Peize Liu

St. Peter's College
University of Oxford

Part B Revision Notes

Functional Analysis

May, 2021

Notes on Functional Analysis

Peize Liu
St Peter's College, University of Oxford

Examinable Syllabus

B4.1 Functional Analysis I

- Brief recall of material from Part A Metric Spaces and Part A Linear Algebra on real and complex normed vector spaces, their geometry and topology and simple examples of completeness. The norm associated with an inner product and its properties. Banach spaces, exemplified by ℓ^p, L^p, $\mathrm{C}(K)$, spaces of differentiable functions. Finite-dimensional normed spaces, including equivalence of norms and completeness. Hilbert spaces as a class of Banach spaces having special properties (illustrations, but no proofs); examples (Euclidean spaces, ℓ^2, L^2).
- Density. Approximation of functions, Stone-Weierstrass Theorem. Separable spaces; separability of subspaces.
- Bounded linear operators, examples (including integral operators). Continuous linear functionals. Dual spaces. Hahn-Banach Theorem (proof for separable spaces only); applications, including density of subspaces and embedding of a normed space into its second dual. Adjoint operators.
- Spectrum and resolvent. Spectral mapping theorem for polynomials.

B4.2 Functional Analysis II

- Hilbert spaces; examples including L^2-spaces. Orthogonality, orthogonal complement, closed subspaces, projection theorem. Riesz Representation Theorem.
- Linear operators on Hilbert space, adjoint operators. Self-adjoint operators, orthogonal projections, unitary operators.
- Baire Category Theorem and its consequences for operators on Banach spaces (Uniform Boundedness, Open Mapping, Inverse Mapping and Closed Graph Theorems). Strong convergence of sequences of operators.
- Weak convergence. Weak precompactness of the unit ball.
- Spectral theory in Hilbert spaces, in particular spectra of self-adjoint and unitary operators.
- Orthonormal sets, Pythagoras, Bessel's inequality. Complete orthonormal sets, Parseval. L^2-theory of Fourier series, including completeness of the trigonometric system. Examples of other orthogonal expansions (Legendre, Laguerre, Hermite etc.).
- Brief contextual comments on the classical theory of Fourier series and modes of convergence; exposition of failure of pointwise convergence of Fourier series of some continuous functions.

Chapter 1

Banach Spaces

In this chapter we discuss the basic concepts of normed vector spaces and bounded linear operators, and investigate how completeness becomes a central idea in the study of linear functional analysis.

In this chapter, the scalar field $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$.

1.1 Definitions and Basic Properties

Definition 1.1. Normed Vector Spaces

Suppose that X is a vector space over $\mathbb{F}$. A norm $\|\cdot\| : X \to \mathbb{R}$ is a map such that for $x, y \in X$, $\lambda \in \mathbb{F}$:

- (Positivity) $\|x\| \geqslant 0$;
- (Definiteness) $\|x\| = 0 \iff x = 0$;
- (Homogeneity) $\|\lambda x\| = |\lambda| \, \|x\|$;
- (Triangular Inequality) $\|x + y\| \leqslant \|x\| + \|y\|$.

The pair $(X, \|\cdot\|)$ is called a normed vector space.

Remark. If $\|\cdot\| : X \to \mathbb{R}$ satisfies everything above expect for definiteness, then it is called a **seminorm**. It induces a proper norm on the quotient space X/X_0 by $\|x + X_0\| := \|x\|$, where $X_0 = \{x \in X : \ \|x\| = 0\}$.

The norm $\|\cdot\|$ induces a metric on X by $d(x, y) := \|x - y\|$. Therefore we have the notions and properties of metric spaces (and topological spaces) on a normed vector space.

Definition 1.2. Banach Spaces

A complete normed vector space is called a Banach space.

Remark. From A2 Metric Spaces, we know that every complete subspace of a Banach space is closed.

Suppose that $\|\cdot\|$ and $\|\cdot\|'$ are two norms on X. They are called equivalent if there exists $C_1, C_2 > 0$ such that for all $x \in X$

$$\|x\| \leqslant C_1 \|x\|', \qquad \|x\|' \leqslant C_2 \|x\|$$

If two norms are equivalent, then they induces the same topology.

Proposition 1.3. Banach Spaces and Absolute Convergence

$(X, \|\cdot\|)$ is a Banach space if and only if every absolutely convergent sequence in X is convergent.

Proof. $\Longrightarrow$ If $\sum_{n=1}^{\infty} \|x_n\| < \infty$ then s_n is a Cauchy sequence in $(X, \|\cdot\|)$ since for $m > n \geqslant N$

$$\|s_n - s_m\| = \left\| \sum_{k=n+1}^{m} x_k \right\| \leqslant \sum_{k=n+1}^{m} \|x_k\| \leqslant \sum_{k=N+1}^{\infty} \|x_k\| \to 0 \text{ as } N \to \infty$$

As $(X, \|\cdot\|)$ is complete we thus obtain that s_n converges to some element $s \in X$.

$\Longleftarrow$ Let (x_n) be a Cauchy sequence. Select a subsequence x_{n_j} so that

$$\left\|x_{n_j} - x_{n_{j+1}}\right\| \leqslant 2^{-j},$$

where the existence of such a subsequence is ensured by the fact that x_n is a Cauchy sequence. Then

$$\sum_{j=1}^{\infty} \left\|x_{n_{j+1}} - x_{n_j}\right\| \leqslant 1 < \infty$$

so by assumption $\sum_{j=1}^{\infty} \left(x_{n_{j+1}} - x_{n_j}\right)$ converges. Hence $x_{n_k} = x_{n_1} + \sum_{j=1}^{k-1} \left(x_{n_{j+1}} - x_{n_j}\right)$ converges, so (x_n) has a convergent subsequence and must thus itself converge. □

1.2 Examples of Banach Spaces

Example 1.4. $(\mathbb{F}^n, \|\cdot\|_p)$

Consider the finite dimensional vector space $\mathbb{F}^n$ equipped with the p-norm:

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} & 1 \leqslant p < \infty \\ \max\limits_{1\leqslant i\leqslant n} |x_i| & p = \infty \end{cases}$$

$(\mathbb{F}^n, \|\cdot\|_p)$ is a Banach space for $1 \leqslant p \leqslant \infty$. This is a consequence of Heine-Borel Theorem.

The p-norms on $\mathbb{F}^n$ are equivalent in the sense that

$$\|x\|_\infty \leqslant \|x\|_p \leqslant n^{1/p} \|x\|_\infty$$

In Corollary 1.24 we shall show that all norms on a finite-dimensional vector space are equivalent.

Example 1.5. Sequence Space $(\ell^p, \|\cdot\|_p)$

The infinite-dimensional analogue of $(F^n, \|\cdot\|_p)$ is the sequence space $(\ell^p, \|\cdot\|_p)$.

$$\ell^p := \begin{cases} \left\{(x_i)_{i\in\mathbb{N}} : \sum_{i\in\mathbb{N}} |x_i|^p < \infty\right\} & 1 \leqslant p < \infty \\ \left\{(x_i)_{i\in\mathbb{N}} : \sup\limits_{i\in\mathbb{N}} |x_i| < \infty\right\} & p = \infty \end{cases}$$

and

$$\|(x_i)\|_p := \begin{cases} \left(\sum_{i\in\mathbb{N}} |x_i|^p\right)^{1/p} & 1 \leqslant p < \infty \\ \sup\limits_{i\in\mathbb{N}} |x_i| & p = \infty \end{cases}$$

Proposition 1.6. Completeness of ℓ^p

$(\ell^p, \|\cdot\|_p)$ is a Banach space for $1 \leqslant p \leqslant \infty$.

Proof.
- $1 \leqslant p < \infty$:

  Suppose that $x^{(n)} = (x_i^{(n)})_{i\in\mathbb{N}}$, $(x^{(n)})_{n\in\mathbb{N}} \subseteq \ell^p$ is a Cauchy sequence. Then for each $i \in \mathbb{N}$,

  $$\left|x_i^{(n)} - x_i^{(m)}\right| \leqslant \left\|x^{(n)} - x^{(m)}\right\|_p^p \to 0$$

  as $n, m \to \infty$. Hence $(x_i^{(n)})_{n\in\mathbb{N}}$ is (pointwise) Cauchy for each $i \in \mathbb{N}$. $x_i^{(n)} \to x_i$ as $n \to \infty$ for some $x_i \in F$. Let $x := (x_i)_{i\in\mathbb{N}}$. We claim that $x \in \ell^p$ and $x^{(n)} \to x$ in ℓ^p as $n \to \infty$.

  Fix $\varepsilon \in (0, 1)$. There exists $N \in \mathbb{N}$ such that for all $m, n > N$, $\left\|x^{(m)} - x^{(n)}\right\| < \varepsilon$. For each $K \in \mathbb{N}$ and $n > N$,

  $$\sum_{i=0}^{K} |x_i^{(n)} - x_i|^p = \lim_{m\to\infty} \sum_{i=0}^{K} |x_i^{(n)} - x_i^{(m)}|^p \leqslant \varepsilon^p$$

  Taking $K \to \infty$, we obtain that $\left\|x^{(n)} - x\right\|_p^p \leqslant \varepsilon^p$. Hence $x^{(n)} \to x$ in ℓ^p. And

  $$\|x\|_p \leqslant \left\|x^{(n)} - x\right\|_p + \left\|x^{(n)}\right\| < \infty$$

  Hence $x \in \ell^p$.

- $p = \infty$: The idea is essentially the same. □

Proposition 1.7. Inclusion Relation for ℓ^p

Suppose that $1 \leqslant p < q \leqslant \infty$. Then $\ell^p \subsetneq \ell^q$.

Proof. For $1 \leqslant p < q < \infty$, note that $(x_i) \in \ell^p$ implies trivially that (x_i) is bounded, which means that $(x_i) \in \ell^\infty$. Furthermore, we must have $|x_i| < 1$ for all but finitely many $i \in \mathbb{N}$. And $|x_i|^q < |x_i|^p$ for $q > p$. Then $(x_i) \in \ell^q$.

The opposite non-inclusions are trivial. □

An important subspace of ℓ^∞ is the space of all sequences converging to 0:

$$c_0 := \left\{(x_i) \in \ell^\infty : \lim_{i\to\infty} x_i = 0\right\}$$

It is a closed subspace of ℓ^∞ and hence a Banach space.

Example 1.8. Function Space $(\mathrm{L}^p, \|\cdot\|_p)$

Let $\Omega \subseteq \mathbb{R}^n$ be a measurable set. [a] The function spaces $\mathcal{L}^p(\Omega)$ are defined by

$$\mathcal{L}^p := \begin{cases} \left\{f : \Omega \to \mathbb{R} \,\middle|\, f \text{ measurable and } \int_\Omega |f|^p < \infty\right\} & 1 \leqslant p < \infty \\ \{f : \Omega \to \mathbb{R} \mid f \text{ measurable and } \exists\, M > 0\ |f| \leqslant M \text{ a.e.}\} & p = \infty \end{cases}$$

and the (semi) p-norms are given by

$$\|f\|_p := \begin{cases} \left(\int_\Omega |f|^p\right)^{1/p} & 1 \leqslant p < \infty \\ \operatorname{ess\,sup} |f| := \inf\{M : |f| \leqslant M \text{ a.e.}\} & p = \infty \end{cases}$$

$\|\cdot\|_p$ is a seminorm on $\mathcal{L}^p(\Omega)$, as $\|f - g\|_p = 0$ if and only if $f = g$ almost everywhere. $\mathrm{L}^p(\Omega)$ is the quotient space of $\mathcal{L}^p$ on which $\|\cdot\|_p$ becomes a proper norm.

[a] All discussions about L^p spaces generalise naturally to the abstract $\mathrm{L}^p(\Omega, \mu)$ with any σ-finite measure μ on the measurable space Ω. So the sequence spaces ℓ^p are also special cases of L^p spaces.

Next we collect some important results from *A3 Integration*.

Theorem 1.9. Hölder's Inequality

Let $1 \leqslant p, q \leqslant \infty$ such that $p^{-1} + q^{-1} = 1$. Such numbers p, q are called a pair of **Hölder conjugates**.

1. For $x, y \in \mathbb{F}^n$, $\sum_{i=1}^{n} |x_i y_i| \leqslant \|x\|_p \|y\|_q = \left(\sum_{i\in\mathbb{N}} |x_i|^p\right)^{1/p} \left(\sum_{i\in\mathbb{N}} |y_i|^q\right)^{1/q}$
2. For $(x_n) \in \ell^p$, $(y_n) \in \ell^q$, $\sum_{n\in\mathbb{N}} |x_n y_n| \leqslant \|x\|_p \|y\|_q = \left(\sum_{n\in\mathbb{N}} |x_n|^p\right)^{1/p} \left(\sum_{n\in\mathbb{N}} |y_n|^q\right)^{1/q}$
3. For $f \in \mathrm{L}^p(\Omega)$, $g \in \mathrm{L}^q(\Omega)$, $\int_\Omega |fg| = \|fg\|_1 \leqslant \|f\|_p \|g\|_q = \left(\int_\Omega |f|^p\right)^{1/p} \left(\int_\Omega |g|^q\right)^{1/q}$

Proof. See *A3 Integration*. The key step is to use the concavity of the function $t \mapsto \log t$ and Jensen's inequality. □

Proposition 1.10. Inclusion Relation for $\mathrm{L}^p(\Omega)$

Suppose that $\Omega \subseteq \mathbb{R}^n$ has finite measure. Let $1 \leqslant p < q \leqslant \infty$. Then $\mathrm{L}^p(\Omega) \supsetneq \mathrm{L}^q(\Omega)$.

Proof. Let m be the Lebesgue measure on $\mathbb{R}^n$. So $m(\Omega) < \infty$. The inclusions follow from the following observation. For $1 \leqslant p < q < \infty$,

$$f \in \mathrm{L}^\infty(\Omega) \implies \|f\|_p = \left(\int_\Omega |f|^p\right)^{1/p} \leqslant m(\Omega)^{1/p} \|f\|_\infty < \infty \implies f \in \mathrm{L}^p(\Omega)$$

$$\begin{aligned} f \in \mathrm{L}^q(\Omega) \implies \|f\|_p &= \left(\int_\Omega |f|^p \cdot 1\right)^{1/p} \\ &\leqslant \left(\left(\int_\Omega (|f|^p)^{q/p}\right)^{p/q} \left(\int_\Omega 1^{(1-p/q)^{-1}}\right)^{1-p/q}\right)^{1/p} && \text{(Hölder's Inequality)} \\ &= m(\Omega)^{\frac{1}{p}-\frac{1}{q}} \|f\|_q \implies f \in \mathrm{L}^p(\Omega) \end{aligned}$$

The opposite non-inclusions are trivial. □

Remark. The proposition is not true if Ω has infinite measure!

Proposition 1.11. Completeness of $\mathrm{L}^p(\Omega)$

$(\mathrm{L}^p(\Omega), \|\cdot\|_p)$ is a Banach space for $1 \leqslant p \leqslant \infty$.

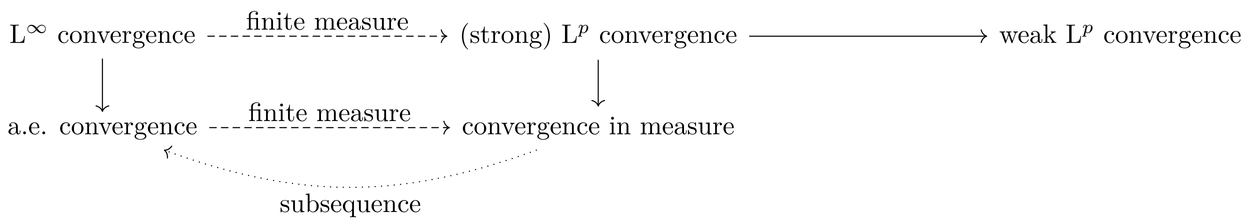
Proof. See *A3 Integration*. □

The convergence in L^p and almost everywhere convergence do not imply each other. However we have the following proposition:

Corollary 1.12. L^p and Almost Everywhere Convergence

Suppose that $(f_n) \subseteq \mathrm{L}^p(\Omega)$ such that $\int_\Omega |f_n - f|^p \to 0$ as $n \to \infty$. Then there exists a subsequence (f_{n_k}) of (f_n) such that $f_{n_k} \to f$ almost everywhere as $k \to \infty$.

Remark. In fact we have the following more general result on the modes of convergence of measurable functions. Here $1 \leqslant p < \infty$.



Example 1.13. Continuous Bounded Functions $(\mathrm{C}_b(\Omega), \|\cdot\|_\infty)$

Consider the space of continuous bounded function on $\Omega \subseteq \mathbb{R}^n$ equipped with the supremum norm: $(\mathrm{C}_b(\Omega), \|\cdot\|_\infty)$. It is a Banach space by the Cauchy criterion of uniform continuity and the fact that the uniform limit of a sequence of continuous functions is continuous.

Example 1.14. Sobolev Space $\mathrm{W}^{1,p}(a,b)$

The Sobolev space

$$\mathrm{W}^{1,p}(a,b) := \{f \in \mathrm{L}^p(a,b) : \; f' \in \mathrm{L}^p(a,b)\}$$

is normed by

$$\|f\|_{\mathrm{W}^{1,p}(a,b)} := \begin{cases} \left(\displaystyle\int_a^b (|f|^p + |f'|^p)\right)^{1/p} & 1 \leqslant p < \infty \\ \max\{\|f\|_\infty, \|f'\|_\infty\} & p = \infty \end{cases}$$

where f' is the **distributional derivative** of f. See B4.3 Distribution Theory for definition. $\mathrm{W}^{1,p}(a,b)$ is a Banach space for $1 \leqslant p \leqslant \infty$.

1.3 Bounded Linear Maps

Definition 1.15. Bounded Linear Maps

Suppose that $T : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is a linear map between normed vector spaces. T is said to be bounded if there exists $M \in \mathbb{R}$ such that for all $x \in X$,

$$\|T(x)\|_Y \leqslant M\|x\|_X$$

The set of bounded linear maps from X to Y is denoted by $\mathscr{B}(X,Y)$.

Definition 1.16. Operator Norm

Suppose that $T : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is a bounded linear map. We define the operator norm

$$\|T\|_{\mathscr{B}(X,Y)} := \sup_{x \in X \setminus \{0\}} \frac{\|T(x)\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|T(x)\|_Y = \sup_{\|x\|_X \leqslant 1} \|T(x)\|_Y$$

Remark. For all $x \in X$, $\|T(x)\|_Y \leqslant \|T\|_{\mathscr{B}(X,Y)} \|x\|_X$.

Proposition 1.17. Continuity and Boundedness

Suppose that $T : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is a linear map. The following are equivalent:

1. T is bounded;
2. T is Lipschitz continuous;
3. T is continuous;
4. T is continuous at 0.

Proof. $1 \implies 2 \implies 3 \implies 4$ is trivial.

"$4 \implies 1$": Suppose that $T \notin \mathscr{B}(X,Y)$. There exists a sequence $(x_n) \subseteq X$ such that $\|T(x_n)\|_Y > n\|x_n\|_X$ for all $n \in \mathbb{N}$. Let $y_n := \dfrac{x_n}{\|T(x_n)\|_Y}$ so that $\|y_n\|_X < \dfrac{1}{n} \to 0$ as $n \to \infty$. By continuity of T at 0 we have $\|T(y_n)\|_Y \to 0$ as $n \to \infty$. However,

$$\|T(y_n)\|_Y = \frac{\|T(x_n)\|_Y}{\|T(x_n)\|_Y} = 1$$

which is a contradiction. □

General procedure of proving $T \in \mathscr{B}(X,Y)$:

- Check that T is well-defined, or $T(x) \in Y$ for all $x \in X$;
- Check that T is linear.
- Find $M > 0$ such that $\|T(x)\|_Y \leqslant M\|x\|_X$ for all $x \in X$.
- Additionally, if we want to prove that $\|T\| = M$, we need to find a sequence $(x_n) \in X$ such that $\dfrac{\|T(x_n)\|_Y}{\|x\|_X} \to M$ as $n \to \infty$.

Proposition 1.18. Completeness of Spaces of Bounded Linear Maps

Suppose that $(X, \|\cdot\|_X)$ is a normed vector space and $(Y, \|\cdot\|_Y)$ is a Banach space. Then the space of bounded linear maps $\left(\mathscr{B}(X,Y), \|\cdot\|_{\mathscr{B}(X,Y)}\right)$ is a Banach space.

Proof. Let (T_n) be a Cauchy sequence in $\mathscr{B}(X,Y)$. Then for every $x \in X$ we have

$$\|T_n(x) - T_m(x)\| \leqslant \|T_n - T_m\| \, \|x\| \to 0$$

as $m, n \to \infty$. So $(T_n(x))$ is a Cauchy sequence in Y and, as Y is complete, thus converges to some element in Y which we call $T(x)$.

We now show that the resulting map $x \mapsto T(x)$ is an element of $\mathscr{B}(X,Y)$ and $T_n \to T$ in $\mathscr{B}(X,Y)$, i.e. $\|T - T_n\| \to 0$. We first note that the linearity of T_n implies that also T is linear. Given any $\varepsilon > 0$ we now let N be so that for $m, n \geqslant N$ we have $\|T_n - T_m\| \leqslant \varepsilon$. Given any $x \in X$ we thus have

$$\|T(x) - T_n(x)\| = \left\| \lim_{m\to\infty} T_m(x) - T_n(x) \right\| = \lim_{m\to\infty} \|T_m(x) - T_n(x)\| \leqslant \varepsilon\|x\|$$

Hence T is bounded (as$\|Tx\| \leqslant (\|T_n\| + \varepsilon)\,\|x\|$ for all x) and so an element of $\mathscr{B}(X,Y)$ with $\|T - T_n\| \leqslant \varepsilon$ for all $n \geqslant N$. We obtain that $T_n \to T$ in $\mathscr{B}(X,Y)$. □

Proposition 1.19. Composition of Bounded Linear Maps

Suppose that $T \in \mathscr{B}(X,Y)$ and $S \in \mathscr{B}(Y,Z)$. Then $S \circ T \in \mathscr{B}(X,Z)$, and

$$\|S \circ T\|_{\mathscr{B}(X,Z)} \leqslant \|S\|_{\mathscr{B}(Y,Z)} \|T\|_{\mathscr{B}(X,Y)}$$

Proof. Trivial. □

A bounded linear operator $T \in \mathscr{B}(X)$ is said to be **invertible in** $\mathscr{B}(X)$, if there exists $S \in \mathscr{B}(X)$ such that $T \circ S = S \circ T = \mathrm{id}_X$.

$T \in \mathscr{B}(X)$ is invertible in $\mathscr{B}(X)$, if it is bijective and

$$\exists\, \delta > 0 \;\forall\, x \in X: \;\|T(x)\|_Y \geqslant \delta\|x\|_X$$

Proposition 1.20. Inverse Mapping Theorem

Suppose that $(X, \|\cdot\|_X)$ is a Banach space. Then every bijective $T \in \mathscr{B}(X)$ is invertible in $\mathscr{B}(X)$.

Proof. See the Inverse Mapping Theorem 4.5 after the Open Mapping Theorem. □

Lemma 1.21. Convergence of Neumann Series

Suppose that X is a Banach space. $T \in \mathscr{B}(X)$ such that $\|T\| \leqslant 1$. Then $\mathrm{id} - T$ is invertible in $\mathscr{B}(X)$ with

$$(\mathrm{id} - T)^{-1} = \sum_{j=0}^{\infty} T^j \in \mathscr{B}(X)$$

Proof. Trivial. □

Corollary 1.22

Suppose that $(X, \|\cdot\|_X)$ is a Banach space. $T \in \mathscr{B}(X)$ is invertible. Then for any $S \in \mathscr{B}(X)$ such that $\|S\| < \left\|T^{-1}\right\|^{-1}$, $T - S$ is invertible.

1.4 Finite-Dimensional Normed Vector Spaces

Now we shall study the most trivial cases of Banach spaces. They are the finite-dimensional normed vector spaces. Here are a collection of the important results.

In a finite-dimensional normed vector space X:

- all norms are equivalent;
- all linear operators are bounded;
- X is complete;
- all bounded closed subspaces are compact;
- the unit sphere is compact.

Proposition 1.23. Equivalence of $\|\cdot\|$ and $\|\cdot\|_2$

Suppose that $\|\cdot\|$ is a norm on $\mathbb{R}^n$. Then $\|\cdot\|$ is equivalent to $\|\cdot\|_2$.

Proof. Let $\{e_1, ..., e_n\}$ be a basis of $\mathbb{R}^n$. For $x = \sum_{i=1}^{n} x_i e_i \in \mathbb{R}^n$, by Cauchy-Schwarz Inequality,

$$\|x\| \leqslant \sum_{i=1}^{n} |x_i| \|e_i\| \leqslant \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^{n} \|e_i\|^2 \right)^{1/2} = \|x\|_2 \left(\sum_{i=1}^{n} \|e_i\|^2 \right)^{1/2}$$

On the other hand, suppose for contradiction that there exists a sequence $(x^{(n)})_{n\in\mathbb{N}}$ such that $\left\|x^{(n)}\right\|_2 \geqslant n\left\|x^{(n)}\right\|$. Let $y^{(n)} := x^{(n)}/\left\|x^{(n)}\right\|_2$. Then $(y^{(n)})_{n\in\mathbb{N}}$ is a sequence in the unit sphere S, which is sequentially compact by Bolzano-Weierstrass Theorem. There exists a subsequence $(y^{(n_k)})_{k\in\mathbb{N}}$ such that $y^{(n_k)} \to y \in S$ in the 2-norm as $k \to \infty$. But

$$\|y\| \leqslant \left\|y - y^{(n_k)}\right\| + \left\|y^{(n_k)}\right\| \leqslant C\left\|y - y^{(n_k)}\right\|_2 + \frac{1}{n_k} \to 0 \qquad \left(C := \left(\sum_{i=1}^{n} \|e_i\|^2 \right)^{1/2} \right)$$

as $k \to \infty$, which is contradictory. Hence there exists $C' > 0$ such that $\|x\|_2 \leqslant C'\|x\|$. We deduce that $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent. □

Corollary 1.24

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

Proposition 1.25

Suppose that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are normed vector spaces. If X is finite-dimensional, then all linear maps $T : X \to Y$ are bounded.

Proof. We define the **graph norm** on X: For $x \in X$,

$$\|x\|_T := \|x\|_X + \|T(x)\|_Y$$

It is easy to check that $\|\cdot\|_T$ is a norm on X. Since it is equivalent to $\|\cdot\|_X$, there exists $C > 0$ such that

$$\|T(x)\|_Y \leqslant \|x\|_T \leqslant C\|x\|_X$$

Hence $T \in \mathscr{B}(X, Y)$. □

Corollary 1.26

Every finite-dimensional normed vector space is homeomorphic to $\mathbb{F}^n$.

Proof. Let X be a finite-dimensional normed vector space. Then there exists a linear isomorphism $T : X \to \mathbb{F}^n$. By the previous theorem T and T^{-1} are continuous. Hence X is homeomorphic to $\mathbb{F}^n$. □

Corollary 1.27

Every finite-dimensional vector space is a Banach space.

Corollary 1.28

Every finite-dimensional subspace of a normed vector space is complete and hence closed.

Lemma 1.29. Riesz Lemma

Let $(X, \|\cdot\|)$ be a normed vector space, and $Y \subsetneq X$ a closed subspace. Then for each $\varepsilon \in (0, 1)$ there exists $x \in S$ (S is the unit sphere) such that

$$\operatorname{dist}(x, Y) := \inf\{\|x - y\| : \ y \in Y\} \leqslant 1 - \varepsilon$$

Proof. As $Y \neq X$ is closed we know that the set $X \setminus Y$ is open and non-empty, so we can choose some $x^* \in X \setminus Y$ and use that $d := \operatorname{dist}(x^*, Y) > 0$, as $X \setminus Y$ must contain some ball $B(x, \delta)$ which ensures that $d \geqslant \delta > 0$.

By the definition of the infimum, we can now select $y^* \in Y$ so that $d \leqslant \|x^* - y^*\| < \dfrac{d}{1-\varepsilon}$ and claim that $x := \dfrac{x^* - y^*}{\|x^* - y^*\|}$ has the desired properties. Clearly $\|x\| = 1$, i.e. $x \in S$ as desired, and we furthermore have that

$$\begin{aligned}\operatorname{dist}(x, Y) &= \inf_{y \in Y} \|x - y\| = \inf_{y \in Y} \left\| \frac{x^*}{\|x^* - y^*\|} - \frac{y^*}{\|x^* - y^*\|} - y \right\| = \inf_{\widetilde{y} \in Y} \left\| \frac{x^*}{\|x^* - y^*\|} - \widetilde{y} \right\| \\ &= \inf_{\widehat{y} \in Y} \left\| \frac{x^* - \widehat{y}}{\|x^* - y^*\|} \right\| = \frac{\operatorname{dist}(x^*, Y)}{\|x^* - y^*\|} \geqslant 1 - \varepsilon\end{aligned}$$

where we used twice that Y is a subspace, to replace the infimum over $y \in Y$ first by an infimum over $\widetilde{y} = \dfrac{y^*}{\|x^* - y^*\|} + y$ and then an infimum over $\widehat{y}$ which is related to $\widetilde{y}$ by $\widetilde{y} = \dfrac{\widehat{y}}{\|x^* - y^*\|}$. □

Proposition 1.30. Equivalent Characterisations of Finite-Dimensional Normed Vector Spaces

Let $(X, \|\cdot\|)$ be a normed vector space. The following are equivalent:

1. $\dim X < \infty$;
2. Every bounded closed subset of X is compact;
3. The unit sphere $S \subseteq X$ is compact.

Proof.

$1 \implies 2$: Suppose that $Y \subseteq X$ is bounded and closed. Let $T : X \to \mathbb{F}^n$ be a linear isomorphism. T and T^{-1} are continuous. Then $T(Y) \subseteq \mathbb{F}$ is bounded and closed, and hence compact by Heine-Borel Theorem. Therefore $Y = T^{-1}(T(Y))$ is also compact.

$2 \implies 3$: This is clear since S is bounded and closed.

$3 \implies 1$: We argue by contradiction and assume that S is compact but $\dim(X) = \infty$. We may choose a sequence of linearly independent elements $y_k \in X$, $k \in \mathbb{N}$. Then the subspace $Y_k := \operatorname{span}\{y_1, \dots, y_k\} \subsetneq Y_{k+1}$ is finite dimensional, and so a closed proper subspace of Y_{k+1}. Applying Riesz Lemma with $\varepsilon = 1/2$ (viewing Y_k as a subspace of Y_{k+1} instead of X), it gives us a sequence of elements $y_k \in Y_{k+1} \cap S$ with $\operatorname{dist}(y_k, Y_k) \geqslant 1/2$. In particular for every $k > \ell$ we have

$$\|y_k - y_\ell\| \geqslant \operatorname{dist}(y_k, Y_{\ell+1}) \geqslant \operatorname{dist}(y_k, Y_k) \geqslant \frac{1}{2}$$

Therefore no subsequence of (y_k) can be a Cauchy sequence. Having constructed a sequence (y_k) in $S \subseteq X$ that does not contain a convergent subsequence, we conclude that S is not sequentially compact and hence not compact, leading to a contradiction. □

1.5 Dual Spaces and Dual Operators

Definition 1.31. Dual Spaces

Let $(X, \|\cdot\|)$ be a normed vector space. The dual space $X^* := \mathscr{B}(X, \mathbb{F})$ is a normed vector space equipped with the operator norm $\|\cdot\|_*$, given by

$$\|\ell\|_* := \sup_{\|x\|=1} |\ell(x)|$$

The notion of dual space in functional analysis is differential from that in linear algebra. Here the dual space X^* is the set of **bounded linear functionals**, contrast to the set of all *linear functionals*, X', which is now called the *algebraic dual space* (of X).

Remark. As a consequence of Proposition 1.18, the dual space X^* is complete regardless of the completeness of X.

Lemma 1.32. Kernels and Linear Functionals

Suppose that $(X, \|\cdot\|)$ is a normed vector space and $\ell : X \to \mathbb{F}$ is a linear functional. Then $\ell \in X^*$ if and only if $\ker \ell$ is closed in X.

Proof. If $\ell \in X^*$, then ℓ is continuous. Hence $\ker \ell = \ell^{-1}(\{0\})$ is closed. Conversely, suppose that $\ker \ell$ is closed. Without loss of generality, $\ell \neq 0$. Let $x_0 \in X$ such that $\ell(x_0) = 1$. Since $\ker \ell$ is closed, $\delta := \operatorname{dist}(x_0, \ker \ell) > 0$. For $x \in X \setminus \ker \ell$, define $y := -\dfrac{x - x_0 \ell(x)}{\ell(x)} \in X$. Then

$$\ell(y) = -\frac{\ell(x) - \ell(x)\ell(x_0)}{\ell(x)} = 0$$

So $y \in \ker \ell$. We have

$$\delta \leqslant \|x_0 - y\| = \left\| x_0 + \frac{x - x_0 \ell(x)}{\ell(x)} \right\| = \frac{\|x\|}{|\ell(x)|}$$

Hence $|\ell(x)| < \delta^{-1}\|x\|$ for all $x \in X$. We deduce that $\ell \in X^*$. □

The dual operator is the same as in linear algebra.

Definition 1.33. Dual Operator

Suppose that $T : X \to Y$ is a linear map between the normed vector spaces X and Y. The dual operator or pull-back of T is $T' : Y' \to X'$, given by $T'(\ell) = \ell \circ T$ for all $\ell \in Y'$.

In case of bounded operator we have the important proposition. The result requires a proposition which is the consequence of Hahn-Banach Theorem in Section 5.1.

Proposition 1.34. Dual of Bounded Operators

Suppose that X and Y are normed vector spaces. Let $T \in \mathscr{B}(X, Y)$. Then $T' \in \mathscr{B}(Y^*, X^*)$ is well-defined and $\|T'\|_{\mathscr{B}(Y^*,X^*)} = \|T\|_{\mathscr{B}(X,Y)}$.

Proof. We have $T'(\ell) \in X^*$ for $\ell \in Y^*$, because

$$\|T'(\ell)\|_{X^*} = \|\ell \circ T\|_{X^*} \leqslant \|T\|_{\mathscr{B}(X,Y)}\|f\|_{Y^*}$$

Hence $T' : Y^* \to X^*$ is well-defined and $\|T'\|_{\mathscr{B}(Y^*,X^*)} \leqslant \|T\|_{\mathscr{B}(X,Y)}$. Conversely, fix $x \in X$. By Proposition 5.4 there exists $\ell \in Y^*$ such that $\|\ell\|_{Y^*} = 1$ and $\ell \circ T(x) = \|T(x)\|$. Hence

$$\|T(x)\| = |\ell \circ T(x)| = |T'(\ell)(x)| \leqslant \|T'\|_{\mathscr{B}(Y^*,X^*)}\|\ell\|_{Y^*}\|x\|_X = \|T'\|_{\mathscr{B}(Y^*,X^*)}\|x\|_X$$

Hence $\|T'\|_{\mathscr{B}(Y^*,X^*)} \geqslant \|T\|_{\mathscr{B}(X,Y)}$. □

Proposition 1.35. Kernels and Images of the Dual Operator

Suppose that X and Y are normed vector spaces. Let $T \in \mathscr{B}(X, Y)$.

1. $\ker T = (\operatorname{im} T')_\circ$, $\ker T' = (\operatorname{im} T)^\circ$;
2. $\overline{\operatorname{im} T} = (\ker T')_\circ$.

(See Definition 5.11 for the annihilators.)

Proof. 1.

$$\begin{aligned} x \in \ker T &\iff T(x) = 0 \\ &\iff \forall\, \ell \in Y^* : \ T'(\ell)(x) = \ell \circ T(x) = 0 \\ &\iff \forall\, f \in \operatorname{im} T' : \ f(x) = 0 \\ &\iff x \in (\operatorname{im} T')_\circ \end{aligned}$$

$$\begin{aligned} f \in \ker T' &\iff T'(f) = f \circ T = 0 \\ &\iff \forall\, x \in X : \ f \circ T(x) = 0 \\ &\iff \forall\, y \in \operatorname{im} T : \ f(y) = 0 \\ &\iff f \in (\operatorname{im} T)^\circ \end{aligned}$$

Note that the second " $\Longleftarrow$ " on the left hand side requires Proposition 5.4.

2. By Corollary 5.13.3, we have $\overline{\operatorname{im} T} = ((\operatorname{im} T)^\circ)_\circ = (\ker T')_\circ$. □

Remark. It is not true in general that $\overline{\operatorname{im} T'} = (\ker T)^\circ$. In fact $(\ker T)^\circ$ is the weak* closure of $\operatorname{im} T'$.

Chapter 2

Hilbert Spaces

We introduce the concept of Hilbert spaces, which generalises the inner product spaces to infinite dimensions.

2.1 Orthogonality

Definition 2.1. Inner Product Spaces

Suppose that X is a vector space over $\mathbb{F}$, An inner product $\langle \cdot, \cdot \rangle : X \to \mathbb{F}$ is a map such that for $x, y, z \in X$, $\lambda \in \mathbb{F}$:

- (Additivity in the Second Slot) $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$;
- (Homogeneity in the Second Sq4lot) $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle$;
- (Symmetry / Conjugate Symmetry) $\langle x, y \rangle = \overline{\langle y, x \rangle}$;
- (Positivity) $\langle x, x \rangle \geqslant 0$;
- (Definiteness) $\langle x, x \rangle = 0 \iff x = 0$.

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

Remark. If $\mathbb{F} = \mathbb{R}$, then $\langle \cdot, \cdot \rangle$ is called **bilinear**; if $\mathbb{F} = \mathbb{C}$, then $\langle \cdot, \cdot \rangle$ is called **sesquilinear**.

Remark. Every inner product on X induces a norm via $\|x\| := \sqrt{\langle x, x \rangle}$.

We collect some results from *A0. Linear Algebra* which holds for general inner product spaces:

- **Cauchy-Schwarz Inequality**: For $x, y \in X$,

$$|\langle x, y \rangle| \leqslant \|x\| \|y\|$$

  with the equality holds if x and y are linearly dependent.

- **Orthogonality**: We say that $x, y \in X$ are orthogonal, if $\langle x, y \rangle = 0$.
- **Pythagorean Theorem**: For $x, y \in X$,

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

  More generally, suppose that $\{x_1, ..., x_n\} \subseteq X$ is an orthonormal set. Then for $x \in X$,

$$\|x\|^2 = \sum_{m=1}^{n} |\langle x_m, x \rangle|^2 + \left\| x - \sum_{m=1}^{n} \langle x_m, x \rangle x_m \right\|^2$$

- **Parallelogram Identity**: For $x, y \in X$,

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

- **Polarisation Identities**: For $x, y \in X$, if $\mathbb{F} = \mathbb{R}$, then

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

if $\mathbb{F} = \mathbb{C}$, then

$$\langle x, y\rangle = \frac{1}{4}\sum_{n=0}^{3} \mathrm{i}^n \|x + \mathrm{i}^n y\|^2$$

- Let $(X, \|\cdot\|)$ be a normed vector space. The norm $\|\cdot\|$ induces an inner product $\langle\cdot,\cdot\rangle$ via the polarisation identity if and only if the parallelogram identity holds in $(X, \|\cdot\|)$.

Lemma 2.2

Suppose that X is an inner product space and $K \subseteq X$ is a convex subset. Let $x \in X$ and $y \in K$. The following are equivalent:

1. $\|x - y\| = \operatorname{dist}(x, K)$;
2. $\operatorname{Re}\langle x - y, z - y\rangle \leqslant 0$ for all $z \in K$.

Proof.

- $\implies$: Fix $z \in K$. Since K is convex, $tz + (1-t)y \in K$ for $t \in [0,1]$. We have

$$\begin{aligned}
&\|x-y\| = \operatorname{dist}(x,K) \leqslant \|x - (tz + (1-t)y)\| \\
\implies &\|x-y\|^2 \leqslant \|(x-y) + t(y-z)\|^2 = \|x-y\|^2 + t^2\|y-z\|^2 - 2t\operatorname{Re}\langle x-y, z-y\rangle \\
\implies &t^2\|y-z\|^2 - 2t\operatorname{Re}\langle x-y, z-y\rangle \geqslant 0
\end{aligned}$$

Let $f(t) := t^2\|y-z\|^2 - 2t\operatorname{Re}\langle x-y, z-y\rangle$. Then $f(t) \geqslant 0$ for $t \in [0,1]$. Note that $f(0) = 0$. Hence we must have

$$f'(0) = -2\operatorname{Re}\langle x-y, z-y\rangle \geqslant 0$$

which implies that $\operatorname{Re}\langle x-y, z-y\rangle \leqslant 0$ as required.

- $\impliedby$: For all $z \in K$,

$$\|x-z\|^2 = \|(x-y)+(y-z)\|^2 = \|x-y\|^2 + \|y-z\|^2 - 2\operatorname{Re}\langle x-y, z-y\rangle \geqslant \|x-y\|^2$$

because $\|y-z\| \geqslant 0$ and $\operatorname{Re}\langle x-y, y-z\rangle \leqslant 0$. Hence $\|x-y\| \leqslant \inf_{z\in K}\|x-z\| = \operatorname{dist}(x,K)$. Since $y \in K$, we have $\|x-y\| = \operatorname{dist}(x,K)$. □

Definition 2.3. Hilbert Spaces

A complete inner product space is called a Hilbert space.

It is clear that all Hilbert spaces are Banach spaces.

Example 2.4. Examples of Hilbert Spaces

We consider the scalar field $\mathbb{F} = \mathbb{C}$.

1. The finite-dimensional vector space $\mathbb{C}^n$ is a Hilbert space with inner product

$$\langle x, y\rangle = \sum_{i=1}^{n} \overline{x_n} y_n$$

2. The sequence space ℓ^2 is a Hilbert space with inner product

$$\langle x, y\rangle = \sum_{n=0}^{\infty} \overline{x_n} y_n$$

3. The function space $L^2(\Omega)$ is a Hilbert space with inner product

$$\langle f, g\rangle = \int_\Omega \overline{f}\cdot g$$

4. The Sobolev space $W^{1,2}(a,b)$ is a Hilbert space with inner product

$$\langle f, g\rangle = \int_a^b \left(\overline{f}\cdot g + \overline{f'}\cdot g'\right)$$

Definition 2.5. Orthogonal Complement

Suppose that X is an inner product space. Let $Y \subseteq X$. Then the orthogonal complement of Y is defined by

$$Y^\perp := \{x \in X : \ \forall\, y \in Y \ \ \langle x, y\rangle = 0\} \subseteq X$$

Proposition 2.6. Properties of Orthogonal Complement

Let X be an inner product space and $Y \subseteq X$.

1. $Y \cap Y^\perp = \{0\}$;
2. $Y^\perp$ is a closed subspace of X;
3. $Y \subseteq Y^{\perp\perp}$;
4. If $Z \subseteq Y$, then $Y^\perp \subseteq Z^\perp$;
5. $(\overline{\operatorname{span} Y})^\perp = Y^\perp$;
6. If $Y, Z \leqslant X$ such that $X = Y + Z$ and $Z \subseteq Y^\perp$, then $Y^\perp = Z$.

Proof.

2. It is straightforward that $Y^\perp$ is a subspace of X. Let $(z_n) \subseteq Y^\perp$ such that $z_n \to z \in X$. By Cauchy-Schwarz Inequality, we have for any $y \in Y$

$$|\langle z_n, y\rangle - \langle z, y\rangle| = |\langle z_n - z, y\rangle| \leqslant \|y\|\|z_n - z\| \to 0$$

as $n \to \infty$. Hence $\langle z, y\rangle = 0$ and $z \in Y^\perp$.

6. Let $c \in Y^\perp$. We write $c = y + z$ for $y \in Y$ and $z \in Z$. Then $y = c - z \in Y^\perp$. But $Y \cap Y^\perp = \{0\}$, we have $c - z = 0$ and hence $c = z \in Z$. □

Proposition 2.7. Closest Point in a Closed Convex Subset, Hilbert Space

Let X be a Hilbert space and $K \subseteq X$ a non-empty closed convex set. Then for all $x \in X$ there exists a unique $y \in K$ such that $\|x - y\| = \operatorname{dist}(x, K)$.

Proof.

- Existence: Let $d : K \to X$ given by $d(z) := \|x - z\|$. For simplicity, write $d_K := \operatorname{dist}(x, K)$. Then there exists a sequence $(y_n) \subseteq K$ such that $d(y_n) \to d_K$ as $n \to \infty$. Applying the Parallelogram Identity to $\frac{1}{2}(x - y_n)$ and $\frac{1}{2}(x - y_m)$,

$$d\left(\frac{1}{2}(y_n + y_m)\right)^2 + \frac{1}{4}\|y_n - y_m\|^2 = \frac{1}{2}\left(d(y_n)^2 + d(y_m)^2\right)$$

Since K is convex, $\frac{1}{2}(y_n + y_m) \in K$ and hence $d_K \leqslant d\left(\frac{1}{2}(y_n + y_m)\right)$. Taking $m, n \to \infty$ we have

$$d_K^2 + \frac{1}{4}\|y_n - y_m\|^2 \leqslant \frac{1}{2}\left(d(y_n)^2 + d(y_m)^2\right) \to d_K^2$$

Hence $\|y_n - y_m\| \to 0$ and (y_n) is a Cauchy sequence. Let $y = \lim_{n\to\infty} y_n$. Since K is closed, by the continuity of the norm we have

$$d(y) = \lim_{n\to\infty} d(y_n) = d_K$$

- Uniqueness: Suppose that $y' \in K$ such that $d(y') = d_K$. Applying the Parallelogram Identity to $\frac{1}{2}(x-y)$ and $\frac{1}{2}(x-y')$,

$$d_K^2 + \frac{1}{4}\|y-y'\|^2 \leqslant d\left(\frac{1}{2}(y+y')\right)^2 + \frac{1}{4}\|y-y'\|^2 = \frac{1}{2}\left(d(y)^2 + d(y')^2\right) = d^2$$

This implies that $\|y-y'\| = 0$ and hence $y = y'$. □

Remark. We shall show that the result can be generalised to uniformly convex Banach spaces in Corollary 5.41.

Theorem 2.8. Projection Theorem

Suppose that X is a Hilbert space and $Y \subseteq X$ is a closed subspace. Then $X = Y \oplus Y^\perp$.

Proof. It suffices to show that $X = Y + Y^\perp$. Fix $x \in X$. By Proposition 2.7 there exists $y_0 \in Y$ such that $\|x - y_0\| = \operatorname{dist}(x, Y)$. For all $y \in Y$ and $t \in \mathbb{R}$, we have

$$\|x-y_0\|^2 \leqslant \|x-(y_0-ty)\|^2 = \|x-y_0\|^2 + 2t\operatorname{Re}\langle x-y_0, y\rangle + t^2\|y\|^2$$

Hence $2t\operatorname{Re}\langle x-y_0, y\rangle + t^2\|y\|^2 \geqslant 0$ for all $t \in \mathbb{R}$. Therefore $\operatorname{Re}\langle x-y_0, y\rangle = 0$. By replacing y with $\mathrm{i}y \in Y$ we obtain $\operatorname{Im}\langle x-y_0, y\rangle = 0$. Hence $\langle x-y_0, y\rangle = 0$ for all $y \in Y$. We conclude that $x - y_0 \in Y^\perp$ and $x = y_0 + (x-y_0) \in Y + Y^\perp$. □

Corollary 2.9

Suppose that X is a Hilbert space and $Y \subseteq X$. Then $X = \overline{\operatorname{span} Y} \oplus Y^\perp$ and $Y^{\perp\perp} = \overline{\operatorname{span} Y}$.

Lemma 2.10. Bessel's Inequality

Suppose that X is an inner product space. Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be an orthonormal sequence. Then for $x \in X$,

$$\sum_{n=0}^{\infty} |\langle x_n, x\rangle|^2 \leqslant \|x\|^2$$

Proof. By Pythagorean Theorem we have

$$\sum_{n=0}^{m} |\langle x_n, x\rangle|^2 \leqslant \sum_{n=0}^{m} |\langle x_n, x\rangle|^2 + \left\| x - \sum_{n=0}^{m} \langle x_n, x\rangle x_n \right\|^2 = \|x\|^2$$

The result follows as we take $m \to \infty$. □

Theorem 2.11. Parseval's Identity

Suppose that X is a Hilbert space. Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be an orthonormal sequence. Then

$$\overline{\operatorname{span}\{x_n\}_{n\in\mathbb{N}}} = \left\{ x = \sum_{n=0}^{\infty} a_n x_n : \ (a_n)_{n\in\mathbb{N}} \in \ell^2 \right\}$$

Furthermore, for $x \in \overline{\operatorname{span}\{x_n\}_{n\in\mathbb{N}}}$, we have $a_n = \langle x_n, x\rangle$ and

$$\|x\|^2 = \sum_{n=0}^{\infty} |\langle x_n, x\rangle|^2$$

Proof. Let $Y := \overline{\operatorname{span}\{x_n\}_{n\in\mathbb{N}}}$ and $Z := \left\{\sum_{n=0}^{\infty} a_n x_n :\ (a_n)_{n\in\mathbb{N}} \in \ell^2\right\}$.

- For $(a_n) \in \ell^2$, by Pythagorean Theorem,

$$\left\|\sum_{k=m}^{n} a_k x_k\right\|^2 = \sum_{k=m}^{n} |a_k|^2 \to 0$$

as $m, n \to \infty$. Hence $(\sum_{k=0}^{n} a_k x_k)_{n\in\mathbb{N}}$ is a Cauchy sequence. The sum $\sum_{n=0}^{\infty} a_n x_n$ converges in norm and hence $Z \subseteq Y$.

- Let $x \in Y$. Put $a_n := \langle x_n, x\rangle$. By Bessel's Inequality, $(a_n) \in \ell^2$. Let $x' = \sum_{n=0}^{\infty} a_n x_n \in Z$. Now

$$\langle x_n, x'\rangle = \left\langle x_n, \sum_{m=0}^{\infty} a_m x_m \right\rangle = \sum_{m=0}^{\infty} a_m \langle x_n, x_m\rangle = a_n = \langle x_n, x\rangle$$

Hence $x - x' \in Y^\perp$. But also $x - x' \in Y$. We deduce that $x = x' \in Z$ and hence $Y \subseteq Z$.

In summary, $Y = Z$. Finally, by the Pythagorean Theorem,

$$\|x\|^2 = \sum_{m=0}^{n} |\langle x_m, x\rangle|^2 + \left\|x - \sum_{m=0}^{n} \langle x_m, x\rangle\right\|^2 \to \sum_{m=0}^{\infty} |\langle x_m, x\rangle|^2$$

as $n \to \infty$. □

Definition 2.12. Orthonormal Basis

Suppose that X is a Hilbert space. A subset $S \subseteq X$ is called an orthonormal basis of X, if $\langle x, y\rangle = \delta_{xy}$ for all $x, y \in S$ and $\overline{\operatorname{span} S} = X$.

Remark. An orthonormal basis of X is not a Hamel basis (i.e. the usual definition of basis for a vector space) because we are taking closure of $\operatorname{span} S$.

Theorem 2.13. Existence of Orthonormal Basis

Every inner product space contains an orthonormal basis.

Proof. Let X be a inner product space. Let $\mathscr{S}$ be the set of all orthonormal subsets of X, partially ordered by set inclusion. By Zorn's Lemma, $\mathscr{S}$ has a maximal element, which is an orthonormal basis of X. □

Remark. It is clear that a Hilbert space has a countable orthonormal basis if and only if it is separable. The backward direction follows from applying the **Gram-Schmidt orthonormalisation process** to a dense countable subset of the Hilbert space.

2.2 Adjoints

Theorem 2.14. Riesz-Fréchet Representation Theorem

Suppose that X is a Hilbert space. For each $\ell \in X^*$, there exists a unique $x_\ell \in X$ such that $\|x_\ell\| = \|\ell\|_*$ and $\ell(x) = \langle x_\ell, x\rangle$ for all $x \in X$.

Proof. Without loss of generality assume that $\ell \neq 0$. Let $Y := \ker \ell$. Since ℓ is bounded, Y is closed by Lemma 1.32. By the Projection Theorem, $X = Y \oplus Y^\perp$. As $Y^\perp \neq \{0\}$, we can take $y^\perp \in Y^\perp \setminus \{0\}$ and assume that $\|y^\perp\| = 1$. Note that $\ell(y^\perp) \neq 0$, For any $x \in X$, let

$$w := x - \frac{\ell(x)}{\ell(y^\perp)} y^\perp \in \ker \ell$$

We have

$$0 = \langle y^\perp, w \rangle = \langle y^\perp, x \rangle - \frac{\ell(x)}{\ell(y^\perp)} \implies \ell(x) = \left\langle \overline{\ell(y^\perp)} y^\perp, x \right\rangle$$

Hence we can take $x_\ell := \overline{\ell(y^\perp)} y^\perp$.

The uniqueness is clear: if $\ell(x) = \langle x_\ell, x \rangle = \langle x'_\ell, x \rangle$ for all $x \in X$, then $\langle x_\ell - x'_\ell, x \rangle = 0$ for all $x \in X$, which implies that $x_\ell = x'_\ell$.

It remains to prove that $\|x_\ell\| = \|\ell\|_*$. On one hand, by Cauchy-Schwarz Inequality,

$$|\ell(x)| = |\langle x_\ell, x \rangle| \leqslant \|x_\ell\| \|x\|$$

which gives $\|\ell\|_* \leqslant \|x_\ell\|$. On the other hand,

$$\|x_\ell\|^2 = \langle x_\ell, x_\ell \rangle = |\ell(x_\ell)| \leqslant \|\ell\|_* \|x_\ell\|$$

which gives $\|x_\ell\| \leqslant \|\ell\|_*$ ☐

Remark. In the case where $\mathbb{F} = \mathbb{R}$, the theorem establishes a cannonical isometric isomorphism $\pi : X^* \to X$ by $\pi(x)(y) = \langle x, y \rangle$. Therefore for real Hilbert spaces we have the natural identification $X = X^*$. For complex Hilbert spaces, the map π is an antilinear bijection, because $\pi(\lambda x) = \overline{\lambda} \pi(x)$ for $\lambda \in \mathbb{C}$.

Similar to the dual operators on Banach spaces, we have the notion of adjoint operators on Hilbert spaces:

Definition 2.15. Adjoint Operators

Suppose that $(X, \langle \cdot, \cdot \rangle_X)$ and $(Y, \langle \cdot, \cdot \rangle_Y)$ are Hilbert spaces. Let $T \in \mathscr{B}(X, Y)$. There exists a unique operator $T^* \in \mathscr{B}(Y, X)$ such that $\|T\|_{\mathscr{B}(X,Y)} = \|T^*\|_{\mathscr{B}(Y,X)}$, and $\langle y, Tx \rangle_Y = \langle T^* y, x \rangle_X$ for all $x \in X$ and $y \in Y$. T^* is called the **adjoint** of T.

Proof. For each $y \in Y$, the map $T_y : x \mapsto \langle y, Tx \rangle_Y$ is a bounded linear functional on X by Cauchy-Schwarz Inequality:

$$|T_y(x)| = |\langle y, Tx \rangle_Y| \leqslant \|y\| \|Tx\| \leqslant \|T\| \|x\| \|y\| \implies \|T_y\|_{X^*} \leqslant \|T\| \|y\|_Y$$

Hence by Riesz-Fréchet Representation Theorem there exists a unique $T^* y \in X^*$ such that $\langle y, Tx \rangle_Y = \langle T^* y, x \rangle_X$ and

$$\|T^* y\|_X = \|T_y\|_{X^*} \leqslant \|T\| \|y\|_Y$$

The map $y \mapsto T^* y$ is clearly linear. We have $\|T^*\| \leqslant \|T\|$ and hence $T^* \in \mathscr{B}(Y, X)$.

Conversely, by definition we have $T^{**} = T$. Hence $\|T\| = \|T^{**}\| \leqslant \|T^*\|$. We conclude that $\|T\| = \|T^*\|$. ☐

Remark. The Riesz-Fréchet Representation Theorem gives two isometric antilinear bijections $\pi_X : X^* \to X$ and $\pi_Y : Y^* \to Y$. We see that the adjoint and the dual operators are related via the following commutative diagram:

$$\begin{array}{ccc} Y & \xrightarrow{T^*} & X \\ {\scriptstyle \pi_Y}\uparrow & & \uparrow{\scriptstyle \pi_X} \\ Y^* & \xrightarrow{T'} & X^* \end{array}$$

Proposition 2.16. Properties of Adjoint Operators

Suppose that $(X, \langle \cdot, \cdot \rangle_X)$ and $(Y, \langle \cdot, \cdot \rangle_Y)$ are Hilbert spaces. Let $T \in \mathscr{B}(X, Y)$.

1. $T^{**} = T$;
2. $\|T\|_{\mathscr{B}(X,Y)} = \|T^*\|_{\mathscr{B}(Y,X)}$;
3. Taking adjoint is antilinear: For $T, S \in \mathscr{B}(X, Y)$ and $\lambda, \eta \in \mathbb{C}$, we have $(\lambda T + \eta S)^* = \overline{\lambda} T^* + \overline{\eta} S^*$;
4. For $T \in \mathscr{B}(X, Y)$ and $S \in \mathscr{B}(Y, Z)$, we have $(ST)^* = T^* S^*$;
5. $T \in \mathscr{B}(X)$ is invertible in $\mathscr{B}(X)$ if and only if $T^* \in \mathscr{B}(X)$ in invertible in $\mathscr{B}(X)$.

Proof. 5. Suppose that $T^* \in \mathscr{B}(X)$ is invertible in $\mathscr{B}(X)$. Let $S := ((T^*)^{-1})^*$. We have

$$\langle TSx, y\rangle = \langle x, S^*T^*y\rangle = \left\langle x, (A^*)^{-1}A^*y\right\rangle = \langle x, y\rangle$$

which is true for all $y \in X$. We have $STx = x$ for all $x \in X$ and hence $ST = \mathrm{id}_X$. Hence T is invertible with bounded inverse $T^{-1} = S$. □

Proposition 2.17. Kernels and Images of Adjoint Operators

Suppose that $(X, \langle\cdot,\cdot\rangle_X)$ and $(Y, \langle\cdot,\cdot\rangle_Y)$ are Hilbert spaces. Let $T \in \mathscr{B}(X, Y)$.

1. $\ker T = (\operatorname{im} T^*)^\perp$;
2. $\overline{\operatorname{im} T} = (\ker T^*)^\perp$.

Proof. Similar to Proposition 1.35. □

Definition 2.18. Self-Adjoint Operators

Suppose that X is an inner product space. $T \in \mathscr{B}(X)$ is called a self-adjoint operator, if $T = T^*$.[a]

[a] If T is a continuous unbounded operator on X, the domain $D(T)$ is a proper dense subset on X. T is said to be **symmetric** if $T = T^*$ on $D(T) \cap D(T^*)$; T is said to be **self-adjoint**, if T is symmetric and $D(T) = D(T^*)$. For bounded operators, the two concepts are equivalent.

Proposition 2.19. Polarisation Identities for Self-Adjoint Operators

Let X be a Hilbert space and $T \in \mathscr{B}(X)$ a self-adjoint operator. For $x, y \in X$,

$$\langle y, Tx\rangle = \frac{1}{4}\left(\langle x + y, T(x + y)\rangle - \langle x - y, T(x - y)\rangle\right)$$

Proof. This is a direct computation. □

Remark. In comparison, if $\mathbb{F} = \mathbb{C}$, then every linear operator $T : X \to X$ satisfies the following Polarisation Identity:

$$\langle y, Tx\rangle = \frac{1}{4}\sum_{n=0}^{3} \mathrm{i}^n \left\langle x + \mathrm{i}^n y, T(x + \mathrm{i}^n y)\right\rangle$$

Proposition 2.20. Norm of Operators on Hilbert Spaces

Let X be a Hilbert space.

1. If $T \in \mathscr{B}(X)$, then
$$\|T\|_{\mathscr{B}(X)} = \sup\{|\langle y, Tx\rangle| : \ \|x\| = \|y\| = 1\}$$
2. In particular, if T is self-adjoint, then
$$\|T\|_{\mathscr{B}(X)} = \sup\{|\langle x, Tx\rangle| : \ \|x\| = 1\}$$

Proof. 1. Let $k = \sup\{|\langle y, Tx\rangle| : \ \|x\| = \|y\| = 1\}$. When $\|x\| = \|y\| = 1$, by Cauchy-Schwarz Inequality,

$$|\langle y, Tx\rangle| \leqslant \|Tx\|\|y\| \leqslant \|T\|\|x\|\|y\| = \|T\|$$

Hence $k \leqslant \|T\|$.

For the other direction, we fix $\varepsilon > 0$ and aim to show that $\|T\| < k + \varepsilon$. We pick $x \in X$ such that $\|x\| = 1$

and $\|Tx\| > \|T\| - \varepsilon$. Let $y = \dfrac{Tx}{\|Tx\|}$. Therefore

$$|\langle y, Tx\rangle| = \|y\|\|Tx\| = \|Tx\| > \|T\| - \varepsilon$$

Hence $\|T\| < k + \varepsilon$. Since ε is arbitrary, we deduce that $\|T\| \leqslant k$.

2. Let $K = \sup\{|\langle x, Tx\rangle| : \ \|x\| = 1\}$. By (1) we have $\|T\| = k \geqslant K$.

   For the other direction, by (1) we can select $x, y \in X$ with $\|x\| = \|y\| = 1$ such that $\|T\| - \varepsilon < |\langle y, Tx\rangle|$. Replacing y by λy with a suitable scalar $|\lambda| = 1$, we may assume that

$$|\langle y, Tx\rangle| = \langle y, Tx\rangle = \operatorname{Re}\langle y, Tx\rangle$$

   We have

$$\begin{aligned} 4\operatorname{Re}\langle y, Tx\rangle &= \langle x+y, T(x+y)\rangle - \langle x-y, T(x-y)\rangle && \text{(Polarisation Identity 2.19)} \\ &\leqslant K\|x+y\|^2 + K\|x-y\|^2 \\ &= 2K\left(\|x\|^2 + \|y\|^2\right) && \text{(Parallelogram Identity)} \\ &= 4K \end{aligned}$$

   Putting things together:

$$\|T\| - \varepsilon < |\langle y, Tx\rangle| = \operatorname{Re}\langle y, Tx\rangle \leqslant K$$

   Since ε is arbitrary, we deduce that $\|T\| \leqslant K$. □

Corollary 2.21

Let X be a Hilbert space and $T \in \mathscr{B}(X)$. Then $\|T^*T\| = \|T\|^2$. In particular, if T is self-adjoint, then $\left\|T^2\right\| = \|T\|^2$.

Proof. Note that T^*T is self-adjoint. By the previous proposition we have

$$\|T^*T\| = \sup_{\|x\|=1} |\langle T^*Tx, x\rangle| = \sup_{\|x\|=1} |\langle Tx, Tx\rangle| = \sup_{\|x\|=1} \|Tx\|^2 = \|T\|^2$$ □

Proposition 2.22. Self-Adjointness of Orthogonal Projections

Let X be a Hilbert space. Let $Y, Z \leqslant X$ be closed subspaces such that $X = Y \oplus Z$. Let $P : X \to Y$ be the projection onto Y along Z. Then the following are equivalent:

1. $Z = Y^\perp$;
2. $P^* = P$;
3. $\|P\| \leqslant 1$ (and in such case either $\|P\| = 1$ or $P = 0$).

Proof. The direct sum projection P satisfies $\ker P = Z$ and $\operatorname{im} P = Y$.

$2 \Longrightarrow 1$: By Proposition 2.17, $Z = \ker P = (\operatorname{im} P^*)^\perp = (\operatorname{im} P)^\perp = Y^\perp$.

$1 \Longrightarrow 3$: For $x \in X$, let $x = y + z$, where $y \in Y$ and $z \in Z = Y^\perp$. By Pythagoras' Theorem, $\|x\|^2 = \|y\|^2 + \|z\|^2$. Then we have

$$\|Px\| = \|P(y+z)\| = \|y\| = \sqrt{\|x\|^2 - \|z\|^2} \leqslant \|x\|$$

We deduce that $\|P\| \leqslant 1$.

$3 \Longrightarrow 2$: Let $x \in \operatorname{im} P$ and $u \in \operatorname{im}(\operatorname{id} - P)$. Note that we have $Pu = 0$ and $Px = x$. Then

$$\|x\| = \|P(u - x)\| \leqslant \|P\|\|u - x\| \leqslant \|u - x\|$$

which holds for all $x \in \operatorname{im} P$ and $u \in \operatorname{im}(\operatorname{id} - P)$. By Lemma 2.2, we have $\operatorname{Re}\langle x, u\rangle \leqslant 0$ for all $x \in \operatorname{im} P$ and $u \in \operatorname{im}(\operatorname{id} - P)$. But $u \in \operatorname{im}(\operatorname{id} - P)$ implies that $-u \in \operatorname{im}(\operatorname{id} - P)$. So we also have $\operatorname{Re}\langle x, u\rangle =$

$-\operatorname{Re}\langle x, -u\rangle \geqslant 0$. Hence $\operatorname{Re}\langle x, u\rangle = 0$. If the scalar field is $\mathbb{C}$, we may replace u by iu to obtain that $\operatorname{Im}\langle x, u\rangle = 0$. Hence $\langle x, u\rangle = 0$ for all $x \in \operatorname{im} P$ and $u \in \operatorname{im}(\operatorname{id} - P)$.

Finally, for $x, y \in X$,

$$\langle y, Px\rangle = \langle y - Py, Px\rangle + \langle Py, Px\rangle = \langle Py, Px\rangle = \langle Py, x - Px\rangle + \langle Py, Px\rangle = \langle Py, x\rangle = \langle y, P^* x\rangle$$

Hence $P = P^*$.

Now suppose that all of the above conditions holds. Suppose that $P \neq 0$. Let $x \in \operatorname{im} P$. Then $Px = x$. So

$$\|P\| \geqslant \frac{\|Px\|}{\|x\|} = 1$$

But we also have $\|P\| \leqslant 1$. Hence $\|P\| = 1$. □

2.3 Isometries

Let (X, d_X) and (Y, d_Y) be metric spaces. A map $T : X \to Y$ is called **isometric**, if $d_Y(T(x), T(y)) = d_X(x, y)$ for all $x, y \in X$. In particular an isometry is injective and continuous.

Now we consider X and Y to be normed vector spaces. Let $T : X \to Y$ be a linear map. T is isometric if and only if $\|T(x)\| = \|x\|$ for all $x \in X$.

Now we consider X and Y to be inner product spaces. Let $T : X \to Y$ be a linear map. By polarisation identities we note that T is isometric if and only if $\langle T(x), T(y)\rangle = \langle x, y\rangle$ for all $x, y \in X$.

Proposition 2.23. Properties of Linear Isometries

Suppose that X and Y be Hilbert spaces. Let $T \in \mathscr{B}(X, Y)$. The following are equivalent:

1. T is isometric;
2. $\|T(x)\| = \|x\|$ for all $x \in X$;
3. $\langle T(x), T(y)\rangle = \|x, y\|$ for all $x, y \in X$;
4. $T^*T = \operatorname{id}_X$.

Theorem 2.24. Mazur-Ulam Theorem

Suppose that X and Y are normed vector spaces. If $T : X \to Y$ is a surjective isometry, and $T(0) = 0$, then T is $\mathbb{R}$-linear.

Proof. Fix $x, y \in X$. We shall show that $T\left(\frac{x+y}{2}\right) = \frac{1}{2}(T(x) + T(y))$. Without loss of generality assume that $x \neq y$. Let $z := \frac{x+y}{2}$.

Construct a descending sequence of sets $(H_n(x, y))_{n\in\mathbb{N}}$ as follows: Let

$$H_0(x, y) := \left\{ w \in X : \ \|w - x\| = \|w - y\| = \frac{1}{2}\|x - y\| \right\}$$

and, inductively,

$$H_{n+1}(x, y) := \left\{ p \in H_n : \ \forall\, w \in H_n(x, y) \ \|p - w\| \leqslant \frac{1}{2} \operatorname{diam} H_n(x, y) \right\}$$

We shall prove by induction that each $H_n(x, y)$ is symmetric about z. We claim that $z \in H_n(x, y)$ and that $w \in H_n(x, y)$ implies $2z - w \in H_n(x, y)$. The base case is trivial. For the induction case, suppose that $H_n(x, y)$ has the desired properties.

For $w \in H_n(x,y)$, since $2z - w \in H_n(x,y)$, we have

$$\|z-w\| = \frac{1}{2}\|(2z-w)-w\| \leqslant \frac{1}{2}\operatorname{diam} H_n(x,y) \implies z \in H_{n+1}(x,y)$$

For $p \in H_{n+1}(x,y)$, since $w, 2z-w \in H_n(x,y)$, we have

$$\|(2z-p)-w\| = \|(2z-w)-p\| \leqslant \frac{1}{2}\operatorname{diam} H_n \implies 2z-p \in H_{n+1}(x,y)$$

which completes the induction. Now we have

$$\left\{\frac{x+y}{2}\right\} = \bigcap_{n=0}^{\infty} H_n(x,y) = \lim_{n\to\infty} H_n(x,y)$$

Since T is surjective and isometric, $T(H_n(x,y)) = H_n(T(x),T(y))$. Since T is continuous, it preserves the limits of those sets. We have

$$\left\{\frac{1}{2}(T(x)+T(y))\right\} = \lim_{n\to\infty} H_n(T(x),T(y)) = \lim_{n\to\infty} T(H_n(x,y)) = T\Big(\lim_{n\to\infty} H_n(x,y)\Big) = \left\{T\left(\frac{x+y}{2}\right)\right\}$$

which gives $T\left(\frac{x+y}{2}\right) = \frac{1}{2}\left(T(x)+T(y)\right)$ as claimed.

Fix $x_0 \in X$. Taking $x = x_0$ and $y = -x_0$ we obtain $T(-x_0) = -T(x_0)$; taking $x = 2x_0$ and $y = 0$ we obtain $T(2x_0) = 2T(x_0)$. Inductively we have $T(2^n x_0) = 2^n T(x_0)$ for all $n \in \mathbb{Z}$. For $k \in \mathbb{Z}$, we expand it in base 2 and inductively we have $T(kx_0) = kT(x_0)$. Hence $T(\alpha x_0) = \alpha T(x_0)$ for all $\alpha \in S := \{2^n k : \ n,k \in \mathbb{Z}\}$. But S is dense in $\mathbb{R}$, by continuity of T we obtain that $T(\lambda x_0) = \lambda T(x_0)$ for all $\lambda \in \mathbb{R}$.

Fix $x_0, y_0 \in X$. Taking $x = 2x_0$ and $y = 2y_0$ we obtain

$$T(x_0+y_0) = \frac{1}{2}(T(2x_0)+T(2y_0)) = \frac{1}{2}(2T(x_0)+2T(y_0)) = T(x_0)+T(y_0)$$

We conclude that T is $\mathbb{R}$-linear. □

Remark. Note that the surjectivity condition can be dropped in the context of inner product spaces.

Proposition 2.25. Linearity of Isometry between Inner Product Spaces

Suppose that X and Y are inner product spaces. If $T: X \to Y$ is a isometry, and $T(0) = 0$, then T is $\mathbb{R}$-linear.

Proof. By the previous proof it suffices to show that $T(z) = \frac{1}{2}(T(x)+T(y))$ for $x, y \in X$, where $z := \frac{x+y}{2}$.

$$\|T(x)-T(y)\| = \|x-y\| = \left\|\frac{y-x}{2}\right\| + \left\|\frac{x-y}{2}\right\| = \|z-x\| + \|z-y\| = \|T(z)-T(x)\| + \|T(z)-T(y)\|$$

Squaring both sides and expanding into inner products:

$$\operatorname{Re}\langle T(z)-T(x), T(z)-T(y)\rangle = \|T(z)-T(x)\|\|T(z)-T(y)\|$$

Hence $T(z)-T(x)$ and $T(z)-T(y)$ are linearly dependent. There exists $\lambda \in \mathbb{F}$ such that $T(z)-T(x) = \lambda\left(T(z)-T(y)\right)$. We have

$$\|T(z)-T(x)\| = \|T(z)-T(y)\| \implies |\lambda| = 1$$
$$\|T(x)-T(y)\| = \|T(z)-T(x)\| + \|T(z)-T(y)\| \implies |\lambda+1| = 2$$

Hence $\lambda = 1$. $T(z) = \frac{1}{2}(T(x)+T(y))$ as claimed. □

Definition 2.26. Unitary Operators

Suppose that X and Y are inner product spaces. $T \in \mathscr{B}(X,Y)$ is called a unitary operator, if T is isometric and surjective.

Proposition 2.27. Properties of Unitary Operators

Suppose that X and Y are Hilbert spaces. Let $U \in \mathscr{B}(X,Y)$. U is unitary if any of the following holds:

1. U is surjective and isometric;
2. Both U and U^* are isometric;
3. $U^*U = \mathrm{id}_X$ and $UU^* = \mathrm{id}_Y$;
4. $U^{-1} = U^*$.

2.4 Radon-Nikodym Theorem*

We outline a proof of Radon-Nikodym Theorem by von Neumman as a consequence of Riesz-Fréchet Representation Theorem.

Let $(\Omega, \mathscr{F})$ be a measurable space. A measure μ on Ω is said to be **σ-finite**, if there exists a sequence $(E_n) \subseteq \mathscr{F}$ such that $\mu(E_n) < \infty$ for each $n \in \mathbb{N}$ and $\Omega = \bigcup\{E_n : n \in \mathbb{N}\}$.

Let μ and ν be two measures on Ω. For any $E \in \mathscr{F}$, $\mu(E) = 0$ implies $\nu(E) = 0$. In such case we say that ν is **absolutely continuous** with respect to μ and write $\nu \ll \mu$.

Theorem 2.28. Radon-Nikodym Theorem

Suppose that $(\Omega, \mathscr{F})$ is a meaurable space. Let μ and ν be σ-finite measures on Ω such that $\nu \ll \mu$. Then there exists a $\mathscr{F}$-measurable function $f : \Omega \to [0, \infty)$ such that

$$\nu(E) = \int_E f \, \mathrm{d}\mu$$

for all $E \in \mathscr{F}$. Moreover, f is unique up to a μ-null set. We say that f is the **Radon-Nikodym derivative** and write $f = \dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

Proof.
- The uniqueness of the Radon-Nikodym derivative is obvious. Suppose tht f_1 and f_2 are such functions. Then
$$\int_E (f_1 - f_2) \, \mathrm{d}\mu = 0$$
for all measurable $E \subseteq \Omega$, and hence $f_1 = f_2$ almost everywhere with respect to μ.
- Now we prove existence. We assume that μ and ν are finite measures.

jj
$$\psi(f) = \int_\Omega f \, \mathrm{d}\mu$$
By Cauchy-Schwarz Inequality,
$$|\psi(f)| = \left| \int_\Omega f \, \mathrm{d}\mu \right| \leqslant \left(\int_\Omega |f|^2 \, \mathrm{d}\mu \right) \left(\int_\Omega 1 \, \mathrm{d}\mu \right) \leqslant \left(\int_\Omega |f|^2 \, \mathrm{d}(\mu + \nu) \right) \left(\int_\Omega 1 \, \mathrm{d}\mu \right) = \mu(\Omega)^{1/2} \|f\|_X$$
Hence $\psi \in X^*$. By Riesz-Fréchet Representation Theorem, there exists $g \in X$ such that
$$\int_\Omega f \, \mathrm{d}\mu = \psi(f) = \langle g, f \rangle = \int_\Omega fg \, \mathrm{d}(\mu + \nu)$$

for all $f \in X$. For a meaurable set $E \subseteq \Omega$, we have

$$\mu(E) = \int_\Omega \mathbf{1}_E \, \mathrm{d}\mu = \int_E g \, \mathrm{d}(\mu + \nu)$$

As g is measurable, the set $A := g^{-1}(-\infty, 0]$ is meaurable, and we have

$$0 \leqslant \mu(A) = \int_A g \, \mathrm{d}\mu + \int_A g \, \mathrm{d}\nu \leqslant 0$$

Hence $\mu(a) = 0$. Since $\nu \ll \mu$, we have $(\mu + \nu)(A) = 0$. In particular, $g > 0$ almost everywhere with respect to $(\mu + \nu)$ on Ω.

Now, for each $n \in \mathbb{Z}_+$ and $E \in \mathscr{F}$, define $f_{E,n} : \Omega \to \mathbb{R}$ by

$$f_{E,n}(x) := \frac{\mathbf{1}_E}{g(x) + 1/n}$$

Note that $f_{E,n} > 0$ a.e.. Then

$$\|f_{E,n}\|_X^2 = \int_\Omega f_{E,n}^2 \, \mathrm{d}(\mu + \nu) \leqslant \int_E n^2 \, \mathrm{d}(\mu + \nu) = n^2 (\mu + \nu)(E)$$

Hence $f_{E,n} \in X$. We have

$$\int_E \frac{1}{g(x) + 1/n} \, \mathrm{d}\mu = \int_E \frac{g(x)}{g(x) + 1/n} \, \mathrm{d}\mu + \int_E \frac{g(x)}{g(x) + 1/n} \, \mathrm{d}\nu$$

Taking the limit $n \to \infty$ we obtain

$$\nu(E) = \int_E \left(\frac{1}{g} - 1 \right) \mathrm{d}\mu$$

So we take the Radon-Nikodym derivative $\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu} = \dfrac{1}{g} - 1$.

- Now we extend the result to σ-finite measures μ and ν.

  Let $\{M_n\}_{n \in \mathbb{N}}$ be a sequence of disjoint measurable sets such that $\bigcup_n M_n = \Omega$ and $\mu(M_n) < \infty$. Let $(N_n)_{n \in \mathbb{N}}$ be a sequence of disjoint measurable sets such that $\bigcup_n N_n = \Omega$ and $\nu(N_n) < \infty$. Taking $\Omega_{mn} := M_n \cap N_n$ and relabelling, we obtain a sequence of disjoint measurable sets Ω_n such that $\bigcup_n \Omega_n = \Omega$ and $\mu(\Omega), \nu(\Omega) < \infty$.

  For $E \in \mathscr{F}$, let $\mu_n(E) := \mu(\Omega_n \cap E)$ and $\nu_n(E) := \nu(\Omega_n \cap E)$. Then μ_n, ν_n are finite measures on Ω and $\nu_n \ll \mu_n$. Then there exist non-negative functions $f_n : \Omega \to [0, \infty)$ such that

$$\nu_n(E) = \int_E f_n \, \mathrm{d}\mu_n$$

  where f_n is unique up to a μ_n-null set. But $\mu_n(\Omega \setminus \Omega_n) = 0$ by definition. We may hence replace f_n by $f_n \mathbf{1}_{\Omega_n}$. We put $f = \sum_{n \in \mathbb{N}} f_n \mathbf{1}_{\Omega_n}$. Therefore

$$\nu(E) = \sum_{n \in \mathbb{N}} \nu(\Omega_n \cap E) = \sum_{n \in \mathbb{N}} \nu_n(E) = \sum_{n \in \mathbb{N}} \int_E f_n \, \mathrm{d}\mu_n = \sum_{n \in \mathbb{N}} \int_{\Omega_n \cap E} f_n \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu$$

  The last equality follows from Monotone Convergence Theorem. □

Let $(\Omega, \mathscr{F})$ be a measurable space. We say that $\nu : \mathscr{F} \to \mathbb{R} \cup \{-\infty, +\infty\}$ is a **signed measure**, if

- $\nu(\varnothing) = 0$;
- ν assumes at most one of $-\infty$ and $+\infty$;
- For any countable collection $\{E_n\}_{n \in \mathbb{N}}$ of disjoint measurable sets,

$$\nu\left(\bigcup_{n \in \mathbb{N}} E_n \right) = \sum_{n \in \mathbb{N}} \nu(E_n)$$

where the series on the right hand side converges absolutely if the left hand side is finite.

From Measure Theory, the **Jordan Decomposition Theorem** tells us that any signed measure has a unique decomposition $\nu = \nu^+ - \nu^-$, where ν^+ and ν^- are mutually singular (positive) measures, and one of which is finite. We say that ν is absolutely continuous respect to μ, if $|\nu| := \nu^+ + \nu^-$ is absolutely continuous respect to μ.

Combining the result with Radon-Nikodym Theorem, we have the following corollary, which is utilised in Example 5.24 to determine the dual spaces of L^p.

Corollary 2.29. Radon-Nikodym Theorem for Signed Measures

Suppose that $(\Omega, \mathscr{F})$ is a meaurable space. Let μ be a σ-finite measure and ν be a σ-finite signed measure on Ω such that $|\nu| \ll \mu$. Then there exists a $\mathscr{F}$-measurable function $f : \Omega \to \mathbb{R}$ such that for all $E \in \mathscr{F}$,

$$\nu(E) = \int_E f \, \mathrm{d}\mu$$

Chapter 3

Metric Spaces

In this chapter we collect three important properties of metric spaces and introduce theorems which is essential to linear functional analysis.

3.1 Density: Stone-Weierstrass Theorem

Recall from *A2 Metric Spaces* that a subset $D \subseteq X$ is said to be **dense** in the metric space X, if $\overline{D} = X$. Equivalently, D is dense in X if

$$\forall\, x \in X \;\forall\, \varepsilon > 0 \;\exists\, y \in D : \; d(x, y) < \varepsilon$$

First we have an extremely useful example.

Example 3.1. Approximation of L^p Functions by Test Functions

Let $\Omega \in \mathbb{R}^n$ be an open set, and $1 \leqslant p < \infty$. Then the space of compactly supported smooth functions $\mathrm{C}_c^\infty(\Omega)$ is dense in $\mathrm{L}^p(\Omega)$.

Proof. Non-examinable. See *B4.3 Distribution Theory* for detail. The central idea is to *mollify* the L^p functions by convolving them with the **standard mollifiers**. ☐

Remark. Note that the result does not hold when $p = \infty$ as one can easily see when trying to approximate step functions by continuous functions.

Proposition 3.2. Unique Extension of Bounded Linear Maps

Suppose that $(X, \|\cdot\|_X)$ is a normed vector space and $(Z, \|\cdot\|_Z)$ is a Banach space. Let $Y \subseteq X$ be a dense subspace. Then every bounded linear operator $T \in \mathscr{B}(Y, Z)$ has a unique extension $\widetilde{T} \in \mathscr{B}(X, Z)$ such that $\widetilde{T}\Big|_Y = T$ and

$$\|T\|_{\mathscr{B}(Y,Z)} = \left\|\widetilde{T}\right\|_{\mathscr{B}(X,Z)}$$

Proof. Fix $x \in X$. Since Y is dense in X, there exists a sequence $(y_n) \subseteq Y$ such that $y_n \to x$. Let

$$\widetilde{T}(x) := z := \lim_{n\to\infty} T(y_n)$$

The limit exists because the (Ty_n) is a Cauchy sequence in the Banach space Z:

$$\|T(y_n) - T(y_m)\| \leqslant \|T\|\|y_n - y_m\| \to 0$$

as $n, m \to \infty$. It is not hard to show that the limit is unique. Hence the map $\widetilde{T} : X \to Z$ is well-defined. It is linear by the algebra of limits. Furthermore

$$\left\|\widetilde{T}(x)\right\| = \lim_{n\to\infty} \|T(y_n)\| \leqslant \|T\| \lim_{n\to\infty} \|y_n\| = \|T\|\|x\|$$

Hence $\left\|\widetilde{T}\right\| \leqslant \|T\|$. Since $T\big|_Y = T$ we must have $\left\|\widetilde{T}\right\| = \|T\|$. □

In order to present the central theorem of Stone-Weierstrass, we need to introduce some definitions. From now on and till the end of this section, K denotes a compact metric space[1] and $\mathrm{C}(K)$ denotes the set of *real-valued* continuous functions on K.

Definition 3.3. Linear Lattices

A subspace $D \subseteq \mathrm{C}(K)$ is called a linear sublattice of $\mathrm{C}(K)$, if

$$f, g \in D \implies \max\{f, g\} \in D$$

or equivalently

$$f \in D \implies |f| \in D$$

Definition 3.4. Separation of Points

We say that $D \subseteq \mathrm{C}(K)$ separates points in K, if

$$\forall\, p, q \in K\, (p \neq q \implies \exists\, f \in D\ f(p) \neq f(q))$$

Definition 3.5. Algebras

A subspace $D \subseteq \mathrm{C}(K)$ is called a subalgebra of $\mathrm{C}(K)$, if D contains the constant functions and is closed under pointwise multiplication:

$$f, g \in D \implies fg \in D$$

Theorem 3.6. Stone-Weierstrass Theorem, Lattice Form

Let K be a compact metric space. A subspace $D \subseteq \mathrm{C}(K)$ is dense in $(\mathrm{C}(K), \|\cdot\|_\infty)$, if

- D is a linear sublattice;
- D contains the constant functions;
- D separates points in K.

Proof. Fix $f \in \mathrm{C}(K)$ and $\varepsilon > 0$. For $p, q \in K$, we construct $f_{p,q} \in D$ as follows:

If $p = q$, simply set $f_{p,p} = f(p)$ to be the constant function. Then $f_{p,p} \in D$ by assumption. If $p \neq q$, since D separates points, there exists $g \in D$ such that $g(p) = 0$ and $g(q) = 1$. We define

$$f_{p,q}(x) = f(p) + (f(q) - f(p))\, g(x)$$

Then in either case, we have $f_{p,q} \in D$ such that $f_{p,q}(p) = f(p)$ and $f_{p,q}(q) = f(q)$. Consider

$$U_{p,q} := (f - f_{p,q})^{-1}\, (B(0, \varepsilon)) \subseteq K$$

Then $U_{p,q}$ is an open subset of K containing p and q. Now $\{U_{p,q}\}_{q \in K}$ is an open cover of K. By compactness of K there exists $q_1, \ldots, q_m \in K$ such that $K = \bigcup_{i=1}^{m} U_{p,q_i}$. Define

$$g_p := \min\{f_{p,q_1}, \ldots, f_{p,q_m}\}$$

Since D is a linear sublattice, $g_p \in D$. We have $g_p(p) = f(p)$. Furthermore, as $|f(x) - f_{p,q}(x)| < \varepsilon$ for $x \in U_{p,q}$, we have $g_p < f + \varepsilon$ on all of K.

[1] More generally, K could be a compact Hausdorff space.

Next, consider the open neighbourbood of p:

$$V_p := (f - g_p)^{-1}(B(0,\varepsilon)) \subseteq K$$

Now $\{V_p\}_{p\in K}$ is an open cover of K. Again by compactness of K there exists $p_1, ..., p_k \in K$ such that $K = \bigcup_{i=1}^{k} V_{p_i}$.

Define

$$g := \max\{g_{p_1}, ..., g_{p_k}\}$$

Then $g \in D$. As $|f(x) - g_p(x)| < \varepsilon$ for $x \in V_p$, we have $g > f - \varepsilon$ on all K. But since $g_p < f + \varepsilon$ for all $p \in K$, we also have $g < f + \varepsilon$ on K.

Now we have constructed $g \in D$ such that $\|f - g\|_\infty < \varepsilon$. We conclude that D is dense in $\mathrm{C}(K)$. □

Theorem 3.7. Stone-Weierstrass Theorem, Subalgebra Form

Let K be a compact metric space. A subspace $D \subseteq \mathrm{C}(K)$ is dense in $(\mathrm{C}(K), \|\cdot\|_\infty)$, if D is a subalgebra that separates points in K.

Proof. Since D is a subalgebra, so is $\overline{D}$. By the previous theorem it suffices to show that $\overline{D}$ is a linear sublattice. Fix $f \in \overline{D}$. We need to show that $|f| \in \overline{D}$.

Without loss of generality we assume that $\|f\|_\infty \leqslant 1$. Consider the sequence $(f_n) \subseteq \overline{D}$, where $f_0 = 0$ and

$$f_{n+1} = f_n + \frac{1}{2}(f^2 - f_n^2)$$

We shall prove that $f_n \nearrow |f|$ pointwise by induction. The base case is trivial. For the induction case,

$$0 \leqslant f_n \leqslant f_{n+1} = f_n + \frac{1}{2}(|f| + f_n)(|f| - f_n) \leqslant f_n + (|f| - f_n) = |f| \qquad \left(f_n \leqslant |f| \leqslant 1 \implies \frac{|f| + f_n}{2} \leqslant 1\right)$$

$$f_{n+2} - f_{n+1} = (f_{n+1} - f_n)\left(1 - \frac{1}{2}(f_{n+1} + f_n)\right) \geqslant f_{n+1} - f_n \geqslant 0 \qquad \left(f_n \leqslant f_{n+1} \leqslant |f| \leqslant 1 \implies \frac{f_n + f_{n+1}}{2} \leqslant 1\right)$$

By Monotone Convergence Theorem, $f_n \to g$ pointwise for some g, where

$$g = g + \frac{1}{2}(f^2 - g^2) \implies g = |f|$$

Hence $f_n \nearrow |f|$ pointwise. Since K is compact, by Dini's Theorem, $f_n \to |f|$ uniformly. Hence $|f| \in \overline{D}$ as claimed. □

Corollary 3.8. Weierstrass Approximation Theorem for Polynomials

Let $K \subseteq \mathbb{R}^n$ be a compact set. For any continuous function $f : K \to \mathbb{R}$, there exists a sequence of polynomials (p_n) such that $p_n \to f$ uniformly in K as $n \to \infty$.

Proof. It is clear that the set of polynomials $P \subseteq \mathrm{C}(K)$ is a subalgebra of $\mathrm{C}(K)$. So P is dense in $\mathrm{C}(K)$. For $f \in \mathrm{C}(K)$ there exists a sequence $(p_n) \subseteq P$ such that $p_n \to f$ in $\mathrm{C}(K)$, i.e., uniformly on K. □

3.2 Separability

Recall from *A2 Metric Spaces* that a metrix space X is said to be **separable**, if it has a countable dense subset.

Proposition 3.9. Linear Span and Separability

Let $(X, \|\cdot\|_X)$ be a normed vector space. If there exists a countable subset $S \subseteq X$ such that $\operatorname{span} \overline{S}$ is dense in X, then X is separable.

Proof. First we consider the scalar field $F = \mathbb{R}$. We enumerate S as $\{s_1, s_2, ...\}$. Consider

$$Y := \left\{ \sum_{i=1}^{N} q_i s_i : \ q_i \in \mathbb{Q},\ s_i \in S,\ N \in \mathbb{N} \right\} \subseteq X$$

There exists a surjection $\bigcup_{N\in\mathbb{N}} \mathbb{Q}^N \to Y$ given by $(q_1, ..., q_N) \mapsto \sum_{i=1}^{N} q_i s_i$. Then

$$\operatorname{card}(Y) \leqslant \operatorname{card}\left(\bigcup_{N\in\mathbb{N}} \mathbb{Q}^N \right) = \aleph_0$$

So Y is countable. Next we claim that Y is dense in $\operatorname{span} \overline{S}$.

Fix $\varepsilon > 0$ and $u \in \operatorname{span} \overline{S}$, where

$$u = \sum_{i=1}^{N} \lambda_i u_i, \qquad \lambda_i \in \mathbb{R},\ u_i \in \overline{S}$$

Since S is dense in $\overline{S}$, for each $i \in \{1, ..., N\}$ there exists $v_i \in S$ such that $\|u_i - v_i\| < \dfrac{\varepsilon}{2N|\lambda_i|}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, if $v_i \neq 0$, then there exists $q_i \in \mathbb{Q}$ such that $|q_i - \lambda_i| < \dfrac{\varepsilon}{2N\|v_i\|}$. Let $v := \sum_{i=1}^{N} q_i v_i \in Y$.

$$\|u - v\| = \left\| \sum_{i=1}^{N} \lambda_i u_i - \sum_{i=1}^{N} q_i v_i \right\| \leqslant \left\| \sum_{i=1}^{N} \lambda_i (u_i - v_i) \right\| + \left\| \sum_{i=1}^{N} (\lambda_i - q_i) v_i \right\| \leqslant \sum_{i=1}^{N} (|\lambda_i| \|u_i - v_i\| + |\lambda_i - q_i| \|v_i\|) < \varepsilon$$

Hence Y is dense in $\operatorname{span} \overline{S}$. But $\operatorname{span} \overline{S}$ is dense in X. We deduce that Y is dense in X. X is separable.

If the scalar field $F = \mathbb{C}$, then instead we consider

$$Y := \left\{ \sum_{i=1}^{N} q_i s_i : \ q_i \in \mathbb{Q}[\mathrm{i}],\ s_i \in S,\ N \in \mathbb{N} \right\} \subseteq X$$

and use the fact that $\mathbb{Q}[\mathrm{i}]$ is countable and dense in $\mathbb{C}$. □

Corollary 3.10

All finite-dimensional normed vector spaces are separable.

Next we need to determine the separability of some usual Banach spaces.

Proposition 3.11. Separability of $\mathrm{C}(K)$

Let K be a compact metric space. Then K and $\mathrm{C}(K)$ are separable.

Proof. First we prove that K is separable. For $n \in \mathbb{Z}_+$, consider the open ball $U_{x,n} := B(x, 1/n)$. Then $\{U_{x,n}\}_{x\in K}$ is an open over of K. By compactness of K, there exists $x_{1,n}, ... x_{k_n,n} \in K$ such that $K = \bigcup_{i=1}^{k_n} U_{x_{i,n},n}$. Now consider

$$E := \bigcup_{n=1}^{\infty} \{x_{1,n}, ..., x_{k_n,n}\}$$

It is clear that E is countable. We claim that it is dense in K. Let $y \in K$ and $\varepsilon > 0$. Pick $n \in \mathbb{Z}_+$ such that $1/n < \varepsilon$. Then $y \in B(x, 1/n)$ for some $x \in E$. And $d(x, y) < 1/n < \varepsilon$. We deduce that E is dense in K.

Next we prove that $\mathrm{C}(K)$ is separable. We enumerate E as $\{y_1, y_2, ...\}$. Let $f_n(x) := d(x, y_n)$. Consider the set of ”monomials”

$$S := \{f_1^{\alpha_1} \cdots f_N^{\alpha_N} : \ \alpha_1, ..., \alpha_N \in \mathbb{N},\ N \in \mathbb{N}\}$$

which is clearly countable. Then $\operatorname{span} S$ is a subalgebra of $\mathrm{C}(K)$.

For $p, q \in K$ with $p \neq q$, let $n \in \mathbb{N}$ such that $d(p,q) > 2/n$. There exists $y_i \in E$ such that $p \in B(y_i, 1/n)$. By triangular inequality,

$$d(y_i, q) > d(p,q) - d(y_i, p) > 1/n > d(y_i, p)$$

Hence $f_i(q) > f_i(p)$. We deduce that $\operatorname{span} S$ separables points. Now by the subalgebra form of Stone-Weierstrass Theorem, $\operatorname{span} S$ is dense in $\mathrm{C}(K)$. By Proposition 3.9, we conlude that $\mathrm{C}(K)$ is separable. ☐

Proposition 3.12. Separability of ℓ^p and L^p

Let $1 \leqslant p < \infty$.

1. The sequence space ℓ^p is separable.
2. Let $\Omega \subseteq \mathbb{R}^n$ be a measurable set. The function space $\mathrm{L}^p(\Omega)$ is separable.

Proof. 1. Let $Y := \operatorname{span} S$ where the countable set $S = \left\{e^{(k)}, k \in \mathbb{N}\right\}$ consists of all sequences $e^{(k)}$ for which $e_j^{(k)} = \delta_{jk}$.

Given any element $x \in \ell^p$, since $\sum_{j \in \mathbb{N}}^{\infty} |x_j|^p$ converges, we obtain that the cut-off sequences $x^{(k)} := (x_1, \ldots, x_k, 0, 0, \ldots)$ approximate x in the sense of ℓ^p, namely

$$\left\| x - x^{(k)} \right\|_p = \left(\sum_{j \geqslant k+1} |x_j|^p \right)^{1/p} \to 0$$

as $k \to \infty$. We thus conclude that Y is dense in ℓ^p and thus obtain from Proposition 3.9 that ℓ^∞ is separable.

2. From Example 3.1 we know that $\mathrm{C}_c(\Omega)$ is dense in $\mathrm{L}^p(\Omega)$. So we shall prove that $\mathrm{C}_c(\Omega)$ is separable.

Consider the set of cubes with rational vertices

$$\mathscr{S} := \{[p_1, q_1] \times \cdots \times [p_n, q_n] : \ p_1, q_1, ..., p_n, q_n \in \mathbb{Q}\}$$

Let $Y := \operatorname{span}\{\mathbf{1}_{I \cap \Omega} : \ I \in \mathscr{S}\}$. Since $\mathscr{S}$ is countable, it suffices to prove that Y is dense in $\mathrm{C}_c(\Omega)$.

Fix $f \in \mathrm{C}_c(\Omega)$ and $\varepsilon > 0$. Let $K := \operatorname{supp} f \subseteq \mathbb{R}^n$ be a compact set. Then f is uniformly continuous on K. There exists $N \in \mathbb{N}$ such that $|f(x) - f(y)| < \varepsilon \cdot m(K)^{-1/p}$ whenever $\|x - y\|_\infty \leqslant 1/N$. Let

$$\mathscr{T} := \left\{ I := \left[\frac{k_1}{N}, \frac{k_1+1}{N}\right] \times \cdots \times \left[\frac{k_n}{N}, \frac{k_n+1}{N}\right] : \ k_1, ..., k_n \in \mathbb{Z}, \ I \cap K \neq \varnothing \right\}$$

Then $\mathscr{T}$ is a finite subset of $\mathscr{S}$ because K is bounded. Consider the "step" function

$$\varphi(x) := \sum_{I \in \mathscr{T}} f(x_I) \mathbf{1}_{I \cap \Omega} \in Y$$

where $x_I \in I$ is some fixed points. We find that

$$\|f - \varphi\|_p^p = \int_K |f - \varphi|^p = \sum_{I \in \mathscr{T}} \int_{I \cap K} |f(x) - f(x_I)|^p \, \mathrm{d}x < \frac{\varepsilon^p}{m(K)} \sum_{I \in \mathscr{T}} \int_{I \cap K} \mathrm{d}x = \varepsilon^p$$

which proves the claim. ☐

Proposition 3.13. Inseparability of ℓ^∞ and L^∞

1. The sequence space ℓ^∞ is inseparable.
2. Let $\Omega \subseteq \mathbb{R}^n$ be an **open** set. The function space $\mathrm{L}^\infty(\Omega)$ is inseparable.

Proof. 1. Note that the set $A := \{0,1\}^{\mathbb{N}} \subseteq \ell^\infty$ is uncountable. And for any distinct $a, b \in A$, $\|a - b\|_\infty = 1$.

Suppose that $D \subseteq \ell^\infty$ is dense in ℓ^∞. For each $a \in A$ there exists $d_a \in D$ such that $\|a - d_a\| < 1/2$. Then $d_a = d_b$ implies that

$$\|a-b\|_\infty \leqslant \|a-d_a\|_\infty + \|b-d_b\|_\infty < 1$$

and hence that $a = b$. Therefore the map $a \mapsto d_a$ is an injective map from A to D. So D is uncountable and ℓ^∞ is inseparable.

2. There exists an open ball $B(x,R) \subseteq \Omega$. Consider the uncountable set $A := \left\{\mathbf{1}_{B(x,r)} : \ 0 < r < R\right\} \subseteq \mathrm{L}^\infty(\Omega)$ and repeat the proof for ℓ^∞, ☐

3.3 Completeness: Baire Category Theorem

A subset $D \subseteq X$ is said to be **nowhere dense** in the metric space X, if there are no open set $U \subseteq X$ for which $D \cap U$ is dense in U, or equivalently, if $\overline{D}$ has an empty interior.

Theorem 3.14. Baire Category Theorem

Let X be a non-empty complete metric space.

1. X is not the countable union of nowhere dense sets.
2. The countable interchapter of dense open sets is non-empty.
3. The countable union of nowhere dense sets has empty interior.
4. The countable interchapter of dense open sets is dense.

In fact the four statements are equivalent. We shall mianly prove (1) and show that other statements are equivalent to it.

Proof. 1. Suppose that $X = \bigcup\limits_{n=1}^{\infty} A_n$ where each A_n is nowhere dense.

Since A_1 is nowhere dense, $\overline{A}_1 \neq X$ and so $X \setminus \overline{A}_1$ is non-empty. Pick $x_1 \in X \setminus \overline{A}_1$. Next, since $X \setminus \overline{A}_1$ is open, there is some closed ball $\overline{B}\left(x_1, r_1\right) \subseteq X \setminus \overline{A}_1$ with $r_1 < 1$. Clearly $B\left(x_1, r_1\right) \cap A_1 = \emptyset$.

Since A_2 is nowhere dense, $\overline{A}_2 \not\supseteq B\left(x_1, r_1\right)$ and so there is some $x_2 \in B\left(x_1, r_1\right) \setminus \overline{A}_2$. We then inductively choose balls $\overline{B}\left(x_n, r_n\right) \subseteq B\left(x_{n-1}, r_{n-1}\right) \setminus \overline{A}_n$ with $r_n < \dfrac{1}{2^{n-1}}$.

Now, (x_n) is a Cauchy sequence, since if $n, m \geqslant N$, then $x_n, x_m \in B\left(x_N, r_N\right)$ and so $d\left(x_m, x_n\right) \leqslant 2r_N \to 0$. Since X is complete, (x_n) converges to some $x \in X$. We have $x \in \overline{B}\left(x_n, r_n\right) \subseteq B\left(x_{n-1}, r_{n-1}\right) \setminus \overline{A}_n$ for all n, which implies that $x \notin A_n$ for any n. This is a contradiction.

2. Note that the complement of a dense open set is a nowhere dense set. The result then follows from (1) and de Morgan's Law.

3. Suppose that $(A_n)_{n=1}^{\infty}$ is a sequence of nowhere dense subsets of X. Suppose that $\bigcup\limits_{n=1}^{\infty} A_n$ contains an open ball $B(x,r)$. Take a closed ball $\overline{B}(x,s) \subseteq B(x,r)$ and consider $\widetilde{A}_n := A_n \cap \overline{B}(x,s)$. Then $\overline{B}(x,s) = \bigcup\limits_{n=1}^{\infty} \widetilde{A}_n$ where each $\widetilde{A}_n$ is nowhere dense in $\overline{B}(x,s)$. This is a contradiction.

4. It follows from (3) and de Morgan's Law. ☐

The following terminology explains the name of the Baire Category Theorem:

Definition 3.15. Meagre Sets, Nonmeagre Sets

Let X be a metric space and $E \subseteq X$.

- E is said to be **of the first category** or **meagre**, if it is the countable union of nowhere dense set.
- E is said to be **of the second category** or **nonmeagre**, if it is not of the first category.

The Baire Category Theorem says that *an open subset of a complete metric space is of the second category.*

In Functional Analysis, the usual notion of the basis of a vector space is called a **Hamel basis**, i.e., a linearly independent subset $S \subseteq X$ such that every vector of X is a *finite linear combination* of the vectors in S.

Proposition 3.16. Cardinality of Hamel Basis (*Off-Syllabus*)

Suppose that X is a Banach space. Then the Hamel basis of X is either finite or uncountable.

Proof. Suppose for that there exists a countable Hamel basis $B = \{x_n : n \in \mathbb{N}\}$ for X. Let $F_n = \operatorname{span}\{x_k : 0 \leqslant k \leqslant n\}$. Then each of the spaces F_n is finite-dimensional and hence complete, and in particular each F_n is closed in X. Since $X = \bigcup_{n \in \mathbb{N}} F_n$, by the Baire Category Theorem, there exists $n \in \mathbb{N}$ such that F_n has non-empty interior. Suppose that $B_(x, \varepsilon) \subseteq F_n$. Since F_n is a vector space and in particular closed under translations, it follows that $B(0, \varepsilon) \subseteq F_n$, and hence $X \subseteq F_n$. In particular, we have $\dim X \leqslant n$, which is a contradiction. ☐

Chapter 4

Fundamental Theorems

This chapter mainly discusses the application of Baire Category Theorem in the theory of Banach spaces. Three fundamental theorems in functional analysis are introduced.

4.1 Uniform Boundedness Principle

Theorem 4.1. Uniform Boundedness Principle / Banach-Steinhaus Theorem

Suppose that $(X, \|\cdot\|_X)$ is a Banach space and $(Y, \|\cdot\|_Y)$ is a normed vector space. Let $\mathscr{F} \subseteq \mathscr{B}(X, Y)$. If for each $x \in X$,

$$\sup \{\|T(x)\|_Y : \ T \in \mathscr{F}\} < \infty$$

Then

$$\sup \left\{\|T\|_{\mathscr{B}(X,Y)} : \ T \in \mathscr{F}\right\} < \infty$$

For a family of bounded linear maps from a Banach space, *pointwise boundedness implies uniform boundedness.*

Proof. Consider the sequence of closed subspaces of X:

$$A_n = \{x \in X : \ \forall T \in \mathscr{F} \ \|T(x)\|_Y \leqslant n\}$$

Then $X = \bigcup_{n \in \mathbb{N}} A_n$. By Baire Category Theorem, there exists $N \in \mathbb{N}$ such that A_N has non-empty interior. Let $B(x_0, r_0) \subseteq A_N$.

For any $x \in B(0, r_0)$ and $T \in \mathscr{F}$, $x_0 + x \in B(x_0, r_0)$. Therefore

$$\|T(x)\|_Y \leqslant \|T(x + x_0)\|_Y + \|T(x_0)\|_Y \leqslant 2N$$

Hence for any $x \in B(0, 1)$ and $T \in \mathscr{F}$, $\|T(x)\|_Y \leqslant 2N/r_0$. We conclude that $\|T\|_{\mathscr{B}(X,Y)} \leqslant 2N/r_0$. □

Proposition 4.2. Uniform Boundedness in Hilbert Spaces

Suppose that X is a Hilbert space and $\mathscr{F} \subseteq \mathscr{B}(X)$. If for each $x, y \in X$,

$$\sup \{|\langle y, Tx\rangle| : \ T \in \mathscr{F}\} < \infty$$

Then $\mathscr{F}$ is uniformly bounded.

Proof. Fix $x \in X$. Define $K_{T,x} \in X^*$ by $K_{T,x}(y) = \langle Tx, y\rangle$ so that $\|Tx\| = \|K_{T,x}\|_*$. Since for each $y \in X$, $\{|K_{T,x}(y)| : T \in \mathscr{F}\}$ is bounded, by the Uniform Boundedness Principle, $\{\|K_{T,x}\|_* : T \in \mathscr{F}\}$ is bounded. We use the Uniform Boundedness Principle again to conclude that $\{\|T\| : T \in \mathscr{F}\}$ is bounded. □

Proposition 4.3. Strong Convergence of Operators

Suppose that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces. Let $(T_n) \subseteq \mathscr{B}(X, Y)$. The following are equivalent:

1. There exists $T \in \mathscr{B}(X, Y)$ such that for each $x \in X$, $T_n(x) \to T(x)$ as $n \to \infty$;
2. For each $x \in X$, $(T_n(x))$ is convergent,
3. There exists $M > 0$ and a dense subset $Z \subseteq X$ such that $\sup_n \|T_n\| \leqslant M$ and $(T_n(z))$ is convergent for each $z \in Z$.

Proof. "$1 \implies 2$": Trivial.

"$2 \implies 3$": Immediate from Uniform Boundedness Principle.

"$3 \implies 1$": Fix $\varepsilon > 0$ and $x \in X$. By density of Z in X, we can choose $z \in Z$ such that $\|x - z\|_X < \varepsilon/4M$. Since $(T_n(z))$ is a Cauchy sequence, there exists $N > 0$ such that for $m, n > N$, $\|T_n(z) - T_m(z)\|_Y < \varepsilon/2$. Then

$$\|T_n(x) - T_m(x)\|_Y \leqslant \|T_n(x) - T_n(z)\|_Y + \|T_m(x) - T_m(z)\|_Y + \|T_n(z) - T_m(z)\|_Y \leqslant 2M\|x - z\|_X + \|T_n(z) - T_m(z)\|_Y \leqslant \varepsilon$$

Hence $(T_n(x))$ is a Cauchy sequence. Since Y is a Banach space, $T(x) := \lim_{n\to\infty} T_n(x)$ exists. It is clear that T is linear. T is bounded because

$$\|T(x)\|_Y = \lim_{n\to\infty} \|T_n(x)\| \leqslant \limsup_{n\to\infty} \|T_n\|_{\mathscr{B}(X,Y)} \|x\|_X \leqslant M\|x\|_X$$

□

Remark. Note that in the proposition, the convergence $T_n \to T$ is a pointwise convergence, or **convergence in the strong operator topology**. This should not be confused with *convergence in the operator norm* $\|T_n - T\|_{\mathscr{B}(X,Y)} \to 0$.

4.2 Open Mappings and Closed Graphs

Theorem 4.4. Open Mapping Theorem

Suppose that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces. $T \in \mathscr{B}(X, Y)$ is *surjective*. Then T is an open map (mapping open sets to open sets).

Proof. Let $U \subseteq X$ be an open subset. To prove that $T(U) \subseteq Y$ is open, we fix $y \in T(U)$ for which $T(x) = y$. Since U is open, there exists $r > 0$ such that $B(x, r) \subseteq U$. By linearity, we have

$$T(B_X(x, r)) = y + \frac{r}{2} T(B_X(0, 2))$$

- $\overline{T(B_X(0,1))}$ contains an open ball $B_Y(0, s)$:

  Since T is surjective, we have
  $$Y = \bigcup_{n=1}^{\infty} \overline{T(B_X(0, n))}$$
  By Baire Category Theorem, there exists $N \in \mathbb{N}$ such that $\overline{T(B_X(0, N))}$ has non-empty interior. Let $w \in Y$ and $s > 0$ such that $B_Y(Nw, Ns) \subseteq \overline{T(B_X(0, N))}$. By linearity, $B_Y(w, s) \subseteq \overline{T(B_X(0, 1))}$. Since $\overline{T(B_X(0, 1))}$ is convex and symmetric about the origin, we have $B_Y(-w, s) \subseteq \overline{T(B_X(0, 1))}$ and hence $B_Y(0, s) \subseteq \overline{T(B_X(0, 1))}$.

- $T(B_X(0, 2))$ contains $\overline{T(B_X(0, 1))}$:

  Fix $z \in \overline{T(B_X(0, 1))}$. First choose $u_1 \in B_X(0, 1)$ such that $\|z - T(u_1)\| < \frac{1}{2}s$. Then
  $$z - T(u_1) \in B_Y\left(0, \frac{1}{2}s\right) = \frac{1}{2}B_Y(0, s) \subseteq \frac{1}{2}\overline{T(B_X(0, 1))} = \overline{T\left(B_X\left(0, \frac{1}{2}\right)\right)}$$

Inductively, we can choose $u_n \in B_X\left(0, 2^{1-n}\right)$ such that $\left\| z - \sum_{k=1}^{n} T(u_k) \right\| < 2^{-n}s$, which impplies that

$$z - \sum_{k=1}^{n} T(u_k) \in \overline{T\left(B_X\left(0, 2^{-n}\right)\right)}$$

Therefore $\sum_{k=1}^{n} T(u_k) \to z$ in Y as $n \to \infty$. It is clear that $u_n \to u \in B_X(0,2)$ as $n \to \infty$. By continuity of T, we deduce that $z = T(u) \in T(B_X(0,2))$. Hence $\overline{T(B_X(0,1))} \subseteq T(B_X(0,2))$.

Now

$$T(B_X(x,r)) = y + \frac{r}{2}T(B_X(0,2)) \supseteq y + \frac{r}{2}B_Y(0,s) = B_Y\left(y,\ \frac{1}{2}rs\right)$$

Hence $y \in (T(U))^{\circ}$. We conclude that $T(U)$ is open. □

Corollary 4.5. Inverse Mapping Theorem

Suppose that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces. $T \in \mathscr{B}(X,Y)$ is *bijective*. Then $T^{-1} \in \mathscr{B}(Y,X)$.

Proof. By Open Mapping Theorem, T is an open map, which implies that T^{-1} is continuous. By Proposition 1.17, $T^{-1} \in \mathscr{B}(X,Y)$. □

Corollary 4.6. Equivalence of Norms

Suppose that X is a Banach space with respect to the norms $\|\cdot\|$ and $\|\cdot\|'$. If there exists $C > 0$ such that $\|x\| \leqslant C\|x\|'$, then $\|\cdot\|$ and $\|\cdot\|'$ are equivalent.

Proof. Directly applying the Inverse Mapping Theorem to the identity operator $\mathrm{id} : (X, \|\cdot\|) \to (X, \|\cdot\|')$. □

Proposition 4.7. Criterion for Closed Image

Suppose that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces. Let $T : X \to Y$ be a linear map. Let $T \in \mathscr{B}(X,Y)$. If $\operatorname{im} T$ has closed linear complement in Y, then $\operatorname{im} T$ is closed in Y.

Proof. Suppose that $Y = \operatorname{im} T \oplus Z$ where $Z \subseteq Y$ is a closed subspace of Y. Let $S : X \times Z \to Y$ given by $S(x,y) = T(x) + y$. It is clear that $X \times Z$ is a Banach space and $S \in \mathscr{B}(X \times Z, Y)$. Since $S(X \times \{0\}) = \operatorname{im} T$, we have $S(X \times (Z \setminus \{0\})) = Y \setminus \operatorname{im} T$. $X \times \{0\}$ is closed in $X \times Z$. So $X \times (Z \setminus \{0\})$ is open in $X \times Z$, and by the Open Mapping Theorem, $Y \setminus \operatorname{im} T$ is open in Y. We conclude that $\operatorname{im} T$ is closed in Y. □

Remark. The proposition has a useful corollary: If $\operatorname{im} T$ has finite codimension, then $\operatorname{im} T$ is closed in Y. It follows from that all finite-dimensional subspaces of a Bana ch space are closed.

Proposition 4.8. Closed Images and Adjoints

Suppose that X and Y are Hilbert spaces, and $T \in \mathscr{B}(X,Y)$. Then $\operatorname{im} T$ is closed if and only if $\operatorname{im} T^*$ is closed.

Proof. It suffices to show only one direction, as $T^{**} = T$. Suppose that $W = \operatorname{im} T^*$ is closed in X. Let $Z = \overline{\operatorname{im} T} \subseteq Y$. Let $S \in \mathscr{B}(X,Z)$ such that $Sx = Tx$ for all $x \in X$. The adjoint $S^* \in \mathscr{B}(Z,X)$. By Proposition 2.17, $Z = \overline{\operatorname{im} S} = (\ker S^*)^{\perp}$. So S^* is injective.

Let P be the orthogonal projection from Y onto Z and compute, for $x \in X$ and $y \in Y$,

$$\langle Tx, y \rangle_Y = \langle Sx, Py \rangle_Y = \langle x, S^*Py \rangle_X$$

This shows that $T^* = S^* \circ P$, and so $\operatorname{im} S^* = W$. Now let $V \in \mathscr{B}(Z,W)$ such that $Vz = S^*z$ for all $z \in Z$. V is bijective. By the Inverse Mapping Theorem, V has a bounded inverse $V^{-1} \in \mathscr{B}(W,Z)$. This implies that V^* is invertible and $(V^*)^{-1} = \left(V^{-1}\right)^* \in \mathscr{B}(Z,W)$.

To conclude, we show that $T \circ (V^*)^{-1} = \mathrm{id}_Z$. This implies that $\operatorname{im} T \supseteq Z$ and so $\operatorname{im} T = Z = \overline{\operatorname{im} T}$ which gives the conclusion. Indeed, pick an arbitrary $z \in Z$, and let $w = (V^*)^{-1} z$. We compute, for $y \in Y$:

$$\langle Tw, y\rangle_Y = \langle Sw, y\rangle_Y = \langle w, S^* y\rangle_X = \langle w, Vy\rangle_X = \langle V^* w, y\rangle_Y = \langle z, y\rangle_Y$$

Since this holds for all $y \in Y$, we deduce that $Tw = z$ and so $T \circ (V^*)^{-1} = \mathrm{id}_Z$ as desired. □

Remark. The technique in the proof is to get rid of the cokernels and to construct a bijective operator, to which we can apply the Inverse Mapping Theorem.

Definition 4.9. Graph of an Operator

Suppose that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are normed vector spaces. $T : X \to Y$ is a linear map. The graph of T is defined by

$$\Gamma(T) := \{(x, T(x)) \in X \times Y : \ x \in X\}$$

Note that if X and Y are both Banach spaces, $X \times Y$ equipped with the 1-norm:

$$\|(x, y)\|_{X\times Y} := \|x\|_X + \|y\|_Y$$

is also a Banach space.

Theorem 4.10. Closed Graph Theorem

Suppose that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces. Let $T : X \to Y$ be a linear map. Then $T \in \mathscr{B}(X, Y)$ if and only if $\Gamma(T) \subseteq X \times Y$ is closed.

Proof. $\Longrightarrow$: Suppose that T is bounded. Let $(x_n) \subseteq X$ such that $x_n \to x \in X$ and $T(x_n) \to y \in Y$ as $n \to \infty$. By continuity of T,

$$y = \lim_{n\to\infty} T(x_n) = T\left(\lim_{n\to\infty} x_n\right) = T(x)$$

Hence $\Gamma(T)$ is closed.

$\Longleftarrow$: Consider the **graph norm** on X: For $x \in X$,

$$\|x\|_T := \|x\|_X + \|T(x)\|_Y$$

The closedness of $\Gamma(T)$ implies that $(X, \|\cdot\|_T)$ is a Banach space. It is clear that $\|x\|_X \leqslant \|x\|_T$. By Corollary 4.6, we deduce that $\|\cdot\|_X$ and $\|\cdot\|_T$ are equivalent. In particular,

$$\|T(x)\|_Y = \|x\|_T - \|x\|_X \leqslant (C - 1)\|x\|_X$$

for some constant $C \geqslant 1$. Hence T is bounded. □

Remark. Following the Closed Graph Theorem, whem we need to prove that a linear map $T : X \to Y$ between Banach spaces is bounded, it suffices to show that (Tx_n) *is a Cauchy sequence in* Y *for every convergent/Cauchy sequence* (x_n) *in* X.

We recall some concepts from Linear Algebra:

Suppose that U is a subspace of the vector space X. V ic called a **linear complement** of U in X, if $X = U \oplus V$.

A linear map $P : X \to X$ is called a **projection**, if $P^2 = P$. Suppose that we have a direct sum decomposition $X = U \oplus V$. The linear map $P : X \to U$ such that $P\big|_U = \mathrm{id}_U$ and $P\big|_V = 0$ is called a **projection of X onto U along V**.

Proposition 4.11. Existence of Closed Linear Complement

Let $Y \subseteq X$ be a closed subspace of the Banach space X. Then Y has a closed linear complement if and only if there exists a continuous projection operator $P : X \twoheadrightarrow Y$.

Proof. $\Longrightarrow$: Suppose that $Y, Z \subseteq X$ are closed subspaces and $X = Y \oplus Z$. Let $P : X \twoheadrightarrow Y$ be the projection of X onto Y along Z. Let $x_n \to x \in X$ and $P(x_n) \to y \in X$ as $n \to \infty$. Since $Y = P(X)$ is closed we have $y \in Y$. Since $Z = (\mathrm{id} - P)(X)$ is closed we also have $x - y = \lim_{n\to\infty} (\mathrm{id} - P)(x_n) \in Z$. Therefore $P(y) = y$ and $P(x - y) = 0$. We deduce that $P(x) = y$. Hence P has closed graph. By the Closed Graph Theorem, P is continuous.

$\Longleftarrow$: Let $P : X \twoheadrightarrow Y$ be a continuous projection. Then we have a direct sum decomposition:

$$X = Y \oplus (\mathrm{id} - P)(X)$$

We claim that $(\mathrm{id} - P)(X)$ is closed in X. Note that $(\mathrm{id} - P)$ is also a continuous projection. For $((\mathrm{id} - P)(x_n)) \subseteq (\mathrm{id} - P)(X)$ such that $(\mathrm{id} - P)(x_n) \to y \in X$, we have

$$y = \lim_{n\to\infty} (\mathrm{id} - P)(x_n) = \lim_{n\to\infty} (\mathrm{id} - P)^2(x_n) = (\mathrm{id} - P)\left(\lim_{n\to\infty} (\mathrm{id} - P)(x_n)\right) = (\mathrm{id} - P)(y) \in (\mathrm{id} - P)(X)$$

Hence $(\mathrm{id} - P)(X)$ is closed. □

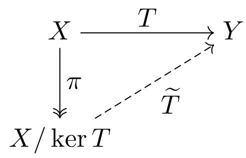
In this section, we have followed the route:

Baire Category Theorem $\Longrightarrow$ Open Mapping Theorem $\Longrightarrow$ Inverse Mapping Theorem $\Longrightarrow$ Closed Graph Theorem

We shall briefly discuss the equivalence of the three theorems in this section by proving:

- *Inverse Mapping Theorem* $\Longrightarrow$ *Open Mapping Theorem*:

  Suppose that X and Y are Banach spaces. Let $T : X \to Y$ be a surjective map. Note that the quotient map $\pi : X \twoheadrightarrow X/\ker T$ is continuous, surjective, and open. T descends to a bounded bijection $\widetilde{T} \in \mathscr{B}(X/\ker T, Y)$ via the following commutative diagram:



  By the Inverse Mapping Theorem, $\widetilde{T}^{-1} \in \mathscr{B}(Y, X/\ker T)$. For an open set $U \subseteq X$, the image is given by

$$T(U) = \widetilde{T} \circ \pi(U) = \left(\widetilde{T}^{-1}\right)^{-1} \circ \pi(U)$$

  which is open in Y. □

- *Closed Graph Theorem* $\Longrightarrow$ *Inverse Mapping Theorem*:

  Suppose that X and Y are Banach spaces. Let $T : X \to Y$ be a bijective map. By the Closed Graph Theorem, $\Gamma(T)$ is closed in $X \times Y$. $\Gamma(T^{-1})$ is just $\Gamma(T)$ swapping two coordinates, and hence is closed in $Y \times X$. By the Closed Graph Theorem, $T^{-1} \in \mathscr{B}(Y, X)$. □

Chapter 5

Duality

In this chapter, we introduce the important Hahn-Banach Theorem and discuss its analytic and geometric consequences. We also briefly discuss the weak and weak* convergence and topology.

We shall frequently use the **sign function** on $\mathbb{C}$ or $\mathbb{R}$, which is given by

$$\operatorname{sgn} z := \begin{cases} z/|z| & z \neq 0 \\ 0 & z = 0 \end{cases}$$

For function $f : X \to \mathbb{C}$, we can define the sign function $\operatorname{sgn} f$ correspondingly. The key property of sign function is that

$$f \cdot \overline{\operatorname{sgn} f} = |f|, \qquad \left|\overline{\operatorname{sgn} f}\right| = 1$$

5.1 Hahn-Banach Theorem

The scalar field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Definition 5.1. Sublinear Functionals

Let X be a vector space. $p : X \to \mathbb{R}$ is called a sublinear functional, if

- (Subadditivity) $\forall\, x, y \in X :\ p(x + y) \leqslant p(x) + p(y)$;
- (Positive Homogeneity) $\forall\, x \in X\ \forall\, \lambda \geqslant 0 :\ p(\lambda x) = \lambda p(x)$.

Theorem 5.2. Hahn-Banach Theorem, general sublinear version

Let X be a **real** normed vector space, $Y \subseteq X$ be a subspace, and $p : X \to \mathbb{R}$ be a sublinear functional. Suppose that $f : Y \to \mathbb{R}$ is a linear functional such that

$$f(y) \leqslant p(y) \text{ for all } y \in Y$$

Then there exists a linear functional $g : X \to \mathbb{R}$ such that

$$g\big|_Y = f \text{ and } g(x) \leqslant p(x) \text{ for all } x \in X$$

Proof. The proof is consist of two parts.

- *One-step Extension Lemma.*

  Suppose that $U \subseteq X$ is a subspace. Let $\varphi : U \to \mathbb{R}$ be a linear functional such that $\varphi \leqslant p$ on U. Let $x_0 \in X \setminus U$ and $V := \operatorname{span}(U \cup \{x_0\})$. We shall extend φ to a linear functional $\psi : V \to \mathbb{R}$ such that $\psi \leqslant p$ on V.

  Note that every element $v \in V$ can be uniquely expressed as $v = u + \lambda x_0$ for some $u \in U$ and $\lambda \in \mathbb{R}$. We consider the family of linear functional $\{\psi_c : V \to \mathbb{R}\}_{c \in \mathbb{R}}$ such that $\psi_c(v) = \varphi(u) + c\lambda$. It is clear that

$\psi_c\big|_U = \varphi$. We shall prove that there exists $c \in \mathbb{R}$ such that $\psi_c \leqslant p$ on V.

For $\lambda = 0$, trivially we have $\psi_c \leqslant p$ on V for all $c \in \mathbb{R}$. For $\lambda > 0$,

$$\psi_c(v) \leqslant p(v) \iff \varphi(u) + c\lambda \leqslant p(u + \lambda x_0) \iff c \leqslant p\Big(\frac{u}{\lambda} + x_0\Big) - \varphi\Big(\frac{u}{\lambda}\Big)$$

We obtain a necessary condition: $c \leqslant \inf\{p(a + x_0) - \varphi(a) : \ a \in U\}$. For $\lambda < 0$, similarly we have

$$\psi_c(v) \leqslant p(v) \iff \varphi(u) + c\lambda \leqslant p(u + \lambda x_0) \iff c \geqslant \varphi\Big(\frac{u}{|\lambda|}\Big) - p\Big(\frac{u}{|\lambda|} - x_0\Big)$$

which gives another necessary condition: $c \geqslant \sup\{\varphi(b) - p(b - x_0) : \ b \in U\}$. Now for the existence of the extension, we need to verify that

$$\sup_{b \in U} (\varphi(b) - p(b - x_0)) \leqslant \inf_{a \in U} (p(a + x_0) - \varphi(a))$$

But it follows from that

$$\varphi(a) + \varphi(b) = \varphi(a + b) \leqslant p(a + b) \leqslant p(a + x_0) + p(b - x_0) \implies \varphi(b) - p(b - x_0) \leqslant p(a + x_0) - \varphi(a)$$

which holds for any $a, b \in U$.

- *Using Zorn's Lemma.*

  Consider the set

$$\mathscr{S} := \big\{(\varphi, D_\varphi) : \ Y \subseteq D_\varphi \subseteq X, \ \varphi : D_\varphi \to \mathbb{R} \text{ is a linear functional such that } \varphi\big|_Y = f \text{ and } \varphi \leqslant p \text{ on } D_\varphi\big\}$$

  equipped with the partial order

$$(\varphi, D_\varphi) \leqslant_{\mathscr{S}} (\psi, D_\psi) \iff D_\varphi \subseteq D_\psi \text{ and } \psi\big|_{D_\varphi} = \varphi$$

  It is clear that $\mathscr{S}$ is non-empty because $(f, Y) \in \mathscr{S}$. We can directly verify (by taking unions) that every chain in $\mathscr{S}$ has an upper bound. Hence by Zorn's Lemma, $\mathscr{S}$ has a maximal element (g, D_g). If $D_g \neq X$, we can pick $x_0 \in X \backslash D_g$ and use the one-step extension lemma to construct an extension of g on $\operatorname{span}(D_g \cup \{x_0\})$, which contradicts the maximality of (g, D_g). Therefore $D_g = X$ and g is the desired extension of f. ☐

Corollary 5.3. Hahn-Banach Extension Theorem

Let X be a normed vector space, $Y \subseteq X$ be a subspace, and $f \in Y^*$. Then there exists $g \in X^*$ such that

$$g\big|_Y = f \text{ and } \|g\|_{X^*} = \|f\|_{Y^*}$$

Proof.

- The scalar field $\mathbb{F} = \mathbb{R}$:

  We set $p(x) := \|f\|_{Y^*}\|x\|_X$ for $x \in X$. Then p is a sublinear functional on X and $f \leqslant p$ on Y. By the sublinear version of Hahn-Banach Theorem, we deduce that there exists an f-extension $g : X \to \mathbb{R}$ such that $g \leqslant p$ on X. Note that

$$\|f\|_{Y^*} \leqslant \|g\|_{X^*} = \sup_{\|x\|=1} |g(x)| \leqslant \sup_{\|x\|=1} |p(x)| = \|f\|_{Y^*}$$

- The scalar field $\mathbb{F} = \mathbb{C}$:

  For $f : Y \to \mathbb{C}$, we write $f = f_1 + \mathrm{i} f_2$, where $f_1, f_2 : Y \to \mathbb{R}$ are both $\mathbb{R}$-linear functionals. For $y \in Y$,

$$f(y) = f_1(y) + \mathrm{i} f_2(y), \qquad f(\mathrm{i}y) = \mathrm{i} f(y) = \mathrm{i} f_1(y) - f_2(y)$$

  Hence

$$f_2(y) = -f_1(\mathrm{i}y) \implies f(y) = f_1(y) - \mathrm{i} f_1(\mathrm{i}y)$$

  We regard $Y = Y_{\mathbb{R}}$ as a subspace of the normed $\mathbb{R}$-vector space X. And $f_1 \in Y_{\mathbb{R}}^*$. By the sublinear version of Hahn-Banach Theorem, f_1 extends to a bounded $\mathbb{R}$-linear functional $g_1 : X \to \mathbb{R}$ such that $g_1(x) \leqslant \|f\|\|x\|$

on X. We claim that $g(x) := g_1(x) - \mathrm{i}g_1(\mathrm{i}x)$ is a bounded $\mathbb{C}$-linear functional on X, which is the desired extension of g.

It is straightforward to check the $\mathbb{C}$-linearity of g. For the estimate of norm, let $x \in X$. Therefore

$$|g(x)| = g(\overline{\operatorname{sgn} g(x)} \cdot x) = g_1(\overline{\operatorname{sgn} g(x)} \cdot x) \leqslant \|g_1\| \left\| \overline{\operatorname{sgn} g(x)} \cdot x \right\| = \|g_1\|\|x\| \leqslant \|f\|\|x\|$$ □

Proposition 5.4

Suppose that $(X, \|\cdot\|_X)$ is a normed vector space. For $x \in X \setminus \{0\}$ there exists $\ell \in X^*$ such that $\|\ell\|_* = 1$ and $\ell(x) = \|x\|$.

Proof. Define $g : \operatorname{span}\{x\} \to F$ by $g(\lambda x) = \lambda\|x\|$. Then $\|g\| = 1$ and $g(x) = \|x\|$. By Hahn-Banach Theorem g extends to $\ell \in X^*$ such that $\|\ell\|_* = \|g\|_{\operatorname{span}\{x\}^*} = 1$. □

Corollary 5.5

Suppose that $(X, \|\cdot\|_X)$ is a normed vector space. Then

$$\|x\|_X = \sup\{|\ell(x)| : \ \ell \in X^*, \ \|\ell\|_* = 1\}, \qquad \|\ell\|_* = \sup\{|\ell(x)| : \ x \in X, \ \|x\|_X = 1\}$$

Remark. Note that the second supremum is in general not attained, whereas the first supremum is always attained by the previous proposition.

Corollary 5.6. Dual Space Separates Points

Suppose that $(X, \|\cdot\|_X)$ is a normed vector space. For $x, y \in X$ with $x \neq y$, there exists $\ell \in X^*$ such that $\ell(x) \neq \ell(y)$.

Proof. Apply Proposition 5.4 to the point $x - y \in X \setminus \{0\}$. □

5.2 Separation of Convex Sets

In this section we look at a geometric application of the Hahn-Banach Theorem.

Definition 5.7. Minkowski Functional (*Off-Syllabus*)

Suppose that X is a normed vector space and $K \subseteq X$ is a convex subset such that $0 \in \operatorname{int}(K)$. The Minkowski functional on K is $p_K : X \to \mathbb{R}$, where

$$p_K(x) = \inf\left\{\lambda > 0 : \ \lambda^{-1}x \in K\right\}$$

Remark. It is clear that $p_K(x) \leqslant 1$ for $x \in K$ and $p_K(x) \geqslant 1$ for $x \notin K$.

Proposition 5.8. Sublinearity of Minkowski Functional (*Off-Syllabus*)

The Minkowski functional $p_K : X \to \mathbb{R}$ is sublinear.

Proof. The homogeneity is trivial. We prove subadditivity. Let $x, y \in X$. Consider $S_x := \{\lambda > 0 : \lambda^{-1}x \in K\}$ and $S_y := \{\mu > 0 : \mu^{-1}y \in K\}$. For $\lambda \in S_x$ and $\mu \in S_y$, by convexity of K,

$$\frac{x+y}{\lambda+\mu} = \frac{\lambda}{\lambda+\mu}\lambda^{-1}x + \frac{\mu}{\lambda+\mu}\mu^{-1}y \in K$$

Therefore $p_K(x+y) \leqslant \lambda + \mu$. Taking the infimum over $\lambda \in S_x$ and $\mu \in S_y$ we deduce that $p_K(x+y) \leqslant p_K(x) + p_K(y)$. □

Remark. We say that $K \subseteq X$ is **symmetric**, if $x \in K$ implies that $\lambda x \in K$ for all $\lambda \in \mathbb{F}$ with $|\lambda| = 1$. We say that $K \subseteq X$ contains no **infinite rays**, if K contains no subset of the form $\{\lambda x : \lambda > 0\}$ for any $x \in K$. Besides the conditions listed in the proposition, if K is symmetric, then p_K is a seminorm on X; if in addition K contains no infinite rays, then p_K is a norm on X.

Theorem 5.9. Hahn-Banach Separation Theorem / Hyperplane Separation Theorem (*Proof Off-Syllabus*)

Suppose that X is a normed vector space. Let $A, B \subseteq X$ be disjoint convex subsets of X, where A is closed and B is compact. Then there exists $\ell \in X^*$ such that

$$\sup_{x\in A} \operatorname{Re} \ell(x) < \inf_{x\in B} \operatorname{Re} \ell(x)$$

Proof.

- The scalar field $\mathbb{F} = \mathbb{R}$ (the Re can be dropped in the theorem).

  Since A is closed and B is compact, from *A2 Metric Spaces* we know that

  $$\operatorname{dist}(A, B) := \inf\{\|x - y\| : x \in A, y \in B\} > 0$$

  So there exists $\varepsilon > 0$ such that $A_\varepsilon \cap B = \varnothing$ where $A_\varepsilon := \{x \in X : \exists y \in A \ \|x - y\| < \varepsilon\}$. Fix $x_0 \in A_\varepsilon$ and $y_0 \in B$. Let $z_0 := x_0 - y_0$ and

  $$M := A_\varepsilon - B + z_0 := \{x - y + z_0 \in X : x \in A_\varepsilon, \ y \in B\}$$

  It is clear that M is convex and $B(0, \varepsilon) \subseteq M$. Then the Minkowski functional p_M is sublinear and $p_M(x) \leqslant \varepsilon^{-1}\|x\|$.

  Let $g : \operatorname{span}\{z_0\} \to \mathbb{R}$ given by $g(\lambda z_0) = \lambda$ for $\lambda \in \mathbb{R}$. Since $z_0 \notin M$, $p_M(z_0) \geqslant 1 = g(z_0)$ and hence $g \leqslant p$ on $\operatorname{span}\{z_0\}$. By Hahn-Banach Theorem, there exists an extension $\ell : X \to \mathbb{R}$ of g such that $\ell \leqslant p_M$ on $\mathbb{R}$. Then $\ell(x) \leqslant p_M(x) \leqslant \varepsilon^{-1}\|x\|$ and hence $\ell \in X^*$.

  Let $\delta > 0$ such that $\delta\|z_0\| < \varepsilon$. Then for any $x \in A$ and $y \in B$, $x + \delta z_0 \in A_\varepsilon$ and $x + \delta z_0 - y + z_0 \in M$. Then $p_M(x + \delta z_0 - y + z_0) \leqslant 1$. We have

  $$\ell(x) + \delta - \ell(y) = \ell(x + \delta z_0 - y + z_0) - \ell(z_0) \leqslant p_M(x + \delta z_0 - y + z_0) - 1 \leqslant 0$$

  Taking supremum over $x \in A$ and $y \in B$ we obtain that

  $$\inf_{y\in B} \ell(y) - \sup_{x\in A} \ell(x) \geqslant \delta > 0$$

- The scalar field $\mathbb{F} = \mathbb{C}$.

  We first consider X as a real normed vector space and proceed as above to obtain a bounded $\mathbb{R}$-linear functional ℓ_0 such that $\inf_{x\in B} \ell_0(x) > \sup_{x\in A} \ell_0(x)$. Then we take $\ell \in X^*$ with $\ell(x) := \ell_0(x) - \mathrm{i}\ell_0(\mathrm{i}x)$, so that $\ell_0 = \operatorname{Re} \ell$. □

Remark. For $\ell \in X^* \setminus \{0\}$, $\{x \in X : \ell(x) = \lambda\}$ is an affine subspace of X of codimension 1, so it is called a **hyperplane** of X. The separation theorem says that we can separate the two disjoint convex sets A and B by a hyperplane $\{x \in X : \ell(x) = \lambda\}$, where $\sup_{x\in A} \ell(x) < \lambda < \inf_{x\in B} \ell(x)$.

By slightly modifying the above proof (in fact simpler) we have the following result.

Theorem 5.10. Hyperplane Separation Theorem (*Proof Off-Syllabus*)

Suppose that X is a normed vector space. Let $A, B \subseteq X$ be disjoint convex subsets of X, one of which has an interior point. Then there exists $\ell \in X^*$ such that

$$\sup_{x\in A} \operatorname{Re} \ell(x) \leqslant \inf_{x\in B} \operatorname{Re} \ell(x)$$

Definition 5.11. Annihilators

Suppose that X is a normed vector space.

- For $M \subseteq X$, the annihilator of M is

$$M^\circ := \{\ell \in X^* : \ell\big|_M = 0\}$$

  For $N \subseteq X^*$, the annihilator of N is

$$N_\circ := \{x \in X : \forall \ell \in N \ \ell(x) = 0\} = \bigcap_{\ell \in N} \ker \ell$$

Proposition 5.12

Suppose that X is a real normed vector space and $Y \subseteq X$ is a closed subspace. Let $x_0 \in X \setminus Y$. Then there exists $\ell \in X^*$ such that $\|\ell\|_* = 1$, $\ell\big|_Y = 0$, and $\ell(x_0) = \operatorname{dist}(x_0, Y)$.

Proof. Define $g : \operatorname{span}\{x_0\} \to \mathbb{R}$ by $g(\lambda x_0) = \lambda \operatorname{dist}(x_0, Y)$. It is dominated by the seminorm $\operatorname{dist}(\cdot, Y) : X \to \mathbb{R}$. Hence we can use Hahn-Banach Theorem to extend it to a linear functional $\ell : X \to \mathbb{R}$, which is bounded, as $\ell(x) \leqslant \operatorname{dist}(x, Y) \leqslant \|x\|$. We know that $\ell\big|_Y \leqslant \operatorname{dist}(\cdot, Y)\big|_Y = 0$ and $\ell(x_0) = g(x_0) = \operatorname{dist}(x_0, Y)$.

Finally we need to show that $\|\ell\| \geqslant 1$. For $\varepsilon \in (0,1)$, there exists $y \in Y$ such that $\|x_0 - y\| < \dfrac{1}{1-\varepsilon} \operatorname{dist}(x_0, Y)$. Then

$$\frac{|\ell(x_0 - y)|}{\|x_0 - y\|} = \frac{\operatorname{dist}(x_0, Y)}{\|x_0, y\|} \geqslant \frac{1}{1-\varepsilon}$$

We conclude the proof by taking $\varepsilon \to 0$. ☐

Corollary 5.13

Suppose that X is a real normed vector space.

1. Let $M \subseteq X$. $\operatorname{span} M$ is dense in X *if and only if* $M^\circ = \{0\}$;
2. Let $N \subseteq X^*$. If $\operatorname{span} N$ is dense in X^*, then $N_\circ = \{0\}$;
3. Let Y be a subspace of X. Then $\overline{Y} = (Y^\circ)_\circ$.

Proof.

1. The forward direction follows from that

$$M^\circ = (\operatorname{span} M)^\circ = (\overline{\operatorname{span} M})^\circ = X^\circ = 0$$

   For the backward direction, suppose that $\operatorname{span} M$ is not dense in X. Then we can take $x_0 \in X \setminus \overline{\operatorname{span} M}$ and use the previous proposition to find $\ell \in X^* \setminus \{0\}$ such that $\ell \in M^\circ$.

2. Suppose that $x_0 \in N_\circ \setminus \{0\}$. By Corollary 5.6, there exists $\ell \in X^*$ such that $\ell(x_0) \neq 0$. By density there exists $(\ell_n) \subseteq \operatorname{span} N$ such that $\ell_n \to \ell$ in X^*. In particular $\ell_n(x) \to \ell(x)$ for any $x \in X$. Since $\ell_n(x_0) = 0$ for all $n \in \mathbb{N}$, we have $\ell(x_0) = 0$, which is contradictory. Hence $N_\circ = \{0\}$.

3. Note that $(Y^\circ)_\circ$ is closed, because it is the interchapter of kernels of some bounded linear functionals. It is clear that $Y \subseteq (Y^\circ)_\circ$, so $\overline{Y} \subseteq (Y^\circ)_\circ$. On the other hand, if $x \notin \overline{Y}$, then by the previous proposition, there exists $\ell \in \overline{Y}^\circ \subseteq Y^\circ$ such that $\ell(x) \neq 0$. Hence $x \notin (Y^\circ)_\circ$. We deduce that $(Y^\circ)_\circ \subseteq \overline{Y}$. ☐

5.3 Biduals and Reflexivity

Let X be a normed vector space. Then X^* is a Banach space. The dual space X^{**} of X^* is called the **bidual** of X.

From *A0. Linear Algebra* we know that a vector space X is naturally (linearly) isomorphic to its algebraic bidual X'' via the map $J_X : X \to X''$, given by $J_X(x)(\ell) := \ell(x)$ for any $x \in X$ and $\ell \in X'$.

It is not necessary that X is isomorphic to X^{**}. It is clear that $J_X : X \to X^{**}$ is injective. By Corollary 5.5, J_X is isometric:

$$\|J_X(x)\|_{**} = \sup\{|J_X(x)(\ell)| : \ \ell \in X^*, \ \|\ell\|_* = 1\} = \sup\{|\ell(x)| : \ \ell \in X^*, \ \|\ell\|_* = 1\} = \|x\|$$

Definition 5.14. Reflexive Normed Vector Space

Let $(X, \|\cdot\|)$ be a normed vector space. We say that X is reflexive if $J_X : X \to X^{**}$ is surjective. In such case X is isometrically isomorphic to X^{**}

Remark. The converse is NOT true! There exists a non-reflexive Banach space that is isometrically isomorphic to its bidual.

Proposition 5.15. Reflexivity of Hilbert Spaces

Every Hilbert space is reflexive.

Proof. Immediate by Riesz-Fréchet Representation Theorem. □

Recall from *A2. Metric Spaces* that (Y, J) is called a **completion** of the metric space X, if Y is a complete metric space, and $J : X \to Y$ is a isometric embedding such that $J(X)$ is dense in Y.

Proposition 5.16. Existence of Linear Completion

Every normed vector space X has a completion (Y, J), where Y is a Banach space and J is a linear isometric embedding.

Proof. Following the above discussion, we choose $J = J_X$ and $Y := \overline{J_X(X)}$. Since Y is a closed subspace of X^{**}, Y is a Banach space. It is trivial that $J_X(X)$ is dense in Y. □

Proposition 5.17. (*Off-Syllabus*)

Suppose that X and Y are isometrically isomorphic normed vector spaces. Then X is reflexive if and only if Y is reflexive.

Proof. It follows directly from the following commutative diagram:

$$\begin{array}{ccc} X^{**} & \xrightarrow{T^{**}} & Y^{**} \\ {\scriptstyle J_X}\big\updownarrow & & {\scriptstyle J_Y}\big\updownarrow \\ X & \xrightarrow{T} & Y \end{array}$$

□

Proposition 5.18. (*Off-Syllabus*)

Let X be a normed vector space. X is reflexive if and only if X is complete and X^* is reflexive.

Proof. $\Longrightarrow$: Suppose that X is reflexive. Then $X \cong X^{**}$. Since X^{**} is complete, so is X. For $\xi \in X^{***}$, let $\ell = \xi \circ J_X \in X^*$. Since J_X is surjective, for $\varphi \in X^{**}$ there exists $x \in X$ such that $\varphi = J_X(x)$. We have

$$\xi(\varphi) = \xi \circ J_X(x) = \ell(x) = J_X(x)(\ell) = \varphi(\ell) = J_{X^*}(\ell)(\varphi)$$

Hence $\xi = J_{X^*}(\ell)$. We deduce that J_{X^*} is surjective and hence X^* is reflexive.

$\Longleftarrow$: Suppose that X is complete and X^* is reflexive. Then $Y := J_X(X)$ is complete and hence closed in X^{**}. Since J_{X^*} is surjective, for $\xi \in Y^\circ$ there exists $\ell \in X^*$ such that $\xi = J_{X^*}(\ell)$. We have

$$(\forall x \in X : \ell(x) = \xi \circ J_X(x) = 0) \implies \ell = 0 \implies \xi = 0 \implies Y^\circ = \{0\}$$

By Corollary 5.13, Y is dense in X^{**}. But Y is closed. We have $Y = J_X(X) = X^{**}$. X is reflexive. $\square$

Definition 5.19. Uniform Convexity (*Off-Syllabus*)

Let X be a normed vector space. Let S_X denotes the unit sphere in X. X is said to be uniformly convex, if

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall x, y \in S_X \, \left(\|x - y\| \geqslant \varepsilon \implies \frac{1}{2}\|x + y\| \leqslant 1 - \delta \right)$$

Proposition 5.20. Uniform Convexity of Inner Product Spaces

Every inner product space is uniformly convex.

Proof. Suppose that X is an inner product space. For $\varepsilon > 0$ and $x, y \in S_X$ with $\|x - y\| < \varepsilon$, by the Parallelogram Identity we have

$$\|x + y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) - \|x - y\|^2 \leqslant 4 - \varepsilon^2$$

We take $\delta = 1 - \dfrac{\sqrt{4 - \varepsilon^2}}{2}$. Then $\dfrac{1}{2}\|x + y\| \leqslant 1 - \delta$. $\square$

Proposition 5.21. (*Off-Syllabus*)

Suppose that X is a uniformly convex Banach space. For each $\ell \in X^* \setminus \{0\}$ there exists a unique $x \in X$ such that $\|x\| = 1$ and $\ell(x) = \|\ell\|_*$.

Proof.

- Existence: Let S_X denotes the unit sphere in X. Let $(x_n) \subseteq S_X$ be such that $f(x_n) \to \|f\|$ as $n \to \infty$. We show that (x_n) is a Cauchy sequence. Suppose not. Then there exist $\varepsilon > 0$ and increasing sequences of integers $(n_k), (m_k)$ such that $\|x_{n_k} - x_{m_k}\| \geqslant \varepsilon$ for all $k \geqslant 1$. Let $y_k = \dfrac{1}{2}(x_{n_k} + x_{m_k})$. Then $\|y_k\| \leqslant 1$ for all $k \geqslant 1$ and $f(y_k) \to \|f\|$ as $k \to \infty$. It follows that $\|y_k\| \to 1$ as $k \to \infty$, so by uniform convexity $\|x_{n_k} - x_{m_k}\| \to 0$ as $k \to \infty$, giving the required contradiction.
- Uniqueness: Suppose that $x, y \in S_X$ are two distinct vectors such that $f(x) = f(y) = \|f\|$. By uniform convexity we must have $\frac{1}{2}\|x + y\| < 1$ and hence

$$\|f\| = f\left(\frac{x + y}{2}\right) \leqslant \|f\| \frac{\|x + y\|}{2} < \|f\|$$

This contradiction completes the proof. $\square$

Remark. The proposition says that in a uniformly convex space every functional attains its norm at a unique vector in the unit sphere. We can compare the result with Proposition 5.4.

Proposition 5.22. Convexity & Reflexivity (*Off-Syllabus*)

A uniformly convex Banach space is reflexive.

Remark. This is a strong result. For the proof one can refer to Question C.4 of Sheet 3 of *C4.1 Further Functional Analysis (2020-2021)*. We will not make use of the result subsequently.

5.4 Dual of Lebesgue Spaces

Next we shall establish the dual space of some particular Banach spaces.

Example 5.23. Dual Space of ℓ^p

Let $1 \leqslant p < \infty$. Let q be the Hölder conjugate of p, i.e. $p^{-1} + q^{-1} = 1$.

1. $(\ell^p)^* \cong \ell^q$;
2. $(c_0)^* \cong \ell^1$.

Proof. 1. We define $\Phi_p : \ell^q \to (\ell^p)^*$ by

$$\Phi_p(x)(y) := \sum_{n\in\mathbb{N}} x_n y_n$$

We shall show that Φ_p is a isometric isomorphism. It is well-defined by Hölder's Inequality:

$$|\Phi_p(x)(y)| \leqslant \sum_{n\in\mathbb{N}} |x_n y_n| \leqslant \|x\|_q \|y\|_p < \infty$$

It is clear that Φ_p is linear.

Surjectivity: Suppose that $f \in (\ell^p)^*$. Let $e^{(j)}$ be the sequence such that $e^{(j)}_k = \delta_{jk}$. Let $x_j := f(e^{(j)})$ and $x = (x_j)_{j\in\mathbb{N}}$. We claim that $x \in \ell^q$.

- $1 < p < \infty$:

  Consider the sequences $x^{(n)} \in \ell^q$ and $y^{(n)} \in \ell^p$ given by

$$x^{(n)} = \sum_{j=0}^{n} x_j e^{(j)}, \qquad y^{(n)} = \sum_{j=0}^{n} |x_j|^{q-1}\, \overline{\operatorname{sgn} x_j}\, e^{(j)}$$

  Then $x^{(n)}y^{(n)} = \sum_{j=0}^{n} |x_j|^q e^{(j)}$ and

$$\left\|y^{(n)}\right\|_p = \left(\sum_{j=0}^{n} |x_j|^{pq-p}\right)^{1/p} = \left(\sum_{j=0}^{n} |x_j|^q\right)^{1-1/q} = \left\|x^{(n)}\right\|_q^{q-1}$$

  And

$$f(y^{(n)}) = \sum_{j=0}^{n} |x_j|^{q-1}\, \overline{\operatorname{sgn} x_j}\, f(e^{(j)}) = \sum_{j=0}^{n} |x_j|^q = \left\|x^{(n)}\right\|_q^q$$

  We have

$$\left(\sum_{j=0}^{n} |x_j|^q\right)^{1/q} = \left\|x^{(n)}\right\|_q = \frac{\left\|x^{(n)}\right\|_q^q}{\left\|x^{(n)}\right\|_q^{q-1}} = \frac{f(y^{(n)})}{\left\|y^{(n)}\right\|_p} \leqslant \|f\|_{p*}$$

  Hence $\|x\|_q \leqslant \|f\|_{p*}$ and $x \in \ell^q$.

- $p = 1$:

  We have

$$x_j = f(e^{(j)}) \leqslant \|f\|_{p*} \left\|e^{(j)}\right\|_p = \|f\|_{p*} < \infty$$

  Hence x is bounded and $x \in \ell^\infty$.

We note that $\Phi(x)(e^{(j)}) = x_j = f(e^{(j)})$. Therefore $\Phi(x) = f$ on $\operatorname{span}\{e^{(j)}\}_{j\in\mathbb{N}}$. But $\operatorname{span}\{e^{(j)}\}_{j\in\mathbb{N}}$ is dense in ℓ^p. We deduce that $\Phi(x) = f$ on ℓ^p. Φ is surjective.

Isometry: For $x \in \ell^q$, by Hölder's Inequality, we have

$$\|\Phi_p(x)\|_{p*} = \sup_{y\in\ell^p\setminus\{0\}} \frac{|\Phi_p(x)(y)|}{\|y\|_p} \leqslant \sup_{y\in\ell^p\setminus\{0\}} \frac{\|x\|_q \|y\|_p}{\|y\|_p} = \|x\|_q$$

For the reverse direction,

- $1 < p < \infty$:

  We construct $y = (y_j)_{j\in\mathbb{N}}$ as in the proof of surjectivity:

  $$y_j = |x_j|^{q-1}\,\overline{\operatorname{sgn} x_j}$$

  So $\|y\|_p = \|x\|_q^{q-1}$ and hence $y \in \ell^p$. Now

  $$\|x\|_q\|y\|_p = \|x\|_q^q = \Phi_p(x)(y) \leqslant \|\Phi_p(x)\|_{p*}\|y\|_p$$

  So $\|x\|_q \leqslant \|\Phi_p(x)\|_{p*}$.

- $p = 1$:

  Fix $\varepsilon > 0$. Let $N \in \mathbb{N}$ such that $\max\limits_{1\leqslant j\leqslant N}|x_j| > \|x\|_\infty - \varepsilon$. Let $y \in \ell^1$ be given by

  $$y = \overline{\operatorname{sgn} x_k}\, e^{(k)}$$

  where $k \in \{1, ..., N\}$ such that $|x_k| = \max\limits_{1\leqslant j\leqslant N}|x_j|$. Then

  $$\|\Phi_1(x)\|_{1*} = \|\Phi_1(x)\|_{1*}\|y\|_1 \geqslant \Phi_1(x)(y) = \max_{1\leqslant j\leqslant N}|x_j| \geqslant \|x\|_\infty - \varepsilon$$

  Since ε is arbitrary we deduce that $\|x\|_\infty \leqslant \|\Phi_1(x)\|_{1*}$

This completes the proof.

2. We define $\Psi : \ell^1 \to (c_0)^*$ by

$$\Psi(x)(y) = \sum_{n\in\mathbb{N}} x_n y_n$$

The proof of surjectivity and isometry are exactly the same as (1), with $q = 1$.

The only thing to note is that $\operatorname{span}\{e^{(j)}\}_{j\in\mathbb{N}}$ is dense in c_0 but not in ℓ^∞. Therefore we cannot prove that $\Phi_\infty : \ell^1 \to (\ell^\infty)^*$ is surjective. □

Example 5.24. Dual Spaces of $\mathrm{L}^p(\Omega)$ (*Proof of Surjectivity Off-Syllabus*)

Let $\Omega \subseteq \mathbb{R}^n$ be a measurable set. Let $1 \leqslant p < \infty$. Let q be the Hölder conjugate of p, i.e. $p^{-1} + q^{-1} = 1$. Then $\mathrm{L}^p(\Omega)^* \cong \mathrm{L}^q(\Omega)$. [a]

[a] If we are considering the general $\mathrm{L}^p(\Omega, \mu)$, then the σ-finiteness of μ is required only for $p = 1$.

Proof. For simplicity we only consider the scalar field $\mathbb{F} = \mathbb{R}$.

As for the sequence spaces, we define $\Phi_p : \mathrm{L}^q(\Omega) \to \mathrm{L}^p(\Omega)^*$ by

$$\Phi_p(f)(g) = \int_\Omega f \cdot g$$

The proof of isometry of Φ_p is essentially the same as for sequence spaces. When $p = 1$, we need to replace the $e^{(k)}$ by suitable indicator functions $\mathbf{1}_{A_k}$. The main work is to prove surjectivity.

- First we consider the case where $m(\Omega) < \infty$.

  Let $\xi \in \mathrm{L}^p(\Omega)^*$. We define $\nu : \mathcal{M}_{\mathtt{Leb}}(\Omega) \to \mathbb{R}$ by $\nu(E) := \xi(\mathbf{1}_E)$. This is well-defined because every measurable set E has finite measure and hence $\mathbf{1}_E \in \mathrm{L}^p(\Omega)$. We claim that ν is a signed measure. Indeed, let $\{E_k\}_{k=0}^\infty$ be a countable disjoint collection of measurable sets and $E = \bigcup_k E_k$. By the countable additivity of the Lebesgue measure,

  $$m(E) = \sum_{k=0}^\infty m(E_k) < \infty \implies \lim_{n\to\infty} \sum_{k=n+1}^\infty m(E_k) = 0$$

Consequently,

$$\lim_{n\to\infty}\left\|\mathbf{1}_E-\sum_{k=1}^{n}\mathbf{1}_{E_k}\right\|_p=\lim_{n\to\infty}\left(\sum_{k=n+1}^{\infty}m\left(E_k\right)\right)^{1/p}=0)$$

But ξ is both linear and continuous on $\mathrm{L}^p(\Omega)$ and hence

$$\xi\left(\chi_E\right)=\sum_{k=0}^{\infty}\xi\left(\chi_{E_k}\right)\implies\nu(E)=\sum_{k=0}^{\infty}\nu\left(E_k\right)$$

To show that ν is a signed measure it must be shown that the series on the right hand side converges absolutely. We set $c_k=\operatorname{sgn}\xi(\mathbf{1}_{E_k})$. Then arguing as above we conclude that the series $\sum_{k=0}^{\infty}\left|\nu\left(E_k\right)\right|=\sum_{k=0}^{\infty}\xi\left(c_k\mathbf{1}_{E_k}\right)$ is Cauchy and thus convergent. Thus ν is a signed measure. Moreover, it is clear that $|\nu|\ll m$.

By Radon-Nikodym Theorem (Corollary 2.29), there exists a function $f:\Omega\to\mathbb{R}$ such that

$$\xi(\mathbf{1}_E)=\nu(E)=\int_E f\,\mathrm{d}x$$

By linearity we have

$$\xi(\varphi)=\int_\Omega f\varphi\,\mathrm{d}x$$

for all simple functions $\varphi:\Omega\to\mathbb{R}$.

- We claim that $f\in\mathrm{L}^q(\Omega)$. Since ξ is bounded, we have

$$\left|\int_\Omega f\varphi\,\mathrm{d}x\right|=|\xi(\varphi)|\leqslant\|\xi\|_{p*}\|\varphi\|_p$$

  - $1<p<\infty$:

    Let (φ_n) be a sequence of non-negative simple functions such that $\varphi_n\nearrow|f|$ pointwise. Let $\psi_n:=\overline{\operatorname{sgn}f}\cdot\varphi_n^{q-1}$. Then ψ_n is also a simple function with $|f|\varphi_n^{q-1}=f\psi_n$ and

$$\|\psi_n\|_p^p=\int_\Omega|\psi_n|^p=\int_\Omega|\varphi_n|^{p(q-1)}=\|\varphi_n\|_q^q$$

    We have

$$\|\varphi_n\|_q^q=\|\varphi_n\|_q^q=\left|\int_\Omega\varphi_n^q\right|\leqslant\left|\int_\Omega|f|\varphi_n^{q-1}\right|=\left|\int_\Omega f\psi_n\right|\leqslant\|\xi\|_{p*}\|\psi_n\|_p=\|\xi\|_{p*}\|\varphi_n\|_q^{q/p}$$

    Therefore

$$\|\varphi_n\|_q=\|\varphi_n\|_q^{q-q/p}\leqslant\|\xi\|_{p*}$$

    Finally, by Fatou's Lemma,

$$\|f\|_q^q=\left|\int_\Omega|f|^q\right|\leqslant\liminf_{n\to\infty}\int_\Omega\varphi_n^q\leqslant\|\xi\|_{p*}^q$$

    We deduce that $\|f\|_q\leqslant\|\xi^*\|_p$. In particular $f\in\mathrm{L}^q(\Omega)$.

  - For $p=1$:

    Supoose that $\|f\|_\infty>\|\xi\|_{1*}$. Then $E:=\{x\in\Omega:\ |f(x)|>\|\xi\|_{1*}+\varepsilon\}$ has positive measure. But

$$m(E)\|\xi\|_{1*}<\left|\int_E|f|\right|=\left|\int_\Omega f\,\overline{\operatorname{sgn}f}\,\mathbf{1}_E\right|\leqslant\|\xi\|_{1*}\left\|\overline{\operatorname{sgn}f}\,\mathbf{1}_E\right\|_1=m(E)\|\xi\|_{1*}$$

    which is contradictory. Hence $\|f\|_\infty\leqslant\|\xi\|_{1*}$ and $f\in\mathrm{L}^\infty(\Omega)$.

Now we use the fact that the set of simple functions is dense in $\mathrm{L}^p(\Omega)$ to conclude that

$$\xi(g) = \int_\Omega fg \,\mathrm{d}x$$

for all $g \in \mathrm{L}^p(\Omega)$. Φ_p is surjective.

- Next we consider general measurable set $\Omega \subseteq \mathbb{R}^n$. We use the fact the Lebesgue measure is σ-finite. So let (Ω_n) be an ascending sequence of measurable sets such that $\Omega = \bigcup_n \Omega_n$ and $m(\Omega_n) < \infty$. From above we have shown that there exists $f_n : \Omega_n \to \mathbb{R}$ such that $\|f_n\|_q \leqslant \|\xi\|_{p*}$, and for $g \in \mathrm{L}^p(\Omega_n) \subseteq \mathrm{L}^p(\Omega)$,

$$\xi(g) = \int_{\Omega_n} f_n g \,\mathrm{d}x$$

From $\Omega_n \subseteq \Omega_{n+1}$ we may take $f_{n+1}\big|_{\Omega_n} = f_n$. Let $f = \bigcup_n f_n$. Then by Fatou's Lemma,

$$\int_\Omega |f|^q \leqslant \liminf_{n\to\infty} \int_{\Omega_n} |f_n|^q \leqslant \|\xi\|_{p*}$$

Hence $f \in \mathrm{L}^q(\Omega)$. For $g \in \mathrm{L}^p(\Omega)$, we take $g_n := g\big|_{\Omega_n} \in \mathrm{L}^p(\Omega_n)$. By Dominated Convergence Theorem,

$$\lim_{n\to\infty} \int_\Omega |g_n - g|^p = \int_\Omega \lim_{n\to\infty} |g|^p |\mathbf{1}_\Omega - \mathbf{1}_{\Omega_n}|^p = 0$$

Hence $g_n \to g$ in $\mathrm{L}^p(\Omega)$. Finally,

$$\begin{aligned} \xi(g) &= \lim_{n\to\infty} \xi(g_n) && \text{(continuity of } \xi) \\ &= \lim_{n\to\infty} \int_\Omega f_n g_n = \lim_{n\to\infty} \int_\Omega f g_n = \lim_{n\to\infty} \int_\Omega fg && \text{(Dominated Convergence Theorem)} \end{aligned}$$

We deduce that $\Phi_p(f) = \xi$. Φ_p is surjective. □

Corollary 5.25. Reflexive Banach Spaces

Let $1 < p < \infty$. Then ℓ^p and $\mathrm{L}^p(\Omega)$ are reflexive.

Proof. We have shown that $\Phi_p : \ell^q \to (\ell^p)^*$ is an isometric isomorphism. Then for $x \in \ell^p$ and $f = \Phi_p(y) \in (\ell^p)^*$

$$J_{\ell^p}(x)(f) = f(x) = \Phi_p(y)(x) = \sum_{n=0}^{\infty} x_n y_n = \Phi_q(x)(y) = \Phi_q(x)(\Phi_p^{-1}(f)) = ((\Phi_p^{-1})' \circ \Phi_q)(x)(f)$$

Hence $J_{\ell^p} = (\Phi_p^{-1})' \circ \Phi_q$ is an isometric isomorphism. ℓ^p is reflexive. The proof for $\mathrm{L}^p(\Omega)$ is identical. □

Proposition 5.26. Dual Spaces and Separability (*Off-Syllabus*)

Suppose that X is a normed vector space. If X^* is separable, then X is separable.

Proof. Let (f_n) be a dense countable subset of S_{X^*}, the unit sphere of X^*. For each $n \in \mathbb{N}$, since

$$1 = \|f_n\|_* = \sup_{\|x\| \leqslant 1} |f_n(x)|$$

We choose $x_n \in X$ with $|f_n(x_n)| > 1/2$ and $\|x_n\| \leqslant 1$. We claim that $Y := \operatorname{span}\{x_n\}_{n\in\mathbb{N}}$ is dense in X. Suppose not. Then by Corollary 5.13.1, $Y^\circ \neq \{0\}$. We take $f \in S_{X^*} \cap Y^\circ$. There exists $n \in \mathbb{N}$ such that $\|f - f_n\|_* < 1/2$. But

$$0 < |f(x_n)| \geqslant |f_n(x_n)| - |(f - f_n)(x_n)| < 1/2 - \|f - f_n\|_* < 1/2 - 1/2 = 0$$

which is contradictory. We conclude that X is separable. □

Corollary 5.27. Non-Reflexive Banach Spaces (*Proof Off-Syllabus*)

$(\ell^\infty)^* \ncong \ell^1$ and $\mathrm{L}^\infty(\Omega)^* \ncong \mathrm{L}^1(\Omega)$. In particular, the spaces ℓ^1, ℓ^∞, $\mathrm{L}^1(\Omega)$, and $\mathrm{L}^\infty(\Omega)$ are non-reflexive.

Proof. We know that ℓ^1 and $\mathrm{L}^1(\Omega)$ are separable, but ℓ^∞ and $\mathrm{L}^\infty(\Omega)$ are not. Therefore by the previous proposition $(\ell^\infty)^* \ncong \ell^1$ and $\mathrm{L}^\infty(\Omega)^* \ncong \mathrm{L}^1(\Omega)$.

If ℓ^1 is reflexive, then $\ell^1 \cong (\ell^1)^{**} \cong (\ell^\infty)^*$, which is contradictory. Similarly, $\mathrm{L}^1(\Omega)$ is not reflexive. Finally, since $(\ell^1)^* \cong \ell^\infty$ and $\mathrm{L}^1(\Omega)^* \cong \mathrm{L}^\infty(\Omega)$, by Proposition 5.18, ℓ^∞ and $\mathrm{L}^\infty(\Omega)$ are not reflexive. □

5.5 Weak and Weak* Convergence

Definition 5.28. Weak Convergence

Suppose that $(X, \|\cdot\|)$ is a normed vector space. For $(x_n)_{n\in\mathbb{N}} \subseteq X$, we say that x_n converges weakly to $x \in X$, if $\ell(x_n) \to \ell(x)$ as $n \to \infty$ for all $\ell \in X^*$. The weak convergence is denoted by a half arrow: $x_n \rightharpoonup x$.

Remark. Note that, since X^* separates points in X, the weak limit of (x_n) is necessarily unique.

Remark. To stress the difference, the usual convergence $\|x_n - x\| \to 0$ is called the **strong convergence**.

The weak convergence of sequences induces a topology on X, namely the weak topology:

Definition 5.29. Weak Topology (*Off-Syllabus*)

The weak topology on X is the coarsest topology on X with respect to which every $\ell \in X^*$ is continuous. The topology is denoted by $\sigma(X, X^*)$.

Remark. More explicitly, the weak topology on X is generated by the open sets

$$\{x \in X : \ |\ell_k(x - x_0)| < \varepsilon, \ 1 \leqslant k \leqslant n\}$$

where $x_0 \in X$, $n \in \mathbb{Z}_+$, and $\varepsilon > 0$.

Proposition 5.30. Strong Convergence $\implies$ Weak Convergence

Suppose that $(X, \|\cdot\|)$ is a normed vector space. If $(x_n) \subseteq X$ such that $x_n \to x \in X$ as $n \to \infty$, then $x_n \rightharpoonup x$ as $n \to \infty$.

Proof. it follows directly from that

$$|\ell(x_n - x)| \leqslant \|\ell\|_* \|x_n - x\| \to 0$$

as $n \to \infty$, for all $\ell \in X^*$. □

Remark. In finite-dimensional normed vector spaces, the strong and weak convergence are equivalent. It is not true in general. For example, suppose that X is a separable Hilbert space with a orthonormal basis (x_n). Then (x_n) converges weakly but not strongly to 0 as $n \to \infty$.

Proposition 5.31. Weak Convergence $\implies$ Boundedness & Lower Semi-Continuity of Norm

Suppose that $(X, \|\cdot\|)$ is a normed vector space. Let $(x_n) \subseteq X$ such that $x_n \rightharpoonup x \in X$. Then $\sup_{n\in\mathbb{N}} \|x_n\| < \infty$. In addition, we have $\|x\| \leqslant \liminf_{n\to\infty} \|x_n\|$.

Proof. Note that each x_n defines a linear functional on X^*: $T_n(\ell) = \ell(x_n)$ for all $\ell \in X^*$. Furthermore, $\|T_n\|_{**} = \|x_n\|$. Now for each $\ell \in X^*$, $\ell(x_n)$ is convergent, and hence bounded. The uniform boundedness principle thus implies that $\|T_n\|$ is bounded. (Note that X^* is complete regardless whether X is complete or not.) Hence (x_n) is uniformly bounded in norm.

Next by Proposition 5.4, there exists $\ell \in X^*$ such that $\|\ell\|_* = 1$ and $\ell(x) = \|x\|$. Then

$$\|x_n\| = \|\ell\|_* \|x_n\| \geqslant |\ell(x_n)| \to |\ell(x)| = \|x\|$$

as $n \to \infty$. Hence $\|x\| \leqslant \liminf_{n\to\infty} \|x_n\|$. □

Example 5.32. Schur Property of ℓ^1

A weakly convergent sequence $(x_n) \subseteq \ell^1$ is strongly convergent.

Proof. Suppose that $(x_n) \subseteq \ell^1$ converges weakly but not strongly to $x \in \ell^1$. Let $y_n := x_n - x$. Then there exists a subsequence (z_n) of (y_n) such that

$$\|z_n\|_1 = \sum_{j\in\mathbb{N}} |z_n(j)| \geqslant \varepsilon$$

for some $\varepsilon > 0$.

We shall construct the increasing sequences (n_k) and (m_k) in the following inductive way:

- Let $m_0 = 0$ and $n_0 = 0$.
- Suppose that we have constructed m_{k-1} and n_{k-1}. We claim that there exists $n_k \geqslant n_{k-1}$ such that

$$\sum_{j\leqslant m_{k-1}} |z_{n_k}(j)| < \frac{\varepsilon}{8}$$

From part (a) we know that $z_n(j) \to 0$ as $n \to \infty$ for each $n \in \mathbb{N}$. Then

$$\forall\, i \in \{0, ..., k-1\}\ \exists\, N_i \in \mathbb{N}\ \forall\, n > N_i\ |z_n(i)| < \frac{\varepsilon}{8(m_{k-1}+1)}$$

Let $n_k := \max\{n_{k-1}, N_0, ..., N_{k-1}\}$. Then

$$\sum_{j\leqslant m_{k-1}} |z_{n_k}(j)| < \sum_{j=0}^{m_{k-1}} \frac{\varepsilon}{8(m_{k-1}+1)} = \frac{\varepsilon}{8}$$

- We claim that there exists $m_k \geqslant m_{k-1}$ such that

$$\sum_{j\geqslant m_k} |z_{n_k}(j)| < \frac{\varepsilon}{8}$$

Since $(z_{n_k}) \in \ell^1$,

$$\|z_{n_k}\|_1 = \sum_{j\in\mathbb{N}} |z_{n_k}(j)| < \infty$$

Choose sufficiently large m_k such that

$$\sum_{j=0}^{m_k} |z_{n_k}(j)| > \|z_{n_k}\|_1 - \frac{\varepsilon}{8}$$

and the result follows.

We identify $(\ell^1)^*$ with ℓ^∞. Consider $b \in \ell^\infty$ such that $b(j) = \overline{\operatorname{sgn} z_{n_k}}$ for $m_{k-1} < j \leqslant m_k$.

For sufficiently large k,

$$\langle b, z_{n_k}\rangle = \sum_{j\in\mathbb{N}} b(j) z_{n_k}(j) = \left(\sum_{j=0}^{m_{k-1}} + \sum_{j=m_k+1}^{\infty}\right) b(j) z_{n_k}(j) + \sum_{j=m_{k-1}+1}^{m_k} |z_{n_k}(j)|$$

where
$$\left|\left(\sum_{j=0}^{m_{k-1}}+\sum_{j=m_k+1}^{\infty}\right)b(j)z_{n_k}(j)\right|\leqslant\sum_{j=0}^{m_{k-1}}|z_{n_k}(j)|+\sum_{j=m_k+1}^{\infty}|z_{n_k}(j)|<\frac{\varepsilon}{8}+\frac{\varepsilon}{8}=\frac{\varepsilon}{4}$$

In addition,
$$\varepsilon<\|x\|=\left(\sum_{j=0}^{m_{k-1}}+\sum_{j=m_k+1}^{\infty}\right)|z_{n_k}(j)|+\sum_{j=m_{k-1}+1}^{m_k}|z_{n_k}(j)|<\frac{\varepsilon}{4}+\sum_{j=m_{k-1}+1}^{m_k}|z_{n_k}(j)|$$

Therefore
$$b(z_{n_k})>\frac{\varepsilon}{4}+\left(\varepsilon-\frac{\varepsilon}{4}\right)=\varepsilon$$

contradicting that $x_n \rightharpoonup 0$ weakly in ℓ^1. We conclude that $x_n \to 0$ in ℓ^1. □

Theorem 5.33. Radon-Riesz Theorem

Suppose that X is a uniformly convex Banach space. Let $(x_n) \subseteq X$ such that $x_n \rightharpoonup x \in X$ and $\|x_n\| \to \|x\|$ as $n \to \infty$. Then $x_n \to x$ strongly in X as $n \to \infty$.

Proof. Suppose that $x = 0$. Then $\|x_n\| \to 0$ directly implies that $x_n \to 0$ in X by definition. Now suppose that $\|x\| \neq 0$. Since $\|x_n\| \to \|x\|$ as $n \to \infty$, there exists $N \in \mathbb{N}$ such that for all $n > N$, $\|x_n\| \geqslant \frac{1}{2}\|x\|$. By discarding the first N terms in the sequence, we assume that $\|x_n\| \neq 0$ for all $n \in \mathbb{N}$. Let $y_n := x_n/\|x_n\|$ and $y := x/\|x\|$. For $\ell \in X^*$,
$$\lim_{n\to\infty}\ell(y_n)=\lim_{n\to\infty}\frac{\ell(x_n)}{\|x_n\|}=\frac{\lim_{n\to\infty}\ell(x_n)}{\lim_{n\to\infty}\|x_n\|}=\frac{\ell(x)}{\|x\|}=\ell(y)$$

Hence $y_n \rightharpoonup y$ in X. Let $z_n := \frac{1}{2}(y_n + y)$. Then $z_n \rightharpoonup y$ in X. By Proposition 5.31,
$$1=\|y\|\leqslant\liminf_{n\to\infty}\|z_n\|=\frac{1}{2}\liminf_{n\to\infty}\|y_n+y\| \tag{$*$}$$

Suppose that $y_n \not\to y$ in X. Then there exists $\varepsilon > 0$ such that $\limsup_{n\to\infty}\|y_n - y\| \geqslant \varepsilon$. By definition of uniform convexity, there exists $\delta > 0$ such that $\|y_n + y\| \leqslant 2(1-\delta)$ whenever $\|y_n - y\| \geqslant \varepsilon$. Hence
$$\liminf_{n\to\infty}\|y_n+y\|\leqslant 2(1-\delta)$$

This contradicts the equation $(*)$. We deduce that $y_n \to y$ in X. Finally, since $\|x_n\| \to \|x\|$ in $\mathbb{R}$ and $x_n = y_n\|x_n\|$, by algebra of limits we conclude that $x_n \to x$ in X. □

Theorem 5.34. Mazur's Theorem

Let X be a normed vector space and $K \subseteq X$ a closed convex set. Let $(x_n) \subseteq K$ be a sequence such that $x_n \rightharpoonup x \in X$ weakly. Then $x \in K$.

Proof. Suppose that $x \notin K$. Then there exists an open ball $B(x, r)$ such that $B(x, r) \cap K = \varnothing$. By Hyperplane Separation Theorem (Theorem 5.10) there exists $\ell_0 \in X^* \setminus \{0\}$ and $c \in \mathbb{R}$ such that
$$\sup_{y\in K}\operatorname{Re}\ell_0(y)\leqslant c\leqslant\inf_{z\in B(x,r)}\operatorname{Re}\ell_0(z)$$

Therefore for $w \in B(0, 1)$,
$$\operatorname{Re}\ell_0(w)=\frac{1}{r}\left(\operatorname{Re}\ell_0(x+rw)-\operatorname{Re}\ell_0(x)\right)\geqslant\frac{1}{r}\left(c-\operatorname{Re}\ell_0(x)\right)$$

Replacing w with $-w$ we obtain

$$\operatorname{Re}\ell_0(w) = -\operatorname{Re}\ell_0(-w) \leqslant -\frac{1}{r}\left(c - \operatorname{Re}\ell_0(x)\right)$$

Hence

$$|\operatorname{Re}\ell_0(w)| \leqslant \frac{1}{r}\left(\operatorname{Re}\ell_0(x) - c\right) \tag{$*$}$$

On the other hand, since $x_n \rightharpoonup x$, we have $\ell_0(x_n) \to \ell_0(x)$. Therefore

$$\operatorname{Re}\ell_0(x_n) \leqslant c \implies \operatorname{Re}\ell_0(x) = \lim_{n\to\infty} \operatorname{Re}\ell_0(x_n) \leqslant c \tag{$**$}$$

Substituting $(**)$ into $(*)$ we obtain $|\operatorname{Re}\ell_0(w)| \leqslant 0$, and hence $\operatorname{Re}\ell_0(w) = 0$. Replacing w with iw we also have $\operatorname{Im}\ell_0(w) = 0$. We deduce that $\ell_0 = 0$, which is contradictory. We conclude that $x \in K$, □

Remark. In the proof of Mazur's Theorem, we see that in fact a sequence of finite linear convex combinations of (x_n) converges strongly to x. For uniformly convex Banach spaces, the result can be substantially improved as follows.

Theorem 5.35. Banach-Saks Theorem

Suppose that X is a uniformly convex Banach space. Any weakly convergent sequence $(x_n) \subseteq X$ has a subsequence (x_{n_k}) which converges strongly in the Cesàro sense:

$$\frac{1}{m}\sum_{k=1}^{m} x_{n_k} \to x \in X \quad \text{as } m \to \infty$$

Proof when X is a Hilbert space:

Without loss of generality we assume that $x_n \rightharpoonup 0$ and $\|x_n\| \leqslant 1$ for all $n \in \mathbb{Z}_+$. We construct subsequence of (x_n) inductively. Let $n_1 = 0$. Suppose that we have $n_k \in \mathbb{N}$. Let $y_k := \sum_{j=0}^{k} x_{n_j}$ and $\|y_k\|^2 \leqslant 3k$ (induction hypothesis). Then we have $\langle y_k, x_n\rangle \to 0$ as $n \to \infty$. Therefore there exists $n_{k+1} > n_k$ such that $\left|\left\langle y_k, x_{n_{k+1}}\right\rangle\right| < 1$. Now we have

$$\left\|y_{n_{k+1}}\right\|^2 = \|y_{n_k}\|^2 + 2\operatorname{Re}\left\langle y_{n_k}, x_{n_{k+1}}\right\rangle + \left\|x_{n_{k+1}}\right\|^2 \leqslant 3k + 2 + 1 = 3(k+1)$$

which completes the induction. Therefore we have

$$\left\|\frac{1}{m}\sum_{k=1}^{m} x_{n_k}\right\|^2 = \frac{1}{m^2}\|y_m\|^2 \leqslant \frac{3}{m} \to 0$$

as $m \to \infty$. Therefore (x_{n_k}) converges strongly to 0 in the Cesàro sense. □

Definition 5.36. Weak* Convergence

Suppose that $(X, \|\cdot\|)$ is a normed vector space. For $(\ell_n)_{n\in\mathbb{N}} \subseteq X^*$, we say that ℓ_n converges weak* to $\ell \in X^*$, if $\ell_n(x) \to \ell(x)$ as $n \to \infty$ for all $x \in X$.

Remark. Similarly we also have the notion of weak* topology. The weak* topology on X^* is $\sigma(X^*, J_X(X))$, which is the coarsest topology on X^* with respect to which every $J_X(x) \in X^{**}$ is continuous.

There are three different modes of convergence on X^*: strong convergence, weak convergence, and weak* convergence. If X is reflexive, then the weak and weak* convergence on X^* are equivalent;, and the corresponding weak topology $\sigma(X^*\ X^{**})$ and weak* topology $\sigma(X^*, J_X(X))$ are equal.

5.6 Weak Sequential Compactness

In Proposition 1.30, we have shown that the closed unit ball is no longer compact with respect to the strong topology for infinite-dimensional normed vector spaces. In this section we shall establish a result asserting that in reflexive Banach spaces the closed unit ball is compact with respect to the weak topology.

Definition 5.37. Weak Sequential Compactness

Suppose that X is a normed vector space. $K \subseteq X$ is called weakly sequentially compact, if every sequence in K has a subsequence weakly convergent to a point in K.

Similarly we also have the notion of weak* sequentially compactness for subsets of X^*.

Definition 5.38. Weak Compactness (*Off-Syllabus*)

Suppose that X is a normed vector space. $K \subseteq X$ is called weakly compact, if K is compact with respect to the weak topology $\sigma(X, X^*)$.

Similarly we also have the notion of weak* compactness for subsets of X^*.

Remark. For normed vector spaces (and more generally metric spaces), compactness is equivalent to sequential compactness. However, it is not true in general that weak compactness and weak sequential compactness are equivalent, as the weak topology may not be metrisable.

Theorem 5.39. Banach-Alaoglu Theorem (*Off-Syllabus*)

Suppose that X is a real normed vector space. The closed unit ball $B_X^* \subseteq X^*$ is weak* compact, i.e., compact with respect to the weak* topology.

Proof. For each $x \in X$ we associate a compact interval

$$K_x := [-\|x\|, \|x\|] \subseteq \mathbb{R}$$

The product space

$$K := \prod_{x \in X} K_x = \{f : X \to \mathbb{R} \mid \forall\, x \in X \ |f(x)| \leqslant \|x\|\} \subseteq \mathbb{R}^X$$

Let $X' \subseteq \mathbb{R}^X$ be the algebraic dual of X. Then the interchapter $X' \cap K$ is closed unit ball $B_{X^*} := \{\ell \in X^* : \ \|\ell\|_* = 1\}$. The weak* topology on B_{X^*} is precisely the induced topology on $\mathbb{R}^X$.

By **Tychonoff's Theorem**[1], K is compact with respect to the product topology. To show that B_{X^*} is compact, it suffices to show that X' is closed in $\mathbb{R}^X$.

For $x, y \in X$ and $\lambda \in \mathbb{R}$, we define $\varphi_{x,y} : \mathbb{R}^X \to \mathbb{R}$ and $\psi_{x,\lambda} : \mathbb{R}^X \to \mathbb{R}$ by

$$\varphi_{x,y}(f) = f(x+y) - f(x) - f(y), \qquad \psi_{x,\lambda} = f(\lambda x) - \lambda f(x)$$

Then

$$X' = \bigcap_{x,y \in X} \varphi_{x,y}^{-1}(\{0\}) \cap \bigcap_{x \in X, \lambda \in \mathbb{R}} \psi_{x,\lambda}^{-1}(\{0\})$$

is closed with respect to the product topology. This completes the proof. □

Theorem 5.40. Eberlein-Šmulian Theorem

Let X be a Banach space. Let $B_X \subseteq X$ be the closed unit ball of X. The following are equivalent:

1. X is reflexive;
2. B_X is weakly sequentially compact;
3. B_X is weakly compact.

Remark. The only examinable proof is the direction 1 $\Longrightarrow$ 2 in the special case of Hilbert spaces.

Proof.

[1]**Tychonoff's Theorem**: The arbitrary product of compact spaces is compact with respect to the product topology.

$1 \Longrightarrow 2$: Suppose that $(x_n) \subseteq B_X$. Let $Y := \overline{\operatorname{span}\{x_n\}_{n\in\mathbb{N}}}$.

- Y is reflexive. (*trivial if X is a Hilbert space*)

  Let $\xi_0 \in Y^{**}$. Let $R_{Y^*} \in \mathscr{B}(X^*, Y^*)$ be the restriction operator and $R'_{Y^*} \in \mathscr{B}(Y^{**}, X^{**})$ be its dual operator . Let $\xi := R'_{Y^*}(\xi_0) \in X^**$. Since X is reflexive, there exists $w \in X$ such that $\xi = J_X(w)$. We claim that $w \in Y^*$. Suppose not. Then by Proposition 5.12 there exists $\ell \in Y^\circ$ such that $\ell(w) = \operatorname{dist}(w, Y) > 0$. But

  $$\ell(w) = \xi(\ell) = R'_{Y^*}(\xi_0)(\ell) = \xi_0 \circ R_{Y^*}(\ell) = \xi_0(0) = 0$$

  This is a contradiction. We have $w \in Y$. In particular $\xi_0 = J_Y(w)$. We deduce that Y is reflexive.

- $(J_Y(x_n))_{n\in\mathbb{N}} \subseteq Y^{**}$ has a weak* convergent subsequence. (*diagonal argument, very important!*)

  Since Y is reflexive and separable, by Proposition 5.26, Y^* is also separable. Suppose that $(\ell_n)_{n=1}^{\infty}$ is dense in Y^*. We construct $S : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ inductively as follows:

  First, let $S(m, 0) := m$. Suppose that we have constructed $S(\cdot, n) : \mathbb{N} \to \mathbb{N}$. Consider the sequence $(J_X(x_{S(m,n)})(\ell_{n+1}))_{m\in\mathbb{N}} = (\ell_{n+1}(x_{S(m,n)}))_{m\in\mathbb{N}} \subseteq \mathbb{R}$, which is bounded in $\mathbb{R}$, because (x_n) is bounded. By Bolzano-Weierstrass Theorem, there exists a subsequence $S(m, n+1)_{m\in\mathbb{N}}$ of $S(m, n)_{m\in\mathbb{N}}$ such that $(J_X(x_{S(m,n+1)})(\ell_{n+1}))_{m\in\mathbb{N}}$ is convergent. Now we take the subsequence $(x_{S(m,m)})$ of (x_m). Then $(J_X(x_{S(m,m)})(\ell_n))_{m\in\mathbb{N}}$ is convergent for each $n \in \mathbb{Z}_+$.

  Note that $(J_X(x_{S(m,m)}))$ is uniformly bounded by 1. Since (ℓ_n) is dense in Y^*, by Proposition 4.3, we have that $(J_X(x_{S(m,m)})(\ell))$ is convergent for each $\ell \in Y^*$. Let $\psi : Y^* \to \mathbb{F}$ such that

  $$\psi(\ell) := \lim_{m\to\infty} J_X(x_{S(m,m)})(\ell)$$

  for all $\ell \in Y^*$. It is straightforward to check that $\psi \in Y^{**}$. Therefore ψ is the weak* limit of $(J_X(x_{S(m,m)}))$ in Y^{**}.

- $(x_n)_{n\in\mathbb{N}} \subseteq X$ has a weak convergent subsequence.

  Let $x := J_X^{-1}(\psi) \in Y$. For $\ell \in X^*$, we have $\ell(x_n) = R_{Y^*} \circ \ell(x_n)$ and $\ell(x) = R_{Y^*} \circ \ell(x)$. We have

  $$\lim_{m\to\infty} \ell(x_{S(m,m)}) = \lim_{m\to\infty} R_{Y^*}\circ\ell(x_{S(m,m)}) = \lim_{m\to\infty} J_X(x_{S(m,m)})(R_{Y^*}(\ell)) = \psi\circ R_{Y^*}(\ell) = R_{Y^*}\circ\ell(x) = \ell(x)$$

  We deduce that $x_{S(m,m)} \rightharpoonup x$ as $m \to \infty$. In particular, B_X is weakly sequentially compact.

$1 \Longrightarrow 3$: Since X is reflexive, $J_X : X \to X^{**}$ is an isometric isomorphism and hence is a homeomorphism with respect to the weak topology of the both spaces. By Proposition 5.18, X^* is also reflexive. Hence the weak topology on X^{**} agrees with the weak* topology. By the Banach-Alaoglu Theorem, $B_{X^{**}} \subseteq X^{**}$ is weakly compact. Therefore $B_X \subseteq X$ is also weakly compact.

$2 \Longrightarrow 1$: Omitted. (Too difficult for an introducotry course like this...)

$3 \Longrightarrow 1$: Omitted. (Too difficult for an introducotry course like this...) □

Corollary 5.41. Closest Point in a Closed Convex Subset, Reflexive Banach Space

Let X be a reflexive Banach space and $K \subseteq X$ a non-empty closed convex set. Then

$$\forall\, x \in X\ \exists\, y \in K : \ \|x - y\| = \operatorname{dist}(x, K)$$

In particular, $y \in K$ is unique if X is uniformly convex.

Proof. Let $(y_n) \subseteq K$ such that $\|x - y_n\| \to \operatorname{dist}(x, K)$ as $n \to \infty$. Since (y_n) is bounded and X is reflexive, by Eberlein-Šmulian Theorem, $y_n \rightharpoonup y$ for some $y \in X$. Since K is closed and convex, by Mazur's Theorem $y \in K$. In particular, $\|x - y\| \geqslant \operatorname{dist}(x, K)$. On the other hand, by the lower semi-continuity of norm,

$$\|x - y\| \leqslant \liminf_{n\to\infty} \|x - y_n\| = \operatorname{dist}(x, K)$$

We conclude that $\|x - y\| = \operatorname{dist}(x, K)$.

Now in addition suppose that X is uniformly convex. Without loss of generality, we assume that $x \notin K$ so that $\operatorname{dist}(x, K) > 0$. Suppose that $y, y' \in K$ such that $\|x - y\| = \|x - y'\| = \operatorname{dist}(x, K)$. Suppose for contradiction that $y \neq y'$. Let $z := \dfrac{x - y}{\operatorname{dist}(x, K)}$ and $z' := \dfrac{x - y'}{\operatorname{dist}(x, K)}$. Then $z, z' \in S_X$ and $z \neq z'$. By uniform convexity of X (*in fact strict convexity would suffice*), there exists $\delta > 0$ such that

$$\frac{1}{2}\|z + z'\| < 1 - \delta \implies \left\| x - \frac{y + y'}{2} \right\| < (1 - \delta) \operatorname{dist}(x, K) < \operatorname{dist}(x, K)$$

But by convexity of K, $(y + y')/2 \in K$, which is contradictory. We deduce that $y = y'$. □

Chapter 6

Spectral Theory

In this chapter, we discuss the spectrum of a linear operator. We categorise different types of spectra and discuss the properties of spectra of bounded operators on Banach spaces. Then we focus on the normal operators on Hilbert spaces.

For convenience, we shall restrict our attension to the complex scalar field (thanks to its algebraic closedness.)

6.1 Spectra and Resolvent Sets

Spectra is the generalisation of eigenvalues of linear operators in infinite-dimensional normed vector spaces. Recall from Linear Algebra that if T is a linear operator on a finite-dimensional vector space X, then

λ is an eigenvalue of $T \iff (\lambda \operatorname{id} - T)$ is not injective $\iff (\lambda \operatorname{id} - T)$ is not invertible $\iff (\lambda \operatorname{id} - T)$ is not surjective

The second and the third equivalence no longer holds for infinite-dimensional spaces. We have the following definitions.

Definition 6.1. Spectra, Resolvent Sets

Suppose that X is a complex normed vector space. Let $T \in \mathscr{B}(X)$.

- The spectrum $\sigma(T)$ of T is the set of $\lambda \in \mathbb{C}$ such that $\lambda \operatorname{id} - T$ has no inverse in $\mathscr{B}(X)$;
- The resolvent set of T is $\rho(T) := \mathbb{C} \setminus \sigma(T)$. For $\lambda \in \rho(T)$, $R_\lambda(T) := (T - \lambda \operatorname{id})^{-1}$ is called the **resolvent** of T at λ.

Definition 6.2. Point Spectra, Continuous Spectra, Residual Spectra

Suppose that X is a complex normed vector space. Let $T \in \mathscr{B}(X)$.

- The point spectrum $\sigma_{\mathrm{p}}(T)$ of T: $\lambda \in \sigma_{\mathrm{p}}(T) \iff \lambda \operatorname{id} - T$ is not injective.

  The elements in $\sigma_{\mathrm{p}}(T)$ are called **eigenvalues** of T; the non-trivial elements in $\ker(\lambda \operatorname{id} - T)$ are called **eigenvectors** of T.
- The continuous spectrum $\sigma_{\mathrm{c}}(T)$ of T: $\lambda \in \sigma_{\mathrm{c}}(T) \iff \lambda \operatorname{id} - T$ is injective and $\operatorname{im}(\lambda \operatorname{id} - T)$ is a proper dense subset of X.
- The residual spectrum $\sigma_{\mathrm{r}}(T)$ of T: $\lambda \in \sigma_{\mathrm{c}}(T) \iff \lambda \operatorname{id} - T$ is injective and $\overline{\operatorname{im}(\lambda \operatorname{id} - T)} \subsetneq X$.

By definition, $\sigma(T)$ is the disjoint union $\sigma_{\mathrm{p}}(T) \cup \sigma_{\mathrm{c}}(T) \cup \sigma_{\mathrm{r}}(T)$.

Definition 6.3. Approximate Point Spectra

$\lambda \in \mathbb{C}$ is called an **approximate eigenvalue** of T, if there exists a sequence $(x_n) \subseteq X$ such that $\|x_n\| = 1$ and $\|\lambda x_n - T(x_n)\| \to 0$ as $n \to \infty$. The approximate point spectrum $\sigma_{\mathrm{ap}}(T)$ of T is the set of approximate eigenvalues of T.

Remark. It is clear from definition that $\sigma_{\mathrm{p}}(T) \subseteq \sigma_{\mathrm{ap}}(T) \subseteq \sigma(T)$.

Proposition 6.4. $\sigma_{\mathrm{c}}(T) \subseteq \sigma_{\mathrm{ap}}(T)$

Suppose that X is a complex Banach space and $T \in \mathscr{B}(X)$. Then $\sigma_{\mathrm{c}}(T) \subseteq \sigma_{\mathrm{ap}}(T)$.

Proof. Suppose for contradiction that $\lambda \in \sigma_{\mathrm{c}}(T) \setminus \sigma_{\mathrm{ap}}(T)$. $\lambda \notin \sigma_{\mathrm{ap}}(T)$ and linearity implies that there exists $c > 0$ such that

$$\|(\lambda \operatorname{id} - T)x\| \geqslant c\|x\|$$

for all $x \in X$. Let $Y := \operatorname{im}(\lambda \operatorname{id} - T)$. Note that $\lambda \operatorname{id} - T$ is bijective from X to Y. Hence it is invertible in $\mathscr{B}(X, Y)$ and its inverse $U \in \mathscr{B}(Y, X)$. Since Y is dense in X, by Proposition 3.2 there exists a unique extension $V \in \mathscr{B}(X, X)$ of U. Now pick $p \in X \setminus Y$ and $(p_n) \subseteq Y$ such that $p_n \to p$. Then $U(p_n) \to V(p)$ and therefore

$$(\lambda \operatorname{id} - T)V(p) = \lim_{n\to\infty} (\lambda \operatorname{id} - T)U(p_n) = \lim_{n\to\infty} p_n = p$$

Hence $p \in Y$, which is contradictory. We conclude that $\sigma_{\mathrm{c}}(T) \subseteq \sigma_{\mathrm{ap}}(T)$. □

Proposition 6.5. Spectra of Dual Operators

Suppose that X is a complex Banach space and $T \in \mathscr{B}(X)$. Then we have

1. $\sigma_{\mathrm{r}}(T) \subseteq \sigma_{\mathrm{p}}(T')$;
2. $\sigma_{\mathrm{c}}(T') \subseteq \sigma_{\mathrm{c}}(T)$.

Proof. 1. Let $\lambda \in \sigma_{\mathrm{r}}(T)$. Let $Y := \lambda \operatorname{id} - T$. Then $\overline{Y} \subseteq X$. By Corollary 5.13.1, there exists $\ell \in Y^\circ$ such that $\|\ell\|_* = 1$. For all $x \in X$,

$$(\lambda \operatorname{id}_{X^*} - T') \circ \ell(x) = \ell((\lambda \operatorname{id}_X - T)x) = 0$$

Hence $(\lambda \operatorname{id}_{X^*} - T')\ell = 0$ and $\lambda \in \sigma_{\mathrm{p}}(T')$.

2. Suppose that $\lambda \in \sigma_{\mathrm{c}}(T')$. By Proposition 1.35 and Corollary 5.13, $\ker(\lambda \operatorname{id}_{X^*} - T') = \{0\}$ implies that $\overline{\operatorname{im}(\lambda \operatorname{id}_X - T)} = X$; $\overline{\operatorname{im}(\lambda \operatorname{id}_{X^*} - T')} = X^*$ implies that $\ker(\lambda \operatorname{id}_X - T) = \{0\}$. Since $\lambda \operatorname{id}_X - T$ is not invertible, we deduce that $\lambda \in \sigma_{\mathrm{c}}(T)$. □

Corollary 6.6. $\sigma(T) = \sigma_{\mathrm{ap}}(T) \cup \sigma_{\mathrm{p}}(T')$

Suppose that X is a complex Banach space and $T \in \mathscr{B}(X)$. Then $\sigma(T) = \sigma_{\mathrm{ap}}(T) \cup \sigma_{\mathrm{p}}(T')$.

Hilbert spaces are reflexive and we can use the adjoints to replace the dual operators. We have the following results.

Proposition 6.7. Spectra of Operators in Hilbert Spaces

Suppose that X is a complex Hilbert space and $T \in \mathscr{B}(X)$. Then we have

1. $\sigma_{\mathrm{p}}(T) \subseteq \sigma_{\mathrm{p}}^*(T^*) \cup \sigma_{\mathrm{r}}^*(T^*)$;
2. $\sigma_{\mathrm{c}}(T) = \sigma_{\mathrm{c}}^*(T^*)$;
3. $\sigma_{\mathrm{r}}(T) \subseteq \sigma_{\mathrm{p}}^*(T^*)$.

where $\sigma_\cdot^*(T) := \{\overline{\lambda} : \lambda \in \sigma_\cdot(T)\}$.

Lemma 6.8

Suppose that X is a complex Hilbert space and $T \in \mathscr{B}(X)$. Then $\sigma(T) \subseteq \overline{\{\langle x, Tx\rangle : \|x\| = 1\}}$.

Proof. Let $\lambda \in \sigma(T) = \sigma_{\mathrm{ap}}(T) \cup \sigma_{\mathrm{p}}^*(T^*)$. If $\lambda \in \sigma_{\mathrm{ap}}(T)$, then there exists $(x_n) \subseteq S_X$ such that $(\lambda \operatorname{id} - T)x_n \to 0$ as

$n \to \infty$. By Cauchy-Schwarz Inequality,

$$|\lambda - \langle x_n, Tx_n\rangle| = \left|\lambda\|x_n\|^2 - \langle x_n, Tx_n\rangle\right| = |\langle x_n, (\lambda \operatorname{id} - T)x_n\rangle| \leqslant \|x_n\|\|(\lambda \operatorname{id} - T)x_n\| \to 0$$

as $n \to \infty$. Hence $\lambda \in \overline{\{\langle x, Tx\rangle : \|x\| = 1\}}$.

If $\lambda \in \sigma_{\mathrm{p}}^*(T^*)$, then $\overline{\lambda} \in \sigma_{\mathrm{p}}(T^*)$ and so there exists $x \in S_X$ with $T^*(x) = \overline{\lambda}(x)$. We have

$$\lambda = \langle \overline{\lambda} x, x\rangle = \langle T^*(x), x\rangle = \langle x, Tx\rangle$$

Hence $\lambda \in \overline{\{\langle x, Tx\rangle : \|x\| = 1\}}$. □

6.2 Bounded Operators on Banach Spaces

Proposition 6.9. Properties of the Spectra of Bounded Operators on Banach Spaces

Suppose that X is a complex Banach space and $T \in \mathscr{B}(X)$.

1. The resolvent set $\rho(\lambda)$ is open in $\mathbb{C}$. The map $\rho(T) \to \mathscr{B}(X)$, $\lambda \mapsto R_\lambda(T)$ is **analytic**, in the sense that for any $\lambda_0 \in \rho(T)$, there exists an open neighbourbood $\Omega \subseteq \rho(T)$ of λ_0 such that the resolvent operator is given by
$$R_\lambda(T) = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R_{\lambda_0}(T)^{n+1}$$
2. $\sigma(T)$ is non-empty, compact, and bounded by $\|T\|_{\mathscr{B}(X)}$.

Proof. 1. Fix $\lambda_0 \in \rho(T)$, so $R_{\lambda_0}(T) = (T - \lambda_0 \operatorname{id})^{-1}$. By Corollary 1.22, $T - \lambda_0 \operatorname{id} - S$ is invertible whenever $\|S\| < \left\|R_{\lambda_0(T)}\right\|^{-1}$. The inverse is given by the Neumann series

$$(T - \lambda_0 \operatorname{id} - S)^{-1} = (T - \lambda_0 \operatorname{id})^{-1} \sum_{n=0}^{\infty} \left((T - \lambda_0 \operatorname{id})^{-1} S\right)^n = R_{\lambda_0}(T) \sum_{n=0}^{\infty} (R_{\lambda_0}(T) S)^n$$

For $\lambda \in B\left(\lambda_0, \left\|R_{\lambda_0(T)}\right\|^{-1}\right)$, we put $S := (\lambda - \lambda_0) \operatorname{id}$. So $T - \lambda \operatorname{id} = T - \lambda_0 \operatorname{id} - S$ is invertible, with inverse given by

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

Hence $\lambda \mapsto R_\lambda(T)$ is analytic. We have also shown that $\rho(T)$ is open.

2. $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is closed. For $\lambda \in \mathbb{C}$ with $|\lambda| > \|T\|$, by Lemma 1.21, $\operatorname{id} - \frac{1}{\lambda} T$ is invertible, and hence so is $\lambda \operatorname{id} - T$. Therefore $\sigma(T) \subseteq B(0, \|T\|)$. Now $\sigma(T)$ is closed and bounded. So it is compact by Heine-Borel Theorem.

Lastly we shall show that $\sigma(T) \neq \varnothing$. Suppose that it is empty. The resolvent map $\lambda \mapsto R_\lambda(T)$ is defined on the whole $\mathbb{C}$. For $\ell \in \mathscr{B}(X)^*$, we define $g_\ell : \mathbb{C} \to \mathbb{C}$ by

$$g_\ell(\lambda) := \ell(R_\lambda(T))$$

We note that g_ℓ is analytic because

$$g_\ell(\lambda) = \ell\left(\sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R_{\lambda_0}(T)^{n+1}\right) = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n \ell\left(R_{\lambda_0}(T)^{n+1}\right)$$

We claim that g_ℓ is bounded. Since g_ℓ is continuous, it is bounded on the closed disc $\overline{B(0, 2\|T\|)}$. On the other hand, for any $\lambda \in \mathbb{C} \setminus \overline{B(0, 2\|T\|)}$, we have

$$\|R_\lambda(T)\| = |\lambda|^{-1} \left\|\operatorname{id} - \frac{1}{\lambda} T\right\| \leqslant |\lambda|^{-1} \sum_{n=0}^{\infty} \left\|\frac{1}{\lambda} T\right\|^n = \frac{1}{|\lambda| - \|T\|} \leqslant \frac{1}{\|T\|}$$

and hence

$$|g_\ell(\lambda)| \leqslant \|\ell\|_* \|R_\lambda(T)\| \leqslant \|\ell\|_* \|T\|^{-1}$$

So g_ℓ is bounded on $\lambda \in \mathbb{C} \setminus \overline{B(0, 2\|T\|)}$.

This shows that g_ℓ is bounded and analytic. By Liouville's Theorem, g_ℓ is constant on $\mathbb{C}$. From the expansion of g_ℓ we deduce that $\ell(R_\lambda(T))^n = 0$ for all $n \geqslant 2$ and $\lambda \in \mathbb{C}$. This holds for all $\ell \in \mathscr{B}(X)^*$. By Proposition 5.4, $R_\lambda(T)^n = 0$ for $n \geqslant 2$ and $\lambda \in \mathbb{C}$. This is impossible, as $R_\lambda(T)$ is invertible. We conclude that $\sigma(T) \neq \varnothing$. □

Definition 6.10. Spectral Radius

Suppose that X is a complex normed vector space and $T \in \mathscr{B}(X)$. The spectral radius of T is defined by

$$\operatorname{rad}\sigma(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}$$

Theorem 6.11. Gelfand's Formula

Suppose that X is a complex Banach space and $T \in \mathscr{B}(X)$. Then

$$\operatorname{rad}\sigma(T) = \lim_{n\to\infty} \|T^n\|^{1/n} = \inf_{n\in\mathbb{N}} \|T^n\|^{1/n}$$

Proof.

- $\operatorname{rad}\sigma(T) \leqslant \inf_{n\in\mathbb{N}} \|T^n\|^{1/n}$.

  For $n \in \mathbb{N}$, consider the operator

  $$\lambda^n \operatorname{id} - T^n = (\lambda \operatorname{id} - T)S = S(\lambda \operatorname{id} - T), \qquad S := \sum_{k=1}^{n-1} \lambda^{k-1} T^{n-k}$$

  For $\lambda \in \mathbb{C}$ with $|\lambda|^n > \|T^n\|$, by Proposition 6.9, $\lambda^n \notin \sigma(T^n)$ and hence $\lambda^n \operatorname{id} - T^n$. Since S commutes with $\lambda \operatorname{id} - T$, we have that $\lambda \operatorname{id} - T$ is invertible. So $\lambda \notin \sigma(T)$. We deduce that every $\lambda \in \sigma(T)$ satisfies $\lambda^n \leqslant \|T^n\|$ for all $n \in \mathbb{N}$. Hence $\operatorname{rad}\sigma(T) \leqslant \inf_{n\in\mathbb{N}} \|T^n\|^{1/n}$.

- $\limsup_{n\to\infty} \|T^n\|^{1/n} \leqslant \operatorname{rad}\sigma(T)$.

  As in the proof of Proposition 6.9, for each $\ell \in \mathscr{B}(X)^*$ we associate the function $g_\ell(\lambda) = \ell(R_\lambda(T))$.

  $$R_\lambda(T) = -\lambda^{-1} \sum_{n=0}^{\infty} \lambda^{-n} T^n \implies g_\ell(\lambda) = -\lambda^{-1} \sum_{n=0}^{\infty} \ell(\lambda^{-n} T^n)$$

  As g_ℓ is analytic on $\rho(T)$, $\sum_{n=0}^{\infty} \ell(\lambda^{-n} T^n)$ converges absolutely whenever $|\lambda| > \operatorname{rad}\sigma(T)$. Fix $|\lambda| = \operatorname{rad}\sigma(T) + \varepsilon$ for some $\varepsilon > 0$. Then the sequence $(\ell(\lambda^{-n} T^n))_{n\in\mathbb{N}}$ is bounded. Consider the isometric embedding $J_{\mathscr{B}(X)} : \mathscr{B}(X) \to \mathscr{B}(X)^{**}$. We have $(J_{\mathscr{B}(X)}(\lambda^{-n} T^n)(\ell))_{n\in\mathbb{N}}$ is bounded for each $\ell \in X*$. By the Uniform Boundedness Principle, $(J_{\mathscr{B}(X)}(\lambda^{-n} T^n))_{n\in\mathbb{N}}$ is bounded in $\mathscr{B}(X)^{**}$ and hence $(\lambda^{-n} T^n)_{n\in\mathbb{N}}$ is bounded in $\mathscr{B}(X)$. There exists $M > 0$ such that

  $$\forall n \in \mathbb{N} : \ |\lambda|^{-n} \|T^n\| \leqslant M$$

  Therefore

  $$\limsup_{n\to\infty} \|T^n\|^{1/n} \leqslant \limsup_{n\to\infty} M^{1/n} |\lambda| = |\lambda| = \operatorname{rad}\sigma(T) + \varepsilon$$

  Since ε is arbitrary, we deduce that $\limsup_{n\to\infty} \|T^n\|^{1/n} \leqslant \operatorname{rad}\sigma(T)$.

Combining the two results above:

$$\inf_{n\in\mathbb{N}} \|T^n\|^{1/n} \leqslant \limsup_{n\to\infty} \|T^n\|^{1/n} \leqslant \operatorname{rad}\sigma(T) \leqslant \inf_{n\in\mathbb{N}} \|T^n\|^{1/n}$$

So all inequalities become eqaulities. In particular the limit of $\|T^n\|^{1/n}$ exists. □

Proposition 6.12. Spectral Mapping Theorem for Polynomials

Suppose that X is a complex Banach space, $T \in \mathscr{B}(X)$, and $p \in \mathbb{C}[x]$. Then

$$\sigma(p(T)) = p(\sigma(T)) := \{p(\lambda) : \ \lambda \in \sigma(T)\}$$

Proof. If $p = c$ is a constant polynomial, then $\sigma(p(T)) = \{c\}$. Since $\sigma(T) \neq \varnothing$, we must have $p(\sigma(T)) = \{c\}$.

Supoose that $\deg p \geqslant 1$. For $\mu \in C$, by the Fundamental Theorem of Algebra we can factorise $p(x) - \mu$ as

$$p(x) - \mu = \alpha(z - \beta_1) \cdots (z - \beta_n)$$

where $\alpha, \beta_1, ..., \beta_n \in \mathbb{C}$. Correspondingly,

$$p(T) - \mu \operatorname{id} = \alpha(T - \beta_1 \operatorname{id}) \cdots (T - \beta_n \operatorname{id})$$

Since the operators on the right hand side mutually commute, we have

$$\begin{aligned} \mu \in \sigma(p(T)) &\iff p(T) - \mu \operatorname{id} \text{ is not invertible in } \mathscr{B}(X) \\ &\iff \exists\, j \in \{1, ..., n\}\ T - \beta_j \operatorname{id} \text{ is not invertible in } \mathscr{B}(X) \\ &\iff \exists\, j \in \{1, ..., n\}\ \beta_j \in \sigma(T) \\ &\iff \exists\, \lambda \in \sigma(T)\ \mu = p(\lambda) \\ &\iff \mu \in p(\sigma(T)) \end{aligned}$$

□

6.3 Normal Operators

Definition 6.13. Normal Operators

Suppose that X is a complex Hilbert space and $T \in \mathscr{B}(X)$. T is called a normal operator, if $TT^* = T^*T$.

We know that self-adjoint operators and unitary operators are normal. From Linear Algebra we have the following elementary properties of normal operators:

- $T \in \mathscr{B}(X)$ is normal if and only if $\|T(x)\| = \|T^*(x)\|$;
- $\lambda \in \sigma_{\mathrm{p}}(T) \iff \overline{\lambda} \in \sigma_{\mathrm{p}}(T^*)$;
- If x, y are eigenvectors of T corresponding to different eigenvalues of T, then $\langle x, y\rangle = 0$.

Proposition 6.14. Spectra of Normal Operators

Suppose that X is a complex Hilbert space and $T \in \mathscr{B}(X)$ is a normal operator. Then

1. $\sigma_{\mathrm{r}}(T) = \varnothing$ and $\sigma(T) = \sigma_{\mathrm{ap}}(T)$;
2. $\operatorname{rad} \sigma(T) = \|T\|$.

Proof.

1. Note that $\lambda \operatorname{id} - T$ is also a normal operator. Then $\lambda \operatorname{id} - T$ is injective if and only if $\overline{\lambda} \operatorname{id} - T^*$ is injective. Thus $\sigma_{\mathrm{p}}(T) = \sigma_{\mathrm{p}}^*(T^*)$. Since $\sigma_{\mathrm{r}}(T) \subseteq \sigma_{\mathrm{p}}^*(T^*)$ and $\sigma_{\mathrm{r}}(T) \cap \sigma_{\mathrm{p}}(T) = \varnothing$, we deduce that $\sigma_{\mathrm{r}}(T) = \varnothing$.
2. We first note that $\|(T^*)^n T^n\| = \|T^n\|^2$ for all $n \in \mathbb{N}$, because

$$\|(T^*)^n T^n\| = \sup_{\|x\|=1} \langle (T^*)^n T^n x, x\rangle = \sup_{\|x\|=1} \langle T^n x, T^n x\rangle = \sup_{\|x\|=1} \|T^n x\|^2 = \|T^n\|^2$$

Since T is normal, T^*T and its powers are self-adjoint. By Corollary 2.21, $\left\|(T^*T)^{2^n}\right\| = \|T^*T\|^{2^n}$

Therefore for $n_k = 2^k$,

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

By the Gelfand's Formula,

$$\operatorname{rad}\sigma(T) = \lim_{n\to\infty} \|T^n\|^{1/n} = \lim_{k\to\infty} \|T^{n_k}\|^{1/n_k} = \lim_{k\to\infty} \|T^*T\|^{1/2} = \|T^*T\|^{1/2} = \|T\| \qquad \square$$

Lemma 6.15. Norm of Resolvent Operator

Suppose that X is a Hilbert space and $T \in \mathscr{B}(X)$ is a normal operator. Let $\lambda \in \rho(T)$. Then

$$\|R_\lambda(T)\| = \frac{1}{\operatorname{dist}(\lambda, \sigma(T))}$$

Proof. First we claim that $\sigma(A)^{-1} = \sigma(A^{-1})$ for invertible operator $A \in \mathscr{B}(X)$. It follows from

$$\begin{aligned} \lambda \in \sigma(A)^{-1} &\iff \lambda^{-1} \in \sigma(A) \\ &\iff \lambda^{-1}\operatorname{id} - A \text{ is not invertible in } \mathscr{B}(X) \\ &\iff \lambda\operatorname{id} - A^{-1} = -\lambda(\lambda^{-1}\operatorname{id} - A)A^{-1} \text{ is not invertible in } \mathscr{B}(X) \\ &\iff \lambda \in \sigma(A^{-1}) \end{aligned}$$

Hence $\sigma(R_\lambda(T)) = \sigma((T - \lambda\operatorname{id})^{-1}) = (\sigma(T) - \lambda)^{-1}$. Since T is normal, so is $R_\lambda(T)$ is also normal. In particular,

$$\operatorname{rad}\sigma(R_\lambda(T)) = \|R_\lambda(T)\|$$

On the other hand, we have

$$\operatorname{rad}\sigma(R_\lambda(T)) = \sup_{\mu\in\sigma(R_\lambda(T))} |\mu| = \sup_{\mu\in(\sigma(T)-\lambda)^{-1}} |\mu| = \sup_{\mu\in\sigma(T)} \frac{1}{|\mu - \lambda|} = \frac{1}{\inf\limits_{\mu\in\sigma(T)} |\mu-\lambda|} = \frac{1}{\operatorname{dist}(\lambda, \sigma(T))} \qquad \square$$

Proposition 6.16. Spectra of Self-Adjoint Operators

Suppose that X is a complex Hilbert space and $T \in \mathscr{B}(X)$ is a self-adjoint operator. Then $\sigma(T) \subseteq [a,b] \subseteq \mathbb{R}$, where

$$a := \inf_{\|x\|=1} \langle x, Tx\rangle, \qquad b := \sup_{\|x\|=1} \langle x, Tx\rangle$$

Furthermore, $a, b \in \sigma(T)$.

Proof. Since T is self-adjoint, for all $x \in X$,

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

Hence $\langle x, Tx\rangle \in \mathbb{R}$. By Lemma 6.8 and the fact that $\sigma(T)$ is compact, $\sigma(T) \subseteq [a,b] \subseteq \mathbb{R}$.

Since $\operatorname{rad}\sigma(T) = \|T\|$, we have $\|T\| \leqslant \max\{|a|,|b|\}$. But by definition of a, b we have $\max\{|a|,|b|\} \leqslant \|T\|$. Hence $\operatorname{rad}\sigma(T) = \|T\| = \max\{|a|,|b|\}$. Since $\sigma(T)$ is closed, either $a \in \sigma(T)$ or $b \in \sigma(T)$. We apply this result to the shifted operator $c\operatorname{id} + T$, where $c \in \mathbb{R}$ is chosen to be sufficiently large and sufficiently small. This shows that both $a \in \sigma(T)$ and $b \in \sigma(T)$. $\square$

Proposition 6.17. Spectra of Unitary Operators

Suppose that X is a complex Hilbert space and $U \in \mathscr{B}(X)$ is a unitary operator. Then $\sigma(U) \subseteq S^1$, where S^1 is the unit circle in $\mathbb{C}$.

Proof. Since $\|U\| = 1$, we have $\operatorname{rad}\sigma(U) = 1$ and hence $\sigma(U) \subseteq \overline{B(0,1)}$. Suppose that $\lambda \in B(0,1) \cap \sigma(U)$. Then $\lambda\operatorname{id} - U$ is not invertible, and so is $\overline{\lambda}\operatorname{id} - U^*$. But then $\overline{\lambda}U - \operatorname{id} = (\overline{\lambda}\operatorname{id} - U^*)U$ is also not invertible. We deduce that $\overline{\lambda}^{-1} \in \sigma(U)$, which is a contradication because $\left|\overline{\lambda}^{-1}\right| > 1$. We conclude that $\sigma(U) \subseteq S^1$. $\square$

Chapter 7

Fourier Series

7.1 Definitions and Basic Properties

Definition 7.1. Complex Fourier Series

Suppose that $f : \mathbb{R} \to \mathbb{C}$ is a 2π-periodic function. We define its Fourier series by

$$\mathcal{F}[f] := \sum_{n\in\mathbb{Z}} c_n \, \mathrm{e}^{inx}, \qquad c_n := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) \, \mathrm{e}^{-inx} \, \mathrm{d}x$$

We also write $f \sim \sum_{n\in\mathbb{Z}} c_n \, \mathrm{e}^{inx}$.

Recall from *A4. Integration* that a function $F : \mathbb{R} \to \mathbb{R}$ is called **absolutely continuous**, if there exists $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$ and $x_0 \in \mathbb{R}$ such that

$$F(x) = \int_{x_0}^{x} f(t) \, \mathrm{d}t$$

F is called the **indefinite integral** of f. Some properties of F are listed below:

- F is continuous by Dominated Convergence Theorem.
- F is differentiable almost everywhere by Lebesgue Differentiation Theorem.
- If F is periodic then so is f.
- Integration by parts: for $\varphi \in \mathrm{C}^\infty(\mathbb{R})$,

$$\int_a^b F\varphi' = F\varphi\Big|_a^b - \int_a^b f\varphi$$

Proposition 7.2. Termwise Differentiation of Fourier Series

Suppose that $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$ and let F be the indefinite integral of f. If F is 2π-periodic and $F \sim \sum_{n\in\mathbb{Z}} c_n \, \mathrm{e}^{inx}$, then f is 2π-periodic and $f \sim \sum_{n\in\mathbb{Z}} inc_n \, \mathrm{e}^{inx}$.

Proof. We have

$$f = F' = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} \quad \text{(a.e.)}$$

Hence f is also 2π-periodic. The 0^{th} Fourier coefficient of f is given by

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) \, \mathrm{d}x = \frac{1}{2\pi}\left(F(\pi) - F(-\pi)\right) = 0$$

For other Fourier coefficients, we integrate by parts:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x = \frac{\mathrm{i}n}{2\pi}\int_{-\pi}^{\pi} F(x)\,\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x = \mathrm{i}nc_n$$

□

Proposition 7.3. Termwise Integration of Fourier Series

Suppose that $f \in \mathrm{L}^1(-\pi,\pi)$ and let F be the indefinite integral of f. If f is 2π-periodic and $F \sim \sum_{n\in\mathbb{Z}} c_n\,\mathrm{e}^{\mathrm{i}nx}$, then $F(x) - c_0 x$ is 2π-periodic and $F(x) - c_0 x \sim C_0 + \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{c_n}{\mathrm{i}n}\,\mathrm{e}^{\mathrm{i}nx}$, where $C_0 \in \mathbb{C}$ is a suitable constant.

Proof. Let $G(x) := F(x) - c_0 x$. We have

$$G(x+2\pi) - G(x) = \int_x^{x+2\pi} f(t)\,\mathrm{d}t - 2\pi c_0 = 2\pi c_0 - 2\pi c_0 = 0$$

So G is 2π-periodic. The Fourier series of $f - c_0$ is obtained by termwise differentiation of the Fourier series of G. The result follows easily. □

7.2 Convergence in $\mathrm{L}^p(-\pi,\pi)$

We aim to study the modes of convergence of Fourier series.

Definition 7.4. Partial Fourier Sum

For a 2π-periodic function $f:\mathbb{R}\to\mathbb{C}$, we define the partial Fourier sum by

$$(S_N f)(x) := \sum_{n=-N}^{N} c_n\,\mathrm{e}^{\mathrm{i}nx} = \sum_{n=-N}^{N} \mathrm{e}^{\mathrm{i}nx}\,\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,\mathrm{e}^{-\mathrm{i}nt}\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\sum_{n=-N}^{N}\mathrm{e}^{\mathrm{i}n(x-t)}\,\mathrm{d}t =: \int_{-\pi}^{\pi} f(t)k_N(x-t)\,\mathrm{d}t$$

where

$$k_N(x) := \frac{1}{2\pi}\sum_{n=-N}^{N}\mathrm{e}^{\mathrm{i}nx} = \frac{1}{2\pi}\frac{\sin\big(N+\frac{1}{2}\big)x}{\sin\frac{x}{2}}$$

is called the **Dirichlet kernel**. The partial Fourier sum can be expressed as a convolution:

$$S_N f = f * k_N$$

The L^2 theory is particular interesting. Let $e_n(x) := \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{\mathrm{i}nx} \in \mathrm{L}^2(-\pi,\pi)$. From *M5. Fourier Series & PDEs* we have already seen that

$$\langle e_m, e_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}nx}\,\mathrm{e}^{-\mathrm{i}mx}\,\mathrm{d}x = \delta_{mn}$$

So $\{e_n\}_{n\in\mathbb{Z}}$ is an orthonormal system in $\mathrm{L}^2(-\pi,\pi)$. We shall show that it is in fact complete. The Fourier series of $f \in \mathrm{L}^2(-\pi,\pi)$ is the expansion with respect to that basis:

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\langle e_n, f\rangle\,, \qquad f \sim \sum_{n\in\mathbb{Z}} c_n\,\mathrm{e}^{\mathrm{i}nx} = \sum_{n\in\mathbb{Z}}\langle e_n, f\rangle\, e_n$$

Theorem 7.5. Completeness of Trigonometric System in $\mathrm{L}^2(-\pi,\pi)$

$\{e_n\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $\mathrm{L}^2(-\pi,\pi)$. Equivalent, for every $f \in \mathrm{L}^2(-\pi,\pi)$, the partial Fourier sum $S_N(f)$ of f converges strongly in $\mathrm{L}^2(-\pi,\pi)$ to f.

Proof using Stone-Weierstrass Theorem:

Let $\mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$ be the set of 2π-periodic continuous functions on $\mathbb{R}$, equipped with the supremum norm. Note that $\mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$ is dense in $\mathrm{L}^2(-\pi,\pi)$. So it suffices to show that $\operatorname{span}\{e_n\}_{n\in\mathbb{Z}}$. Since $\mathbb{R}/2\pi\mathbb{Z}$ is compact, the proof can be concluded by using Stone-Weierstrass Theorem. We need to show that $\operatorname{span}\{e_n\}_{n\in\mathbb{Z}}$ forms a subalgebra.

The linear span of $\{e_n\}_{n\in\mathbb{Z}}$ clearly contains constant functions and is closed under pointwise multiplication (which is because $\{e_n\}_{n\in\mathbb{Z}}$ is closed under pointwise multiplication: $e_n(x)e_m(x) = e_{n+m}(x)$). Hence $\operatorname{span}\{e_n\}_{n\in\mathbb{Z}}$ is a subalgebra of $\mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$. In addition, $e_1(x) = \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}x}$ is injective on $\mathbb{R}/2\pi\mathbb{Z}$, and hence separates points. This proves the claim. □

Proof using Projection Theorem:

By the Projection Theorem we have

$$\mathrm{L}^2(-\pi,\pi) = \overline{\operatorname{span}\{e_n\}_{n\in\mathbb{Z}}} \oplus (\operatorname{span}\{e_n\}_{n\in\mathbb{Z}})^\perp$$

It suffices to show that $(\operatorname{span}\{e_n\}_{n\in\mathbb{Z}})^\perp = \{0\}$. Let $f \in \mathrm{L}^2(-\pi,\pi)$ of which all Fourier coefficients vanish. We need to show that $f = 0$ a.e..

- f is continuous.

  We only prove the case where f is real-valued. Suppose for contradication that $f \neq 0$. Since $|f|$ is periodic and continuous, it attains maximum value $M > 0$ at some $x_0 \in [-\pi,\pi]$. Replacing f by $-f$ if necessary, we may assume that $f(x_0) = M > 0$. Using a translation we may further assume that $x_0 = 0$. Let $\delta > 0$ such that $|f| > M/2$ in $(-\delta,\delta)$. Consider the trigonometric polynomial

  $$g(x) = 1 + \cos x - \cos\delta$$

  Note that

  $$g(x)\begin{cases} > 1 & x \in (-\delta,\delta) \\ \leqslant 1 & x \in (-\pi,\pi)\setminus(-\delta,\delta) \end{cases}$$

  Therefore

  $$\begin{aligned} \int_{-\pi}^{\pi} f(x)g^n(x)\,\mathrm{d}x &\geqslant \int_{-\delta/2}^{\delta/2} f(x)g^n(x)\,\mathrm{d}x - \int_{(-\pi,\pi)\setminus(-\delta,\delta)} |f(x)||g^n(x)|\,\mathrm{d}x \\ &\geqslant \frac{M}{2}\left(1 + \cos\frac{\delta}{2} - \cos\delta\right)^n \delta - 2\pi M \to \infty \end{aligned}$$

  as $n \to \infty$. However, on the other hand, since g^n is a trigonometric polynomial, that the Fourier coefficients of f are zero implies that

  $$\int_{-\pi}^{\pi} f(x)g^n(x)\,\mathrm{d}x = 0$$

  for all $n > 0$, which leads to a contradiction.

- f is only square-integrable.

  Let F be an indefinite integral of f. Since $\langle 1, f\rangle = 0$,

  $$F(x+2\pi) - F(x) = \int_x^{x+2\pi} f(t)\,\mathrm{d}t = 0$$

  Hence F is 2π-periodic. By termwise integration, all Fourier coefficients except for possibly the 0^{th} are zero. Note that F ia continuous. So by the previous result $F(x) = C_0$ for some constant $C_0 \in \mathbb{C}$. Hence $f = 0$ a.e. as desired. □

The second proof above only uses the integrability of f. So we have the following corollary:

Corollary 7.6

Let $f \in \mathrm{L}^1(-\pi,\pi)$. If all Fourier coefficients of f vanish, then $f = 0$ alomost everywhere.

The first proof above generalises to L^p easily:

Corollary 7.7

Let $1 \leqslant p < \infty$. Then $\operatorname{span}\{e_n\}_{n\in\mathbb{Z}}$ is dense in $\mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$ and hence in $\mathrm{L}^p(-\pi,\pi)$.

Theorem 7.8. Carleson-Hunt Theorem

Let $1 < p < \infty$. The partial Fourier sum S_N converges strongly in $\mathrm{L}^p(-\pi,\pi)$ to id. That is, $S_N f \to f$ in $\mathrm{L}^p(-\pi,\pi)$ as $N \to \infty$ for each $f \in \mathrm{L}^p(-\pi,\pi)$.

Furthermore, the convergence $S_N f \to f$ holds almost everywhere.

Proof. Omitted. This is a very difficult result. □

For L^1 divergence, we have the following result.

Theorem 7.9. Kolmogorov's Theorem

There exists a function $f \in \mathrm{L}^1(-\pi,\pi)$ such that $(S_N f)_{N\in\mathbb{N}}$ diverges everywhere. In fact, there exists a set E of positive measure and a sequence (N_k) such that $|S_{N_k} f(x)| \to \infty$ as $k \to \infty$ for $x \in E$. In particular, $(S_N f)$ does not converge in $\mathrm{L}^1(-\pi,\pi)$.

Proof. Omitted. □

We discuss the L^∞ (specifically $\mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$) divergence in the next chapter.

7.3 Pointwise Convergence and Divergence

Theorem 7.10. Riemann-Lebesgue Lemma

For $f \in \mathrm{L}^1(-\pi,\pi)$, we have

$$\lim_{|n|\to\infty} \int_{-\pi}^{\pi} f(x)\, \mathrm{e}^{inx}\ \mathrm{d}x = 0$$

In other words, the Fourier coefficients (c_n) of f satisfies $c_n \to 0$ as $|n| \to \infty$.

Proof. Since $\mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$ is dense in $\mathrm{L}^1(-\pi,\pi)$, we split $f = g + h$, where $g \in \mathrm{C}(\mathbb{R}/2\pi\mathbb{Z}) \subseteq \mathrm{L}^2(-\pi,\pi)$ and $\|h\|_1 \leqslant \varepsilon$.

For the function $g \in \mathrm{L}^2(-\pi,\pi)$, by Bessel's Inequality we have

$$\sum_{n=-\infty}^{\infty} |\langle e_n, g\rangle| \leqslant \|g\|_2$$

and hence

$$\lim_{|n|\to\infty} |\langle e_{-n}, g\rangle| = \lim_{|n|\to\infty} \left| \int_{-\pi}^{\pi} g(x)\, \mathrm{e}^{inx}\ \mathrm{d}x \right| = 0$$

For h we have

$$\left| \int_{-\pi}^{\pi} h(x)\, \mathrm{e}^{inx}\ \mathrm{d}x \right| \leqslant \int_{-\pi}^{\pi} |h(x)|\, \mathrm{d}x \leqslant \varepsilon$$

Since ε is arbitrary, the result follows. □

Definition 7.11. Hölder Continuity

For some $\alpha \in (0,1]$, we say that f is α-Hölder continuous at $x_0 \in \mathbb{R}$, if there exist $A > 0$ and $\delta > 0$ such that $|f(x) - f(y)| \leqslant A|x-y|^\alpha$ for all $x, y \in B(x_0, \delta)$.

We can accordingly define the left and right α-Hölder continuity of a function.

Theorem 7.12. Dirichlet's Theorem

Suppose that $f \in L^1(-\pi,\pi)$ is 2π-periodic. If f is left and right α-Hölder continuous at $x_0 \in \mathbb{R}$ for some $\alpha \in (0,1]$. Then

$$\lim_{N\to\infty} S_N f(x_0) = \frac{1}{2}\left(f(x_0^+) + f(x_0^-)\right)$$

Proof. We prove the theorem in the case where f is α-Hölder continuous at x_0. The theorem holds obviously for constant function f. Therefore we may assume that $f(x_0) = 0$. We fix $\delta > 0$ such that $|f(x_0+h)| \leqslant A|h|^\alpha$ for $|h| < \delta$.

Recall that

$$S_N f(0) = \int_{-\pi}^{\pi} f(t) k_N(0-t)\,\mathrm{d}t = \int_0^{\pi} (f(t) + f(-t)) k_N(t)\,\mathrm{d}t$$

For $0 < t < \delta$, we have

$$\sin\frac{t}{2} \geqslant \frac{t}{\pi} \implies |k_N(t)| \leqslant \frac{1}{2\pi}\frac{1}{\sin\frac{t}{2}} \leqslant \frac{1}{2t}$$

Therefore

$$\left|\int_0^\delta (f(t)+f(-t))k_N(t)\,\mathrm{d}t\right| \leqslant \int_0^\delta |f(t)+f(-t)||k_N(t)|\,\mathrm{d}t \leqslant \int_0^\delta 2At^\alpha \cdot \frac{1}{2t}\,\mathrm{d}t = A\alpha^{-1}\delta^\alpha$$

It remains to consider

$$\begin{aligned} J_{N,\delta} :&= \int_\delta^\pi (f(t)+f(-t))k_N(t)\,\mathrm{d}t = \int_\delta^\pi (f(t)+f(-t))\left(\cos Nt + \cot\frac{t}{2}\sin Nt\right)\,\mathrm{d}t \\ &= \int_{-\pi}^{\pi} \left(g_\delta(t)\cos Nt + h_\delta \sin Nt\right)\,\mathrm{d}t \end{aligned}$$

where

$$g_\delta(t) := \mathbf{1}_{(\delta,\pi)}(t)(f(t)+f(-t)), \qquad h_\delta(t) = \mathbf{1}_{(\delta,\pi)}(t)(f(t)+f(-t))\cot\frac{t}{2}$$

Since $g_\delta, h_\delta \in L^1(-\pi,\pi)$, by Riemann-Lebesgue Theorem, we have $J_{N,\delta} \to 0$ as $N \to \infty$. Combining the two intervals we have

$$\limsup_{N\to\infty} |S_N f(0)| \leqslant A\alpha^{-1}\delta^\alpha$$

We take $\delta \to 0$ and conclude that $S_N f(0) \to 0$ as $N \to \infty$. □

Corollary 7.13. Localisation Property of Fourier Series

Suppose that $f, g \in C(\mathbb{R}/2\pi\mathbb{Z})$ are equal in an open interval of $x_0 \in (-\pi,\pi)$. Then the Fourier series of f and g either both converge at x_0 or both diverge at x_0.

Proof. We have $f, g \in C(\mathbb{R}/2\pi\mathbb{Z}) \subseteq L^1(-\pi,\pi)$ and then $f - g \in L^1(-\pi,\pi)$. $f - g = 0$ in the interval (a,b) where $a < x_0 < b$. In particular $f - g$ is α-Hölder continuous at x_0 for any $\alpha \in (0,1]$. We have

$$\lim_{N\to\infty} (S_N f(x_0) - S_N g(x_0)) = \lim_{N\to\infty} S_N(f-g)(x_0) = 0$$

So $S_N f(x_0)$ and $S_N g(x_0)$ either both converge or both diverge. □

Theorem 7.14. Cesàro Convergence of Fourier Series

For $f \in C(\mathbb{R}/2\pi\mathbb{Z})$, it holds that

$$\sigma_N f := \frac{1}{N+1}\sum_{n=0}^{N} S_n f \to f \text{ in } C(\mathbb{R}/2\pi\mathbb{Z}) \text{ as } N \to \infty$$

Proof. We write the partial Fourier sum as a convolution: $S_N f = f * k_N$. Then $\sigma_N f = f * F_N$, where

$$F_N(x) = \frac{1}{N+1}\sum_{n=0}^{N} k_n(x) = \frac{1}{2\pi(N+1)}\frac{1-\cos(N+1)x}{1-\cos x}$$

F_N is called the **Féjer kernel**. We can verify that it has the following good properties:

- $F_N \geqslant 0$;
- $\|F_N\|_1 = 1$;
- For $\delta < x < \pi$, $0 \leqslant F_N(x) \leqslant \dfrac{1}{\pi(N+1)(1-\cos\delta)}$.

We write

$$|\sigma_N f(x) - f(x)| = \left|\int_{-\pi}^{\pi} f(t)F_N(x-t)\,\mathrm{d}t - f(x)\int_{-\pi}^{\pi} F_N(t)\,\mathrm{d}t\right| = \left|\int_{-\pi}^{\pi}(f(x-t)-f(x))F_N(t)\,\mathrm{d}t\right|$$

Fix $\varepsilon > 0$. Since f is uniformly continuous, there exists $\delta > 0$ such that $|f(x-t)-f(x)| < \varepsilon$ for $|t| < \delta$. Therefore

$$\left|\int_{-\delta}^{\delta}(f(x-t)-f(x))F_N(t)\,\mathrm{d}t\right| < \varepsilon\left|\int_{-\delta}^{\delta} F_N(t)\,\mathrm{d}t\right| < \varepsilon$$

For $|t| > \delta$, we have

$$\left|\int_{(-\pi,\pi)\setminus(-\delta,\delta)}(f(x-t)-f(x))F_N(t)\,\mathrm{d}t\right| \leqslant 2\|f\|_\infty \cdot \frac{1}{\pi(N+1)(1-\cos\delta)} \to 0$$

as $N \to \infty$. Combining the two intervals we have

$$\limsup_{N\to\infty} |\sigma_N f(x) - f(x)| \leqslant \varepsilon$$

which holds for all $x \in (-\pi,\pi)$ and $\varepsilon > 0$. We conclude that $\|\sigma_N f - f\|_\infty \to 0$ as $N \to \infty$. □

The properties of Féjer kernels can be generalised to the so-called good kernels.

Corollary 7.15. Good Kernels and Uniform Convergence

We say that $(K_N)_{N\in\mathbb{N}} \subseteq \mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$ is a family of **good kernels**, if they satisfy the following properties:

1. (K_N) is uniformly bounded in $\mathrm{L}^1(-\pi,\pi)$;
2. $\displaystyle\int_{-\pi}^{\pi} K_N(x)\,\mathrm{d}x = 1$;
3. For any $\delta > 0$, we have $\displaystyle\lim_{N\to\infty}\int_{(-\pi,\pi)\setminus(-\delta,\delta)} |K_N(x)|\,\mathrm{d}x = 0$.

Then for $f \in \mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$, we have $f * K_N \to f$ in $\mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$.

Next we discuss the divergence of Fourier series of continuous functions.

Proposition 7.16. Divergence of Fourier Series in $\mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$

For every $x_0 \in \mathbb{R}$, there exists $f \in \mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$ such that $(S_N f(x_0))_{N\in\mathbb{N}}$ is divergent. In particular S_N does not converges strongly in $\mathscr{B}(\mathrm{C}(\mathbb{R}/2\pi\mathbb{Z}))$.

Proof. We define $A_N \in \mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})^*$ by

$$A_N f := S_N f(0) = \int_{-\pi}^{\pi} f(x)k_N(x)\,\mathrm{d}x$$

Suppose for contradiction that the Fourier series of all $f \in \mathrm{C}(\mathbb{R}/2\pi\mathbb{Z})$ converges at 0. Then $A_N f$ is bounded for every f. By the Uniform Boundedness Principle, $(\|A_N\|_*)_{N\in\mathbb{N}}$ is bounded.

Take $\varepsilon > 0$. Consider the continuous function f defined by

$$f = \begin{cases} \operatorname{sgn}(k_N(x)) & |k_N(x)| \geqslant \varepsilon \\ \text{linear} & |k_N(x)| \leqslant \varepsilon \end{cases}$$

We have $\|f\|_\infty = 1$ and

$$A_N f \geqslant \int_{\{|k_N(x)|\geqslant\varepsilon\}} |k_N(x)|\,\mathrm{d}x - \int_{\{|k_N(x)|<\varepsilon\}} |f(x)||k_N(x)|\,\mathrm{d}x \geqslant \|k_N\|_1 - 2\pi\varepsilon$$

Hence we have $\|A_N\|_* \geqslant |A_N f| \geqslant \|k_N\|_1 - 2\pi\varepsilon$. Taking $\varepsilon \to 0$ we deduce that $\|A_N\|_* \geqslant \|k_N\|_1$.

Finally we need to estimate $\|k_N\|_1$:

$$\begin{aligned}
\|k_N\|_1 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|\sin\left(N+\frac{1}{2}\right)x|}{|\sin\frac{x}{2}|}\,\mathrm{d}x \geqslant \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{|\sin\left(N+\frac{1}{2}\right)x|}{|x|}\,\mathrm{d}x = \frac{2}{\pi}\int_0^{(N+\frac{1}{2})\pi} \frac{|\sin x|}{|x|}\,\mathrm{d}x \\
&\geqslant \frac{2}{\pi}\int_{\{x\in(0,(N+\frac{1}{2})\pi):\ |\sin x|>\frac{1}{2}\}} \frac{|\sin x|}{|x|}\,\mathrm{d}x \geqslant \frac{1}{\pi}\sum_{k=0}^{N}\int_{(k+\frac{1}{3})\pi}^{(k+\frac{1}{2})\pi} \frac{\mathrm{d}x}{|x|} \\
&= \frac{1}{\pi}\sum_{k=0}^{N} \ln\frac{k+\frac{1}{2}}{k+\frac{1}{3}}
\end{aligned}$$

For large k,

$$\ln\frac{k+\frac{1}{2}}{k+\frac{1}{3}} = \ln\left(1+\frac{1}{6k+2}\right) \sim \frac{1}{6k+2}$$

Hence

$$\|k_N\|_1 \geqslant C\sum_{k=0}^{N}\frac{1}{6k+2} \sim C\ln N \to \infty$$

as $N \to \infty$. We deduce that $\|A_N\|_* \to \infty$ as $N \to \infty$, which is contradictory.

Hence there exists a continuous function with divergent Fourier series at 0. By translation there exist continuous functions with divergent Fourier series at any given $x_0 \in \mathbb{R}$. □

The idea above can be used to construct an explicit example.

Example 7.17. Continuous Function with Divergent Fourier Series at 0

For $x \in [0, \pi]$, we set

$$f(x) = \sum_{p=1}^{\infty}\frac{1}{p^2}\sin\left(\left(2^{p^3}+1\right)\frac{x}{2}\right)$$

We extend the domain to $\mathbb{R}$ such that f is a 2π-periodic even function. Then f is continuous with Fourier series divergent at 0.

Proof. By Weierstrass M-test, the series that defines f is uniformly convergent on $\mathbb{R}$. Hence f is continuous on $\mathbb{R}$. The Fourier series of f is given by

$$f(x) \sim \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n\cos nx$$

The coefficients are given by

$$\begin{aligned}
a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}x \\
&= \frac{2}{\pi}\int_0^{\pi}\sum_{p=1}^{\infty}\frac{1}{p^2}\sin\left(\left(2^{p^3}+1\right)\frac{x}{2}\right)\cos nx\,\mathrm{d}x \\
&= \frac{2}{\pi}\sum_{p=1}^{\infty}\frac{1}{p^2}\int_0^{\pi}\sin\left(\left(2^{p^3}+1\right)\frac{x}{2}\right)\cos nx\,\mathrm{d}x \qquad \text{(by uniform convergence)}
\end{aligned}$$

$$= \frac{2}{\pi} \sum_{p=1}^{\infty} \frac{1}{p^2} \alpha\left(2^{p^3-1}, n\right)$$

where $\alpha(m,n) := \displaystyle\int_0^\pi \sin\left(\frac{2m+1}{2}x\right) \cos nx \, \mathrm{d}x = \frac{1}{2}\left(\frac{1}{m+n+\frac{1}{2}} + \frac{1}{m-n+\frac{1}{2}}\right)$.

To show that the Fourier series of f diverges at 0, it suffices to show that

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty}\sum_{p=1}^{\infty} \frac{1}{p^2} \alpha\left(2^{p^3-1}, n\right) = \sum_{p=1}^{\infty} \frac{1}{p^2} \sum_{n=0}^{\infty} \alpha\left(2^{p^3-1}, n\right) = \infty$$

For $m, N \geqslant 1$,

$$\sum_{n=0}^{N} \alpha(m,n) = \frac{1}{2}\sum_{n=0}^{N}\left(\frac{1}{m+n+\frac{1}{2}} + \frac{1}{m-n+\frac{1}{2}}\right) = \frac{1}{2}\left(\frac{1}{m+\frac{1}{2}} + \sum_{i=m-N}^{m+N} \frac{1}{i+\frac{1}{2}}\right) \geqslant 0$$

When $m = N$,

$$\sum_{n=0}^{m} \alpha(m,n) = \frac{1}{2}\left(\frac{1}{m+\frac{1}{2}} + \sum_{i=0}^{2m} \frac{1}{i+\frac{1}{2}}\right) \sim \frac{1}{2}\ln m$$

for large m. Therefore

$$\sum_{p=1}^{\infty} \frac{1}{p^2} \sum_{n=0}^{\infty} \alpha\left(2^{p^3-1}, n\right) \geqslant \frac{1}{p^2} \sum_{n=0}^{2^{p^3-1}} \alpha\left(2^{p^3-1}, n\right) \sim \frac{1}{2p^2} \ln\left(2^{p^3-1}\right) = \frac{p^3-1}{2p^2} \ln 2 \to \infty$$

as $p \to \infty$. This completes the argument. □

References

- Melanie Rupflin, *B4.1 Functional Analysis I* (2020-2021).
- Luc Nguyen, *B4.2 Functional Analysis II* (2020-2021).
- Stuart White, *C4.1 Further Functional Analysis* (2020-2021).
- Zhongmin Qian, *A3 Integration* (2016-2017).
- Jan Kristensen, *B4.3 Distribution Theory* (2020-2021).
- Theo Bühler & Dietmar A. Salamon, *Functional Analysis* (lecture notes for the 3rd-year course *Functional Analysis I* at ETH Zürich).
- H. L. Royden, *Real Analysis* (4th Edition).
- Gerald B. Folland, *Real Analysis - Modern Techniques and Their Applications* (2nd Edition)
- 张恭庆, 林源渠, 泛函分析讲义（上册）.
- 周民强, 实变函数论.
- 于品, 数学分析讲义（清华丘成桐英才班）.