overline{\text{span}}\{e_n : n \in \mathbb{N}\}$, it follows that, given $\varepsilon > 0$, there exists $y \in \text{span}\{e_n : n \in \mathbb{N}\}$ such that $\|x - y\| < \varepsilon$. Pick $N \in \mathbb{N}$ such that $y \in \text{span}\{e_1, \dots, e_N\}$. Then for $n \geq N$, $P_n x - y \in \text{span}\{e_1, \dots, e_n\}$ and so $x - P_n x \perp P_n x - y$. By Pythagoras,

$$\|x - P_n x\|^2 \leq \|x - P_n x\|^2 + \|P_n x - y\|^2 = \|x - y\|^2 \leq \varepsilon^2$$

Hence, for all $n \geq N$, $\|x - P_n x\| \leq \varepsilon$. \blacksquare

Remark. $P_n: X \rightarrow X$ is the orthogonal projection onto $\text{span}\{e_1, \dots, e_n\}$.



Definition Fourier coefficients and Fourier series

We call $\langle x, e_n \rangle$ the n^{th} Fourier coefficient of x wrt (e_n) . The sum $\sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$ is called the Fourier series of x .

Example. In $C(\mathbb{T})$,

$$\langle f, e_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt = \hat{f}_n$$

Moreover, by Theorem 6.13, $\sum_{n \in \mathbb{Z}} \hat{f}_n e_n$ converges to f in $\|\cdot\|_2$.

Theorem 6.14 Parseval's identities

Let (e_n) be an orthonormal basis of an inner product space X .

- (i) For every $x \in X$, $\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2$.
- (ii) For every $x, y \in X$, $\langle x, y \rangle = \sum_{n=1}^{\infty} \langle x, e_n \rangle \overline{\langle y, e_n \rangle}$.

Proof. (i): We have

$$\|x\|^2 \stackrel{6.13}{=} \left\| \lim_{n \rightarrow \infty} \sum_{i=1}^n \langle x, e_i \rangle e_i \right\|^2 \stackrel{\text{cty of } \|\cdot\|}{=} \lim_{n \rightarrow \infty} \left\| \sum_{i=1}^n \langle x, e_i \rangle e_i \right\|^2 \stackrel{6.5}{=} \lim_{n \rightarrow \infty} \sum_{i=1}^n |\langle x, e_i \rangle|^2$$

(ii): Using Theorem 6.13 and continuity of the inner product,

$$\begin{aligned} \langle x, y \rangle &= \left\langle \sum_n \langle x, e_n \rangle e_n, \sum_m \langle y, e_m \rangle e_m \right\rangle \\ &= \sum_{m,n} \langle x, e_n \rangle \overline{\langle y, e_m \rangle} \langle e_n, e_m \rangle \\ &= \sum_n \langle x, e_n \rangle \overline{\langle y, e_n \rangle} \end{aligned}$$

■

By Parseval above, we see that $(\langle x, e_n \rangle)_n$ is in ℓ_2 . Does every $c \in \ell_2$ arise in this way?

Theorem 6.15 Riesz-Fischer theorem

Let $(e_n)_{n \in \mathbb{N}}$ be an orthonormal basis of a Hilbert space \mathcal{H} . Then, for every $c = (c_n) \in \ell_2$, there exists $x \in \mathcal{H}$ such that $\langle x, e_n \rangle = c_n$ for all n .

Proof. Fix $c \in \ell_2$. For each $n \in \mathbb{N}$, let $s_n = \sum_{i=1}^n c_i e_i$. For $n \geq m$,

$$\|S_n - S_m\|^2 = \left\| \sum_{i=m+1}^n c_i e_i \right\|^2 = \sum_{i=m+1}^n |c_i|^2 \rightarrow 0 \quad \text{as } m \rightarrow \infty$$

so (s_n) is Cauchy and hence converges to some $x \in \mathcal{H}$. Finally, observe that

$$\forall n \in \mathbb{N} \quad \langle x, e_n \rangle = \lim_{k \rightarrow \infty} \underbrace{\langle s_k, e_n \rangle}_{=c_n \ \forall k \geq n} = c_n$$

as required. ■

Theorem 6.16

Let \mathcal{H} be a separable, infinite-dimensional Hilbert space. Then $\mathcal{H} \cong \ell_2$.

Proof. Consider $T: \mathcal{H} \rightarrow \ell_2, x \mapsto (\langle x, e_n \rangle)_{n \in \mathbb{N}}$. By Parseval, T is well-defined and an isometric linear map. By Theorem 6.15, T is surjective. ■

6.4 Operators on Hilbert spaces

Let $(e_n)_{n=1}^{\infty}$ be an orthonormal basis of a Hilbert space \mathcal{H} . Let $T \in \mathcal{B}(\mathcal{H})$. Define $a_{ij} = \langle Te_j, e_i \rangle$. Then $A = (a_{ij})_{i,j=1}^{\infty}$ is the matrix of T wrt (e_n) . If $x = \sum x_i e_i$ and $y = Tx = \sum y_j e_j$, then

$$y_i = \langle Tx, e_i \rangle = \left\langle \sum_j x_j Te_j, e_i \right\rangle = \sum_j x_j \langle Te_j, e_i \rangle = \sum_j a_{ij} x_j$$

Example

- The matrix of $T = I$ wrt any orthonormal basis is

$$A = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \end{bmatrix}$$

- The matrix of the right shift operator $R \in \mathcal{B}(\ell_2)$ with respect to the unit vector basis is

$$A = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & 0 & \\ & & 1 & 0 \\ & & & \ddots & \ddots \end{bmatrix}$$

- For $a = (a_i) \in \ell_\infty$, the matrix of $T \in \mathcal{B}(\ell_2)$ defined by $T(\sum x_i e_i) = \sum a_i x_i e_i$ is

$$A = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & a_3 & \\ & & & \ddots \end{bmatrix}$$

Remark. Not every matrix is a matrix of an operator. A necessary (but *not* sufficient) condition is that each column and row is in ℓ_2 . However, the matrix of $T \in \mathcal{B}(\mathcal{H})$ does uniquely determine T .

Theorem 6.17

Let \mathcal{H} be a Hilbert space and $T \in \mathcal{B}(\mathcal{H})$. Then there exists a unique map $T^* : \mathcal{H} \rightarrow \mathcal{H}$ such that

$$\forall x, y \in \mathcal{H} \quad \langle Tx, y \rangle = \langle x, T^*y \rangle$$

Moreover, $T^* \in \mathcal{B}(\mathcal{H})$.

Remark. We call T^* the *adjoint* of T .

Proof of Theorem 6.17. We first prove uniqueness. Suppose $S_1, S_2 : \mathcal{H} \rightarrow \mathcal{H}$ are such that, for every $x, y \in \mathcal{H}$, $\langle Tx, y \rangle = \langle x, S_1 y \rangle = \langle x, S_2 y \rangle$. Setting $x = S_1 y - S_2 y$, we get $\langle S_1 y - S_2 y, S_1 y \rangle = \langle S_1 y - S_2 y, S_2 y \rangle$, which implies that $\|S_1 y - S_2 y\|^2 = 0$. Hence, $S_1 = S_2$.

Next, we show existence. Fix $y \in \mathcal{H}$. For each $x \in \mathcal{H}$, define $f(x) = \langle Tx, y \rangle$. Note that f is linear and bounded:

$$|f(x)| \leq \|Tx\| \|y\| \leq \|T\| \|x\| \|y\|$$

so $f \in \mathcal{H}^*$ with $\|f\| \leq \|T\| \|y\|$. By Riesz representation theorem, there exists $z \in \mathcal{H}$ such that

$$f(x) = \langle Tx, y \rangle = \langle x, z \rangle$$

for all $x \in \mathcal{H}$. Set $T^*y = z$. This defines a map $T^* : \mathcal{H} \rightarrow \mathcal{H}$ such that, for every $x, y \in \mathcal{H}$, $\langle Tx, y \rangle = \langle x, T^*y \rangle$.

Finally, we check linearity and boundedness:

- T^* linear

$$\langle x, T^*(\lambda y + z) \rangle = \langle Tx, \lambda y + z \rangle = \bar{\lambda} \langle Tx, y \rangle + \langle Tx, z \rangle = \bar{\lambda} \langle x, T^*y \rangle + \langle x, T^*z \rangle = \langle x, \lambda T^*y + T^*z \rangle$$

This holds for all $x \in \mathcal{H}$, so $T^*(\lambda y + z) = \lambda T^*y + T^*z$.

- T^* bounded

$$\|T^*y\|^2 = \langle T^*y, T^*y \rangle = \langle TT^*y, y \rangle \leq \|TT^*y\|\|y\| \leq \|T\|\|T^*y\|\|y\|$$

which implies that $\|T^*y\| \leq \|T\|\|y\|$. Hence, T^* is bounded with $\|T^*\| \leq \|T\|$. \blacksquare

Remarks

- We had the dual operator $T^*: \mathcal{H}^* \rightarrow \mathcal{H}^*$ sending $f \mapsto f \circ T$. After identifying \mathcal{H} with \mathcal{H}^* via the Riesz representation theorem, we get that T^* is the adjoint.
- If (e_n) is an orthonormal basis of \mathcal{H} , then

$$\langle T^*e_j, e_i \rangle = \overline{\langle e_i, T^*e_j \rangle} = \overline{\langle Te_i, e_j \rangle}$$

Thus, if A is the matrix of T wrt (e_n) , then the matrix of T^* wrt (e_n) is \overline{A}^{tr} .

Example

- $I^* = I$ since $\lambda Ix, y\rangle = \langle x, y\rangle = \langle x, Iy\rangle$
- Recall that right-shift operator R and left-shift operator L on ℓ_2 . By considering the matrix representations, we have $R^* = L$.
- On $\mathcal{H} = \ell_2$, the adjoint of $T(\sum x_i e_i) = \sum a_i x_i e_i$ ($a \in \ell_\infty$) is $T^*(\sum x_i e_i) = \sum \overline{a}_i x_i e_i$.

Proposition 6.18

Let \mathcal{H} be a Hilbert space, $S, T \in \mathcal{B}(\mathcal{H})$ and λ, μ scalars.

- (i) $(\lambda S + \mu T)^* = \overline{\lambda}S^* + \overline{\mu}T^*$
- (ii) $(ST)^* = T^*S^*$
- (iii) $(T^*)^* = T$
- (iv) $\|T^*\| = \|T\|$
- (v) $\|T^*T\| = \|T\|^2$

Proof. (i) to (iii) are easy to prove via similar manipulations we've done.

(iv): The proof of Theorem 6.17 gives us $\|T^*\| \leq \|T\|$. But then $\|T\| = \|(T^*)^*\| \leq \|T^*\|$.

(v): $\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$ and, for $x \in \mathcal{H}$,

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, T^*Tx \rangle \leq \|x\|\|T^*Tx\| \leq \|T^*\|\|T\|\|x\|^2$$

which gives us $\|T\|^2 \leq \|T^*\|\|T\|$. \blacksquare

Definition Hermitian operator

We say that $T \in \mathcal{B}(\mathcal{H})$ is hermitian or self-adjoint if $T^* = T$.

Remark. If A is the matrix of T wrt some orthonormal basis, then

$$T \text{ hermitian} \iff \overline{A}^{\text{tr}} = A$$

Example Some hermitian operators

- I is hermitian
- On $\mathcal{H} = \ell_2$, $T(\sum x_i e_i) = \sum a_i x_i e_i$ with $a \in \ell_\infty$ is hermitian iff $a_i \in \mathbb{R}$ for all i
- Let Y be a closed subspace of \mathcal{H} and $P: \mathcal{H} \rightarrow \mathcal{H}$ the orthogonal projection onto Y along Y^\perp . This is hermitian since

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

- For any $T \in \mathcal{B}(\mathcal{H})$, we have $(T^*T)^* = T^*T^{**} = T^*T$, so T^*T is hermitian.

Proposition 6.19

Let \mathcal{H} be a complex Hilbert space and $T \in \mathcal{B}(\mathcal{H})$. Then there exists unique hermitian operators T_1, T_2 such that $T = T_1 + iT_2$.

Proof. We first prove uniqueness. Suppose $T = T_1 + iT_2$ with T_1, T_2 hermitian. Then $T^* = T_1^* - iT_2^*$, so

$$T_1 = \frac{1}{2}(T + T^*), \quad T_2 = \frac{1}{2i}(T - T^*)$$

To prove uniqueness, we simply define T_1 and T_2 as above. It is easy to check that they are hermitian and satisfy $T = T_1 + iT_2$. \blacksquare

Definition Unitary operator

We say that $T \in \mathcal{B}(\mathcal{H})$ is unitary if $T^*T = TT^* = I$.

Example. On $\mathcal{H} = \ell_2$, the operator $T(\sum x_i e_i) = \sum a_i x_i e_i$ is unitary iff $|a_i| = 1$ for all i .

Proposition 6.20

Let $T \in \mathcal{B}(\mathcal{H})$. Then T is unitary iff T is isometric and surjective.

Proof.

(\implies) $\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, T^*Tx \rangle = \|x\|^2$ so T is isometric. Moreover, T is invertible and thus surjective.

(\impliedby) $\langle Tx, Tx \rangle = \langle x, x \rangle$ for all $x \in \mathcal{H}$. By the polarisation identities, $\langle Tx, Ty \rangle = \langle x, y \rangle$ for all $x, y \in \mathcal{H}$. So $\langle x, T^*Ty \rangle = \langle x, y \rangle$ for all $x, y \in \mathcal{H}$ and thus $T^*T = I$. Thus, T is injective. But we are given that T is also surjective. Since T is a continuous linear bijection, we deduce from the inversion theorem that T is invertible with $T^{-1} = (T^*T)T^{-1} = T^*$. \blacksquare

Remark. In general, $T^*T = I$ does not imply T is invertible, e.g., the right-shift operator.

Definition Normal operator

We say that $T \in \mathcal{B}(\mathcal{H})$ is normal if $T^*T = TT^*$.

Example. Hermitian and unitary operators are normal.

7 Spectral theory

Let X be a Banach space. Recall that $\mathcal{B}(X) = \{T: X \rightarrow X \mid T \text{ linear and bounded}\}$ is a Banach space (Theorem 1.7). In fact, $\mathcal{B}(X)$ is a Banach algebra with composition as ‘multiplication’: for every $S, T, U \in \mathcal{B}(X)$ and scalar λ ,

- $(ST)U = S(TU)$
- $S(T + U) = ST + SU$ and $(S + T)U = SU + TU$
- $\lambda(ST) = (\lambda S)T = S(\lambda T)$
- $\|ST\| \leq \|S\|\|T\|$ (Proposition 1.6)

Definition Invertible operator

Let $T \in \mathcal{B}(X)$. We say that T is invertible if there exists $S \in \mathcal{B}(X)$ such that $ST = TS = I$.

Note that

$$T \text{ invertible} \iff T \text{ linear bijection and } T^{-1} \text{ bdd} \stackrel{\text{Thm 4.6}}{\iff} T \text{ linear bijection}$$

Theorem 7.1

Let $T \in \mathcal{B}(X)$. If $\|I - T\| < 1$, then T is invertible and

$$\|T^{-1}\| \leq \frac{1}{1 - \|I - T\|}$$

Remark. 1 is optimal because of 0. The idea behind the proof is to take inspiration from the binomial expansion.

Proof of Theorem 7.1. Define $H = I - T$. Then $T = I - H$ and $\|H\| < 1$. For every $n \in \mathbb{N}$, $\|H^n\| \leq \|H\|^n$ so $\sum_{n=0}^{\infty} \|H^n\|$ converges by comparison with geometric series. Therefore, $\sum_{n=0}^{\infty} H^n$ converges absolutely and thus converges. To see that it converges to the inverse of T , let $S = \sum_{n=0}^{\infty} H^n$. Then $HS = SH = \sum_{n=1}^{\infty} H^n = S - I$. Rearranging, we get $S(I - H) = (I - H)S = I$. Hence, $T^{-1} = S$. Finally, we have

$$\|T^{-1}\| \leq \sum_{n=0}^{\infty} \|H^n\| \leq \sum_{n=0}^{\infty} \|H\|^n = \frac{1}{1 - \|H\|} = \frac{1}{1 - \|I - T\|}$$

as required. ■

Let $\mathcal{G}(X) = \{T \in \mathcal{B}(X) : T \text{ invertible}\}$. Having shown that it is non-empty, it is now easy to see that $\mathcal{G}(X)$ forms a group under composition.

Theorem 7.2

Let X be a Banach space. Then

- (i) $\mathcal{G}(X)$ is an open subset of $\mathcal{B}(X)$. (In particular, $T \in \mathcal{G}(X) \implies D_{1/\|T^{-1}\|}(T) \subset \mathcal{G}(X)$.)
- (ii) The map $\mathcal{G}(X) \rightarrow \mathcal{G}(X), T \mapsto T^{-1}$ is continuous.
- (iii) For any sequence (T_n) in $\mathcal{G}(X)$ that converges to some $T \in \mathcal{B}(X) \setminus \mathcal{G}(X)$, then $\|T_n^{-1}\| \rightarrow \infty$.

Proof. (i): Fix $T \in \mathcal{G}(X)$. Let $S \in \mathcal{B}(X)$ such that $\|S - T\| < 1/\|T^{-1}\|$. Note that

$$\|I - T^{-1}S\| = \|T^{-1}(T - S)\| \leq \|T^{-1}\| \|T - S\| < 1$$

so $T^{-1}S \in \mathcal{G}(X)$ by Theorem 7.1. Since $\mathcal{G}(X)$ is a group, $S = T(T^{-1}S) \in \mathcal{G}(X)$.

(ii): Fix $T \in \mathcal{G}(X)$. Let $S \in \mathcal{B}(X)$ such that $\|S - T\| < 1/(2\|T^{-1}\|)$. Then $S \in \mathcal{G}(X)$ (from (i)) and

$$\|S^{-1} - T^{-1}\| = \|S^{-1}(T - S)T^{-1}\| \leq \|S^{-1}\| \|T - S\| \|T^{-1}\| \quad (*)$$

Note that $\|S^{-1}\| \leq \|S^{-1} - T^{-1}\| + \|T^{-1}\| \leq \|S^{-1}\| \|T - S\| \|T^{-1}\| + \|T^{-1}\| \leq \frac{1}{2} \|S^{-1}\| + \|T^{-1}\|$, so $\|S^{-1}\| \leq 2\|T^{-1}\|$. Then (*) becomes

$$\|S^{-1} - T^{-1}\| \leq 2\|T^{-1}\|^2\|T - S\| \rightarrow 0 \quad \text{as } S \rightarrow T$$

(iii): From (i), $\|S - T_n\| < 1/\|T_n^{-1}\| \implies S \in \mathcal{G}(X)$. Since $T \notin \mathcal{G}(X)$, we have $1/\|T_n^{-1}\| \leq \|T - T_n\| \rightarrow 0$, so $\|T_n^{-1}\| \rightarrow \infty$. \blacksquare

Intuition: Invertible operators are ‘large’, whereas finite-rank operators and compact operators are ‘small’.

Remarks

- If $\dim X < \infty$, then $\mathcal{B}(X) = \mathcal{K}(X) = \mathcal{F}(X)$.
- If $\dim X = \infty$, then $\mathcal{G}(X) \cap \mathcal{K}(X) = \emptyset$
- If $T \in \mathcal{K}(X)$, then the dual operator $T^*: X^* \rightarrow X^*$ is also compact.
If \mathcal{H} is a Hilbert space and $T \in \mathcal{K}(\mathcal{H})$, then the adjoint operator $T^*: \mathcal{H} \rightarrow \mathcal{H}$ is also compact.
Direct proof: Let (x_n) be a bounded sequence in \mathcal{H} . Fix $C \geq 0$ such that, for every n , $\|x_n\| \leq C$. Then for all n , $\|T^*x_n\| \leq \|T^*\| \|x_n\| \leq C\|T\|$. Since T is compact, $(T(T^*x_n))$ has a convergent subsequence. Passing to a subsequence if necessary, we can take $(T(T^*x_n))$ to be convergent. Then

$$\begin{aligned} \|T^*x_m - T^*x_n\| &= \langle T^*(x_m - x_n), T^*(x_m - x_n) \rangle \\ &= \langle TT^*(x_m - x_n), x_m - x_n \rangle \\ &\leq \|x_m - x_n\| \|T(T^*x_m) - T(T^*x_n)\| \\ &\leq 2C\|T(T^*x_m) - T(T^*x_n)\| \\ &\rightarrow 0 \quad \text{as } m, n \rightarrow \infty \end{aligned}$$

7.1 Operators on complex Banach spaces

From now on, the scalar field is \mathbb{C} and we fix a nonempty Banach space X .

Definition Spectrum of an operator

Let $T \in \mathcal{B}(X)$. The spectrum of T is the set

$$\sigma(T) = \{\lambda \in \mathbb{C}: \lambda I - T \notin \mathcal{G}(X)\}$$

Example

- If $\dim X < \infty$, then $\sigma(T)$ is the set of eigenvalues of T . (By the Fundamental Theorem of Algebra, $\sigma(T)$ is nonempty and finite.)
- Consider the right-shift operator R . Observe that R has no eigenvalues. Indeed, suppose that $Rx = \lambda x$ for some $\lambda \in \mathbb{C}, x \in \ell_2 \setminus \{0\}$. Then $(0, x_1, x_2, \dots) = (\lambda x_1, \lambda x_2, \lambda x_3, \dots)$. So $\lambda \neq 0$. By induction, we see that $x_n = 0$ for all $n \in \mathbb{N}$. (Later on, we will see that $\sigma(R) = \Delta = \{z \in \mathbb{C}: |z| \leq 1\}$.)
- $\sigma(I) = \{1\}$ and $\sigma(\lambda I) = \{\lambda\}$.

Proposition 7.3

Let X be a non-empty, compact Banach space and $T \in \mathcal{B}(X)$. Then $\sigma(T)$ is a closed subset of $\{\lambda \in \mathbb{C}: |\lambda| \leq \|T\|\}$. Hence, $\sigma(T)$ is compact.

Proof. Consider the map $f: \mathbb{C} \rightarrow \mathcal{B}(X)$, $\lambda \mapsto \lambda I - T$. This is continuous, so $\sigma(T) = f^{-1}(\mathcal{B}(X) \setminus \mathcal{G}(X))$ is closed. Moreover, if $|\lambda| > \|T\|$, then $\|T/\lambda\| < 1$, which implies that $I - T/\lambda \in \mathcal{G}(X)$ and thus $\lambda I - T \in \mathcal{G}(X)$. \blacksquare

Definition Resolvent set and function

Let $T \in \mathcal{B}(X)$. The resolvent set of T is

$$\rho(T) = \mathbb{C} \setminus \sigma(T)$$

The resolvent function of T is

$$\begin{aligned} R_T: \rho(T) &\rightarrow \mathcal{B}(X) \\ \lambda &\mapsto (\lambda I - T)^{-1} \end{aligned}$$

Remarks

- $\rho(T)$ is open in \mathbb{C}
- R_T is continuous by Theorem 7.2(ii).

If $|\lambda| > \|T\|$, we have

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

Example. Consider the left-shift operator $L: \ell_2 \rightarrow \ell_2$. If $|\lambda| < 1$, then $x = (1, \lambda, \lambda^2, \dots) \in \ell_2$ and $x \neq 0$, with

$$Lx = (\lambda, \lambda^2, \lambda^3, \dots) = \lambda x$$

so λ is an eigenvalue of L . Thus, $\lambda I - L$ is not injective and hence $\lambda \in \sigma(L)$. We've therefore deduced that $\sigma(L) \supset D = \{z \in \mathbb{C}: |z| < 1\}$. By Proposition 7.3, $\Delta = \overline{D} \subset \sigma(L)$ and $\sigma(L) \subset \{\lambda \in \mathbb{X}: |\lambda| \leq \|L\| = 1\}$. Thus, $\sigma(L) \subset \Delta$. Hence, $\sigma(L) = \Delta$.

Definition Point spectrum

The point spectrum of $T \in \mathcal{B}(X)$ is the set

$$\sigma_p(T) = \{\lambda \in \mathbb{C}: \lambda \text{ eigenvalue of } T\}$$

Remarks

- $\sigma_p(T) \subset \sigma(T)$ so $\overline{\sigma_p(T)} \subset \sigma(T)$.
- If $\dim X < \infty$, then $\sigma_p(T) = \sigma(T)$ by rank-nullity theorem.
- Let $R, L: \ell_2 \rightarrow \ell_2$ be the right-shift and left-shift operators, respectively. We've already seen that $\sigma_p(R) = \emptyset$, and we will later see that $\sigma(L) \neq \emptyset$. On the other hand, observe that $\sigma_p(L) = D$ since

$$|\lambda| = 1, (x_2, x_3, \dots) = \lambda(x_1, x_2, \dots) \implies \forall n \ |x_{n+1}| = |\lambda x_n| = |x_n| \rightarrow 0$$

Moreover, we had $\sigma(L) = \Delta$ from the previous example.

Theorem 7.4

Let $T \in \mathcal{B}(X)$. Then $\sigma(T) \neq \emptyset$.

Proof. Recall that $R_T: \rho(T) \rightarrow \mathcal{B}(X)$, $\lambda \mapsto (\lambda I - T)^{-1}$ is continuous by Theorem 7.2. In fact, it is holomorphic:

$$R_T(\lambda) - R_T(\mu) = R_T(\lambda)[(\mu I - T) - (\lambda I - T)]R_T(\mu)$$

$$\begin{aligned}
 &= (\mu - \lambda)R_T(\lambda)R_T(\mu) \\
 \implies \frac{R_T(\lambda) - R_T(\mu)}{\lambda - \mu} &= -R_T(\lambda)R_T(\mu) \rightarrow -R_T(\mu)^2 \quad \text{as } \lambda \rightarrow \mu
 \end{aligned}$$

Moreover, it is bounded: for $|\lambda| > \|T\|$,

$$\|R_T(\lambda)\| \leq \left\| \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}} \right\| \leq \sum_{n=0}^{\infty} \left\| \frac{T^n}{\lambda^{n+1}} \right\| \leq \sum_{n=0}^{\infty} \frac{\|T\|^n}{\lambda^{n+1}} = \frac{1}{|\lambda|} \frac{1}{1 - \|T/\lambda\|} = \frac{1}{|\lambda| - \|T\|} \rightarrow 0$$

as $|\lambda| \rightarrow \infty$. (Concretely, the above tells us that R_T is bounded outside some compact disc, but the extreme value theorem tells us that R_T is also bounded on the compact disc.)

Now, suppose, on the contrary, that $\sigma(T) = \emptyset$. Then $R_T: \mathbb{C} \rightarrow \mathcal{B}(X)$ is holomorphic and bounded. For the remainder of this proof, we will assume X is some Hilbert space \mathcal{H} . (See non-examinable remark below for the general case.) Fix $x, y \in \mathcal{H}$ and define

$$\begin{aligned}
 f: \mathbb{C} &\rightarrow \mathbb{C} \\
 \lambda &\mapsto \langle R_T(\lambda)x, y \rangle
 \end{aligned}$$

Note that

$$\frac{f(\lambda) - f(\mu)}{\lambda - \mu} = \left\langle \frac{R_T(\lambda) - R_T(\mu)}{\lambda - \mu} x, y \right\rangle \rightarrow \langle -R_T(\mu)^2 x, y \rangle \quad \text{as } \lambda \rightarrow \mu$$

so f is holomorphic. But also

$$|f(\lambda)| \leq \|R_T(\lambda)x\| \|y\| \leq \|R_T(\lambda)\| \|x\| \|y\| \rightarrow 0 \quad \text{as } |\lambda| \rightarrow \infty$$

so f is bounded. By Liouville's theorem, $f \equiv 0$. Note that this holds for all $x, y \in \mathcal{H}$. Finally, fix $\lambda \in \mathbb{C}$. For every $x, y \in \mathcal{H}$, $\langle R_T(\lambda)x, y \rangle = 0$ and hence $R_T(\lambda) = 0$ — a contradiction! ■

Proof for general Banach space. Fix $\varphi \in \mathcal{B}(X)^*$ and consider $f = \varphi \circ R_T$. By Liouville, we can again conclude that $f \equiv 0$. Then for all $\lambda \in \mathbb{C}$ and $\varphi \in \mathcal{B}(X)^*$, $\varphi(R_T(\lambda)) = 0$. By Hahn-Banach, we conclude that $R_T(\lambda) = 0$ for all $\lambda \in \mathbb{C}$ — a contradiction! ■

Example. Consider $T: \ell_2 \rightarrow \ell_2$, $\sum x_n e_n \mapsto \sum 2^{-n} x_n e_n$. Note that T is injective, so $0 \notin \sigma_p(T)$. But $T e_n = 2^{-n} e_n \rightarrow 0$ as $n \rightarrow \infty$.

The above example motivates the following definition:

Definition Approximate eigenvalue

Let $T \in \mathcal{B}(X)$ and $\lambda \in \mathbb{C}$. We say that λ is an approximate eigenvalue of T if there exists (x_n) in S_X such that $(\lambda I - T)x_n \rightarrow 0$. The sequence (x_n) is called an approximate eigenvector of T .

Proposition 7.5

Let $T \in \mathcal{B}(X)$ and $\lambda \in \mathbb{C}$.

- (i) If λ is an eigenvalue of T , then it is an approximate eigenvalue of T .
- (ii) If λ is an approximate eigenvalue of T , then $\lambda \in \sigma(T)$.

Proof. (i): Pick a nonzero $x \in X$ such that $(\lambda I - T)x = 0$. Take $x_n = x/\|x\|$ for all n . Then $(\lambda I - T)x_n = 0 \rightarrow 0$.

(ii): Suppose, on the contrary, that $\lambda \notin \sigma(T)$. Let $S = (\lambda I - T)^{-1}$. Fix a sequence (x_n) in S_X . Then

$$1 = \|x_n\| = \|S(\lambda I - T)x_n\| \leq \|S\| \|(\lambda I - T)x_n\|$$

so $\|(\lambda I - T)x_n\| \geq 1/\|S\|$ and $\|(\lambda I - T)x_n\| \not\rightarrow 0$. Hence, λ is not an approximate eigenvalue of T . \blacksquare

Notation. Let $\sigma_{\text{ap}}(T) = \{\lambda \in \mathbb{C} : \lambda \text{ approximate eigenvalue of } T\}$.

Remarks

- Proposition 7.5 tells us that $\sigma_p(T) \subset \sigma_{\text{ap}}(T) \subset \sigma(T)$.
- If $\dim X < \infty$, then $\sigma_p(T) = \sigma_{\text{ap}}(T) = \sigma(T)$.
- For $T : \ell_2 \rightarrow \ell_2$, $\sum x_n e_n \mapsto \sum 2^{-n} x_n e_n$,

$$\sigma_p(T) \subsetneq \sigma_{\text{ap}}(T)$$

since $0 \in \sigma_{\text{ap}}(T) \setminus \sigma_p(T)$.

- For the right-shift operator $R : \ell_2 \rightarrow \ell_2$,

$$\sigma_{\text{ap}}(T) \subsetneq \sigma(T)$$

since T not surjective $\implies 0 \in \sigma(T)$ but T isometric $\implies 0 \notin \sigma_{\text{ap}}(T)$.

Recall from IB Analysis and Topology that, for a subset A of a topological space Z , the *boundary of A in Z* is $\partial A = \overline{A} \setminus \text{int } A$. Recall also that $\partial A = \partial(Z \setminus A) = \overline{A} \cap \overline{Z \setminus A}$.

Example. In \mathbb{C} , $\partial\Delta = \Delta \setminus D = \mathbb{T}$.

Theorem 7.6

Let $T \in \mathcal{B}(X)$. Then $\partial\sigma(T) \subset \sigma_{\text{ap}}(T)$. Hence $\sigma_{\text{ap}}(T) \neq \emptyset$.

Proof. Let $\lambda \in \partial\sigma(T)$. Then we can pick a sequence (λ_n) in $\rho(T)$ such that $\lambda_n \rightarrow \lambda$. So $\lambda_n I - T \in \mathcal{G}(X)$ for all $n \in \mathbb{N}$, with $\lambda_n I - T \rightarrow \lambda I - T \notin \mathcal{G}(X)$. By Theorem 7.2, $\|(\lambda_n I - T)^{-1}\| \rightarrow \infty$ as $n \rightarrow \infty$. Thus, we can pick a sequence (y_n) in S_X such that $\|(\lambda_n I - T)^{-1}y_n\| \rightarrow \infty$. Set $x_n = \frac{(\lambda_n I - T)^{-1}y_n}{\|(\lambda_n I - T)^{-1}y_n\|}$. Then $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and

$$(\lambda I - T)x_n = (\lambda_n I - T)x_n + (\lambda - \lambda_n)x_n \rightarrow 0$$

since

$$\|(\lambda_n I - T)x_n\| = \frac{1}{\|(\lambda_n I - T)^{-1}x_n\|} \rightarrow 0, \quad \|(\lambda - \lambda_n)x_n\| = |\lambda - \lambda_n| \rightarrow 0$$

Thus, $\lambda \in \sigma_{\text{ap}}(T)$. Since $\sigma(T)$ is closed, non-empty proper subset of \mathbb{C} , it is not open. Hence, $\partial\sigma(T) = \sigma(T) \setminus \text{int } \sigma(T) \neq \emptyset$. \blacksquare

Example. Consider the right-shift operator $R : \ell_2 \rightarrow \ell_2$. Note that $\|T\| = 1$, so $\sigma(T) \subset \Delta$. Moreover, $0 \in \sigma(T)$ since T is not surjective. If $|\lambda| < 1$ and $\|x\| = 1$, then $\|(\lambda I - T)x\| \geq \|Tx\| - |\lambda| = \|x\| - |\lambda| = 1 - |\lambda| > 0$, so $\lambda \notin \sigma_{\text{ap}}(T)$. Now suppose, on the contrary, that there exists some $\lambda \in \Delta \setminus \{0\}$ such that $\lambda \notin \sigma(T)$. Then we can pick $s \in [0, 1)$ such that $\mu = s\lambda$ satisfies $\mu \in \partial\sigma(T) \subset \sigma_{\text{ap}}(T)$ — a contradiction! Hence, $\sigma(T) = \Delta$.

How big is the spectrum of an operator?

Definition Spectral radius

The spectral radius of $T \in \mathcal{B}(X)$ is

$$r(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}$$

By Proposition 7.3 and Theorem 7.4, $r(T)$ is well-defined, with $r(T) \leq \|T\|$.

Example. For both the right-shift and left-shift operators $R, L \in \mathcal{B}(\ell_2)$, we have $r(R) = r(L) = 1 = \|R\| = \|L\|$.

Example. For $T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathcal{B}(\ell_2^2)$, $r(T) = 0 < \|T\|$.

Let $a, b \in \mathcal{R}$ with \mathcal{R} a unital ring. If a, b are both invertible, then ab is invertible. On the other hand, the converse is false in general, e.g., $a = L, b = R$. However, the converse does hold if $ab = ba$. Indeed, let $z = (ab)^{-1}$ so $zab = abz = 1$ and thus $zba = baz = 1$. Then $(zb)a = 1$ and $a(zb) = azbabz = abz = 1$.

Theorem 7.7 Spectral mapping theorem

Let $T \in \mathcal{B}(X), p = a_0 + a_1z + \cdots + a_nz^n$ a complex polynomial. Define $p(T) = a_0I + a_1T + a_2T^2 + \cdots + a_nT^n$. Then $\sigma(p(T)) = \{p(\lambda) : \lambda \in \sigma(T)\}$.

Proof. If $n = 0$, then $p(T) = a_0I$ so the result holds since $\sigma(a_0I) = \{a_0\}$. Now let $n \geq 1$ and $a_n \neq 0$. Fix $\mu \in \mathbb{C}$. Write

$$\mu - p(z) = c \prod_{k=1}^n (\lambda_k - z)$$

for some $c, \lambda_1, \dots, \lambda_n \in \mathbb{C}$ with $c \neq 0$. Note that $p(\lambda) = \mu \iff \lambda \in \{\lambda_1, \dots, \lambda_n\}$. So

$$\begin{aligned} \mu \in \sigma(p(T)) &\iff \mu I - p(T) = c \prod_{k=1}^n (\lambda_k I - T) \notin \mathcal{G}(X) \\ &\iff \exists k \ \lambda_k I - T \notin \mathcal{G}(X) && \text{(see text above theorem)} \\ &\iff \exists \lambda \in \sigma(T) \ \mu = p(\lambda) \end{aligned}$$

as required. ■

Corollary 7.8

$$r(T) \leq \inf_{n \in \mathbb{N}} \|T^n\|^{1/n}$$

Proof. If $\lambda \in \sigma(T)$ and $n \in \mathbb{N}$, then $\lambda^n \in \sigma(T^n)$ (by Theorem 7.7) and hence $|\lambda^n| \leq \|T^n\|$, i.e., $|\lambda| \leq \|T^n\|^{1/n}$. Taking sup over all $\lambda \in \sigma(T)$ and then inf over all $n \in \mathbb{N}$ gives the result. ■

Theorem 7.9 Spectral radius formula

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

Proof. Recall that the resolvent function R_T is holomorphic. Note that $\rho(T) \supset \{\lambda \in \mathbb{C} : |\lambda| > r(T)\} \supset \{\lambda \in \mathbb{C} : |\lambda| > \|T\|\}$. Moreover, recall that, for $|\lambda| > \|T\|$, we have $R_T(\lambda) = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}$.

Fix $\varphi \in \mathcal{B}(X)^*$. Then $\varphi \circ R_T : \rho(T) \rightarrow \mathbb{C}$ is holomorphic and, for $|\lambda| > \|T\|$, $(\varphi \circ R_T)(\lambda) = \sum_{n=0}^{\infty} \frac{\varphi(T^n)}{\lambda^{n+1}}$ is a Laurent expansion which must be valid on $\{\lambda \in \mathbb{C} : |\lambda| > r(T)\}$. Then

$$\varphi\left(\frac{T^n}{\lambda^n}\right) \rightarrow 0 \quad \text{for } |\lambda| > r(T)$$

By Theorem 4.2, $(T^n/\lambda^n)_{n=0}^{\infty}$ is bounded, i.e., for $|\lambda| > r(T)$, there exists $M = M(\lambda)$ such that $\|\frac{T^n}{\lambda^n}\| \leq M$ for all n . So $\|T^n\|^{1/n} \leq M^{1/n}|\lambda|$. Hence, $\limsup_{n \rightarrow \infty} \|T^n\|^{1/n} \leq |\lambda|$ and

$$r(T) \leq \inf_{n \in \mathbb{N}} \|T^n\|^{1/n} \leq \liminf_{n \rightarrow \infty} \|T^n\|^{1/n} \leq \limsup_{n \rightarrow \infty} \|T^n\|^{1/n} \leq r(T)$$

as required. ■

Theorem 7.10

Let $T \in \mathcal{K}(X)$ and $\lambda \in \sigma_{\text{ap}}(T) \setminus \{0\}$. Then $\lambda \in \sigma_p(T)$.

Proof. Let (x_n) be an approximate eigenvalue, i.e., $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $(\lambda I - T)x_n \rightarrow 0$. Since $T \in \mathcal{K}(X)$, there exists $k_1 < k_2 < k_3 < \dots$ such that (Tx_n) converges to some $y \in X$. Then

$$x_{k_n} = \frac{1}{\lambda}(\lambda I - T)x_{k_n} + \frac{1}{\lambda}Tx_{k_n} \rightarrow \frac{1}{\lambda}y$$

Since $\|\cdot\|$ is continuous, $\|y\| = |\lambda|$. In particular, $y \neq 0$. Finally, $Tx_{k_n} \rightarrow T(\frac{1}{\lambda}y) = \frac{1}{\lambda}Ty$ (by continuity of T) and so $z = \frac{1}{\lambda}Ty$. \blacksquare

Remark. It follows that $\sigma(T) = \{0\}$ or $\sigma_p(T) \neq \emptyset$. Indeed, if $\sigma(T) \neq \{0\}$, then there exists $\lambda \in \partial\sigma(T) \setminus \{0\}$, so $\lambda \in \sigma_{\text{ap}}(T)$ and $\lambda \neq 0$.

Definition Invariant subspace

Let $T \in \mathcal{B}(X)$ and $Y \subset X$ a subspace. We say that T acts on Y or that Y is an invariant subspace of T if $T(Y) \subset Y$.

Example

- $\{0\}$ and X are invariant subspaces for any $T \in \mathcal{B}(X)$.
- If $\lambda \in \sigma_{\text{ap}}(T)$, then the eigenspace $E(\lambda) = E_T(\lambda) = \ker(\lambda I - T)$ of λ is a closed invariant subspace: $\forall x \in E(\lambda) \quad Tx = \lambda x$.
- For the right-shift operator R on ℓ_2 , $\sigma_p(R) = \emptyset$ but $Y = \overline{\text{span}}\{e_n : n \geq 2\}$ is a nontrivial invariant subspace.

Does every $T \in \mathcal{B}(X)$ have a nontrivial closed invariant subspace? This is known as the *invariant subspace problem*. As we've seen, if $T \in \mathcal{K}(X)$ and $\sigma(T) = \{0\}$, then yes! In fact, it exists even for every $T \in \mathcal{K}(X)$. However, the answer in general is 'no'.⁶ Nevertheless, the problem remains open for Hilbert spaces.

7.2 Operators on complex Hilbert spaces

Now, let us study the spectral properties of operators on a non-empty complex Hilbert space \mathcal{H} .

Proposition 7.11

For $T \in \mathcal{B}(\mathcal{H})$, $\sigma(T^*) = \{\bar{\lambda} : \lambda \in \sigma(T)\}$.

Proof. Observe that $T \in \mathcal{G}(\mathcal{H}) \iff T^* \in \mathcal{G}(\mathcal{H})$:

(\Rightarrow) Let $S = T^{-1}$. Then $ST = TS = I$, so $T^*S^* = S^*T^* = I$. So $T^* \in \mathcal{G}(\mathcal{H})$ with $(T^*)^{-1} = S^* = (T^{-1})^*$.

(\Leftarrow) If $T^* \in \mathcal{G}(\mathcal{H})$, then by above, $T^{**} = T \in \mathcal{G}(\mathcal{H})$.

Now, for $\lambda \in \mathbb{C}$, $\lambda I - T \in \mathcal{G}(\mathcal{H}) \iff \bar{\lambda}I - T^* = (\lambda I - T)^* \in \mathcal{G}(\mathcal{H})$. Hence, $\lambda \in \sigma(T) \iff \bar{\lambda} \in \sigma(T^*)$. \blacksquare

Remark. This lets us compute the spectrum of the right-shift operator $R \in \mathcal{B}(\ell_2)$ much more quickly. However, our more direct proof works more generally for ℓ_p .

Theorem 7.12

If $T \in \mathcal{B}(\mathcal{H})$ is hermitian, then $\sigma(T) \subset \mathbb{R}$.

⁶If you're interested, click [here](#) for a solution by Charles Read in 1984.

Proof. For $x \in \mathcal{H}$, $\langle Tx, x \rangle = \langle x, Tx \rangle = \overline{\langle Tx, x \rangle}$, so $\langle Tx, x \rangle \in \mathbb{R}$. Fix $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Set $S = \lambda I - T$. Let $\delta = |\operatorname{Im} \lambda|$, which is strictly positive. For $x \in \mathcal{H}$ with $\|x\| = 1$, we have $\delta \leq |\lambda - \langle Tx, x \rangle| = |\lambda\langle x, x \rangle - \langle Tx, x \rangle| = |\langle Sx, x \rangle| \leq \|Sx\|$. By homogeneity, $\delta\|x\| \leq \|Sx\|$ for every $x \in \mathcal{H}$. In particular, S is injective. Let $Y = \operatorname{im} S$. Let us show that Y is closed. Indeed, suppose (y_n) is a sequence in Y with $y_n \rightarrow y \in \mathcal{H}$. Pick (x_n) such that $y_n = Sx_n$. Then $\delta\|x_m - x_n\| \leq \|y_m - y_n\| \rightarrow 0$ as $m, n \rightarrow \infty$. So (x_n) is Cauchy. Let $x = \lim_{n \rightarrow \infty} x_n$. Then $Sx = \lim_{n \rightarrow \infty} Sx_n = y \in Y$. As Y is closed, $\mathcal{H} = Y \oplus Y^\perp$. If $z \in Y^\perp$ has $\|z\| = 1$, then $\delta \leq |\langle Sz, z \rangle| = 0$ — a contradiction! Hence, $Y^\perp = \{0\}$ and S is surjective. ■

Remarks

- The above result implies that, if λ is an eigenvalue of T , then $\lambda \in \mathbb{R}$. But there is a much simpler direct proof: if $Tx = \lambda x$ for some $x \neq 0$, then

$$\lambda\langle x, x \rangle = \langle Tx, x \rangle = \langle x, Tx \rangle = \bar{\lambda}\langle x, x \rangle$$

- If T is compact hermitian, then $\sigma(T) \subset \mathbb{R}$ has an alternative proof: using theorems 7.6 and 7.10,

$$\partial\sigma(T) \subset \sigma_{\operatorname{ap}}(T) \subset \sigma_p(T) \cup \{0\} \subset \mathbb{R}$$

Since S is compact, we conclude that $\sigma(T) \subset \mathbb{R}$. Indeed, suppose, on the contrary, that there exists some $y \in \sigma(T) \setminus \mathbb{R}$. Since S is compact, $\lambda = \max\{t \in \mathbb{R} : ty \in \sigma(T)\}$ is well-defined and $\lambda y \in \partial\sigma(T)$ but $\lambda y \notin \mathbb{R}$.

Corollary 7.13

If $T \in \mathcal{B}(\mathcal{H})$ is hermitian, then $\sigma(T) = \sigma_{\operatorname{ap}}(T)$.

Proof. Since $\sigma(T) \subset \mathbb{R}$, $\sigma(T) = \partial\sigma(T) \subset \sigma_{\operatorname{ap}}(T) \subset \sigma(T)$. ■

Corollary 7.14

Let $T \in \mathcal{B}(\mathcal{H})$ be hermitian and $\lambda, \mu \in \mathbb{C}$ be distinct eigenvalues of T . For any $x \in E(\lambda)$ and $y \in E(\mu)$, we have $x \perp y$.

Proof. $\lambda\langle x, y \rangle = \langle Tx, y \rangle = \langle x, Ty \rangle = \mu\langle x, y \rangle$. Since $\lambda \neq \mu$, $\langle x, y \rangle = 0$. ■

Theorem 7.15

Let $T \in \mathcal{B}(\mathcal{H})$ be hermitian. Then $r(T) = \|T\|$.

Proof (using spectral radius formula). Note that $\|T^2\| = \|T^*T\| = \|T\|^2$. By induction, $\|T^{2^n}\| = \|T\|^{2^n}$ for all $n \in \mathbb{N}$. Then

$$r(T) = \lim_{n \rightarrow \infty} \|T^n\|^{1/n} = \lim_{k \rightarrow \infty} \|T^{2^k}\|^{1/2^k} = \|T\|$$

as required. ■

Proof (more elementary). By scaling, we may assume $\|T\| = 1$ WLOG. Then $r(T) \leq 1$. Let us show that either $+1$ or -1 is in $\sigma_{\operatorname{ap}}(T)$. (In fact, at least one of them *must* be in $\sigma_{\operatorname{ap}}(T)$.)

Choose a sequence (x_n) in \mathcal{H} such that $\|x_n\| = 1$ for each $n \in \mathbb{N}$ and $\|Tx_n\| \rightarrow \|T\| = 1$. Observe that

$$\begin{aligned} \|(I - T^2)x_n\| &= \langle x_n - T^2x_n, x_n - T^2x_n \rangle \\ &= 1 - 2\|Tx_n\|^2 + \|T^2x_n\| \\ &\leq 2 - 2\|Tx_n\|^2 \end{aligned}$$

$$\rightarrow 0 \quad \text{as } n \rightarrow \infty$$

so $(I - T)(I + T)x_n \rightarrow 0$ as $n \rightarrow \infty$. We now consider two cases:

- $(I + T)x_n \rightarrow 0$ In this case, -1 is an approximate eigenvalue.
- $(I + T)x_n \not\rightarrow 0$ In this case, we can pick $k_1 < k_2 < k_3 < \dots$ in \mathbb{N} and $\varepsilon > 0$ such that $\|(I + T)x_n\| \geq \varepsilon$ for every $n \in \mathbb{N}$. Set $y_n = \frac{(I+T)x_n}{\|(I+T)x_n\|}$. Note that $\|y_n\| = 1$ for all $n \in \mathbb{N}$ and $(I - T)y_n \rightarrow 0$. Thus, $+1$ is an approximate eigenvalue. ■

Theorem 7.16

Let $T \in \mathcal{B}(\mathcal{H})$ be compact hermitian. Then

- (i) If $\lambda \in \sigma(T) \setminus \{0\}$, then λ is an eigenvalue of T and $\dim E(\lambda) < \infty$.
- (ii) For every $\delta > 0$, $\{\lambda \in \sigma(T) : |\lambda| > \delta\}$ is finite.
- (iii) $\sigma(T)$ is uncountable and $\sigma(T) \setminus \{0\}$ is a sequence $\lambda_1, \lambda_2, \lambda_3, \dots$ that is finite or tends to zero and that satisfies $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots$

Proof. (i): $\lambda \in \sigma(T) = \sigma_{\text{ap}}(T)$ (by Corollary 7.13). Since λ is a non-zero approximate eigenvalue and T is compact, we deduce by Theorem 7.10 that λ is an eigenvalue of T . Next, suppose, on the contrary, that $\dim E(\lambda) = \infty$. Then we can pick an orthonormal sequence $(e_n)_{n=1}^{\infty}$ of eigenvectors for λ . Note that, for every $m \neq n$, we have $\|Te_m - Te_n\| = |\lambda| \|e_m - e_n\| = \sqrt{2}|\lambda| > 0$, so (Te_n) has no convergent subsequence — a contradiction!

(ii): Suppose not. Then for some $\delta > 0$, we can pick a sequence $(\lambda_n)_{n \in \mathbb{N}}$ of pairwise distinct eigenvalues such that $|\lambda_n| \geq \delta$ for all $n \in \mathbb{N}$. Now, for each $n \in \mathbb{N}$, choose x_n such that $Tx_n = \lambda_n x_n$ and $\|x_n\| = 1$. Note that (x_n) is pairwise orthogonal by Corollary 7.14. Thus, for every $m \neq n$,

$$\|Tx_m - Tx_n\| = \|\lambda_m x_m - \lambda_n x_n\| = \sqrt{\|\lambda_m x_m\|^2 + \|\lambda_n x_n\|^2} \geq \delta\sqrt{2}.$$

(iii): Immediate from (ii). ■

Lemma 7.17

Let $T \in \mathcal{B}(\mathcal{H})$ and $Y \subset \mathcal{H}$ a closed subspace. If T acts on Y , then T^* acts on Y^\perp .

Proof. Fix $z \in Y^\perp$. Note that, for every $y \in Y$, we have $\langle T^*z, y \rangle = \langle z, Ty \rangle = 0$ since $Ty \in Y$. Hence, $T^*z \in Y^\perp$. ■

Remark. If T hermitian, then T acts on $Y \implies T^*$ also acts on Y . Moreover, both $T|_Y$ and $T|_{Y^\perp}$ are hermitian since $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in \mathcal{H}$.

Example. $T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ acts on $Y = \text{span}\{e_1\}$ but does not act on $Y^\perp = \text{span}\{e_2\}$.

Theorem 7.18 Spectral theorem for compact hermitian operators

Let $T \in \mathcal{B}(\mathcal{H})$ be compact hermitian. Let $Y = (\ker T)^\perp$. Then there exists an orthonormal basis e_1, e_2, \dots of Y and $\lambda_1, \lambda_2, \dots$ in \mathbb{R} that is finite or tends to zero and that satisfies $|\lambda_1| \geq |\lambda_2| \geq \dots$ and

$$T \left(\sum_{n \geq 1} x_n e_n + z \right) = \sum_{n \geq 1} \lambda_n x_n e_n \quad \forall z \in \ker T$$

Proof. By Theorem 7.16, $\sigma(T) \setminus \{0\}$ is a sequence of eigenvalues μ_1, μ_2, \dots that is finite or tends to zero and satisfies $|\mu_1| \geq |\mu_2| \geq \dots$. Pick an orthonormal basis of each $E(\mu_n)$ (finite by

Theorem 7.16) and put these together to get an orthonormal sequence e_1, e_2, \dots (Corollary 7.13) of eigenvectors. Set $Y = \overline{\text{span}}\{e_n : n \geq 1\}$. Then

$$T \left(\sum_{n \geq 1} x_n e_n \right) = \sum_{n \geq 1} \lambda_n x_n e_n$$

where each e_n is μ_m for some m and (λ_n) is a sequence that is finite or tends to zero and that satisfies $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$. Note that $Y \cap \ker T = \{0\}$, so $\ker T \subset Y^\perp$. If $Y^\perp = \{0\}$, then we are done. Otherwise, the operator $S = T|_{Y^\perp}$ has spectrum $\sigma(S) = \{0\}$ since, by Theorem 7.17, we have that S is a compact hermitian operator on Y^\perp with no nonzero eigenvalue. It follows that $\|S\| = r(S) = 0$ and hence $Y^\perp = \ker T$, which completes the proof. ■

Remarks

- In some references, the spectral theorem reads $T = \sum_{n \geq 1} \mu_n P_n$ where P_n is the orthogonal projection onto $E(\mu_n)$.
- A similar result holds for a compact normal operator T — simply consider T^*T .

Application

We say that $T \in \mathcal{B}(\mathcal{H})$ is *positive* if $\langle Tx, x \rangle \geq 0$ for all $x \in \mathcal{H}$. Note that, by the polarisation identity, T is hermitian.

- If T is compact positive, then there exists a unique positive operator S such that $S^2 = T$. We then write $T^{1/2} := S$.
- For a compact normal operator T , the operator $|T| := (TT^*)^{1/2}$ is well-defined and there exists a unitary operator U such that $T = |T|U$.