Function Spaces and Applications*

Thomas Walker

Autumn 2023

Contents

*Chapter 9 of these notes, along with some preliminaries provided in Chapter 10.5, contains material from Chapter 10 and 11 of [1]. This is not examinable, however, it supplements the existing material.

Introduction

What is functional analysis? Essentially, it is linear algebra in infinite dimensions. There are two main sources of differences that arise as we move to infinite dimensions.

1. Norms are no longer equivalent.
   - Recall, that a norm is a function $\|\cdot\|$ on a vector space satisfying the following.
     (a) $\|\lambda x\| = |\lambda|\|x\|$.
     (b) $\|x+y\| \leq \|x\| + \|y\|$.
     (c) $\|x\| = 0$ if and only if $x = 0$.
   - Norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if there exists a constant c such that
   $$\frac{1}{c}\|\cdot\|_2 \leq \|\cdot\|_1 \leq c\|\cdot\|_2.$$
2. Linear operators. We can represent linear operators as matrices acting on vectors.
   $$\begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & \ddots & \\ \vdots & & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{\infty} a_{1k}x_k \\ \vdots \\ \vdots \end{pmatrix}.$$

   From which questions about convergence arise.

1 Topological and Metric Spaces

1.1 Topological Spaces

Let X be a set.

Definition 1.1.1. *A subset $\mathcal{O}$ of $\mathcal{P}(X)$ is a topology if the following hold.*

1. $\emptyset, X \in \mathcal{O}$.
2. *For a family* $(O_i)_{i\in\mathcal{I}} \subseteq \mathcal{O}$ *we have* $\bigcup_{i\in\mathcal{I}} O_i \in \mathcal{O}$.
3. *For a family* $(O_i)_{i=1}^n \subseteq \mathcal{O}$ *we have* $\bigcap_{i=1}^n O_i \in \mathcal{O}$.

Elements of the topology are called open.

Definition 1.1.2. *For a topology $(X, \mathcal{O})$, a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ converges to x if for all $O \in \mathcal{O}$ with $x \in O$, there exists an $N \in \mathbb{N}$ such that $x_n \in O$ for $n \geq N$.*

Example 1.1.3. *Topologies for a set X include the following.*

- $\mathcal{O} = \{\emptyset, X\}$.
- $\mathcal{O} = \mathcal{P}(X)$.

1.2 Metric Spaces

Let X be a set.

Definition 1.2.1. *A metric is an application $d : X \times X \to [0, \infty)$ with the following properties.*

1. *Definite. That is, $d(x, y) = 0$ if and only if $x = y$.*
2. *Symmetric. That is, $d(x, y) = d(y, x)$.*
3. *Satisfies the triangle inequality. That is, $d(x, y) \leq d(x, z) + d(z, y)$.*

A set X with a metric d is called a metric space, denoted (X, d).

Definition 1.2.2. *The ball with centre $x \in X$ and radius $r \geq 0$ is the set*

$$B_r(x) = B(x, r) = \{y \in X : d(x, y) < r\}.$$

Definition 1.2.3. *A set $O \subseteq X$ is open if for all $x \in O$ there exists an $r > 0$ such that $B(x, r) \subseteq O$.*

Definition 1.2.4. *A set is closed if its complement is an open set.*

Example 1.2.5. *Some examples of sets with metrics are the following.*

- $\mathbb{R}^n$ *and* $d(x, y) = \sum_{i=1}^n |x_i - y_i|$.
- $\mathcal{C}([0, 1]; \mathbb{R})$, *the set of continuous functions from* $[0, 1] \to \mathbb{R}$, *and* $d(f, g) = \sup_{x\in[0,1]} |f(x) - g(x)|$.

Proposition 1.2.6. *Let (X, d) be a metric space, and let $\mathcal{O}$ be the set of open sets. Then $\mathcal{O}$ is a topology.*

Proof.

- Clearly $X \in \mathcal{O}$, as for any $r > 0$ and $x \in X$ we have that $B_r(x) \subseteq X$. Moreover, $\emptyset \in \mathcal{O}$, as Definition 1.2.3 holds trivially for the set has no elements.

- Let $(O_i)_{i\in I} \subseteq \mathcal{O}$. Then for any $x \in \bigcup_{i\in I} O_i$ we have $x \in O_i$ for some $i \in I$, and so there exists an $r > 0$ such that $B_r(x) \subseteq O_i \subseteq \bigcup_{i\in I} O_i$. Therefore, $\bigcup_{i\in I} O_i \in \mathcal{O}$.

- Let $(O_i)_{i=1}^n \subseteq \mathcal{O}$. Then for any $x \in \bigcap_{i=1}^n O_i$ there exists an $r_i > 0$ such that $B_{r_i}(x) \subseteq O_i$ for each $i = 1, \dots, n$. Let $r = \min(r_1, \dots, r_n) > 0$, then $B_r(x) \subseteq \bigcap_{i=1}^n O_i$. Therefore, $\bigcap_{i=1}^n O_i \in \mathcal{O}$.

With each of these, we conclude that $\mathcal{O}$ is a topology. □

In Definition 1.1.2, the notion of convergence is formulated in a topology. In a metric space $x_n \overset{n\to\infty}{\longrightarrow} x$ if and only if $d(x_n, x) \overset{n\to\infty}{\longrightarrow} 0$.

1.2.1 Sets

Definition 1.2.7. *Let (X, d) be a metric space with $S \subseteq X$.*

1. *S is closed if S^c is open.*
2. *The closure of S denoted $\bar{S}$, is the smallest closed set which contains S. One can formulate this as $\bar{S} = \bigcap_{C \text{ closed}, C\supseteq S} C$.*
   - *Equivalently, we can say that for any $x \in \bar{S}$ there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq S$ such that $x_n \to x$.*
3. *The interior of S denoted $\mathring{S}$, is the largest open set contained in S. One can formulate this as $\mathring{S} = \bigcup_{O \text{ open}, O\subseteq S} O$.*
   - *Equivalent, for every $x \in \mathring{S}$ there exists an $r > 0$ such that $B(x, r) \subseteq S$.*

Definition 1.2.8. *A subset $A \subseteq X$ is dense if $\bar{A} = X$.*

Example 1.2.9. *The property of being dense is dependent on extrinsic factors, namely the parent set.*

- *For $\mathbb{Q} \subseteq \mathbb{R}$ we have $\bar{\mathbb{Q}} = \mathbb{R}$.*
- *For $\mathbb{Z} \subseteq \mathbb{R}$ we have $\bar{\mathbb{Z}} = \mathbb{Z}$.*

Proposition 1.2.10. *Let $A \subseteq X$. Then $A = \mathring{A}$ if and only if A is open in (X, d).*

Proof. ($\Rightarrow$) If $A = \mathring{A}$ then A is open as $\mathring{A}$ is open.
($\Leftarrow$) If A is open then

$$\mathring{A} = A \cup \bigcup_{V \text{ open}, V\subseteq A} V$$

which implies that $A \subseteq \mathring{A}$. Therefore, as $\mathring{A} \subseteq A$ it follows that $A = \mathring{A}$. □

Proposition 1.2.11. *Let $A \subseteq X$. Then $A = \bar{A}$ if and only if A is closed in (X, d).*

Proof. $(\Rightarrow)$ If $A = \bar{A}$ then A is closed as $\bar{A}$ is closed.
$(\Leftarrow)$ If A is closed then

$$\bar{A} = A \cap \bigcap_{F \text{ closed}, A \subseteq F} F$$

which implies that $\bar{A} \subseteq A$ and hence $A = \bar{A}$. □

Definition 1.2.12. *A subset $S \subseteq X$ is bounded if there exists an $x \in X$ and $r > 0$ such that $S \subseteq B(x, r)$.*

1.2.2 Continuity

Let (X, d) and (Y, d') be metric spaces and let $f : X \to Y$.

Proposition 1.2.13. *For $x_0 \in X$ the following are equivalent.*

1. *For all $\epsilon > 0$ there exists a $\delta > 0$ such that $d(x_0, y) < \delta$ implies that*

$$d'(f(x_0), f(y)) < \epsilon.$$

2. *If $(x_n)_{n\in\mathbb{N}} \subseteq X$ converges to x_0 then $f(x_n) \to f(x_0)$.*

Proof. $(1) \Rightarrow (2)$. Let $(x_n)_{n\in\mathbb{N}}$ be such that $x_n \to x_0$. Given an $\epsilon > 0$ let $\delta > 0$ be such that $d(x_0, y) < \delta$ implies $d'(f(x_0), f(y)) < \epsilon$. There exists an $N \in \mathbb{N}$ such that $d(x_n, x_0) < \delta$ for $n \geq N$. Hence, $d'(f(x_0), f(x_n)) < \epsilon$ for $n \geq N$. Therefore, $f(x_n) \to f(x_0)$.
$(2) \Rightarrow (1)$. Suppose that for $\epsilon > 0$ no $\delta > 0$ exists such that $d(x_0, y) < \delta$ implies $d'(f(x_0), f(y)) < \epsilon$. Then for each $n \in \mathbb{N}$ there exists an $x_n \in X$ such that $d(x_0, x_n) < \frac{1}{n}$ and $d'(f(x_0), f(x_n)) \geq \epsilon$. In particular, $(x_n)_{n\in\mathbb{N}}$ is a sequence converging to x_0, and so by assumption $f(x_n) \to f(x_0)$. However, this contradicts the construction of the sequence. □

Remark 1.2.14. *If either of the conditions of Proposition 1.2.13 hold, then f is said to be continuous at x_0.*

Proposition 1.2.15. *The following are equivalent.*

1. *For any open set $O \subseteq Y$, the set $f^{-1}(O)$ is open in X.*
2. *f is continuous at any $x_0 \in X$.*

Proof. $(1) \Rightarrow (2)$. For $x_0 \in X$, given an $\epsilon > 0$ let $O = B_\epsilon(f(x_0))$. Then the set $f^{-1}(O)$ is open and such that $x \in f^{-1}(O)$. Hence, there exists a $\delta > 0$ such that $B_\delta(x) \subseteq f^{-1}(O)$. In particular, this means that if $d(x_0, y) < \delta$ then $d'(f(x_0), f(y)) < \epsilon$. Therefore, we conclude that f is continuous at $x_0 \in X$ by statement 1 of Proposition 1.2.13.
$(2) \Rightarrow (1)$. Consider $x_0 \in f^{-1}(O)$. Since $f(x_0) \in O$ and O is open, there exists an $\epsilon > 0$ such that $B_\epsilon(f(x_0)) \subseteq O$. As f is continuous at x_0, there exists a $\delta > 0$ such that $f(B_\delta(x_0)) \subseteq B_\epsilon(f(x_0)) \subseteq O$, statement 1 of Proposition 1.2.13. Therefore, $B_\delta(x_0) \subseteq f^{-1}(O)$ which means that $f^{-1}(O)$ is open. □

Remark 1.2.16. *If either of the conditions of Proposition 1.2.15 hold, then f is said to be continuous on X.*

Proposition 1.2.13 provides a local viewpoint of continuity, whilst Proposition 1.2.15 global viewpoint.

Definition 1.2.17. *A map f is uniformly continuous on X if for any $\epsilon > 0$ there exists a $\delta > 0$ such that for any $(x, y) \in X \times X$ with $d(x, y) < \delta$ we have that $d'(f(x), f(y)) < \epsilon$.*

Example 1.2.18.

1. *Consider* $f : [1, \infty) \to \mathbb{R}$ *given by* $f(x) = x^2$. *Then for any* $\delta > 0$ *we have*

$$|f(x) - f(x + \delta)| = \left|(x + \delta)^2 - x^2\right| = \left|2x\delta + \delta^2\right|.$$

As the right-hand side tends to infinity as $x \to \infty$, *the function* f *is not uniformly continuous.*

2. *Consider* $f : [1, \infty) \to \mathbb{R}$ *given by* $f(x) = \log(x)$. *By the fundamental theorem of calculus we know that*

$$|f(x) - f(y)| = \left| \int_x^y \frac{1}{t} \, \mathrm{d}t \right| \leq |x - y|.$$

Therefore, f *is uniformly continuous as given any* $\epsilon > 0$ *we have that*

$$|f(x) - f(y)| < \epsilon$$

for any $x, y \in [1, \infty)$ *satisfying* $|x - y| < \epsilon$.

1.2.3 Completeness

Definition 1.2.19. *A sequence* $(x_n)_{n \in \mathbb{N}} \subseteq X$ *is convergent if there exists an* $x \in X$ *such that* $d(x_n, x) \overset{n \to \infty}{\longrightarrow} 0$.

Definition 1.2.20. *A sequence* $(x_n)_{n \in \mathbb{N}} \subseteq X$ *is a Cauchy sequence if for any* $\epsilon > 0$ *there exists an* $N \in \mathbb{N}$ *such that for* $n, m > N$ *we have* $d(x_n, x_m) < \epsilon$.

Remark 1.2.21. *By the triangle inequality, a convergent sequence is a Cauchy sequence.*

Definition 1.2.22. *A metric space* (X, d) *is complete if Cauchy sequences in* X *are convergent with respect to the metric* d.

Example 1.2.23.

1. *The set* $\mathbb{Q}$ *with* $d(x, y) = |x - y|$ *is not complete as there exists a sequence* $(r_n)_{n \in \mathbb{N}} \subseteq \mathbb{Q}$ *such that* $\left| r_n - \sqrt{2} \right| \to 0$, *but* $\sqrt{2} \notin \mathbb{Q}$.

2. *The set* $\mathbb{R}$ *with* $d(x, y) = |x - y|$ *is complete. Indeed, let* $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ *be a Cauchy sequence. Then there exists an* $N \in \mathbb{N}$ *such that*

$$|x_N - x_m| \leq 1$$

for $m \geq N$. *Therefore,*

$$|x_n| \leq C := \max\left(|x_1|, \ldots, |x_{N-1}|, |x_N| + 1\right)$$

for every $n \in \mathbb{N}$, *which means that* $(x_n)_{n \in \mathbb{N}}$ *is bounded. Hence, the sequences* $(y_n)_{n \in \mathbb{N}}$ *given by* $y_n = \inf_{m \geq n}(x_m)$ *and* $(z_n)_{n \in \mathbb{N}}$ *given by* $z_n = \sup_{m \geq n}(x_m)$ *are well-defined. In particular, these sequences are convergent as they are monotonic and bounded. Let* $y_n \to y$ *and* $z_n \to z$. *Then for* $\epsilon > 0$ *there exists an* $N_1 \in \mathbb{N}$ *such that*

$$|z_n - z| + |y_n - y| \leq \frac{\epsilon}{3}$$

for $n \geq N_1$. *Furthermore, there exists an* $N_2 \in \mathbb{N}$ *such that*

$$|x_n - x_m| \leq \frac{\epsilon}{3} \tag{1.2.1}$$

for $m \geq n \geq N_2$. Taking the supremum of (1.2.1) *we deduce that $|x_n - z_n| \leq \frac{\epsilon}{3}$ for $n \geq N_2$. Similarly, taking the infimum of* (1.2.1) *we deduce that $|x_n - y_n| \leq \frac{\epsilon}{3}$ for $n \geq N_2$. Therefore, for $n \geq \max(N_1, N_2)$ we deduce that*

$$\begin{aligned} |z-y| &\leq |z-z_n| + |z_n - x_n| + |x_n - y_n| + |y_n - y| \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon. \end{aligned}$$

Therefore, $y = z =: x$ which implies $x_n \to x$, and so $\mathbb{R}$ with $d(x,y) = |x-y|$ is complete.

Theorem 1.2.24. *If (X,d) is a metric space then there exists a metric space (Y,d') such that*

1. *Y is complete,*
2. *there is an injection $i : X \to Y$, and*
3. *$d(x,y) = d'(i(x), i(y))$.*

Theorem 1.2.25 (Banach Fixed Point Theorem)**.** *Let (X,d) be a complete metric space. Let $f : X \to X$ be a contraction, that is there exists a $\kappa \in (0,1)$ such that $d(f(x), f(y)) \leq \kappa d(x,y)$ for any $x, y \in X$. Then f has a unique fixed point, that is there exists a unique $x_0 \in X$ such that $f(x_0) = x_0$.*

Proof. Let $x_1 \in X$ and consider the sequence $(x_n)_{n \in \mathbb{N}} \subseteq X$ given by $x_n = f(x_{n-1})$ for $n \geq 2$. Then

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq \kappa d(x_{n-1}, x_n).$$

Proceeding by induction we deduce that $d(x_n, x_{n+1}) \leq \kappa^{n-1} d(x_1, x_2)$. Let $N \in \mathbb{N}$ and consider $l > k > N$. Then by the triangle inequality, it follows that

$$\begin{aligned} d(x_l, x_k) &\leq d(x_l, x_{l-1}) + d(x_{l-1}, x_{l-2}) + \cdots + d(x_{k+1}, x_k) \\ &\leq \left(\kappa^{l-2} + \kappa^{l-3} + \cdots + \kappa^{k-1}\right) d(x_1, x_2) \\ &\leq \left(\kappa^{l-2} + \kappa^{l-3} + \dots\right) d(x_1, x_2) \\ &= \frac{\kappa^{l-2}}{1-\kappa} d(x_1, x_2) \\ &\leq \frac{\kappa^{N}}{1-\kappa} d(x_1, x_2) \\ &\xrightarrow{N \to \infty} 0. \end{aligned}$$

Therefore, the sequence is Cauchy, and hence convergent to some $x_0 \in X$ as (X,d) is a complete metric space. The contractive property of f implies it is continuous. As $x_n \to x_0$ it follows by the continuity of f that $f(x_n) \to f(x_0)$ and so by the uniqueness of limits $x_0 = f(x_0)$. Now suppose that there exists another fixed point $y \in X$ of f. Then

$$d(f(x_0), f(y)) = d(x_0, y)$$

which contradicts the contracting property of f. Therefore, the fixed point x_0 is unique. $\square$

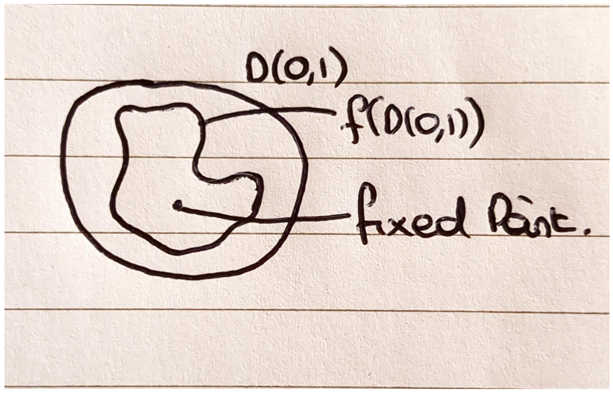


Figure 1: An illustration of the conditions required for Theorem 1.2.25

Example 1.2.26.

1. *Translations do not satisfy the conditions of Theorem 1.2.25 as $\kappa = 1$. For example, $f(x) = x + 1$ is such that $|f(x) - f(y)| = |x - y|$, indeed $f(x) = x$ has no solutions.*

2. *For $a \in (0, 1)$, consider the metric space (X, d) where $X = \mathbb{R} \setminus \left\{\frac{b}{1-a}\right\}$ and $d(x, y) = |x - y|$. Then $h : X \to X$ given by $h(x) = ax + b$ is well-defined and a contraction as*

$$|h(x) - h(y)| = a|x - y|.$$

However, h does not have a fixed point in X. Indeed, (X, d) is not a complete metric space, so one cannot apply Theorem 1.2.25.

1.2.4 Compactness

Theorem 1.2.27 (Bolzano-Weierstrass)**.** *A bounded sequence of real numbers has a convergent subsequence. That is, if $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ is such that $|x_n| \leq R$ for some $R > 0$, then there exists an extraction φ and $y \in \mathbb{R}$ such that $x_{\varphi(n)} \to y$.*

Remark 1.2.28. *An extraction $\varphi : \mathbb{N} \to \mathbb{N}$ is a strictly increasing function and can be used to index a subsequence.*

For a metric space (X, d) and $S \subseteq X$, the Bolzano-Weierstrass property says that for all sequences $(x_n)_{n\in\mathbb{N}} \subseteq S$ there exists a $y \in S$ and extraction φ such that $x_{\varphi(n)} \to y$ as $n \to \infty$.

Definition 1.2.29. *A collection of sets $(O_i)_{i\in I}$ is an open cover of $S \subseteq X$ if each O_i is open and $S \subseteq \bigcup_{i\in I} O_i$.*

Definition 1.2.30. *A sub-cover of an open cover $(O_i)_{i\in I}$ of S is a subset $J \subseteq I$ such that $S \subseteq \bigcup_{i\in J} O_i$.*

The finite open cover property says that for any open cover, one can extract a finite sub-cover.

Example 1.2.31. *Let $X = \mathbb{R}$ and $S = \mathbb{Z}$. Then $\mathbb{Z}$ does not satisfy the finite open cover property. Choose $O_i = \left(i - \frac{1}{10}, i + \frac{1}{10}\right)$ for $i \in \mathbb{N}$. Then $(O_i)_{i\in\mathbb{N}}$ is an open cover of $\mathbb{Z}$ with no finite sub-cover.*

Theorem 1.2.32. *The Bolzano-Weierstrass property and the finite open cover property are equivalent.*

Definition 1.2.33. *If either the Bolzano-Weierstrass property or the finite cover property holds, then S is called compact.*

Example 1.2.34.

1. *$[a, b] \subseteq \mathbb{R}$ is compact. By Theorem 1.2.27, any sequence in $[a, b]$ has a convergent subsequence. In particular, the limit of this subsequence is in $[a, b]$ as $[a, b]$ is closed.*
2. *$(a, b) \subseteq \mathbb{R}$ is not compact. The cover $\left(\left(a + \frac{b-a}{n}, b - \frac{b-a}{n}\right)\right)_{n\in\mathbb{N}}$ has no finite subcover.*
3. *Any finite subset $S \subseteq \mathbb{R}$ is compact. For any open cover $(O_i)_{i\in\mathbb{N}}$ one can extract a finite subcover $(O_{i_k})_{k\in\{1,\ldots,|S|\}}$ where O_{i_k} is such that the k^{th} element of S is in O_{i_k}.*
4. *$\mathbb{Q} \subseteq \mathbb{R}$ is not compact. A sequence in $\mathbb{Q}$ converging to $\sqrt{2}$ has no convergent subsequence.*

Lemma 1.2.35. *If $S \subseteq X$ is compact then it is closed.*

Proof. Note that $S \subseteq X$ is closed if and only if $S = \bar{S}$. By construction $S \subseteq \bar{S}$ and so it suffices to show that $\bar{S} \subseteq S$. Choose $x \in \bar{S}$, then there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq S$ such that $x_n \to x$. By the Bolzano-Weierstrass property, there exists an extraction φ and $y \in S$ such that $x_{\varphi(n)} \to y$. However, it must also be the case that $x_{\varphi(n)} \to x$, as any subsequence of a convergent sequence converges to the same limit. Therefore, $x = y \in S$, which implies that $\bar{S} \subseteq S$ which completes the proof. □

Lemma 1.2.36. *If $S \subseteq X$ is compact then it is bounded.*

Proof. Suppose that S were not bounded. Then one can construct a sequence $(x_n)_{n\in\mathbb{N}}$ such that $d(x_0, x_n) \geq n$ for $n \in \mathbb{N}$ and some fixed $x_0 \in S$. The sequence $(x_n)_{n\in\mathbb{N}}$ has no convergent subsequence as

$$d\left(x, x_n\right) \geq \left|d\left(x_0, x_n\right) - d(x, x_0)\right| \overset{n\to\infty}{\longrightarrow} \infty$$

for any $x \in S$. Therefore, S does not satisfy the Bolzano-Weierstrass property, which contradicts S being compact. □

Theorem 1.2.37 (Heine-Borel)**.** *In the metric space $(\mathbb{R}^n, d)$ where $d(x, y) = \sum_{i=1}^{n} |x_i - y_i|$, the compact sets are precisely the closed and bounded sets.*

Remark 1.2.38.

1. *As all norms are equivalent in finite dimensions, the conclusions of Theorem 1.2.37 hold in any finite-dimensional normed vector spaces.*
2. *From Lemma 1.2.35 and Lemma 1.2.36, we see that being closed and bounded is a necessary condition for a set to be compact. Theorem 1.2.37 then asserts that in finite-dimensional spaces, being closed and bounded is a sufficient condition for a set to be compact.*
3. *Compact sets are the same for equivalent metrics.*

Theorem 1.2.39. *If $S \subseteq X$ is compact, then the following hold.*

1. *Any continuous function $f : S \to \mathbb{R}$ achieves its supremum.*
2. *Any continuous function $f : S \to \mathbb{R}$ is uniformly continuous.*

Proof.

1. Let $M = \sup_{x \in S} f(x)$ where $f : S \to \mathbb{R}$ is a continuous function.
   (a) If $M = \infty$, then there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq S$ such that $f(x_n) \to \infty$. However, by compactness, we know there exists an extraction φ and $y \in S$ such that $x_{\varphi(n)} \to y$. Therefore, by continuity we have that $f(x_{\varphi(n)}) \to f(y) \in \mathbb{R}$ which contradicts $f(x_n) \to \infty$. Hence, we must have $M < \infty$.
   (b) If $M < \infty$, then choose $(x_n)_{n\in\mathbb{N}} \subseteq S$ such that $f(x_n) \to M$. Then by compactness there exists an extraction φ and $y \in S$ such that $x_{\varphi(n)} \to y$. By continuity, we have that $f(x_{\varphi(n)}) \to f(y)$ and so by the uniqueness of limits we conclude that $f(y) = M$.
2. Let $f : S \to \mathbb{R}$ be a continuous function. Suppose f is not uniformly continuous. Then, there exists an $\epsilon > 0$ such that for $\delta = \frac{1}{n}$, for any $n \in \mathbb{N}$, there exists $x_n, y_n \in S$ with $d(x_n, y_n) < \frac{1}{n}$ such that $|f(x_n) - f(y_n)| \geq \epsilon$.
   - By compactness, there exists an extraction φ and $\tilde{x} \in S$ such that $x_{\varphi(n)} \to \tilde{x}$. Similarly, there exists an extraction ψ of $\left(x_{\varphi(n)}\right)_{n\in\mathbb{N}}$ and $\tilde{y} \in S$ such that $y_{\psi(n)} \to \tilde{y}$.
   - Given any $\tilde{\epsilon} > 0$, it follows for N sufficiently large with $n, m \geq N$ that

$$\begin{aligned} d\left(\tilde{x}, \tilde{y}\right) &\leq d\left(\tilde{x}, x_{\psi(n)}\right) + d\left(\tilde{y}, x_{\psi(n)}\right) \\ &\leq d\left(\tilde{x}, x_{\psi(n)}\right) + d\left(x_{\psi(n)}, y_{\psi(m)}\right) + d\left(y_{\psi(m)}, \tilde{y}\right) \\ &\leq \frac{\tilde{\epsilon}}{3} + \frac{\tilde{\epsilon}}{3} + \frac{\tilde{\epsilon}}{3} \\ &= \tilde{\epsilon}. \end{aligned}$$

   Therefore, $d\left(\tilde{x}, \tilde{y}\right) = 0$ which implies that $\tilde{x} = \tilde{y}$. On the other hand, by the continuity of f we have that $f(x_{\psi(n)}) \to f(\tilde{x})$ and $f(y_{\psi(n)}) \to f(\tilde{y})$ which implies that $|f(\tilde{x}) - f(\tilde{y})| \geq \epsilon$, which gives rise to a contradiction. Therefore, f is uniformly continuous.

□

Exercise 1.2.40. *Provide an alternative proof of statement 1 of Theorem 1.2.39 by using open covers.*

Example 1.2.41. *The compactness condition of Theorem 1.2.39 is essential. Consider the space $\mathcal{C}((0,1), \mathbb{R})$ and the function $f(x) = \sin\left(\frac{1}{x}\right) \in \mathcal{C}((0,1), \mathbb{R})$. The function $f(x)$ is bounded and continuous on $(0,1)$ but it is not uniformly continuous. Similarly, the continuity condition of statement 1 is essential. Consider $f : [0,1] \to \mathbb{R}$ given by*

$$f(x) = \begin{cases} 2x & x \in \left[0, \frac{1}{2}\right) \\ 0 & x = \frac{1}{2} \\ 2 - 2x & x \in \left(\frac{1}{2}, 1\right]. \end{cases}$$

Then $\sup_{x\in[0,1]} f(x) = 1$ but $f(x) \neq 1$ for any $x \in [0,1]$.

1.3 Solution to Exercises

Exercise 1.2.40

Solution. Let $\epsilon > 0$. Then for each $x \in S$ there exists a $\delta_x > 0$ such that $|x - y| < \delta_x$ implies $|f(x) - f(y)| < \frac{\epsilon}{2}$. Note that $\left(B_{\frac{\delta_x}{2}}(x)\right)_{x \in S}$ is an open cover for S. Therefore, by the compactness of S, we can extract a finite subcover, say $\left(B_{\frac{\delta_{x_i}}{2}}(x_i)\right)_{i=1,\ldots,n}$. Let $\delta = \min_{i=1,\ldots,n}\left(\frac{\delta_{x_i}}{2}\right)$. Then for $x \in S$ there exists an x_i such that $|x - x_i| < \delta < \delta_{x_i}$. Therefore, for y such that $|y - x| < \delta$ we have

$$|y - x_i| \leq |y - x| + |x - x_i| < 2\delta < \delta_{x_i},$$

and so

$$|f(x) - f(y)| \leq |f(x) - f(x_i)| + |f(x_i) - f(y)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence, f is uniformly continuous. □

2 The Lebesgue Measure

2.1 Measure Spaces

Let X be a set.

Definition 2.1.1. *A σ-algebra is a collection of subsets $\mathcal{A} \subseteq \mathcal{P}(X)$, that satisfies the following.*

1. $X \in \mathcal{A}$.
2. *If $S \in \mathcal{A}$, then $A^c \in \mathcal{A}$.*
3. *If $(S_i)_{i\in\mathbb{N}} \subseteq \mathcal{A}$ then $\bigcup_{i\in\mathbb{N}} S_i \in \mathcal{A}$.*

Remark 2.1.2. *By combining statements 2 and 3 in Definition 2.1.1, it follows that a σ-algebra is closed under countable intersections.*

Definition 2.1.3. *A function $\mu : \mathcal{A} \to [0, \infty]$ is a measure if it satisfies the following.*

1. $\mu(\emptyset) = 0$.
2. *If $(S_i)_{i\in\mathbb{N}} \subseteq \mathcal{A}$ are such that $S_i \cap S_j = \emptyset$ for $i \neq j$ then*

$$\mu\left(\bigcup_{i\in\mathbb{N}} S_i\right) = \sum_{i\in\mathbb{N}} \mu(S_i).$$

Remark 2.1.4. *Property 2 of Definition 2.1.3 is referred to as countable additivity, and can be thought of as a continuity property.*

1. *The countable additivity property implies that if $(S_j)_{j\in\mathbb{N}} \subseteq \mathcal{A}$ is an increasing sequence of sets then*

$$\lim_{j\to\infty} \mu(S_j) \to \mu\left(\bigcup_{j\in\mathbb{N}} S_j\right).$$

   This can be proved by applying countable additivity to the sets $E_j = S_{j+1} \setminus S_j$.

2. *A similar result holds for a decreasing sequence of sets. Namely, if $(S_j)_{j\in\mathbb{N}} \subseteq \mathcal{A}$ is a decreasing sequence of sets and $\mu(S_0) < \infty$, then*

$$\lim_{j\to\infty} \mu(S_j) \to \mu\left(\bigcap_{j\in\mathbb{N}} S_j\right).$$

   This is shown by using statement 1 on the complements of the S_j.

2.2 The Lebesgue Measure on $\mathbb{R}^d$

Theorem 2.2.1. *There exists a σ-algebra $\mathcal{A} \subseteq \mathcal{P}\left(\mathbb{R}^d\right)$, and a measure μ such that the following hold.*

1. *Open sets of $\mathbb{R}^d$, under the canonical metric, are in $\mathcal{A}$.*
2. *The rectangle $R = \prod_{i=1}^d (a_i, b_i)$ has measure $\mu(R) = \prod_{i=1}^d (b_i - a_i)$.*
3. *If $A \in \mathcal{A}$ with $\mu(A) = 0$ and $B \subseteq A$ then $B \in \mathcal{A}$ with $\mu(B) = 0$.*

Remark 2.2.2.

- *The σ-algebra and measure of Theorem 2.2.1 are known as the Lebesgue σ-algebra and Lebesgue measure respectively.*
- *The countable intersection of open sets gives rise to many interesting sets, and so by countable additivity our σ-algebra captures a rich collection of sets.*
- *Statement 2 of Theorem 2.2.1 tells us that μ extends our intuition on the size of sets in $\mathbb{R}^d$.*
- *Statement 3 of Theorem 2.2.1 emphasises that the measure space is complete.*
- *Sets in the Lebesgue σ-algebra are called measurable sets.*
- *The Lebesgue measure is invariant under translations, that is for $x \in \mathbb{R}^d$ and A a measurable set we have*

$$\mu(A + x) = \mu(A).$$

- *For $\lambda \in \mathbb{R}$ and A a measurable set, the Lebesgue measure has the following scaling property,*

$$\mu(\lambda A) = \lambda^d \mu(A).$$

Example 2.2.3. *With the measure and σ-algebra of Theorem 2.2.1 we can understand why the requirement that $\mu(S_0) < \infty$ in statement 2 of Remark 2.1.4 is necessary. Indeed, suppose $S_j = (j, \infty)$ for $j \in \mathbb{N}$. Then the sequence of sets $(S_j)_{j \in \mathbb{N}}$ is decreasing, however*

$$\lim_{j \to \infty} \mu(S_j) = \infty,$$

whereas,

$$\mu\left(\bigcap_{j \in \mathbb{N}} S_j\right) = \mu(\emptyset) = 0.$$

Proposition 2.2.4. *A hyperplane in $\mathbb{R}^d$ has zero Lebesgue measure.*

Proof. A hyperplane in $\mathbb{R}^d$ is of the form

$$A_b = \left\{x \in \mathbb{R}^d : a_1 x_1 + \cdots + a_d x_d = b\right\},$$

where $a_1, \ldots, a_d, b \in \mathbb{R}$ are fixed. Due to the translational invariance of the Lebesgue measure, we can consider

$$A := A_0 = \left\{x \in \mathbb{R}^d : a_1 x_1 + \cdots + a_d x_d = 0\right\}.$$

We will assume without loss of generality that $a_d \neq 0$. We can isolate the graph of x_d by considering the continuous function

$$f(x_1, \ldots, x_{d-1}) = \frac{-(a_1 x_1 + \cdots + a_{d-1} x_{d-1})}{a_d}.$$

Consider the compact set $K_j = \prod_{i=1}^{d-1} [-j, j] \subseteq \mathbb{R}^{d-1}$. Then as f is continuous, it is uniformly continuous on K_j. Therefore, for a given $\epsilon > 0$ we can partition K_j such that in each partition the variation of f is at most $\frac{\epsilon}{2^{j+d-1} j^{d-1}}$. Then

$$\mu(f(K_j)) = \frac{\epsilon}{2^{j+d-1} j^{d-1}} \mu(K_j) = \frac{\epsilon}{2^{j+d-1} j^{d-1}} (2j)^{d-1} = \frac{\epsilon}{2^j}.$$

As $A \subseteq \bigcup_{j=1}^{\infty} f(K_j)$ it follows that

$$\mu(A) \leq \sum_{j=1}^{\infty} \frac{\epsilon}{2^j} = \epsilon.$$

Therefore, $\mu(A) = 0$ as $\epsilon > 0$ was arbitrary. $\square$

Definition 2.2.5. *A function $f : \mathbb{R}^d \to \mathbb{R}$ is measurable if $f^{-1}((-\infty, a))$ is a measurable set for all $a \in \mathbb{R}$.*

Proposition 2.2.6.

1. *The composition of measurable functions is measurable.*
2. *If $(f_n)_{n\in\mathbb{N}}$ is a sequence of measurable functions such that $f_n(x) \to f(x)$ for all x, then f is measurable. In other words, the function $\lim_{n\to\infty}(f_n)$ is measurable. Moreover, $\sup_{n\in\mathbb{N}}(f_n)$, $\inf_{n\in\mathbb{N}}(f_n)$, $\limsup_{n\to\infty}(f_n)$ and $\liminf_{n\to\infty}(f_n)$ are all measurable.*
3. *Sums and products of measurable functions are measurable.*
4. *Continuous functions are measurable.*

Definition 2.2.7. *A property is true almost everywhere or for almost any x if it is true on the complement of a zero-measure set.*

2.3 The Lebesgue Integral

2.3.1 The Integral of Simple Functions

Definition 2.3.1. *A simple function is of the form*

$$f = \sum_{i=1}^{N} c_i \mathbf{1}_{A_i}$$

where for each $i = 1, \ldots, N$ the $c_i \in \mathbb{R}$ and the A_i is a measurable set of $\mathbb{R}^d$ of finite measure.

The integral of a simple function is

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \sum_{i=1}^{N} c_i \mu(A_i).$$

Similarly, for a measurable set S the integral of a simple function on S is

$$\int_S f(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\mathbf{1}_S(x)\,\mathrm{d}x.$$

Henceforth, we will often use the abbreviated notation

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \int f\,\mathrm{d}x$$

2.3.2 The Integral of Non-Negative Functions

Let $f : \mathbb{R}^d \to [0, \infty]$ be a non-negative function on $\mathbb{R}^d$. The integral of f is taken to be

$$\int f\,\mathrm{d}x = \sup\left(\left\{\int s\,\mathrm{d}x : 0 \le s \le f,\ s \text{ a simple function}\right\}\right).$$

Proposition 2.3.2. *Let f be a non-negative function on $\mathbb{R}^d$.*

1. *If $\int f\,\mathrm{d}x < \infty$ then $f < \infty$ almost everywhere.*

2. *If* $\int f \, \mathrm{d}x = 0$ *then* $f = 0$ *almost everywhere.*

Example 2.3.3. *The converse of statement 1 of Proposition 2.3.2 does not hold. Consider* $f(x) \equiv 1$. *Then* $f < \infty$ *almost everywhere but* $\int f \, \mathrm{d}x = \infty$.

2.3.3 The Integral of Real-Valued Functions

A measurable function $f : \mathbb{R}^d \to (-\infty, \infty)$ admits the representation $f = f_+ - f_-$ where

- $f_+ = \max(0, f)$, and
- $f_- = \max(0, -f)$.

We say that f is integrable, written $f \in L^1\left(\mathbb{R}^d\right)$, if $\int f_+ < \infty$ and $\int f_- < \infty$. The integral of an integrable function is taken to be

$$\int f \, \mathrm{d}x = \int f_+ \, \mathrm{d}x - \int f_- \, \mathrm{d}x.$$

Proposition 2.3.4.

1. *For* $\alpha, \beta \in \mathbb{R}$ *and* $f, g \in L^1\left(\mathbb{R}^d\right)$ *we have that*

$$\int \alpha f + \beta f \, \mathrm{d}x = \alpha \int f \, \mathrm{d}x + \beta \int g \, dx.$$

2. *For* $f \in L^1\left(\mathbb{R}^d\right)$ *we have that*

$$\left| \int f \, \mathrm{d}x \right| \leq \int |f| \, \mathrm{d}x.$$

3. *A function* $f \in L^1\left(\mathbb{R}^d\right)$ *is zero almost everywhere if and only if* $\int_S f \, \mathrm{d}x = 0$ *for all measurable sets* S.

Proposition 2.3.5. *Let* $f, g : S \to \mathbb{R}$ *be integrable measurable functions that satisfy* $f \leq g$ *almost everywhere in* S. *Then,*

$$\int_S f \leq \int_S g.$$

Proof. Suppose that f and g are non-negative measurable functions. Then for any simple function s such that $0 \leq s \leq f$, there exists an equal almost everywhere simple function $\tilde{s}$ such that $0 \leq \tilde{s} \leq g$ and $\int s = \int \tilde{s}$. Therefore,

$$\left\{ \int s \, \mathrm{d}x : 0 \leq s \leq f, \ s \text{ a simple function} \right\} \subseteq \left\{ \int s \, \mathrm{d}x : 0 \leq s \leq g, \ s \text{ a simple function} \right\}$$

which implies that

$$\sup \left(\left\{ \int s \, \mathrm{d}x : 0 \leq s \leq f, \ s \text{ a simple function} \right\} \right) \leq \sup \left(\left\{ \int s \, dx : 0 \leq s \leq g, \ s \text{ a simple function} \right\} \right)$$

which then implies that $\int f \leq \int g$. For arbitrary integrable measurable functions f and g we can write $f = f_+ - f_-$ and $g = g_+ - g_-$ where f_+, f_-, g_+, g_- are non-negative. As $f \leq g$ almost everywhere it follows that $f_+ \leq g_+$ almost everywhere and $g_- \leq f_-$ almost everywhere. Hence,

$$\int f = \int f_+ - \int f_- \leq \int g_+ - \int g_- = \int g.$$

$\square$

In light of Proposition 2.3.4, a reasonable suggestion for a distance on L^1 is $d(f,g) = \int |f-g| \, \mathrm{d}x$. However, this is not a metric as if $f, g \in L^1$ are such that $d(f,g) = 0$ then we can only say that $f(x) = g(x)$ for almost all x.

- For continuous functions f and g such that $d(f,g) = 0$ it is possible to conclude that $f(x) = g(x)$ for all x.
- However, for $f \equiv 0$ and $g = \mathbf{1}_{\{0\}}$ we have $d(f,g) = 0$ but $f(0) \neq g(0)$.

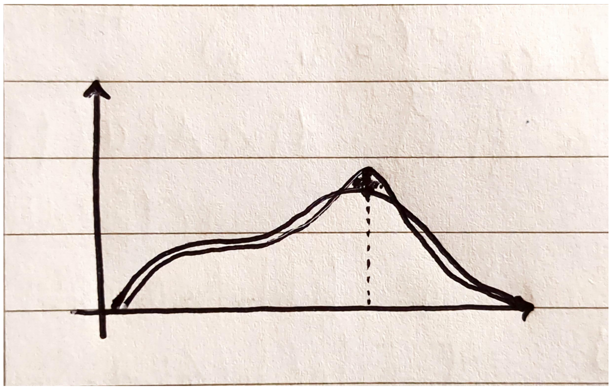


Figure 2: An illustration as to why continuous functions equal almost everywhere must be equal exactly.

To overcome this issue, we use equivalence classes. For $f \in L^1$ let

$$[f] = \left\{ g \in L^1 : f(x) = g(x) \text{ almost everywhere} \right\}.$$

Consequently, $d(f,g) = 0$ if and only if $[f] = [g]$. Abusing notation we will still speak of "functions" rather than "equivalence classes".

2.3.4 Connection to the Riemann Integral

Throughout let $I = [a,b]$ where $-\infty < a < b < \infty$.

Definition 2.3.6. *A set of points $\mathcal{P} = (x_i)_{i=0}^N$, for $N \in \mathbb{N}$, is a partition of I if*

$$a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b.$$

Definition 2.3.7. *A function $F : I \to \mathbb{R}$ is a step function if there exists a partition $\mathcal{P}$ such that*

$$F(x) = \sum_{i=0}^{N-1} a_i \mathbf{1}_{[x_i, x_{i+1})}$$

where each $a_i \in \mathbb{R}$.

Definition 2.3.8. *For $f : I \to \mathbb{R}$ a bounded function and $\mathcal{P} = (x_i)_{i=0}^N$ a partition of I let*

- *the upper sum of f with respect to $\mathcal{P}$ be*

$$U_{\mathcal{P},I}(f) = \sum_{i=0}^{N-1} \left(\sup_{t \in [x_i, x_{i+1})} f(t) \right) (x_{i+1} - x_i),$$

- *and the lower sum of f with respect to $\mathcal{P}$ be*

$$L_{\mathcal{P},I}(f) = \sum_{i=0}^{N-1} \left(\inf_{t \in [x_i, x_{i+1})} f(t) \right) (x_{i+1} - x_i).$$

Definition 2.3.9. *A bounded function $f : I \to \mathbb{R}$ is Riemann integrable if for every $\epsilon > 0$ there exists a partition $\mathcal{P}$ of I such that*

$$|U_{\mathcal{P},I}(f) - L_{\mathcal{P},I}(f)| < \epsilon.$$

Proposition 2.3.10. *If f is Riemann integrable then*

$$\inf_{\mathcal{P}} U_{\mathcal{P},I}(f) = \sup_{\mathcal{P}} L_{\mathcal{P},I}(f).$$

We denote the Riemann integral on I of a Riemann integrable function f as

$$\int_a^b f = \inf_{\mathcal{P}} U_{\mathcal{P},I}(f) = \sup_{\mathcal{P}} L_{\mathcal{P},I}(f).$$

Theorem 2.3.11. *Every Riemann integrable function on I is Lebesgue integrable and*

$$\int_a^b f(x)\,\mathrm{d}x = \int_I f(x)\,\mathrm{d}x.$$

Remark 2.3.12. *All the facts and techniques for Riemann integration extend to Lebesgue integrals of Riemann integrable functions.*

With this equivalence, we can characterise the set of Riemann integrable functions using measure theory.

Theorem 2.3.13. *Let f be bounded on I. Then f is Riemann integrable on I if and only if it is continuous almost everywhere.*

One can readily extend Riemann integration to unbounded domains. In this case, a function is Riemann integrable if the upper and lower sums are absolutely convergent and coincide. Similarly, an unbounded function on a finite or infinite domain is Riemann integrable if the upper and lower sums are absolutely convergent and coincide. We refer to both cases as improper Riemann integration.

Proposition 2.3.14. *For a function f, if the improper Riemann integral absolutely converges, then f is also Lebesgue integrable and the integrals coincide.*

2.4 Convergence of Functions and Convergence of Integrals

Example 2.4.1.

1. *Let $f_n = \mathbf{1}_{[n,n+1]}$ on $\mathbb{R}$. Then $\int f_n = 1$ and $f_n(x) \to f(x) = 0$ for all $x \in \mathbb{R}$. However, $\int f_n \not\to \int f$.*
2. *Let $f_n = n\mathbf{1}_{\left(0,\frac{1}{n}\right)}$ on $\mathbb{R}$. Then $\int f_n = 1$ and $f_n(x) \to f(x) = 0$ for all $x \in \mathbb{R}$. However, $\int f_n \not\to \int f$.*

Lemma 2.4.2. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions such that $\operatorname{supp}(f_n) \subseteq K$ for every $n \in \mathbb{N}$ with K a* *compact set independent of* $n \in \mathbb{N}$*. Moreover, suppose* $\sup_x |f_n(x) - f(x)| \overset{n\to\infty}{\longrightarrow} 0$*. Then,*

$$\int f_n \,\mathrm{d}x \overset{n\to\infty}{\longrightarrow} \int f \,\mathrm{d}x.$$

Proof. We note that as K is compact, $\mu(K) < \infty$. Therefore,

$$\begin{aligned}
\left|\int f_n \,\mathrm{d}x - \int f \,\mathrm{d}x\right| &= \left|\int f_n - f \,\mathrm{d}x\right| \\
&\leq \int |f_n - f| \,\mathrm{d}x \\
&\leq \int_K \sup_y |f_n(y) - f(y)| \,\mathrm{d}x \\
&= \mu(K) \sup_y |f_n(y) - f(y)| \\
&\overset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Therefore,

$$\int f_n \,\mathrm{d}x \overset{n\to\infty}{\longrightarrow} \int f \,\mathrm{d}x.$$

□

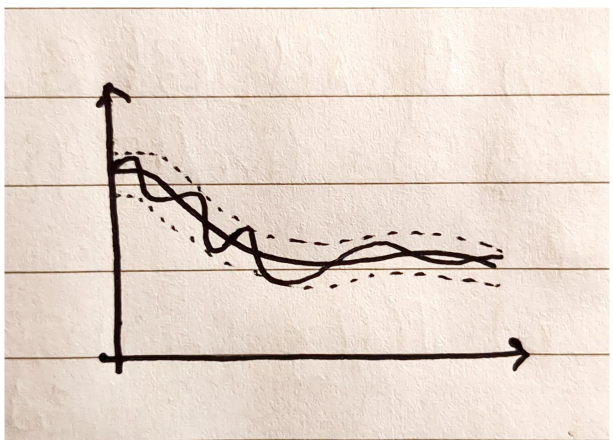


Figure 3: For the supremum between a sequence of functions and its limit to converge it must be the case that the functions lie within an ever-decreasing bounded region of the limit function.

Example 2.4.3. *Consider* $f_n = \frac{1}{n}\mathbf{1}_{[0,n]}$*. Even though* $\sup_x |f_n(x)| = \frac{1}{n} \to 0$ *as* $n \to \infty$*, there does not exist a compact set* K *such that* $\mathrm{supp}(f_n) \subseteq K$ *for every* $n \in \mathbb{N}$*. Thus, Lemma 2.4.2 cannot be applied, indeed* $\int f_n \,\mathrm{d}x = 1 \not\to 0$*.*

Theorem 2.4.4 (Monotone Convergence Theorem)**.** *Let* $(f_n)_{n\in\mathbb{N}}$ *be a sequence of non-negative measurable functions such that* $f_{n+1}(x) \geq f_n(x)$ *for almost all* x*.*

1. *Then* $f_n(x) \to f(x) = \sup_{n\in\mathbb{N}} f_n(x)$ *almost everywhere.*

2. *Furthermore,* $\int f_n \to \int f$*. In particular, if the right-hand side is finite, then we also have convergence*

in L^1. That is,

$$\int |f_n - f| \overset{n\to\infty}{\longrightarrow} 0.$$

Example 2.4.5. *The sequence of functions $(f_n)_{n\in\mathbb{N}}$ must be non-decreasing to apply Theorem 2.4.4. Indeed, let $f_n = \frac{1}{n}\mathbf{1}_{[0,n]}$, then $f_n(x) \to 0$ but $\int f_n \not\to \int f$.*

Remark 2.4.6. *Note that the monotonicity condition is only required to hold almost everywhere. The zero measure sets on which monotonicity may not hold can depend on n. What's more, the zero-measure set on which monotonicity does not can depend on the function f_n since the countable union of zero-measure sets is still a zero-measure set.*

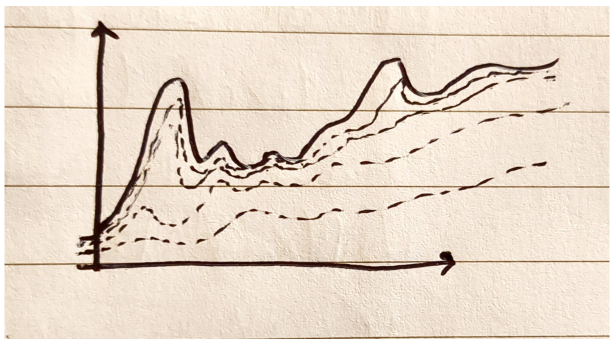


Figure 4: For a sequence of functions to converge monotonically from below to its limit, the graph of a function in the sequence must lie between the limiting function and the graph of the previous function in the sequence.

Theorem 2.4.7 (Dominated Convergence Theorem)**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions such that the following hold.*

1. *$f_n(x) \to f(x)$ for almost all x.*
2. *There exists a $g \in L^1$ such that $|f_n(x)| \le g(x)$ for almost any x.*

Then,

$$\int f_n \to \int f.$$

Example 2.4.8. *Recall Example 2.4.1 where we had pointwise convergence but not the convergence of the integrals.*

1. *To apply Theorem 2.4.7 we would need $g(x) = \sup_{n\in\mathbb{N}}(f_n(x)) = \mathbf{1}_{[0,\infty)}$ to be integrable, which it is not the case.*
2. *To apply Theorem 2.4.7 we would need $g(x) = \sup_{n\in\mathbb{N}}(f_n(x))$ to be integrable, however it is bounded below by n on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$ for every $n \in \mathbb{N}$. Consequently,*

$$\int g \ge \sum_{n=1}^{\infty}\left(\frac{1}{n} - \frac{1}{n+1}\right)n = \sum_{n=1}^{\infty}\left(\frac{1}{n(n+1)}\right)n = \sum_{n=1}^{\infty}\frac{1}{n+1} = \infty.$$

Therefore, no $g \in L^1$ exists such that $|f_n(x)| \le g(x)$ for almost any x.

Theorem 2.4.4 and Theorem 2.4.7 imply convergence in L^1 starting from pointwise convergence.

Example 2.4.9. *Pick a sequence $(x_n)_{n\in\mathbb{N}}$ such that the following hold.*

1. x_n *is increasing.*
2. $x_{n+1} - x_n \to 0$ *as* $n \to \infty$.
3. $x_n \to \infty$.

For example, $x_n = \sqrt{n}$. Let $y_n \in [0,1)$ be such that $x_n - y_n \in \mathbb{Z}$, for instance $y_n = x_n - \lfloor x_n \rfloor$, then let $f_n = \mathbf{1}_{(y_n, y_{n+1})}$. Note that a correction needs to be made when $y_{n+1} < y_n$. From this we have that

$$\int f_n = y_{n+1} - y_n = x_{n+1} - x_n \overset{n\to\infty}{\longrightarrow} 0,$$

and so convergence in the L^1 sense. However, $f_n(x) \not\to 0$ for all x as the y_n continually traverse the interval $[0,1)$.

Proposition 2.4.10. *If $f_n \to f$ in L^1, then there exists an extraction φ such that $f_{\varphi(n)}(x) \to f(x)$ for almost all x.*

Theorem 2.4.11 (Fatou's Lemma)**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions, then*

$$\liminf_{n\to\infty}\left(\int f_n\right) \geq \int \liminf_{n\to\infty}(f_n).$$

Example 2.4.12. *The sequence of functions $(f_n)_{n\in\mathbb{N}}$ must be non-negative to apply Theorem 2.4.11. Indeed, consider $f_n(x) = \mathbf{1}_{\left[0,1-\frac{1}{n}\right]}(x) - (n-1)\mathbf{1}_{\left(1-\frac{1}{n},1\right]}$. Then,*

$$\begin{aligned}
\liminf_{n\to\infty}\left(\int f_n\right) &= \liminf_{n\to\infty}(0) \\
&= 0 \\
&\not\geq 1 \\
&= \int \mathbf{1}_{[0,1)} \\
&= \int \liminf_{n\to\infty}(f_n).
\end{aligned}$$

3 Banach Spaces

3.1 Norms

Throughout let E be a vector space over $\mathbb{R}$ or $\mathbb{C}$. For simplicity, we will assume it to be $\mathbb{R}$.

Definition 3.1.1. *A norm $\|\cdot\| : E \to [0, \infty)$ satisfies the following.*

1. $\|x\| = 0$ *if and only if* $x = 0$.
2. $\|\lambda x\| = |\lambda|\|x\|$ *for all* $x \in E$ *and* $\lambda \in \mathbb{R}$.
3. $\|x + y\| \leq \|x\| + \|y\|$ *for all* $x, y \in E$.

Example 3.1.2. *The following are examples of norms on vector spaces.*

1. *On $\mathbb{R}$, the map $|\cdot|$ is a norm.*
2. *On $\mathbb{R}^d$ the following are norms.*
   (a) $\|x\|_1 = \sum_{i=1}^d |x_i|$.
   (b) $\|x\|_\infty = \max_{i=1,\ldots,d} |x_i|$.

Definition 3.1.3. *A vector space endowed with a norm is called a normed vector space.*

Remark 3.1.4. *For a norm $\|\cdot\|$, the application $d(x, y) = \|x - y\|$ is a metric, referred to as the induced metric by the norm.*

Definition 3.1.5. *A Banach space is a normed vector space that is complete with respect to the induced metric.*

Definition 3.1.6. *Norms $\|\cdot\|_1$ and $\|\cdot\|_2$, are equivalent if there exists a constant $C > 0$ such that*

$$\frac{1}{C}\|\cdot\|_1 \leq \|\cdot\|_2 \leq C\|\cdot\|_1.$$

Remark 3.1.7. *From a norm, we get a metric, from which we define a topology, and thus establish a notion of convergence. Equivalent norms induce the same topology and notion of convergence.*

Theorem 3.1.8. *In finite dimensions, norms are equivalent. In other words, if $\dim(E) < \infty$ then any norms on E are equivalent in the sense of Definition 3.1.6.*

Proof. Let $(e_i)_{1\leq i\leq d}$ be a basis of E. Consider the norm

$$\left\|\sum_{i=1}^d x_i e_i\right\|_2 = \left(\sum_{i=1}^d |x_i|^2\right)^{\frac{1}{2}}.$$

Then consider another norm $\|\cdot\|$ on E. On the one hand,

$$
\begin{aligned}
\|x\| &= \left\| \sum_{i=1}^{d} x_i e_i \right\| \\
&\overset{\text{T.I}}{\leq} \sum_{i=1}^{d} \|x_i e_i\| \\
&\overset{\text{Homo.}}{=} \sum_{i=1}^{d} |x_i| \|e_i\| \\
&\leq d \max_{1\leq i\leq d} (|x_i|) \max_{1\leq i\leq d} (\|e_i\|) \\
&\leq \left(d \max_{1\leq i\leq d} (\|e_i\|) \right) \|x\|_2 \\
&\leq M\|x\|_2. \qquad (3.1.1)
\end{aligned}
$$

On the other hand, consider the set

$$S = \{x \in E : \|x\|_2 = 1\}.$$

Then S is clearly bounded, and it is closed as $\|\cdot\|_2$ is a continuous function. Therefore, S is compact by Theorem 1.2.37. Note that the map $x \mapsto \|x\|$ is continuous on $(E, \|\cdot\|_2)$ as from (3.1.1) it follows that the map is bounded with respect to $\|\cdot\|_2$. Therefore, this map reaches its infimum on S, say m. Observe that $m \neq 0$ as otherwise there would exist an $x \in S$ such that $\|x\| = 0$ which implies $x = 0$, however, $0 \notin S$. Hence, $\|x\| \geq m > 0$ for $\|x\|_2 = 1$. Applying this to $y = \frac{x}{\|x\|_2}$ we deduce that

$$\|x\| \geq m\|x\|_2$$

for all $x \in E$. Combining this with (3.1.1) we conclude that

$$m\|x\|_2 \leq \|x\| \leq M\|x\|_2$$

for all $x \in E$. Thus the norms $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent. □

Example 3.1.9. *In infinite dimensions, metrics no longer ought to be equivalent. Let $X = \mathcal{C}([0,1], \mathbb{R})$. Then the following are metrics.*

- $d_1(f,g) = \sup_{x\in[0,1]} (|f(x) - g(x)|)$.
- $d_2(f,g) = \int_0^1 |f(x) - g(x)| \, \mathrm{d}x$.

These are not equivalent, as for $f_n(x) = x^n$ and $g \equiv 0$ we have

- $d_1(f_n, 0) = 1$, *but*
- $d_2(f_n, 0) = \frac{1}{n}$.

With d_1 the space X is complete but with d_2 the space X is not complete. Figure 5 is an example of a sequence of functions in X that converge in d_2 to something not in X.

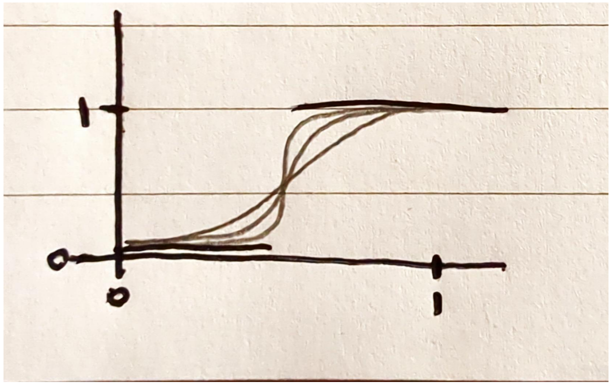


Figure 5: An example of how metrics in infinite dimensions need not be equivalent.

3.2 Spaces of Continuous Functions

We will consider functions on $\mathbb{R}^d$ or on open sets $\Omega \subseteq \mathbb{R}^d$.

Definition 3.2.1.

- *The set of bounded functions $\Omega \to \mathbb{R}$ is denoted $\mathcal{B}(\Omega, \mathbb{R})$.*
- *The set of continuous and bounded functions $\Omega \to \mathbb{R}$ is denoted $\mathcal{C}^0(\Omega, \mathbb{R})$.*

Remark 3.2.2.

- *As we will only work with real functions, we will simply denote these spaces as $\mathcal{B}(\Omega)$ and $\mathcal{C}^0(\Omega)$ respectively. Moreover, when the context is clear these function spaces may be denoted by $\mathcal{B}$ and $\mathcal{C}^0$ respectively. Sometimes $\mathcal{C}^0$ is also written as $\mathcal{C}$.*
- *The function spaces $\mathcal{B}$ and $\mathcal{C}$ are vector spaces, usually equipped with the uniform norm.*

Definition 3.2.3. *The uniform norm is the map*

$$\|f\|_\infty = \sup_{x \in \Omega} (|f(x)|)$$

on $\mathcal{B}(\Omega)$ and $\mathcal{C}(\Omega)$.

Definition 3.2.4. *If $f_n \overset{n\to\infty}{\longrightarrow} f$ with respect to $\|\cdot\|_\infty$, then we say the sequence $(f_n)_{n\in\mathbb{N}}$ converges to f uniformly.*

Theorem 3.2.5. *The uniform limit of continuous functions is continuous.*

Proof. Given $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ we have

$$\|f_n - f\|_\infty < \frac{\epsilon}{3}.$$

For fixed x, as f_N is continuous, there exists a $\delta > 0$ such that if $|x - y| < \delta$ then

$$|f_N(x) - f_N(y)| < \frac{\epsilon}{3}.$$

Therefore, for $|x-y|<\delta$ we have that

$$\begin{aligned}|f(x)-f(y)| &\leq |f(x)-f_N(x)|+|f_N(x)-f_N(y)|+|f_N(y)-f(y)| \\ &< \frac{\epsilon}{3}+\frac{\epsilon}{3}+\frac{\epsilon}{3} \\ &= \epsilon,\end{aligned}$$

where the first and the third differences are bounded by the uniform convergence, and the second difference is bounded by the continuity of f_N. This shows that f is continuous at x. ☐

Theorem 3.2.6. *With the uniform norm, the spaces $\mathcal{B}(\Omega)$ and $\mathcal{C}(\Omega)$ are Banach spaces.*

Proof. We will only carry out the proof for $\mathcal{C}(\Omega)$. Let $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}(\Omega)$ be a Cauchy sequence.
Step 1. Find a candidate for the limit.
For any $x\in\Omega$ consider the sequence $(f_n(x))_{n\in\mathbb{N}} \subseteq \mathbb{R}$. As

$$|f_n(x)-f_m(x)| \leq \|f_n-f_m\|_\infty \overset{n\to\infty}{\longrightarrow} 0$$

we deduce that the sequence $(f_n(x))_{n\in\mathbb{N}} \subseteq \mathbb{R}$ is a Cauchy sequence and hence convergent as $\mathbb{R}$ is complete. Let $f(x) := \lim_{n\to\infty} f(x)$. Note that $f\in\mathcal{B}(\Omega)$.
Step 2. Show that $(f_n)_{n\in\mathbb{N}}$ converges to f uniformly.
Choose $\epsilon>0$. Then there exists an $N\in\mathbb{N}$ such that for $n,m>N$ we have that $\|f_n-f_m\|<\epsilon$. Therefore, for all $x\in\Omega$ we have

$$|f_n(x)-f_m(x)|<\epsilon.$$

Sending $m\to\infty$ we deduce that $|f_n(x)-f(x)|<\epsilon$, which implies that

$$\|f_n-f\|_\infty<\epsilon.$$

Step 3. Show that $f\in\mathcal{C}(\Omega)$.
Using step 2 we can apply Theorem 3.2.5 to conclude. ☐

3.3 Spaces of Differentiable Functions

For $\alpha=(\alpha_1,\dots,\alpha_d)\in\mathbb{N}_0^d$ let

- $|\alpha|=\alpha_1+\cdots+\alpha_d$, and
- $\partial_x^\alpha=\partial_{x_1}^{\alpha_1}\dots\partial_{x_d}^{\alpha_d}$.

Definition 3.3.1. *The function space $\mathcal{C}^k(\Omega)$ contains functions on Ω which are k times differentiable with continuous derivatives $\partial_x^\alpha f$ for all $|\alpha|\leq k$.*

The space $\mathcal{C}^k(\Omega)$ is a vector space which we endow with the norm

$$\|f\|_{\mathcal{C}^k}=\max_{|\alpha|\leq k}\|\partial_x^\alpha f\|_\infty.$$

With this norm, $\mathcal{C}^k(\Omega)$ is a normed vector space.

Theorem 3.3.2. *The vector space $\mathcal{C}^k(\Omega)$ with $\|\cdot\|_{\mathcal{C}^k}$ is a complete normed vector space, that is a Banach space.*

Proof. Consider a Cauchy sequence $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}^k(\Omega)$.
Step 1. Find a candidate for the limit.
For $|\alpha| \le k$, the sequence $(\partial_x^\alpha f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence for $(\mathcal{C}, \|\cdot\|_\infty)$. Therefore, by Theorem 3.2.6 there exists a limit $f \in \mathcal{C}$ for $(f_n)_{n\in\mathbb{N}}$ and there exists a limit $g_\alpha \in \mathcal{C}$ for $(\partial_x^\alpha f_n)_{n\in\mathbb{N}}$.
Step 2. Claim that $f \in \mathcal{C}^k$ and $g_\alpha = \partial_x^\alpha f$.

- For the case $k = 1$ and $d = 1$. We know that $f_n(x) \to f(x)$ and $\partial_x f_n(x) \to g(x)$ in $\|\cdot\|_\infty$. By the fundamental theorem of calculus, we have that

$$f_n(x) - f_n(y) = \int_y^x \partial_x f_n(t)\,\mathrm{d}t.$$

  Recall that the integral of uniformly convergent function converges to the integral of the limit. Hence, as $n \to \infty$ we get that

$$f(x) - f(y) = \int_y^x g(t)\,\mathrm{d}t.$$

  Applying the fundamental theorem of calculus once again, it follows that f is differentiable with derivative g.

- For the case $k \ge 2$ and $d = 1$ we proceed by induction and use a similar approach to the previous case for the inductive step.

- For the case $k = 1$ and $d \ge 2$. The case follows analogously to the first case, where we instead apply the fundamental theorem of calculus component-wise. That is,

$$f_n(x) - f_n(x + te_j) = \int_0^t \partial_j f_n(x + se_j)\,\mathrm{d}s$$

  where e_j is the canonical j^{th} unit vector.

- For the case $k \ge 2$ and $d \ge 2$ we proceed by induction.

Step 3. f_n converges to f in $\mathcal{C}^k$.
Given $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\|f_n - f_m\|_{\mathcal{C}^k} \le \epsilon$ for $n, m \ge N$. This means that

$$\max_{|\alpha|\le k} \|\partial_x^\alpha f_n - \partial_x^\alpha f_m\|_\infty \le \epsilon.$$

Letting $m \to \infty$ we deduce that

$$\|\partial_x^\alpha f_n - g_\alpha\|_\infty \le \epsilon.$$

Previously we showed that $g_\alpha = \partial_n^\alpha f$. Therefore,

$$\|f_n - f\|_{\mathcal{C}^k} = \max_{|\alpha|\le k} \|\partial_x^\alpha f_n - \partial_x^\alpha f\|_\infty \le \epsilon.$$

□

Example 3.3.3. *Consider functions in $\mathcal{C}^1((-1,1))$. The map*

$$\|f\| = \|\partial_x f\|_\infty$$

is not a norm, as it is not definite. For example, $\|1\| = 0$. The map

$$\|f\| = \|\partial_x f\|_\infty + |f(0)|$$

is a norm. The fundamental theorem of calculus tells us $f(x) = f(0) + \int_0^x f'(t)\,\mathrm{d}t$. Hence, $\|f\| = 0$ if and only if $f = 0$. With this norm the space $\mathcal{C}^1((-1,1))$ is a Banach space.

3.4 Function Spaces on Compact Sets

In the previous sections, we considered spaces of real-valued functions defined on open sets $\Omega \subseteq \mathbb{R}^d$. Here we will suppose that $\Omega \subseteq \mathbb{R}^d$ is open and bounded, and then consider spaces of real-valued functions defined on $\bar{\Omega}$.

Theorem 3.4.1. *The space $\mathcal{B}\left(\bar{\Omega}\right)$, with norm $\|\cdot\|_\infty$ is a Banach space.*

Proof. Let $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{B}\left(\bar{\Omega}\right)$ be a Cauchy sequence. Observe that there exists an $N \in \mathbb{N}$ such that for every $m \geq N$ we have

$$\|f_N - f_m\|_\infty < 1.$$

As f_N is a bounded function there is an $M > 0$ such that $|f_N(x)| \leq M$ for all $x \in \bar{\Omega}$. Therefore, for sufficiently large m we have

$$|f_m(x)| \leq M + 1 \tag{3.4.1}$$

for all $x \in \bar{\Omega}$. As

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty \overset{n,m\to\infty}{\longrightarrow} 0$$

the sequence $(f_n(x))_{n\in\mathbb{N}} \subseteq \mathbb{R}$ is a Cauchy sequence, and hence convergent as $\mathbb{R}$ is complete. Let $f(x)$ be this limit. By (3.4.1), we deduce that $f(x) \leq M + 1$, as inequalities are preserved under limits. This holds for all $x \in \bar{\Omega}$ thus $f \in \mathcal{B}\left(\bar{\Omega}\right)$, hence, $\mathcal{B}\left(\bar{\Omega}\right)$ with the uniform norm is complete. $\square$

Theorem 3.4.2. *The space $\mathcal{C}^k\left(\bar{\Omega}\right)$, for $k \in \mathbb{N}$, with norm $\|\cdot\|_{\mathcal{C}^k}$ is a Banach space.*

Proof. The proof proceeds in the same way as the proof of Theorem 3.3.2, however, need to additionally check that the candidate limits are continuous up to the boundary of Ω. We check continuity up to the boundary for f, with the understanding that the other cases follow similarly. More specifically, consider $x \in \partial\Omega$ and a sequence $(y_n)_{n\in\mathbb{N}} \subseteq \Omega$ converging to x. Observe that for any $n, m \in \mathbb{N}$ we have

$$|f(y_n) - f(x)| \leq |f_m(y_n) - f(y_n)| + |f_m(y_n) - f_m(x)| + |f_m(x) - f(x)|.$$

Therefore, given $\epsilon > 0$ let $m \in \mathbb{N}$ be such that

$$\sup_{x\in\bar{\Omega}} |f_m(x) - f(y)| < \frac{\epsilon}{3},$$

which we can do as we have already established that $f_n \to f$ uniformly. Then let $n \in \mathbb{N}$ be such that

$$|f_m(y_n) - f_m(x)| < \frac{\epsilon}{3},$$

which we can do as $f_m \in \mathcal{C}^0\left(\bar{\Omega}\right)$. Consequently, we deduce that

$$|f(y_n) - f(x)| < \epsilon,$$

and so $f(y_n) \to f(x)$, which means that f is continuous up to the boundary of $\bar{\Omega}$. $\square$

Remark 3.4.3. *A function f is in $\mathcal{C}^k\left(\bar{\Omega}\right)$ if for any points $x \in \partial\Omega$ and $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq k$, the sequence $(\partial^\alpha f(y_n))_{n\in\mathbb{N}}$ admits a limit when $(y_n)_{n\in\mathbb{N}} \subseteq \Omega$ is such that $y_n \to x$. That is, there exists a $\beta \in \mathbb{R}$ such that for every sequence $(y_n)_{n\in\mathbb{N}} \subseteq \Omega$ with $y_n \to x$ we have $\partial^\alpha f(y_n) \to \beta$.*

Corollary 3.4.4. *The space $\mathcal{C}^0\left(\bar{\Omega}\right)$ is a closed subset of $\mathcal{B}\left(\bar{\Omega}\right)$.*

Proof. Continuous functions on compact domains are bounded so that $\mathcal{C}^0\left(\bar{\Omega}\right) \subseteq \mathcal{B}\left(\bar{\Omega}\right)$. As $\|\cdot\|_\infty = \|\cdot\|_{\mathcal{C}^0}$ we know by Theorem 3.4.2 that $\left(\mathcal{C}^0\left(\bar{\Omega}\right), \|\cdot\|_\infty\right)$ is a Banach space. It is clear then that $\mathcal{C}^0\left(\bar{\Omega}\right)$ is a closed subset of $\mathcal{B}\left(\bar{\Omega}\right)$. $\square$

4 L^p Spaces

4.1 The L^p Norm

Functions throughout this section are defined on $\mathbb{R}^d$ or $\Omega \subseteq \mathbb{R}^d$ open.

Definition 4.1.1. *If f is a measurable function, its L^p norm is*

$$\|f\|_{L^p} = \left(\int |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}}$$

for $1 \leq p < \infty$ and

$$\|f\|_{L^\infty} = \inf \{ M > 0 : |f(x)| < M \text{ almost everywhere} \} .$$

Remark 4.1.2. *Integrals are of non-negative functions, and so are well-defined despite taking potentially infinite value.*

Definition 4.1.3. *The set L^p, more specifically $L^p\left(\mathbb{R}^d, \mathbb{R}\right)$, is the set of measurable functions $f : \mathbb{R}^d \to \mathbb{R}$, such that $\|f\|_{L^p} < \infty$.*

Remark 4.1.4.

- *For $\Omega \subseteq \mathbb{R}^d$ open, we can similarly define the $L^p(\Omega)$ norm as*

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} .$$

- *Note that $L^p(\Omega)$ is a space of equivalence classes rather than functions. That is, f and g are equivalent if and only if $f = g$ is almost everywhere.*

Proposition 4.1.5 (Young's Inequality)**.** *If $\frac{1}{p} + \frac{1}{q} = 1$, for $1 \leq p, q \leq \infty$, then for all $x, y > 0$ we have*

$$xy \leq \frac{1}{p} x^p + \frac{1}{q} y^q .$$

Proof. Using the fact that $\log(\cdot)$ is a concave function we deduce that

$$\log \left(\frac{1}{p} x^p + \frac{1}{q} y^q \right) \geq \frac{1}{p} \log \left(x^p \right) + \frac{1}{q} \log \left(y^q \right) .$$

Exponentiating both sides we get

$$\frac{1}{p} x^p + \frac{1}{q} x^q \geq xy .$$

□

Proposition 4.1.6 (Hölder's Inequality)**.** *For $\Omega \subseteq \mathbb{R}^d$ open, let $p, q, r \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. Then*

$$\|fg\|_{L^r(\Omega)} \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)} .$$

Proof.

- Consider $p = r$ so that $q = \infty$. As

$$\|g\|_{L^\infty(\Omega)} = \inf\{M > 0 : |g(x)| < M \text{ a.e in } \Omega\}$$

we have that $|g| \leq \|g\|_{L^\infty(\Omega)}$ almost everywhere in Ω. Hence

$$\left(\int_\Omega |fg|^r\right)^{\frac{1}{r}} \leq \|g\|_{L^\infty(\Omega)} \left(\int_\Omega |f|^r\right)^{\frac{1}{r}} = \|f\|_{L^r(\Omega)} \|g\|_{L^\infty(\Omega)}.$$

- For $r = 1$ and $1 < p < \infty$, it is clear from Young's Inequality that

$$\begin{aligned}\int_\Omega |fg| &\leq \int_\Omega \frac{1}{p}|f|^p + \frac{1}{q}|g|^q \\ &\leq \frac{1}{p}\|f\|^p_{L^p(\Omega)} + \frac{1}{q}\|g\|^q_{L^q(\Omega)}.\end{aligned}$$

Therefore, if $\|f\|_{L^p(\Omega)} = 1$ and $\|g\|_{L^q(\Omega)} = 1$, it follows that

$$\int_\Omega |fg| \leq= \frac{1}{p} + \frac{1}{q} = 1.$$

Hence, for arbitrary $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$ we have that

$$\int_\Omega \left| \frac{f}{\|f\|_{L^p(\Omega)}} \frac{g}{\|g\|_{L^q(\Omega)}} \right| \leq 1$$

which implies that

$$\int_\Omega |fg| \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}$$

which is equivalent to $\|fg\|_{L^1(\Omega)} \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}$.

- For $r \neq 1$, note that $\frac{1}{(\frac{p}{r})} + \frac{1}{(\frac{q}{r})} = 1$. Let $\tilde{p} = \frac{p}{r}$ and $\tilde{q} = \frac{q}{r}$. Then using the result for $r = 1$ we can deduce that

$$\begin{aligned}\||fg|^r\|_{L^1(\Omega)} &\leq \||f|^r\|_{L^{\tilde{p}}(\Omega)} \, \||g|^r\|_{L^{\tilde{q}}(\Omega} \\ &= \left(\int_\Omega |f|^{r\tilde{p}}\right)^{\frac{1}{\tilde{p}}} \left(\int_\Omega |g|^{r\tilde{q}}\right)^{\frac{1}{\tilde{q}}}.\end{aligned}$$

Therefore,

$$\left(\int_\Omega |fg|^r\right)^{\frac{1}{r}} \leq \left(\int_\Omega |f|^p\right)^{\frac{1}{p}} \left(\int_\Omega |g|^q\right)^{\frac{1}{q}}$$

and thus

$$\|fg\|_{L^r(\Omega)} \leq \|f\|_{L^p(\Omega)} \|g\|_{L^r(\Omega)}.$$

$\square$

Example 4.1.7. *If $p = q = 2$ and $r = 1$. Then*

$$\int |fg| \leq \left(\int f^2\right)^{\frac{1}{2}} \left(\int g^2\right)^{\frac{1}{2}}$$

and we recover the Cauchy-Schwartz inequality.

Proposition 4.1.8 (Minkowski's Inequality)**.** *For $\Omega \subseteq \mathbb{R}^d(\Omega)$, if $f, g \in L^p(\Omega)$, then $f + g \in L^p(\Omega)$ and*

$$\|f + g\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}.$$

Proof. For $1 \leq p < \infty$ we have that

$$\begin{aligned} \|f+g\|^p_{L^p(\Omega)} &= \int_\Omega |f+g|^p \\ &\overset{\text{T.I}}{\leq} \int_\Omega |f||f+g|^{p-1} + \int_\Omega |g||f+g|^{p-1} \\ &\overset{\text{Prop. 4.1.6}}{\leq} \|f\|_{L^p(\Omega)}\|f+g\|^{p-1}_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}\|f+g\|^{p-1}_{L^p(\Omega)}. \end{aligned}$$

Dividing both sides by $\|f+g\|^{p-1}_{L^p(\Omega)}$ we conclude that

$$\|f + g\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}.$$

When $p = \infty$ we note that if

$$m_f \in \{M > 0 : |f(x)| < M \text{ a.e. in } \Omega\}$$

and

$$m_g \in \{M > 0 : |g(x)| < M \text{ a.e. in } \Omega\}$$

then

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| < m_f + m_g.$$

Taking infimums we conclude that

$$\|f + g\|_{L^\infty(\Omega)} \leq \|f\|_{L^\infty(\Omega)} + \|g\|_{L^\infty(\Omega)}.$$

□

Theorem 4.1.9. *For $1 \leq p \leq \infty$ the map $\|\cdot\|_{L^p(\Omega)}$ is a norm on $L^p(\Omega)$.*

Proof. Note $\|f\|_{L^p(\Omega)} = 0$ if and only if f is zero almost everywhere, and thus equivalent to zero. Furthermore, for $\lambda \in \mathbb{R}$ we have $\|\lambda f\|_{L^p(\Omega)} = |\lambda|\|f\|_{L^p(\Omega)}$. The triangle inequality is Proposition 4.1.8. Therefore, $\|\cdot\|_{L^p(\Omega)}$ is a norm on $L^p(\Omega)$. □

Proposition 4.1.10 (Generalised Minkowski Inequality)**.**

$$\left\| \int f(x,y)\,\mathrm{d}y \right\|_{L^p_x} \leq \int \|f(x,y)\|_{L^p_x}\,\mathrm{d}y.$$

Remark 4.1.11. *In Proposition 4.1.10 the y can be thought of as the summation variable and x as the variable with respect to which we are computing the norm.*

Example 4.1.12. *Consider the function $f : \mathbb{R}^d \to \mathbb{R}$ given by*

$$f(x) = \frac{\mathbf{1}_{B_1(0)}}{|x|^\alpha}$$

where $\alpha \in \mathbb{R}^d$. Recall that

$$\int_0^1 \frac{1}{|x|^{\alpha p}}\,\mathrm{d}x \begin{cases} = \infty & \alpha p \geq 1 \\ < \infty & \alpha p < 1. \end{cases}$$

This implies that $f \in L^p(\mathbb{R})$ if and only if $\alpha < \frac{d}{p}$. More generally, in $\mathbb{R}^d$ as f is a radial function we know that $\mathrm{d}x = Cr^{d-1}\,\mathrm{d}r$ where C is the volume of the unit sphere in $\mathbb{R}^d$. Therefore,

$$\left(\int_{B_1(0)} \frac{1}{|x|^{\alpha p}}\,\mathrm{d}x\right)^{\frac{1}{p}} = C^{\frac{1}{p}} \left(\int_0^1 r^{d-1-\alpha p}\,\mathrm{d}r\right)^{\frac{1}{p}}.$$

Consequently, $f \in L^p\left(\mathbb{R}^d\right)$ if and only if $\alpha < \frac{d}{p}$

The space L^p can contain surprisingly exotic functions as its regularity is only formulated as an integral, which disregards behaviour at individual points.

Exercise 4.1.13.

1. *Find a function in $L^p(\mathbb{R})$ which is essentially unbounded on any $[n, n+1]$ for $n \in \mathbb{Z}$.*
2. *Find a function in $L^p((0,1))$ which is unbounded on any (a,b) for $a, b \in (0,1)$.*

4.2 Convergence

We have established that $(L^p, \|\cdot\|_{L^p})$ is a normed vector spaces. Consequently, we can start asking questions about convergence in this space, and how spaces with different values of p are related.

Theorem 4.2.1. *The space L^p with norm $\|\cdot\|_{L^p}$ is a Banach space.*

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence with respect to $|\cdot|$. Then we can extract a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ such that

$$|f_{n_k} - f_{n_{k+1}}| < \frac{1}{2^k}.$$

As the sequence $(f_n)_{n\in\mathbb{N}}$ is convergent, the limit of $(f_{n_k})_{k\in\mathbb{N}}$ coincides with the limit of $(f_n)_{n\in\mathbb{N}}$. Hence, it suffices to consider a convergent Cauchy sequence $(f_n)_{n\in\mathbb{N}} \subseteq L^p$ such that

$$\|f_{n+1} - f_n\|_{L^p} \leq \frac{1}{2^n}.$$

With this consider the following.

- $f = f_0 + \sum_{n=0}^{\infty}(f_{n+1} - f_n)$.
  - This is only formal now as we have no way to make sense of the convergence.
- $g = |f_0| + \sum_{n=0}^{\infty}|f_{n+1} - f_n|$.
  - The convergence here has a pointwise meaning as we are dealing with non-negative functions.
- $S_k f = f_0 + \sum_{n=0}^{k}(f_{n+1} - f_n)$.
- $S_k g = |f_0| + \sum_{n=0}^{k}|f_{n+1} - f_n|$.

Step 1: Show the candidate f is well-defined and in L^p.
By Minkowski's inequality we have that

$$\|S_k g\|_{L^p} \leq \|f_0\|_{L^p} + \sum_{n=0}^{k}\|f_{n+1} - f_n\|_{L^p} \leq C + \sum_{n=0}^{k}\frac{1}{2^n} \leq \tilde{C} < \infty.$$

As $S_k g \nearrow g$ pointwise, we can conclude by the monotone convergence theorem that

$$\int |g|^p = \lim_{k\to\infty}\int |S_k g|^p \leq \tilde{C}.$$

This implies that $g \in L^p$, and $g < \infty$ almost everywhere. Consequently $\sum_{n=0}^{\infty} |f_{n+1} - f_n|$ is absolutely convergent which implies that f is absolutely convergent. Therefore, as $|f| \leq |g|$ we conclude that $f \in L^p$.

Step 2. Show f_n converges to f in L^p.

Note that

$$|f - S_k f| \leq |f| + |S_k f| \leq 2g$$

so that $|f - S_k f|^p \leq 2^p g^p$. Therefore, as $|f - S_k f|^p \to 0$ pointwise almost everywhere, by step 1 we can conclude by the dominated convergence theorem that

$$\|f - S_k f\|_{L^p}^p = \int |f - S_k f|^p \to 0.$$

□

Proposition 4.2.2. *If $\Omega \subseteq \mathbb{R}^d$ is bounded, then $L^p(\Omega) \subseteq L^q(\Omega)$ whenever $p \geq q$.*

Proof. Let $f \in L^p(\Omega)$. Note that $\frac{1}{q} = \frac{1}{p} + \frac{1}{\frac{pq}{p-q}}$. Let $r := \frac{pq}{p-q}$, then $\|\mathbf{1}\|_{L^r(\Omega)} < \infty$ as Ω is bounded. Therefore, by Hölder's inequality

$$\|f\|_{L^q(\Omega)} = \|f\mathbf{1}\|_{L^q(\Omega)} \leq \|f\|_{L^p(\Omega)} \|\mathbf{1}\|_{L^r(\Omega)} < \infty.$$

Therefore, $f \in L^q(\Omega)$. □

Example 4.2.3. *The condition that Ω is bounded in Proposition 4.2.2 is necessary for the inclusion to hold. Consider $\Omega = (1, \infty)$ and $f(x) = \frac{1}{x}$. Then*

$$\|f\|_{L^2((1,\infty))} = \left(\int_1^{\infty} \frac{1}{|x|^2} \, \mathrm{d}x \right)^{\frac{1}{2}} < \infty,$$

however,

$$\|f\|_{L^1((1,\infty))} = \int_1^{\infty} \frac{1}{|x|} \, \mathrm{d}x = \infty.$$

Therefore, $L^2((1,\infty)) \not\subseteq L^1((1,\infty))$.

4.3 Convolution

Throughout, we will only be dealing with functions defined on $\mathbb{R}^d$. Let $\mathcal{C}_c^0$ denote the set of compactly supported continuous functions, with analogous definitions for $\mathcal{C}_c^k$ and $\mathcal{C}_c^{\infty}$.

Definition 4.3.1. *For $f \in L^1$ and $\phi \in \mathcal{C}_c^0$, their convolution is*

$$(f \star \phi)(x) = \int_{\mathbb{R}^d} f(y) \phi(x - y) \, \mathrm{d}y.$$

Remark 4.3.2.

- *The integral of Definition 4.3.1 makes sense as the integrand is in L^1. Note $L^p \subseteq L^1$ locally. That is, if $f \in L^p$ and K is a compact set, then*

  $$\int f \mathbf{1}_K \, \mathrm{d}x \overset{\text{Hölder's}}{\leq} \|f\|_{L^p} \|\mathbf{1}_K\|_{L^q}$$

  for $\frac{1}{p} + \frac{1}{q} = 1$. Therefore, as $\|\mathbf{1}_K\|_{L^q} < \infty$ we conclude that on K we have $f \in L^1$. Consequently, convolutions still make sense for $f \in L^p$ when ϕ has compact support.

- *If both $f, \phi \in \mathcal{C}_c^0$, then $f \star \phi = \phi \star f$.*
- *The convolution operation $(f, \phi) \mapsto f \star \phi$ is bilinear.*

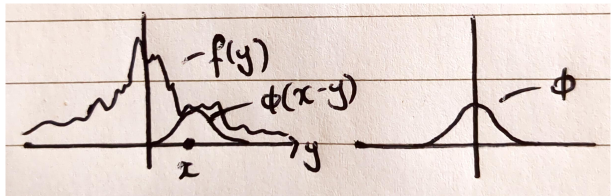


Figure 6: An illustration of how the convolution can be interpreted as a smoothing operation for a rough function f, by taking a weighted average at x over the compact support of ϕ.

Definition 4.3.3. *For $f \in L^1$, the support of f denoted $\operatorname{supp}(f)$ is the smallest closed set such that $f = 0$ almost everywhere in $\mathbb{R} \setminus \operatorname{supp}(f)$.*

Definition 4.3.4. *For sets A and B let*

$$A + B = \{a + b : a \in A, b \in B\}.$$

Lemma 4.3.5. *For $f \in L^1$ and $\phi \in \mathcal{C}_c^0$ we have*

$$\operatorname{supp}(f \star \phi) \subseteq \operatorname{supp}(f) + \operatorname{supp}(\phi).$$

Intuition. Ideally, one would say that if $\int f(y)\phi(x-y)\,\mathrm{d}y = (f \star \phi)(x) \neq 0$ then there exists a y such that $f(y) \neq 0$ and $\phi(x-y) \neq 0$. Therefore, $x = y + (x-y) \in \operatorname{supp}(f) + \operatorname{supp}(\phi)$. However, f here is an equivalence class, and it doesn't make sense to talk about evaluating f at points. One instead has to work with small open sets. □

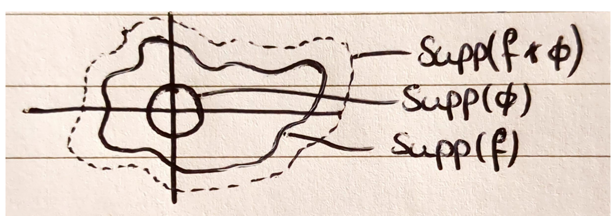


Figure 7: Thinking about a convolution as a weighted sum over a compact support, then graphically this is what we would expect the support of $f \star \phi$ to be.

Proposition 4.3.6. *If $f \in L^p$ and $\phi \in \mathcal{C}_c^0$ then $f \star \phi \in L^p$ and*

$$\|f \star \phi\|_{L^p} \leq \|f\|_{L^p} \|\phi\|_{L^1}.$$

Proof. For $p = 1$ we can write

$$\begin{aligned}\int |(f \star \phi)(x)| \,\mathrm{d}x &= \int \left| \int f(y)\phi(x-y)\,\mathrm{d}y \right| \mathrm{d}x \\ &\overset{\text{T.I}}{\leq} \iint |f(y)||\phi(x-y)| \,\mathrm{d}y\,\mathrm{d}x \\ &\overset{\text{Fubini.}}{=} \int |f(y)| \int |\phi(x-y)| \,\mathrm{d}x\,\mathrm{d}y \\ &= \|f\|_{L^1} \|\phi\|_{L^1}.\end{aligned}$$

For the case when $p > 1$ by Proposition 4.1.10 we deduce that

$$\begin{aligned}\left\| \int f(x-y)\phi(y)\,\mathrm{d}y \right\|_{L^p_x}^p &\leq \int \|f(x-y)\phi(y)\|_{L^p_x} \,\mathrm{d}y \\ &= \int |\phi(y)| \|f(x-y)\|_{L^p_x} \,\mathrm{d}y \\ &= \|\phi\|_{L^1} \|f\|_{L^p}.\end{aligned}$$

Where in the last inequality we have pulled out $\|f\|_{L^p}$ as by translational invariance $\|f(x-y)\|_{L^p_x} = \|f(x)\|_{L^p_x}$, and so is independent of y. $\square$

Exercise 4.3.7.

1. *Show that if $f \in L^p$ and $\phi \in L^1$ then $f \star \phi \in L^p$.*
2. *Show that if $f \in L^1_{\mathrm{loc}}$ and $\phi \in \mathcal{C}^0_c$ then $f \star \phi \in \mathcal{C}^0$.*
   - *The space L^p_{loc} is the space of functions for which on every compact set K we have $\|f\mathbf{1}_K\|_{L^p} < \infty$.*

Proposition 4.3.8. *If $f \in L^1_{loc}$ and $\phi \in \mathcal{C}^k_c$, then $f \star \phi \in \mathcal{C}^k_c$. What's more*

$$\partial^\alpha (f \star \phi) = f \star \partial^\alpha \phi$$

if $|\alpha| \leq k$.

Proof. We proceed for $k = 1$. Let $G(x) := (\phi \star f)(x)$. Fix $i \in \{1, \ldots, d\}$ and $x \in \mathbb{R}^d$. Consider

$$\frac{G(x + h_n \mathbf{e}_i) - G(x)}{h_n} = \int \underbrace{\frac{\phi(x + h_n \mathbf{e}_i - y) - \phi(x-y)}{h_n}}_{F^x_n(y)} f(y)\,\mathrm{d}y$$

where $h_n \to 0$. We know that $F^x_n(y)$ is supported on $B_R(x)$ for R sufficiently large. As $\phi \in \mathcal{C}^k_c\left(\mathbb{R}^d\right)$ we know that

$$f(y)F^x_n(y) \overset{n\to\infty}{\longrightarrow} f(y)\partial_i\phi(x-y)$$

pointwise almost everywhere. Moreover,

$$|f(y)F^x_n(y)| \leq |f(y)| \, \|\phi\|_{\mathcal{C}^1(B_R)}.$$

As $f \in L^1_{\mathrm{loc}}$ we know the right-hand side of the above is in $L^1(B_R)$. Hence, by the dominated convergence theorem

$$\lim_{n\to\infty} \left(\frac{G(x + h_n \mathbf{e}_i) - G(x)}{h_n} \right) = \int \partial_i \phi(x-y) f(y)\,\mathrm{d}y.$$

It follows that $\partial_i (f \star \phi)(x) = ((\partial_i \phi) \star f)(x)$ for all $i \in \{1, \ldots, d\}$ and all $x \in \mathbb{R}^d$. As $\partial_i \phi \in \mathcal{C}^0_c$ it follows that $\partial_i (f \star \phi) \in \mathcal{C}^0_c\left(\mathbb{R}^d\right)$. Proceed by induction to complete the proof. $\square$

4.4 Mollifer

For a function $\varphi \in \mathcal{C}_c^\infty$ with $\int \varphi = 1$ we define the sequence of mollifiers $(\varphi_n)_{n\in\mathbb{N}}$ where

$$\varphi_n(x) = n^d \varphi(nx).$$

Note that, $\operatorname{supp}(\varphi_n) = \frac{1}{n}\operatorname{supp}(\varphi)$, whilst $\int \varphi_n = 1$. Intuitively, $f \star \varphi_n$ should converge in some sense to f. As $f \star \varphi_n$ can be thought of as a weighted average of f over $\operatorname{supp}(\varphi_n)$, thus we are performing an increasingly concentrated average.

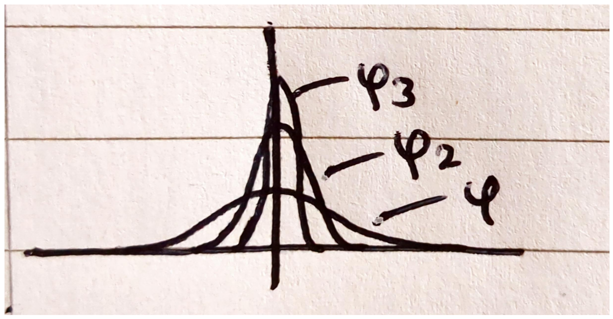


Figure 8: A graphical representation of a mollifer, φ, and subsequent φ_n.

Theorem 4.4.1.

1. *If $f \in \mathcal{C}_c^0$, then $f \star \varphi_n \overset{n\to\infty}{\longrightarrow} f$ under the uniform topology on $\mathcal{C}$.*
2. *If $f \in L^p$, for $1 \le p < \infty$, then $f \star \varphi_n \overset{n\to\infty}{\longrightarrow} f$ in L^p.*

Proof.

1. Given an $\epsilon > 0$, there exists a $\delta > 0$ such that for $|x-y| < \delta$ we have that $|f(x) - f(y)| < \epsilon$. Note that

$$\begin{aligned}(f \star \varphi_n)(x) - f(x) &= \int \varphi_n(x-y) f(y)\,\mathrm{d}y - f(x)\\ &= \int \varphi_n(x-y)(f(y) - f(x))\,\mathrm{d}y.\end{aligned}$$

The last equality follows from the fact that $f(x)$ is independent of y and $\int \varphi_n(x-y)\,\mathrm{d}y = 1$. We can choose $N \in \mathbb{N}$ such that for $x, y \in \operatorname{supp}(\varphi_N)$ we have $|x-y| < \delta$. Then for $n \ge N$ we have

$$|(f \star \varphi_n)(x) - f(x)| \le \epsilon \int |\varphi_n(x-y)|\,\mathrm{d}y \le \epsilon C.$$

Hence, we have uniform convergence.

2. Let $f \in L^p$. Using the fact that $\mathcal{C}_c^0$ is dense in L^p, given an $\epsilon > 0$ there exists a $g \in \mathcal{C}_c^0$ and $h \in L^p$ such that $f = g + h$, and $\|h\|_{L^p} \le \epsilon$. Hence,

$$f \star \varphi_n - f = g \star \varphi_n - g + h \star \varphi_n - h$$

so that

$$\|f \star \varphi_n - f\|_{L^p} \le \|g \star \varphi_n - g\|_{L^p} + \|h \star \varphi_n\|_{L^p} + \|h\|_{L^p}.$$

The second and third terms on the right-hand side are less than or equal to ϵ by construction. The function in the first term has compact support, that is independent of n, and so $g \star \varphi_n - g \to 0$ uniformly. Therefore, $\|g \star \varphi_n - g\|_{L^p} \overset{n\to\infty}{\longrightarrow} 0$.

$\square$

Corollary 4.4.2. *For $1 \leq p < \infty$, the space $\mathcal{C}_c^\infty$ is dense in L^p.*

Theorem 4.4.1 breaks down for $p = \infty$. If it were true then we could choose $f \in L^\infty \setminus \mathcal{C}^0$ and find a sequence $(f_n) \subseteq \mathcal{C}^\infty$ such that $f_n \to f$ in L^∞. However, for continuous functions $\|f\|_{L^\infty} = \|f\|_{\mathcal{C}^0}$. Therefore, the sequence is convergent in $\mathcal{C}^0$ with the uniform topology, which implies that $f \in \mathcal{C}^0$, which is a contradiction.

4.5 Solution to Exercises

Exercise 4.1.13

Solution.

1. For a given p let $f_n(x) = \mathbf{1}_{[n,n+1]} \frac{1}{|x-n|^{\frac{1}{2p}}}$. By Example 4.1.12 the L^p-norm of f_n is finite and is independent of $n \in \mathbb{Z}$ since the measure is translationally invariant. Therefore,

$$f(x) = \sum_{n \in \mathbb{Z}} \frac{1}{n^2} f_n(x)$$

is absolutely convergent. As $L^p(\mathbb{R})$ is complete we deduce that $f \in L^p(\mathbb{R})$. Notice that f is unbounded at $n \in \mathbb{Z}$ and so satisfies the requirements of the exercise.

2. As $\mathbb{Q} \cap [0,1]$ is countable we can enumerate it as $(q_n)_{n \in \mathbb{N}}$. As before we consider

$$f(x) = \mathbf{1}_{[0,1]} \sum_{n=1}^{\infty} \frac{1}{n^2} \frac{1}{|x - q_n|^{\frac{1}{2p}}}.$$

Note that $f \in L^p(\mathbb{R})$ by similar arguments and satisfies the requirement of the exercise as $\mathbb{Q} \cap [0,1]$ is dense in $[0,1]$ and f is unbounded at each q_n.

□

Exercise 4.3.7

Solution.

1. Observe that

$$\|\phi \star f\|_{L_x^p}^p = \left\| \int \phi(y) f(x-y)\, \mathrm{d}y \right\|_{L_x^p}^p.$$

Applying the generalised Minkowski inequality we deduce that

$$\begin{aligned} \|\phi \star f\|_{L_x^p}^p &\leq \int \|\phi(y) f(x-y)\|_{L_x^p}\, \mathrm{d}y \\ &= \int |\phi(y)| \|f\|_{L_x^p}\, \mathrm{d}y \\ &= \|\phi\|_{L^1} \|f\|_{L^p}. \end{aligned}$$

Therefore, $\phi \star f \in L^p$ as $\|\phi\|_{L^1}$ and $\|f\|_{L^p}$ are finite by assumption.

2. Fix $x \in \mathbb{R}^d$. For $z \in \mathbb{R}^d$ observe that

$$|(\phi \star f)(x) - (\phi \star f)(z)| \leq \int |f(y)| |\phi(x-y) - \phi(z-y)|\, \mathrm{d}y.$$

Assume that $\mathrm{supp}(\phi) \subseteq B_R$, so that $\mathrm{supp}(\phi(\cdot - x)) \subseteq B_R(x)$ and $\mathrm{supp}(\phi(\cdot - z)) \subseteq B_R(z)$. Suppose that $|z - x| = \delta$ so that

$$\mathrm{supp}(\phi(\cdot - x) - \phi(\cdot - z)) \subseteq B_{R+2\delta}(x).$$

Then
$$|(\phi \star f)(x) - (\phi \star f)(z)| \leq \int_{B_{R+2\delta}(x)} |f(y)||\phi(x-y) - \phi(z-y)| \, \mathrm{d}y.$$
As $\phi(\cdot - x) - \phi(\cdot - z)$ is continuous and compactly supported, it is also uniformly continuous on $B_{R+2\delta}(x)$. Therefore, for $\epsilon > 0$ there exists a $\delta > \delta_0 > 0$ such that $|\tilde{y} - \bar{y}| < \delta_0$ implies that
$$|\phi(\tilde{y}) - \phi(\bar{y})| < \epsilon.$$
Hence,
$$|(\phi \star f)(x) - (\phi \star f)(z)| \leq \epsilon \int_{B_{R+2\delta}(x)} |f(y)| \, \mathrm{d}y.$$
Thus we have continuity, but we do not have uniform continuity as the right-hand side is dependent on x.
□

5 ℓ^p Spaces

In this section, we will briefly explore ℓ^p spaces which can be thought of as a discrete analogue of L^p, but with some key differences.

5.1 ℓ^p Norm

Definition 5.1.1. *For $1 \leq p < \infty$ define the real vector space*

$$\ell^p = \left\{ x = (x_k)_{k\in\mathbb{N}} \subseteq \mathbb{R} : \sum_{k\in\mathbb{N}} |x_k|^p < \infty \right\}.$$

When $p = \infty$ define the real vector space

$$\ell^\infty = \left\{ (x_k)_{k\in\mathbb{N}} \subseteq \mathbb{R} : \sup_{k\in\mathbb{N}} (|x_k|) < \infty \right\}.$$

For $1 \leq p < \infty$ the space ℓ^p consists of absolutely summable sequences. Whereas ℓ^∞ deals with bounded sequences, which is a significant distinction between the spaces.

Definition 5.1.2. *For $1 \leq p < \infty$ let $\|\cdot\|_{\ell^p} : \ell^p \to \mathbb{R}$ be given by*

$$\|x\|_{\ell^p} = \left(\sum_{k\in\mathbb{N}} |x_k|^p \right)^{\frac{1}{p}}.$$

For $p = \infty$ let $\|\cdot\|_{\ell^\infty} : \ell^\infty \to \mathbb{R}$ be given by

$$\|x\|_{\ell^\infty} = \sup_{k\in\mathbb{N}} (|x_k|).$$

Remark 5.1.3. *If $f = \sum_{k=0}^{\infty} c_k \mathbf{1}_{[k,k+1]}$, for $(c_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$, then*

$$\|f\|_{L^p} = \|(c_k)_{k\in\mathbb{N}}\|_{\ell^p}.$$

Proposition 5.1.4 (Hölder's Inequality)**.** *Let $1 \leq p, q \leq \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then,*

$$\|fg\|_{\ell^1} \leq \|f\|_{\ell^p} \|g\|_{\ell^q}$$

for $x \in \ell^p$ and $y \in \ell^q$.

Proof.

- For $1 \leq p < \infty$, from Young's Inequality we get that

$$\begin{aligned} \sum_{k=1}^{\infty} |x_k y_k| &\leq \sum_{k=1}^{\infty} \left(\frac{1}{p}|x_k|^p + \frac{1}{q}|y_k|^q \right) \\ &\leq \frac{1}{p} \sum_{k=1}^{\infty} |x_k|^p + \frac{1}{q} \sum_{k=1}^{\infty} |y_k|^q. \end{aligned}$$

  If $\frac{1}{p} + \frac{1}{q} = 1$, $\|x\|_{\ell^p} = 1$ and $\|y\|_{\ell^q} = 1$, then

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \frac{1}{p}\|x\|_{\ell^p} + \frac{1}{q}\|y\|_{\ell^q} = 1.$$

Therefore, for arbitrary $x \in \ell^p$ and $y \in \ell^q$ we have that

$$\sum_{k=1}^{\infty} \left| \frac{x_k}{\|x\|_{\ell^p}} \frac{y_k}{\|y\|_{\ell^q}} \right| \leq 1$$

which implies that

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \|x\|_{\ell^p} \|y\|_{\ell^q}.$$

Thus, $\|xy\|_{\ell^1} \leq \|x\|_{\ell^p} \|y\|_{\ell^q}$.

- When $p = \infty$, then $q = 1$ and

$$\begin{aligned} \|xy\|_{\ell^1} &= \sum_{k=1}^{\infty} |x_k y_k| \\ &\leq \sum_{k=1}^{\infty} \left(\sup_{k \in \mathbb{N}} |x_k| \right) |y_k| \\ &= \|x\|_{\ell^\infty} \sum_{k=1}^{\infty} |y_k| \\ &= \|x\|_{\ell^\infty} \|y\|_{\ell^1}. \end{aligned}$$

□

Proposition 5.1.5 (Minkowski's Inequality). *If $x, y \in \ell^p$ then $x + y \in \ell^p$ and*

$$\|x + y\|_{\ell^p} \leq \|x\|_{\ell^p} + \|y\|_{\ell^p}.$$

Proof.

- If $1 \leq p < \infty$, then

$$\begin{aligned} \|x + y\|_{\ell^p}^p &= \sum_{k=1}^{\infty} |x_k + y_k|^p \\ &\overset{\text{T.I}}{\leq} \sum_{k=1}^{\infty} |x_k| |x_k + y_k|^{p-1} + \sum_{k=1}^{\infty} |y_k| |x_k + y_k|^{p-1} \\ &\overset{\text{Prop. } 5.1.4}{\leq} \|x\|_{\ell^p} \|x + y\|_{\ell^p}^{p-1} + \|y\|_{\ell^p} \|x + y\|_{\ell^p}^{p-1}. \end{aligned}$$

Dividing both sides by $\|x + y\|_{\ell^p}^{p-1}$ we conclude that

$$\|x + y\|_{\ell^p} \leq \|x\|_{\ell^p} + \|y\|_{\ell^p}.$$

- If $p = \infty$, then

$$\begin{aligned} \|x + y\|_{\ell^\infty} &= \sup_{k \in \mathbb{N}} (|x_k + y_k|) \\ &\leq \sup_{k \in \mathbb{N}} |x_k| + \sup_{k \in \mathbb{N}} |y_k| \\ &= \|x\|_{\ell^\infty} + \|y\|_{\ell^\infty}. \end{aligned}$$

□

Theorem 5.1.6. *For $1 \leq p \leq \infty$ the map $\|\cdot\|_{\ell^p}$ is a norm on ℓ^p.*

Proof. Clearly, $\|x\|_{\ell^p} = 0$ if and only if $x_k = 0$ for all $k \in \mathbb{N}$. Furthermore, for $\lambda \in \mathbb{R}$ we have $\|\lambda x\|_{\ell^p} = |\lambda|\|x\|_{\ell^p}$. The triangle inequality is Proposition 5.1.5. Therefore, $\|\cdot\|_{\ell^p}$ is a norm on ℓ^p. $\square$

Consequently, we can consider ℓ^p as a normed vector space with norm $\|\cdot\|_{\ell^p}$.

5.2 Convergence

Theorem 5.2.1. *For $1 \leq p \leq \infty$, the space ℓ^p is a Banach space.*

Proof.

- Consider the case when $1 \leq p < \infty$. Let $\left(x^{(n)}\right)_{n\in\mathbb{N}} \subseteq \ell^p$ be a Cauchy sequence. Then given an $\epsilon > 0$, there exists an $\tilde{N} \in \mathbb{N}$ such that for all $n, m \geq \tilde{N}$ we have

$$\left\|x^{(n)} - x^{(m)}\right\|_{\ell^p} < \epsilon$$

so that for any $k \in \mathbb{N}$ we have

$$\left|x_k^{(n)} - x_k^{(m)}\right|^p < \epsilon^p.$$

Hence the sequence $\left(x_k^{(n)}\right)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ is a Cauchy sequence and therefore converges to a limit we will denote $x_k^{(\infty)}$. The sequence $\left(x^{(n)}\right)_{n\in\mathbb{N}}$ is Cauchy and thus bounded so that for some $M > 0$ we have

$$\left\|x^{(n)}\right\|_{\ell^p} \leq M$$

for all $n \in \mathbb{N}$. Therefore, for any $N \in \mathbb{N}$ we have

$$\left(\sum_{k=1}^{N} \left|x_k^{(\infty)}\right|^p\right)^{\frac{1}{p}} = \lim_{n\to\infty} \left(\sum_{k=1}^{N} \left|x_k^{(n)}\right|^p\right)^{\frac{1}{p}} \leq \lim_{n\to\infty} \left\|x^{(n)}\right\|_{\ell^p} \leq M.$$

Sending $N \to \infty$ gives

$$\left\|x^{(\infty)}\right\|_{\ell^p} \leq M,$$

meaning $x^{(\infty)} \in \ell^p$. Recall, that

$$\left\|x^{(n)} - x^{(m)}\right\|_{\ell^p} < \epsilon$$

for any $n, m \geq \tilde{N}$. Therefore, for any $N \in \mathbb{N}$ we have

$$\left(\sum_{k=1}^{N} \left|x_k^{(n)} - x_k^{(m)}\right|^p\right)^{\frac{1}{p}} < \epsilon.$$

Sending $m \to \infty$ gives

$$\left(\sum_{k=1}^{N} \left|x_k^{(n)} - x_k^{(\infty)}\right|^p\right)^{\frac{1}{p}} < \epsilon.$$

Sending $N \to \infty$ we conclude that, $x^{(n)} \to x^{(\infty)}$ in ℓ^p. Hence $(\ell^p, \|\cdot\|_{\ell^p})$ is a Banach space when $1 \leq p < \infty$.

- Now consider the case when $p = \infty$. Let $\left(x^{(n)}\right)_{n\in\mathbb{N}} \subseteq \ell^\infty$ be a Cauchy sequence. As $\left|x_k^{(n)} - x_k^{(m)}\right| \leq \left\|x^{(n)} - x^{(m)}\right\|_{\ell^\infty}$, it follows that $\left(x_k^{(n)}\right)_{k\in\mathbb{N}} \subseteq \mathbb{R}$ is Cauchy. Therefore, as before, we can construct the sequence $x^{(\infty)}$, where $x_k^{(\infty)} = \lim_{n\to\infty}\left(x_k^{(n)}\right)$. For any $N \in \mathbb{N}$ we have

$$\sup_{k=1,\dots,N}\left|x_k^{(\infty)}\right| = \lim_{n\to\infty}\sup_{k=1,\dots,N}\left|x_k^{(n)}\right| \leq \lim_{n\to\infty}\left\|x^{(n)}\right\|_{\ell^\infty}.$$

As the sequence $\left(x^{(n)}\right)_{n\in\mathbb{N}}$ is Cauchy it is bounded, hence $x^{(\infty)}$ is bounded and thus is in ℓ^∞. Furthermore, as $\left(x_k^{(n)}\right)_{n\in\mathbb{N}}$ is Cauchy there exists an $N \in \mathbb{N}$ such that

$$\left|x_k^{(n)} - x_k^{(m)}\right| < \frac{\epsilon}{2},$$

sending $m \to \infty$ gives

$$\left|x_k^{(n)} - x_k^{(\infty)}\right| \leq \frac{\epsilon}{2}.$$

Taking the supremum over $k \in \mathbb{N}$ we deduce that

$$\left\|x^{(n)} - x^{(\infty)}\right\|_{\ell^\infty} < \epsilon$$

which shows that $x^{(n)} \to x^{(\infty)}$ in ℓ^∞. Hence, $(\ell^\infty, \|\cdot\|_{\ell^\infty})$ is a Banach space.

□

Proposition 5.2.2. *If $p \leq q$ then $\ell^p \subseteq \ell^q$.*

Proof. If $p = \infty$ then $q = \infty$ and so $\ell^p \subseteq \ell^q$. Similarly, $\ell^p \subseteq \ell^\infty$ for all $p \in [1, \infty)$ as absolutely summable sequences are bounded. For $1 \leq p < \infty$ let $x \in \ell^p$ and consider $p \leq q < \infty$. As

$$\|x\|_{\ell^p}^p = \sum_{k=0}^{\infty}|x_k|^p < \infty,$$

it must be the case that $|x_k|^p \to 0$ as $k \to \infty$. More specifically there exists a $K \in \mathbb{N}$ such that $|x_k| < 1$ for $k \geq K$ which implies that $|x_k|^q \leq |x_k|^p$ for $k \geq K$. Thus,

$$\begin{aligned}
\|(x_k)_{k\in\mathbb{N}}\|_{\ell^q}^q &= \sum_{k=0}^{\infty}|x_k|^q \\
&= \sum_{k=0}^{K-1}|x_k|^q + \lim_{N\to\infty}\sum_{k=K}^{N}|x_k|^q \\
&\leq \sum_{k=0}^{K-1}|x_k|^q + \lim_{N\to\infty}\sum_{k=K}^{N}|x_k|^p \\
&\leq \sum_{k=0}^{K-1}|x_k|^q + \|(x_k)_{k\in\mathbb{N}}\|_{\ell^p}^p \\
&< \infty.
\end{aligned}$$

Therefore, $x \in \ell^q$. □

Remark 5.2.3. *Note the difference between Proposition 5.2.2 and Proposition 4.2.2.*

6 Linear Maps

6.1 Continuous Maps

Let E and F be normed vector spaces. The set of continuous linear maps from E to F is denoted by $\mathcal{L}(E, F)$.

Proposition 6.1.1. *Let E and F be normed vector spaces, and consider $T \in \mathcal{L}(E, F)$. Then the following are equivalent.*

- *T is continuous at zero.*
- *T is continuous on E.*
- *T is bounded, that is*
$$\|T\|_{E\to F} := \sup_{x\in E\setminus\{0\}} \frac{\|Tx\|_F}{\|x\|_E} < \infty.$$

Proposition 6.1.2. *The space $\mathcal{L}(E, F)$ endowed with $\|\cdot\|_{E\to F}$ is a normed vector space. Moreover, if F is a Banach space, then $\mathcal{L}(E, F)$ is a Banach space.*

Proof. Let $(T_n)_{n\in\mathbb{N}} \subseteq \mathcal{L}(E, F)$ be a Cauchy sequence. Fix $x \in E \setminus \{0\}$. Given an $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\|T_n - T_m\|_{\mathcal{L}(E,F)} < \frac{\epsilon}{\|x\|_E}$ for all $n, m \geq N$. Hence,

$$\|T_n(x) - T_m(x)\|_F \leq \|T_n - T_m\|_{\mathcal{L}(E,F)}\|x\|_E < \epsilon.$$

Therefore, the sequence $(T_n(x))_{n\in\mathbb{N}} \subseteq F$ is Cauchy which implies that $T_n(x) \to y_x \in F$. Let $T : E \to F$ be given by $T(x) = y_x$. For $x_1, x_2 \in E$ and $\lambda \in \mathbb{R}$, note that

$$\begin{aligned}
T(x_1 + \lambda x_2) &= \lim_{n\to\infty} T_n(x_1 + \lambda x_2) \\
&= \lim_{n\to\infty} T_n(x_1) + \lambda \lim_{n\to\infty} T_n(x_2) \\
&= T(x_1) + \lambda T(x_2).
\end{aligned}$$

Therefore, $T : E \to F$ is linear. As the sequence $(T_n)_{n\in\mathbb{N}} \subseteq \mathcal{L}(E, F)$ is Cauchy it is bounded. That is, there exists a $M > 0$ such that for all $n \in \mathbb{N}$ we have

$$\|T_n\|_{\mathcal{L}(E,F)} \leq M.$$

Moreover, for any $x \in E$, with $\|x\|_F = 1$, and $\epsilon > 0$, there exists an $N_x \in \mathbb{N}$ such that

$$\|T_n(x) - T(x)\|_F < \epsilon$$

for $n \geq N_x$. Therefore, for $n \geq N_x$ we deduce that

$$\|Tx\|_F \leq \|T_n(x) - T(x)\|_F + \|T_n(x)\|_F \leq \epsilon + M,$$

which implies that $\|T\|_{\mathcal{L}(E,F)} < \infty$ and so $T \in \mathcal{L}(E, F)$ as T is linear. Moreover, as $(T_n)_{n\in\mathbb{N}} \subseteq \mathcal{L}(E, F)$ is Cauchy, given an $\epsilon > 0$ we have

$$\sup_{\|x\|=1} \|T_n(x) - T_m(x)\|_F = \|T_n - T_m\|_{\mathcal{L}(E,F)} < \epsilon.$$

Hence, sending $m \to \infty$ gives

$$\sup_{\|x\|=1} \|T_n(x) - T(x)\|_F = \|T_n - T\|_{\mathcal{L}(E,F)} < \epsilon$$

so that $T_n \to T$ in $\mathcal{L}(E, F)$. $\square$

6.2 Dual Spaces

Throughout, let E be a Banach space.

Definition 6.2.1. *A linear form on E is a linear map of the form $E \to \mathbb{R}$ (or $\mathbb{C}$).*

Definition 6.2.2. *The dual of E denoted E', is the set of continuous linear forms. That is, $E' = \mathcal{L}(E, \mathbb{R})$.*

Example 6.2.3. *Let $E = \mathbb{R}^d$. Then $\varphi : E \to \mathbb{R}$ given by $(x_1, \dots, x_d) \mapsto x_i$ is a linear form. In fact, any linear form on $\mathbb{R}^d$ can be written as*

$$x = (x_1, \dots, x_d) \mapsto x \cdot y = \sum_{i=1}^{d} x_i y_i$$

for some $y \in \mathbb{R}^d$. Note that $\varphi(x) = x \cdot y$ where $y = (\underbrace{0, \dots, 1}_{i}, \dots, 0)$.

Exercise 6.2.4. *Show that for $p \in (1, \infty)$, we have that $(\ell^p)' = \ell^q$ where $\frac{1}{p} + \frac{1}{q} = 1$.*

Theorem 6.2.5 (Hahn-Banach)**.** *Let $G \subseteq E$ be a linear subspace, and $g \in \mathcal{L}(G, \mathbb{R})$ be bounded. Then there exists an extension $f \in E'$ such that*

- *$f = g$ on G, and*
- *$\|f\|_{E \to \mathbb{R}} = \|g\|_{G \to \mathbb{R}}$.*

Proof. Let $P = \{h : D(h) \subseteq E \to \mathbb{R}$, satisfying $1 - 5\}$.

1. $D(h)$ is a linear subspace.
2. $h \in \mathcal{L}(D(h), \mathbb{R})$.
3. $G = D(g) \subseteq D(h)$.
4. $h = g$ on G.
5. $\|h\|_{D(h) \to \mathbb{R}} = \|g\|_{G \to \mathbb{R}}$.

Let us introduce an order relation $\leq$ on P where $h_1 \leq h_2$ if and only if the following hold.

1. $D(h_1) \subseteq D(h_2)$.
2. $h_2 = h_1$ on $D(h_1)$.

Step 1: P is inductive.
Let $Q \subseteq P$ be a totally ordered subset. Then let $(h, D(h))$ be given by $D(h) = \bigcup_{q \in Q} D(q)$ and $h(x) = q(x)$ if $x \in D(q)$. This is well-defined, and h is an upper bound of Q, implying P is inductive.
Step 2: Apply Zorn's Lemma.
By Lemma 10.1.5 there exists a maximal element f.
Step 3: Show that $D(f) = E$.
Proceed by contradiction, and assume that $D(f) \neq E$. Then choose $x_0 \in E \setminus D(f)$. Let $(h, D(h))$ be given by $D(h) = D(f) + \mathbb{R}x_0$ and $h(x + tx_0) = f(x) + \alpha t$ for $(x, t) \in D(f) \times \mathbb{R}$. Let $C_0 = \|g\|_{G \to \mathbb{R}}$. We want to choose α such that

$$|f(x) + t\alpha| \leq C_0 \|x + tx_0\|.$$

By positive homogeneity we note that $|f(x)+t\alpha| = |t|\left|f\left(\frac{x}{t}\right)+\alpha\right|$, so it suffices to consider $t=\pm 1$. Thus, it suffices to require that

$$\begin{cases} f(x)+\alpha \le C_0\|x+x_0\| \\ f(x)-\alpha \le C_0\|x+x_0\| \end{cases}$$

which is equivalent to

$$\sup_{y\in D(h)} \left(f(y) - C_0\|y+x_0\|\right) \le \alpha \le \left(\inf_{z\in D(h)} C_0\|z+x_0\| - f(z)\right).$$

For such an α to exists we need

$$f(y) - C_0\|y+x_0\| \le C_0\|z+x_0\| - f(z)$$

for all y and z, which happens if and only if

$$f(y-z) = f(y)-f(z) \le C_0\|z+x_0\| + C_0\|y+x_0\|.$$

This holds since

$$f(y-z) \le C_0\|y-z\| \le C_0(\|y+x_0\| + \|z+x_0\|)$$

by the triangle inequality. Therefore, by the construction of f it follows that $\|h\|_{D(h)} = \|g\|_{G\to\mathbb{R}}$ and so $h\in P$. In particular, $f\le h$ which contradicts f being a maximal element. $\square$

6.3 Applications of the Hahn-Banach Theorem

Theorem 6.3.1. *If E is a normed vector space and $x\in E$, then there exits a $\rho\in E'$ such that*

$$\|x\|_E = \frac{\rho(x)}{\|\rho\|_{E'}}$$

where $\|\rho\|_{E'} = \|\rho\|_{E\to\mathbb{R}}$.

Proof. Define ρ on $\mathbb{R}x$ by $\rho(tx)=t$. Note that

$$\frac{|\rho(tx)|}{\|x\|_E} = \frac{|t|}{\|tx\|_E} \le \frac{1}{\|x\|_E},$$

with equality at $x\in\mathbb{R}x$. Thus, we can extend ρ to E using Theorem 6.2.5 such that $\|\rho\|_{E'} = \frac{1}{\|x\|_E}$. Then,

$$\rho(x) = 1 = \frac{\|x\|_E}{\|x\|_E} = \|x\|_E\|\rho\|_{E'}$$

so that

$$\|x\|_E = \frac{\rho(x)}{\|\rho\|_{E'}}.$$

$\square$

Remark 6.3.2.

- *Equivalently, we can say that there exists a $\rho\in E'$ with $\|\rho\|_{E'}=1$ such that $\rho(x)=\|x\|_E$.*
- *In finite dimensions, say with $E=\mathbb{R}^d$, any linear form can written as $\rho_y:\mathbb{R}^d\to\mathbb{R}$ where $x\mapsto x\cdot y = \sum_{i=1}^d x_iy_i$. Note that*

$$\|\rho_y\| = \sup_{x\in\mathbb{R}^d\setminus\{0\}} \frac{|x\cdot y|}{\|x\|} \le \|y\|$$

by Cauchy-Schwartz. More specifically,

$$\frac{|x \cdot y|}{\|y\|} = \frac{|\rho_y(x)|}{\|y\|} = \|x\|$$

if and only if y is parallel to x.

Theorem 6.3.3. *Let E be a normed vector space with $F \subseteq E$ a linear subspace. Then if $\bar{F} \neq E$, it follows that there exists a $\rho \in E'$ such that $\rho \neq 0$ and*

$$\rho(x) := \langle \rho, x \rangle = 0$$

for all $x \in F$.

Proof. Let $v \in E \setminus \bar{F}$ and define $\tilde{F} = F + \operatorname{span}(v)$. Note that for each $u \in \tilde{F}$ we can write $u = f + \lambda v$ uniquely, for $f \in F$ and $\lambda \in \mathbb{R}$. Let $g : \tilde{F} \to \mathbb{R}$ be given by

$$u \mapsto \lambda.$$

Note that $g(u) = 0$ for all $u \in F$. As $v \notin \bar{F}$ there exists an $\epsilon > 0$ such that $\|v - f\|_E \geq \epsilon > 0$ for all $f \in F$. As F is a linear subspace we note that $f \in F$ if and only if $-\frac{f}{\lambda} \in F$. So we can equivalently say that $\left\| v + \frac{f}{\lambda} \right\|_E \geq \epsilon > 0$ for all $f \in F$. Hence, for $u \in \tilde{F}$ we have that

$$\begin{aligned}
\|g\|_{(\tilde{F})'} &= \sup_{u \in \tilde{F} \setminus \{0\}} \frac{|g(u)|}{\|u\|_E} \\
&= \sup_{u \in \tilde{F} \setminus \{0\}} \frac{|\lambda|}{\|\lambda v + f\|_E} \\
&= \sup_{u \in \tilde{F} \setminus \{0\}} \frac{1}{|\lambda|} \frac{|\lambda|}{\left\| v + \frac{f}{\lambda} \right\|_E} \\
&\leq \frac{1}{\epsilon}.
\end{aligned}$$

As g is linear, it follows that $g \in \left(\tilde{F}\right)'$. Therefore, by Theorem 6.2.5 this can be extended to $\rho \in E'$. $\square$

6.4 Riesz Representation Theorem

For $p, q \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$, we say that p and q are dual, and usually write $q = p'$. Let $f \in L^{p'}$ and consider the linear form $\rho_f : L^p \to \mathbb{R}$ where

$$\varphi \mapsto \int f \varphi \, \mathrm{d}x.$$

Note that by Hölder's inequality this is well-defined and bounded,

$$|\rho_f(\varphi)| = \left| \int f \varphi \right| \leq \|f\|_{L^{p'}} \|\varphi\|_{L^p}.$$

Consequently, $\rho_f \in (L^p)'$ with

$$\|\rho_f\|_{(L^p)'} \leq \|f\|_{L^{p'}}.$$

Exercise 6.4.1. *Show that $\|\rho_f\|_{(L^p)'} = \|f\|_{L^{p'}}$.*

Theorem 6.4.2 (Riesz Representation Theorem)**.** *If $1 \le p < \infty$, then any element of $(L^p)'$ can be represented as ρ_f for some $f \in L^{p'}$.*

Remark 6.4.3. *The same holds if L^p is replaced with ℓ^p.*

The statement of Theorem 6.4.2 breaks down for $p = \infty$. One can see how for the space ℓ^p. Observe that

$$|\rho_y(x)| = \left|\sum_{n=0}^{\infty} x_n y_n\right| \le \|x\|_{\ell^\infty}\|y\|_{\ell^1}.$$

Which means that ℓ^1 provides linear forms on ℓ^∞, namely $\rho_y \in (\ell^\infty)'$ for $y \in \ell^1$. Now let $X \subseteq \ell^\infty$ be the sequences with a limit. Then define ρ on X by $\rho((x_n)_{n\in\mathbb{N}}) = \lim_{n\to\infty}(x_n)$. By the Theorem 6.2.5, ρ can be extended to ℓ^∞. Hence, we get a $\rho \in (\ell^\infty)'$ such that $\rho(x) = \lim_{n\to\infty}(x_n)$ if $(x_n)_{n\in\mathbb{N}}$ converges. Suppose $\rho(x) = \rho_y(x) = \sum_{n\in\mathbb{N}} x_n y_n$ for some $y \in \ell^1$. As $y \in \ell^1$, given an $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\sum_{n\ge N} |y_n| < \epsilon$. Let $x \in \ell^\infty$ be given by

$$x_n = \begin{cases} 0 & n < N \\ 1 & n \ge N. \end{cases}$$

Then as $\lim_{n\to\infty}(x_n) = 1$ we have

$$1 = \rho(x) = |\rho_y((x_n)_{n\in\mathbb{N}})| = \left|\sum_{n\ge N} y_n\right| < \epsilon.$$

Therefore, ρ cannot be equal to ρ_y for any $y \in \ell^1$, and so the statement of Theorem 6.4.2 cannot hold.

Exercise 6.4.4. *Show that the f in the statement of Theorem 6.4.2 is unique, up to equality almost everywhere.*

Example 6.4.5.

1. *Consider $T : \ell^2 \to \mathbb{R}$ given by*

   $$T(x) = \sum_{n\in\mathbb{N}} x_n e^{an}$$

   where $a \in \mathbb{R}$. Since

   $$\sum_{n\in\mathbb{N}} |e^{an}|^2 = \frac{e^{2a}}{1 - e^{2a}},$$

   it follows by Theorem 6.4.2 and Exercise 6.4.1 that

   $$\|T\|_{(\ell^2)'} = \left(\frac{e^{2a}}{1 - e^{2a}}\right)^{\frac{1}{2}}.$$

2. *For $I = (0, \infty)$ and $p = (1, \infty)$ consider the operator $T : L^p(I) \to \mathbb{R}$ given by*

   $$T(f) = \int_I \arctan(y) f(y)\,\mathrm{d}y.$$

   As T is linear, if it were were bounded then $T \in (L^p)'$. So by Theorem 6.4.2 there would exist $g \in L^{p'}$ such that

   $$T(f) = \int_I g(y) f(y)\,\mathrm{d}y.$$

However, this would imply that $\arctan(y) \in L^{p'}(I)$, which is not the case. Therefore, T cannot be a bounded operator.

6.5 Bi-dual Space

For E a Banach space the bi-dual of E is the dual of E', namely E''. On E'' we have the norm

$$\|f\|_{E''} = \sup_{\rho \in E' \setminus \{0\}} \frac{|f(\rho)|}{\|\rho\|_{E'}}.$$

There is a natural map from $\Phi : E \to E''$ given by $x \mapsto f_x$ where $f_x : E' \to \mathbb{R}$ is such that $\rho \mapsto \rho(x)$.

Exercise 6.5.1. *Verify that f_x is linear.*

Observe that

$$\begin{aligned} \|f_x\|_{E''} &= \sup_{\rho \in E' \setminus \{0\}} \frac{|f_x(\rho)|}{\|\rho\|_{E'}} \\ &= \sup_{\rho \in E' \setminus \{0\}} \frac{|\rho(x)|}{\|\rho\|_{E'}} \\ &\overset{(1)}{=} \|x\|_E. \end{aligned}$$

To justify (1) recall that $|\rho(x)| \leq \|\rho\|_{E'} \|x\|_E$ and note that by Theorem 6.2.5 we can construct a ρ that achieves this upper bound. Thus, $f_x \in E''$ and so Φ is well-defined. In particular, we deduce that Φ is an isometry, which implies that Φ is an injective linear operator. If Φ is also surjective, we call E a reflexive space.

Example 6.5.2.

1. *On $\mathbb{R}^d$ with the Euclidean norm, any linear form is bounded and can be represented as*

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

for some $y \in \mathbb{R}^d$. Furthermore,

$$\|\rho_y\|_{(\mathbb{R}^d)'} = \sup_{x \in \mathbb{R}^d \setminus \{0\}} \frac{(y, x)}{\|x\|} = \|y\|.$$

It is easy to check then that Φ is an isomorphism. Consequently, $\mathbb{R}^d$ with the Euclidean norm is reflexive.

2. *Consider L^p for $1 < p < \infty$. By Theorem 6.4.2, $(L^p)' \simeq L^{p'}$. Consequently,*

$$(L^p)'' \simeq \left(L^{p'}\right)' \simeq L^p. \tag{6.5.1}$$

Therefore, L^p is reflexive for $1 < p < \infty$.

3. *For $p \in \{1, \infty\}$, the space L^p is not reflexive. Note that although the first equality in (6.5.1) holds for $p = 1$, the second inequality does not hold as $p' = \infty$.*

4. *The same conclusions made for L^p hold for ℓ^p.*

6.6 Solution to Exercises

Exercise 6.2.4

Solution. Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. For $v \in \ell^q$ let $T(v) : \ell^p \to \mathbb{R}$ be given by

$$u \mapsto \sum_{n \in \mathbb{N}} v_n u_n.$$

Step 1: Show that the map $T(v) : \ell^p \to \mathbb{R}$ is well-defined for $v \in \ell^q$.
Observe that

$$\begin{aligned}
|T(v)(u)| &= \left|\sum_{n\in\mathbb{N}} v_n u_n\right| \\
&\leq \sum_{n\in\mathbb{N}} |v_n u_n| \\
&\leq \|v\|_{\ell^q}\|u\|_{\ell^p} \\
&< \infty.
\end{aligned}$$

Therefore, $T(v)$ is well-defined.
Let $T : \ell^q \to (\ell^p)'$ be given by $v \mapsto T(v)$.
Step 2: Show that $T : \ell^q \to (\ell^p)'$ is well-defined and continuous.
The map $v \mapsto T(v)$ is well-defined as from step 1 we know that $T(v) \in (\ell^p)'$. For $v^1, v^2 \in \ell^p$, $\lambda \in \mathbb{R}$, and fixed $u \in \ell^p$ we have that

$$\begin{aligned}
T\left(v^1 + \lambda v^2\right)(u) &= \sum_{n\in\mathbb{N}} \left(v_n^1 + \lambda v_n^2\right) u_n \\
&= \sum_{n\in\mathbb{N}} v_n^1 u_n + \lambda \sum_{n\in\mathbb{N}} v_n^2 u_n \\
&= T\left(v^1\right)(u) + \lambda T\left(v^2\right)(u).
\end{aligned}$$

Hence $v \mapsto T(v)$ is linear. Next observe that for $u \in \ell^p \setminus \{0\}$ we have that

$$\frac{|T(v)(u)|}{\|u\|_{\ell^p}} \leq \frac{\sum_{n\in\mathbb{N}} |v_n u_n|}{\|u\|_{\ell^p}} \leq \frac{\|v\|_{\ell^q}\|u\|_{\ell^p}}{\|u\|_{\ell^p}} = \|v\|_{\ell^q}.$$

Hence

$$\|T(v)\|_{(\ell^p)'} \leq \|v\|_{\ell^q}.$$

Therefore,

$$\|T\|_{\ell^q \to (\ell^p)'} = \sup_{v\in\ell^q\setminus\{0\}} \frac{\|T(v)\|_{(\ell^q)'}}{\|v\|_{\ell^p}} \leq 1$$

which implies that the map is bounded and hence continuous as it is also linear.
Step 3: Show that T is injective.
Suppose that for $u, v \in \ell^q$ we have that $T(u) = T(v)$. For $i \in \mathbb{N}$, consider $e^i \in \ell^p$ where

$$e_n^i = \begin{cases} 1 & n = i \\ 0 & \text{otherwise.} \end{cases}$$

Then $u_i = T(u)\left(e^i\right) = T(v)\left(e^i\right) = v_i$. Therefore, $u = v$ and so $v \mapsto T(v)$ is injective.
Step 4: Show that T is surjective.
Let $\xi \in (\ell^p)'$ and consider $v = (v_n)_{n\in\mathbb{N}}$ where $v_n = \xi(e_n)$. For $u \in \ell^p$ let $u^N = \left(u_n \mathbf{1}_{n\leq N}\right)_{n\in\mathbb{N}}$. Observe that

$$\begin{aligned}
T(v)\left(u^N\right) &= \sum_{n=1}^{N} v_n u_n \\
&= \sum_{n=1}^{N} \xi(e_n) u_n \\
&= \xi\left(\sum_{n=1}^{N} u_n e_n\right) \\
&= \xi\left(u^N\right)
\end{aligned}$$

which implies $\left|T(v)\left(u^N\right)\right| \leq \|\xi\|_{(\ell^p)'} \left\|u^N\right\|_{\ell^p}$. Moreover,

$$\left\|u^N - u\right\|_{\ell^p}^p = \sum_{n=N+1}^{\infty} |u_n|^p \overset{N\to\infty}{\longrightarrow} 0.$$

Hence,

$$\left|T(v)\left(u^N\right) - \xi\left(u\right)\right| = \left|\xi\left(u^N - u\right)\right| \leq \|\xi\|_{(\ell^p)'} \left\|u^N - u\right\|_{\ell^p} \overset{N\to\infty}{\longrightarrow} 0.$$

Therefore, $T(v)\left(u^N\right) \to \xi(u)$ in $\mathbb{R}$ as $N \to \infty$. As $T(v)\left(u^N\right) \to T(v)(u)$ as $N \to \infty$ by the continuity of T, it follows using the uniqueness of limits that $T(v)u = \xi(u)$. As this holds for any $u \in \ell^p$ it follows that $T(v) = \xi$ in the $(\ell^p)'$ sense. As $\xi \in (\ell^p)'$ was arbitrary we conclude that T is surjective.

Step 5: Deduce that $(\ell^p)' = \ell^q$.

The map T is a bijective and continuous map, so $(\ell^p)' = \ell^q$. □

Exercise 6.4.1

Solution. Let $\varphi(x) = \operatorname{sgn}(f(x))|f(x)|^{p'-1}$. Then

$$\begin{aligned}\|\varphi\|_{L^p}^p &= \int |f|^{p(p'-1)} \\ &= \int |f|^{p'},\end{aligned}$$

so that $f \in L^p$. Therefore, as

$$\begin{aligned}\frac{|\rho_f(\varphi)|}{\|\varphi\|_{L^p}} &= \frac{\int |f|^{p'}}{\left(\int |f|^{p'}\right)^{\frac{1}{p}}} \\ &= \left(\int |f|^{p'}\right)^{1-\frac{1}{p}} \\ &= \|f\|_{L^{p'}}.\end{aligned}$$

it follows that $\|\rho_f\|_{(L^p)'} = \|f\|_{L^{p'}}$. □

Exercise 6.4.4

Solution. Suppose that for $f, g \in L^{p'}$ we have $\rho_f = \rho_g$. Then

$$\int f\varphi \,\mathrm{d}x = \int g\varphi \,\mathrm{d}x$$

for all $\varphi \in L^p$. In particular,

$$\int (f-g)\varphi \,\mathrm{d}x = 0$$

for all $\varphi \in L^p$. Letting $\varphi = \operatorname{sgn}(f-g)\mathbf{1}_{[-n,n]^d}$ we deduce that $h_n = |f-g|\mathbf{1}_{[-n,n]^d} = 0$ almost everywhere. As $h_n \to |f-g|$ pointwise almost everywhere we deduce using the dominated convergence theorem that

$$0 = \lim_{n\to\infty} \int h_n \,\mathrm{d}x = \int |f-g| \,\mathrm{d}x,$$

which implies that $f = g$ almost everywhere. □

Exercise 6.5.1

Solution. Note that for $\rho_1, \rho_2 \in E'$ and $\lambda \in \mathbb{R}$ we have

$$\begin{aligned}f_x(\rho_1 + \lambda\rho_2)(x) &= (\rho_1 + \lambda\rho_2)(x) \\ &= \rho_1(x) + \lambda\rho_2(x) \\ &= f_x(\rho_1) + \lambda f_x(\rho_2).\end{aligned}$$

Hence, f_x is linear. □

7 Compactness in Normed Vector Spaces

7.1 Compact Sets

In metric spaces the equivalent Bolzano-Weierstrass property and the open-covering property characterise compactness. For finite-dimensional vector spaces, Theorem 1.2.37 identifies compact sets. For infinite-dimensional vector spaces, the identification of compact sets is not as straightforward.

Lemma 7.1.1. *Let E be a normed vector space, and let $M \subseteq E$ be a closed linear subspace where $M \neq E$. Then for all $\epsilon > 0$ there exists $u \in E$ such that,*

1. $\|u\| = 1$, *and*
2. $\operatorname{dist}(u, M) \geq 1 - \epsilon$.

Proof. Pick $v \in E \setminus M$. Then $d := \operatorname{dist}(v, M) > 0$ as $v \notin M$ and M is closed. So there exists an $m_0 \in M$ such that

$$d \leq \|v - m_0\| \leq \frac{d}{1-\epsilon}.$$

Now let $u = \frac{v - m_0}{\|v - m_0\|}$. It is clear that $\|u\| = 1$. Moreover, for $m \in M$ we have

$$\begin{aligned}
\|u - m\| &= \left\| \frac{v - m_0}{\|v - m_0\|} - m \right\| \\
&= \frac{1}{\|v - m_0\|} \|v - m_0 - \|v - m_0\| m\| \\
&\geq \frac{1-\epsilon}{d} \|v - m'\|
\end{aligned}$$

where m' is some element of M. Hence as $\|v - m'\| \geq d$ we have that

$$\|u - m\| \geq 1 - \epsilon.$$

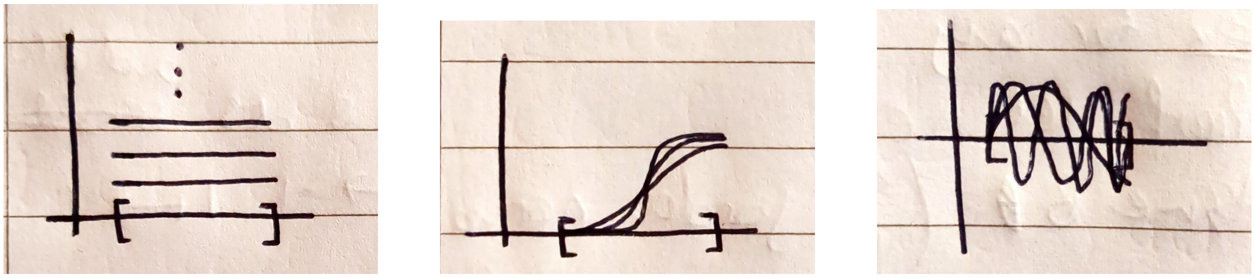
$\square$

Example 7.1.2. *Let $E = \mathbb{R}^d$ with the Euclidean norm, and let $M \subseteq E$ be a linear subspace with $M \neq E$. Then one considers the line orthogonal to M passing through the origin. Choosing a point where this line intersects the unit ball will provide a satisfactory vector u.*

Theorem 7.1.3 (Riesz)**.** *Let E be a normed vector space of infinite dimension. Then the closed unit ball is not compact.*

Proof. Let $u_0 \in E$ be of unit norm. Then, by Lemma 7.1.1, for $\epsilon \in (0, 1)$ there exists a unit vector such that $\|u_1\| = 1$ and $\operatorname{dist}(u_1, \operatorname{span}(u_0)) \geq 1 - \epsilon$. As E is infinite-dimensional, we can continue to find a unit vector u_n such that $\operatorname{dist}(u_n, \operatorname{span}(u_0, \ldots, u_{n-1})) \geq 1 - \epsilon$. The sequence $(u_n)_{n \in \mathbb{N}}$ is such that $\|u_n - u_m\| \geq 1 - \epsilon$ for all $n \neq m$. Therefore, the sequence has no convergent subsequence and so does not satisfy the Bolzano-Weierstrass property. Therefore, the closed-unit ball is not compact. $\square$

Theorem 7.1.3, shows that extending our notions of compactness to infinite dimensions fails rather fundamentally. Theorem 7.1.7 will give us a characterisation of compactness for the set of continuous functions on the closure of open and bounded sets Ω, denoted $\mathcal{C}^0\left(\bar{\Omega}\right)$.



(a) A sequence of functions that is unbounded.

(b) A sequence of functions converging to a step function.

(c) A sequence of functions that oscillate at an ever-increasing rate.

Figure 9: Examples illustrating some necessary conditions for sequences of functions to admit convergent subsequences.

From Figure 9a we note that we must require a sequence of functions to be bounded to admit a convergence subsequence. Similarly, Figures 9b and 9c show that we must have a condition which ensures the derivatives of these functions are bounded.

Definition 7.1.4. *A sequence $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}^0\left(\bar{\Omega}\right)$ is bounded with constant C, if*

$$\|f_n\|_\infty \leq C$$

for every $n \in \mathbb{N}$.

Definition 7.1.5. *A sequence $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}^0\left(\bar{\Omega}\right)$ is equicontinuous if for all $\epsilon > 0$ there exists a $\delta > 0$ such that for $x, y \in \bar{\Omega}$ with $|x - y| < \delta$ we have $|f_n(x) - f_n(y)| < \epsilon$ for all $n \in \mathbb{N}$.*

Example 7.1.6. *Let $f_n : \overline{B_{\mathbb{R}^d}(0,1)} \to \mathbb{R}$ be given by $f_n(x) = e^{-n\|x\|}$. As $x \mapsto e^{-x}$ and $x \mapsto \|x\|$ are continuous their composition $f_n(x)$ is continuous. Moreover, the sequence of functions $(f_n)_{n\in\mathbb{N}}$ is bounded. However, let $x = 0$ and $\epsilon = \frac{1}{2}$. Then for any $\delta > 0$ let $y \in \overline{B_{\mathbb{R}^d}(0,1)}$ be such that $\|y\| = \frac{\delta}{2}$. Then*

$$|f_n(0) - f_n(y)| = \left|1 - e^{-\frac{n\delta}{2}}\right| \overset{n\to\infty}{\longrightarrow} 1.$$

Therefore, there exists an $n \in \mathbb{N}$ such that

$$|f_n(0) - f_n(y)| \geq \frac{1}{2}$$

and so the sequence $(f_n)_{n\in\mathbb{N}}$ is not equicontinuous.

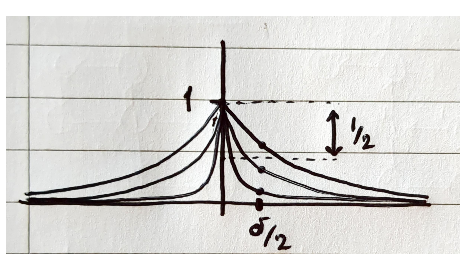


Figure 10: Intuitively the functions referenced in Example 7.1.6 are not equicontinuous as the gradients of the function near the origin diverge as n gets large.

Theorem 7.1.7 (Arzela-Ascoli). *Let $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}^0\left(\bar{\Omega}\right)$ be a sequence that is bounded, with constant C, and equicontinuous. Then the sequence $(f_n)_{n\in\mathbb{N}}$ admits a convergent subsequence.*

Proof. To simplify the proof we suppose $(f_n)_{n\in\mathbb{N}}$ is uniformly equicontinuous. Uniform equicontinuity says that for all $\epsilon > 0$ there exists a $\delta > 0$ such that for all $x, y \in \bar{\Omega}$ and $n \in \mathbb{N}$ then $|x-y| < \delta$ implies $|f_n(x) - f_n(y)| < \epsilon$.
Step 1: Finding a dense set of points.
Arrange the rational numbers in $\bar{\Omega}$ into a sequence $(r_n)_{n\in\mathbb{N}}$.
Step 2: Apply the Cantor diagonal argument.
Let $\varphi_1 : \mathbb{N} \to \mathbb{N}$ be such that $\left(f_{\varphi_1(n)}(r_1)\right)_{n\in\mathbb{N}}$ converges. This is possible since the sequence $(f_n(r_1))_{n\in\mathbb{N}}$ is bounded and so has a convergent subsequence. Now let $(f_{\varphi_2(n)})_{n\in\mathbb{N}}$ be a subsequence of $\left(f_{\varphi_1(n)}\right)_{n\in\mathbb{N}}$ such that $\left(f_{\varphi_2(n)}(r_2)\right)_{n\in\mathbb{N}}$ converges. Again we can do this as the sequences are bounded and so admit convergent subsequences. Note that $\left(f_{\varphi_2(n)}(r_1)\right)_{n\in\mathbb{N}}$ converges as $(f_{\varphi_2(n)})_{n\in\mathbb{N}} \subseteq (f_{\varphi_1(n)})_{n\in\mathbb{N}}$. Continue in this way to determine $\varphi_k : \mathbb{N} \to \mathbb{N}$ such that $\left(f_{\varphi_k(n)}\right)_{n\in\mathbb{N}} \subseteq \left(f_{\varphi_{k-1}(n)}\right)_{n\in\mathbb{N}}$ and $\left(f_{\varphi_k(n)}(r_k)\right)_{n\in\mathbb{N}}$ converges. Again note that $\left(f_{\varphi_k(n)}(r_j)\right)_{n\in\mathbb{N}}$ converges for all $j = 1, \ldots, k-1$. Now set $\varphi(n) = \varphi_n(n)$. Then $\left(f_{\varphi(n)}(r_j)\right)_{n\in\mathbb{N}}$ converges for any $j \in \mathbb{N}$ as $\left(f_{\varphi(n)}\right)_{n\in\mathbb{N}} \subseteq \left(f_{\varphi_j(n)}\right)_{n\in\mathbb{N}}$ for all $j \in \mathbb{N}$.
Step 3: The candidate limit.
Let $f(r) = \lim_{n\to\infty} f_{\varphi(n)}(r)$ for all $r \in \mathbb{Q} \cap \bar{\Omega}$.
Step 4: Extend f using uniform equicontinuity.
For any $\epsilon > 0$, by uniform equicontinuity, there exists a $\delta > 0$ such that $|x-y| < \delta$ implies $|f_{\varphi(n)}(x) - f_{\varphi(n)}(y)| < \epsilon$ for all $n \in \mathbb{N}$. Thus, we can extend f to $\bar{\Omega}$ by letting $f(x) = \lim_{r\to x} f(r)$.
Step 5: $f_{\varphi(m)}$ converges to f in $\|\cdot\|_\infty$.
Fix $\epsilon > 0$.

- Choose $\delta > 0$ such that $|x-y| < \delta$ implies $|f_n(x) - f_n(y)| < \frac{\epsilon}{3}$ for all $n \in \mathbb{N}$.
- Choose $N \in \mathbb{N}$ such that for all $x \in \bar{\Omega}$ there exists a $j \in \{1, \ldots, N\}$ such that $|x - r_j| < \delta$.
- Choose $M \in \mathbb{N}$ such that for all $j \in \{1, \ldots, N\}$ if $m > M$, then $\left|f_{\varphi(m)}(r_j) - f(r_j)\right| < \frac{\epsilon}{3}$.

For $x \in \bar{\Omega}$, choose j_0 such that $|r_{j_0} - x| < \delta$. If $n > M$ then

$$\begin{aligned} \left|f(x) - f_{\varphi(n)}(x)\right| &\leq |f(x) - f(r_j)| + \left|f(r_j) - f_{\varphi(n)}(r_j)\right| + \left|f_{\varphi(n)}(r_j) - f_{\varphi(n)}(x)\right| \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &\leq \epsilon. \end{aligned}$$

□

Example 7.1.8. *Consider the sequence $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}^0\left(\overline{B_{\mathbb{R}^d}(0,1)}\right)$ from Example 7.1.6. Suppose that $f_{\varphi(n)} \to f$ in $\mathcal{C}^0(0,1)$. Then $f_{\varphi(n)}(x) \to f(x)$ for each $x \in (0,1)$. However,*

$$f_{\varphi(n)}(x) = e^{-\varphi(n)\|x\|} \longrightarrow \begin{cases} 1 & x = 0 \\ 0 & \textit{otherwise,} \end{cases}$$

which is not a continuous function. Therefore, there cannot exist a convergent subsequence $\left(f_{\varphi(n)}\right)_{n\in\mathbb{N}} \subseteq \mathcal{C}^0(0,1)$. Recall, that the sequence $(f_n)_{n\in\mathbb{N}}$ was shown not to be equicontinuous. Hence, the requirement of equicontinuity in Theorem 7.1.7 is necessary.

7.2 Compact Operators

Let E and F be Banach spaces. Recall that $\mathcal{L}(E,F)$ is the set of bounded linear operators $E \to F$. Moreover,

$$\|T\|_{E\to F} = \sup_{x\in E\setminus\{0\}} \frac{\|Tx\|_F}{\|x\|_E}.$$

Thus,

$$\|Tx\|_F \leq \|T\|_{E\to F}\|x\|_E.$$

Definition 7.2.1. *A set $S \subseteq X$ is pre-compact if $\bar{S}$ is compact.*

Definition 7.2.2. *The operator $T \in \mathcal{L}(E,F)$ is compact if $T\left(\bar{B}^E\right)$ is pre-compact, where*

$$\bar{B}^E := \{x \in E : \|x\| \leq 1\}.$$

Example 7.2.3.

1. *Using Theorem 7.1.3 it follows that for a Banach space E, the operator $\mathrm{Id} : E \to E$ is compact if and only if $\dim(E) < \infty$. Therefore, in some sense, compact operators must shrink sets on which they are applied.*

2. *Consider $\mathrm{Id} : \mathcal{C}^1\left(\bar{\Omega}\right) \to \mathcal{C}^0\left(\bar{\Omega}\right)$. The unit ball consists of functions $f \in \mathcal{C}^1\left(\bar{\Omega}\right)$ such that $\|f\|_\infty + \sum_{i=1}^d \|\partial_i f\|_\infty \leq 1$. In particular,*
$$|f(x) - f(y)| \leq C\|x-y\|$$
for any $x, y \in \bar{\Omega}$ by the mean value theorem. Therefore, using Theorem 7.1.7 we deduce that the image of the unit ball is compact.

3. *For $T : E \to F$ where $\dim(F) < \infty$, the image of the unit ball is bounded and so by Theorem 1.2.37 its closure is compact and hence the set is pre-compact. Therefore, T is compact.*

4. *Let $T : L^p(0,1) \to \mathcal{C}^0(0,1)$ where $f \mapsto \int K(x,y)f(y)\,\mathrm{d}y$ for $K \in \mathcal{C}^1\left([0,1]^2\right)$. This is well-defined by Hölder's inequality. Moreover,*
$$\begin{aligned}|Tf(x) - Tf(x')| &= \left|\int_0^1 \left(K(x,y) - K(x',y)\right) f(y)\,\mathrm{d}y\right| \\ &\overset{(1)}{\leq} \int_0^1 \left|K(x,y) - K(x',y)\right| |f(y)|\,\mathrm{d}y \\ &\overset{(2)}{\leq} C\,|x-x'|\,\|f\|_{L^p},\end{aligned}$$
where (1) is the generalised triangle inequality, and (2) follows from Hölder's inequality and the mean value theorem applied to K. Therefore, by Theorem 7.1.7 the operator is compact.

5. *Let $T : \ell^p \to \ell^p$ be given by*
$$T(e_i) = \begin{cases} 0 & i \text{ even} \\ e_{i+1} & i \text{ odd.}\end{cases}$$
As
$$\|T(x)\|_{\ell^p} \leq \|x\|_{\ell^p},$$
we have that $T \in \mathcal{L}(\ell^p)$. However, T is not compact since for the sequence $(e_{2i+1})_{i\in\mathbb{N}} \subseteq \bar{B}^{\ell^p}$ the sequence $\left(T\left(e_{2i+1}\right)\right)_{i\in\mathbb{N}} = \left(e_{2i}\right)_{i\in\mathbb{N}}$ has no convergent subsequence as
$$\|e_{2i} - e_{2j}\|_{\ell^p} = 2^{\frac{1}{p}}\delta_{ij}$$
which implies that any subsequence is not Cauchy.

Theorem 7.2.4. *The set of compact operators denoted $\mathcal{K}(E,F)$, is closed in $\mathcal{L}(E,F)$.*

Proof. Let $(T_i)_{i\in\mathbb{N}} \subseteq \mathcal{K}(E,F)$ be a sequence converging to $T \in \mathcal{L}(E,F)$. Let $(x_j)_{j\in\mathbb{N}} \subseteq \bar{B}^E$. We can use a diagonal argument to find a $\varphi : \mathbb{N} \to \mathbb{N}$ such that $\left(T_i\left(x_{\varphi(j)}\right)\right)_{j\in\mathbb{N}}$ converges for each $i \in \mathbb{N}$. We can write

$$\left\|Tx_{\varphi(n)} - Tx_{\varphi(m)}\right\| \leq \left\|Tx_{\varphi(n)} - T_kx_{\varphi(n)}\right\| + \left\|T_kx_{\varphi(n)} - T_kx_{\varphi(m)}\right\| + \left\|T_kx_{\varphi(m)} - Tx_{\varphi(m)}\right\|$$

where the first term can be made small for large k as $\|Tx_{\varphi(n)} - T_kx_{\varphi(n)}\| \leq \|T - T_k\| \left\|x_{\varphi(n)}\right\|$ where $\|T - T_k\| \to 0$ and $\left\|x_{\varphi(n)}\right\| \leq 1$, similarly for the third term. The second term can be made small by the fact that $\left(T_kx_{\varphi(n)}\right)_{n\in\mathbb{N}}$ is convergent. Hence, we deduce that $\left(Tx_{\varphi(n)}\right)_{n\in\mathbb{N}}$ is Cauchy, and thus it converges as F is a Banach space. Therefore, $T\left(\bar{B}^E\right)$ is pre-compact and thus $T \in \mathcal{K}(E,F)$. ☐

Definition 7.2.5. *Let $T \in \mathcal{L}(H)$. The range of T is $\mathrm{Ran}(T) := T(H)$. If $\dim(\mathrm{Ran}(T)) < \infty$, then T is said to be a finite range or a finite rank operator.*

Exercise 7.2.6. *Let E be a Banach space and consider $T \in \mathcal{L}(E)$ a finite range operator. Show that $T \in \mathcal{K}(E)$.*

Corollary 7.2.7. *Let $T_n : E \to F$ be a sequence of finite range operators. If $T_n \to T$, then T is compact.*

Proof. Using Exercise 7.2.6 we know that T_n is compact. Therefore, if $T_n \to T$ exists, Theorem 7.2.4 says that T is compact. ☐

Example 7.2.8. *Let $T : \ell^2 \to \ell^2$ be given by $(x_n)_{n\in\mathbb{N}} \mapsto (c_nx_n)_{n\in\mathbb{N}}$. One can think of this operator as the matrix*

$$\begin{pmatrix} c_1 & & & 0 \\ & c_2 & & \\ & & c_3 & \\ 0 & & & \ddots \end{pmatrix}.$$

- *T is bounded if and only if $|c_n| \leq C$ for all $n \in \mathbb{N}$.*
- *T is compact if and only if $c_n \to 0$. To see the suppose that $c_n \to 0$. Then consider the operator T_k given by the matrix*

$$T_k = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_k \end{pmatrix}.$$

Observe that

$$\begin{aligned} \|T - T_k\|_{\ell^2\to\ell^2} &= \sup_{x\in\ell^2\setminus\{0\}} \frac{\|(T - T_k)(x)\|_{\ell^2}}{\|x\|_{\ell^2}} \\ &= \sup_{x\in\ell^2\setminus\{0\}} \frac{\sqrt{\sum_{m=k+1}^{\infty} |c_mx_m|^2}}{\|x\|_{\ell^2}} \\ &\leq \sup_{x\in\ell^2\setminus\{0\}} \frac{\sup_{m\geq k+1} |c_m|\|x\|_{\ell^2}}{\|x\|_{\ell^2}} \\ &= \sup_{m\geq k+1} |c_m|. \end{aligned}$$

Hence, $T_k \to T$ and so by Corollary 7.2.7, the operator T is compact. For the converse assume T is compact and suppose that $c_n \not\to 0$ as $n \to \infty$. Then for some $\epsilon > 0$ there exists an extraction $\varphi(n)$ such that $\left|c_{\varphi(n)}\right| \geq \epsilon$ for all $n \in \mathbb{N}$. Let $\left(x^{(n)}\right)_{n\in\mathbb{N}}$ be the sequence where $x_i^{(n)} = \delta_{i\varphi(n)}$. It follows that $\left\|x^{(n)}\right\|_{\ell^2} = 1$ for all $n \in \mathbb{N}$ and

$$\left\|Tx^{(n)} - Tx^{(m)}\right\|_{\ell^2} \geq \sqrt{2}\epsilon$$

for all $n \neq m$. Hence, the sequence $\left(Tx^{(n)}\right)_{n\in\mathbb{N}} \subseteq T\left(\bar{B}^E\right)$ has no convergent subsequence and so $T\left(\bar{B}^E\right)$ is not pre-compact. This contradicts T being compact, therefore, we must have that $c_n \to 0$ as $n \to \infty$.

7.3 Solution to Exercises

Exercise 7.2.6

Solution. As $T\left(\bar{B}^E\right) \subseteq \mathrm{Ran}(T)$ it follows that $\dim\left(T\left(\bar{B}^E\right)\right) < \infty$. Moreover, $T\left(\bar{B}^E\right)$ is bounded as $T \in \mathcal{L}(E)$. In particular, $\overline{T\left(\bar{B}^E\right)}$ is a closed and bounded finite-dimensional set, which implies that it is compact. Therefore, $T\left(\bar{B}^E\right)$ is pre-compact, meaning that T is compact. □

8 Hilbert Spaces

Throughout let H be a real vector space.

8.1 Inner Product

Definition 8.1.1. *An inner product on H is an application $(\cdot,\cdot) : H \times H \to \mathbb{R}$ that satisfies the following.*

1. *It is bilinear. That is,*
$$(ax + by, z) = a(x, z) + b(y, z)$$
and
$$(z, ax + by) = a(z, x) + b(z, y)$$
for all $x, y, z \in H$ and $a, b \in \mathbb{R}$.
2. *It is symmetric. That is, $(x, y) = (y, x)$ for all $x, y \in H$.*
3. *It is positive definite. That is $(x, x) \geq 0$ for all $x \in H$ and $(x, x) = 0$ if and only if $x = 0$.*

Remark 8.1.2. *Elements $x, y \in H$ are orthogonal if $(x, y) = 0$.*

Lemma 8.1.3 (Cauchy-Schwartz)**.** *For $x, y \in H$ we have that*
$$|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}. \tag{8.1.1}$$

Proof. The map $t \mapsto (x+ty, x+ty)$ is a non-negative polynomial in t. Hence, its discriminant is negative. Thus,
$$(2(x, y))^2 - 4(y, y)(x, x) \leq 0,$$
which is equivalent to
$$|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}.$$
□

Remark 8.1.4. *Note that equality in* (8.1.1) *holds if and only if $x = \lambda y$ for some $\lambda \in \mathbb{R}$.*

Proposition 8.1.5. *If $(\cdot,\cdot)$ is an inner product on H, then*
$$\|x\| = \sqrt{(x, x)} \tag{8.1.2}$$
is a norm on H.

Proof. By the positive definiteness of the inner product, $\|x\| = 0$ if and only if $x = 0$. By the bilinearity of the inner product, homogeneity of the norm follows. Moreover, using the Cauchy-Schwartz inequality
$$\begin{aligned}
\|x + y\|^2 &= (x + y, x + y) \\
&= (x, x) + 2(x, y) + (y, y) \\
&= \|x\|^2 + 2(x, y) + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2 .
\end{aligned}$$
Hence, $\|\cdot\|$ is a norm on H. □

For a norm $\|\cdot\|$ given by (8.1.2) for some inner product $(\cdot,\cdot)$, the following identities hold.

- Parallelogram law,
$$\left\|\frac{u+v}{2}\right\|^2 + \left\|\frac{u-v}{2}\right\|^2 = \frac{\|u\|^2+\|v\|^2}{2}.$$
- Polarization identity,
$$(u,v) = \frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right).$$

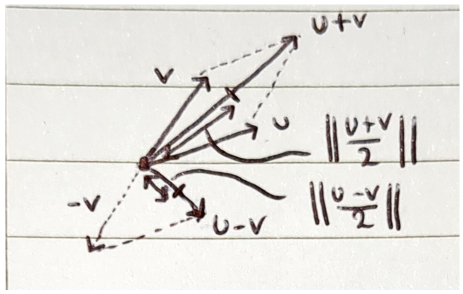


Figure 11: Parallelogram law

Definition 8.1.6. *A Hilbert space is a complete normed vector space whose norm is given by an inner product as in* (8.1.2).

Remark 8.1.7. *We only consider real Hilbert spaces, however, the theory can be extended to complex vector spaces by replacing symmetry in Definition 8.1.1 with conjugate symmetry. That is,*

- $x \mapsto (x,y)$ *for all* y *is linear, and*
- $y \mapsto (x,y)$ *for all* x *is anti-linear.*

In other words, $(x,y) = \overline{(y,x)}$.

Example 8.1.8.

1. *The space* $\mathbb{R}^d$ *with the Euclidean inner product*
$$(x,y) = \sum_{i=1}^{d} x_i y_i$$
is a real Hilbert space. Similarly, $\mathbb{C}^d$ *with inner product*
$$(x,y) = \sum_{i=1}^{d} x_i \overline{y}_i$$
is a complex Hilbert space.
2. *The space* ℓ^2 *with the inner product*
$$((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) = \sum_{n=1}^{\infty} x_n y_n$$

is a real Hilbert space.

3. *The space $L^2(\Omega)$ with the inner product*

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

is a real Hilbert space.

- *$L^p(\Omega)$ for $p \neq 2$ is not a Hilbert space.*

8.2 Projection

Theorem 8.2.1. *Let H be a Hilbert space. Let $K \subseteq H$ be a closed and convex set. Then for every $f \in H$ there exists a unique $u \in K$ such that*

$$\|f - u\| = \min_{v \in K} \|f - v\| = \operatorname{dist}(f, K). \tag{8.2.1}$$

Moreover, u is characterised by the property that $u \in K$ and

$$(f - u, v - u) \leq 0 \tag{8.2.2}$$

for all $v \in K$.

Proof. Step 1: Existence of $\min_v \|f - v\|$.
Consider a sequence $(v_n)_{n\in\mathbb{N}} \subseteq K$ such that

$$d_n := \|f - v_n\| \to d := \min_{v \in K} \|f - v\|.$$

Applying the parallelogram identity to $\|f - v_n\|$ and $\|f - v_m\|$ we deduce that

$$\left\| f - \frac{v_n + v_m}{2} \right\|^2 + \left\| \frac{v_n - v_m}{2} \right\|^2 = \frac{1}{2}\left(d_n^2 + d_m^2\right)$$

which implies that

$$\left\| \frac{v_n - v_m}{2} \right\|^2 \leq \frac{1}{2}\left(d_n^2 + d_m^2\right) - d^2 \overset{n,m\to\infty}{\longrightarrow} 0.$$

Hence $(v_n)_{n\in\mathbb{N}}$ is Cauchy, which implies that it is convergent to some $u \in H$. Passing to the limit we conclude that

$$\|f - u\| = \min_{v \in K} \|f - v\|.$$

Step 2: Equivalence of the characterisations.
Assume that u satisfies (8.2.1) and consider a $v \in K$. By the convexity of K it follows that

$$(1 - t)u + tv \in K$$

for all $t \in [0, 1]$. Therefore,

$$\|f - ((1 - t)u + tv)\|^2 \geq \|f - u\|^2.$$

The left-hand side is polynomial in t and can be expanded as

$$\|f - u\|^2 - 2t(f - u, v - u) + O\left(t^2\right).$$

As $t \to 0$, the assumption of (8.2.1) can only hold if $(f - u, v - u) \leq 0$. Conversely, suppose that (8.2.2) holds, then for all $v \in K$ it follows that

$$\|u - f\|^2 - \|v - f\|^2 = 2(f - u, v - u) - \|u - v\|^2 \leq 0$$

which implies that $\|u - f\| \leq \|v - f\|$ for all $v \in K$.
Step 3: Uniqueness.
Suppose u_1 and u_2 satisfy (8.2.2), then

1. $(f - u_1, v - u_1) \leq 0$ for all $v \in K$, and
2. $(f - u_2, v - u_2) \leq 0$ for all $v \in K$.

Choosing $v = u_2$ and $v = u_1$ in the first and second conditions respectively it follows that

1. $(f - u_1, u_2 - u_1) \leq 0$, and
2. $(f - u_2, u_1 - u_2) \leq 0$.

Adding these together it follows that $\|u_1 - u_2\|^2 \leq 0$ which implies that $u_1 = u_2$. □

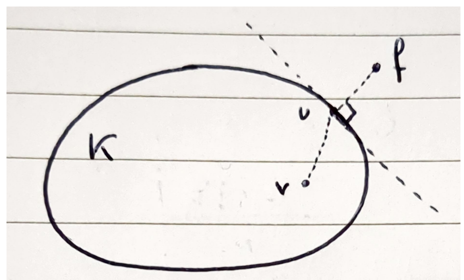


Figure 12: An illustration of the condition stated in (8.2.2).

Proposition 8.2.2. *An alternative characterisation of u in Theorem 8.2.1 when K is additionally a linear subspace of H, is $u \in K$ and*

$$(f - u, v) = 0 \tag{8.2.3}$$

for all $v \in K$.

Proof. Suppose that $u \in K$ satisfies (8.2.3). Then for $v \in K$ we have $u - v \in K$ so that

$$\begin{aligned}
\|f - v\|^2 &= \|f - u + u - v\|^2 \\
&= \|f - u\|^2 + 2(f - u, u - v) + \|u - v\|^2 \\
&\overset{(8.2.3)}{=} \|f - u\|^2 + \|u - v\|^2.
\end{aligned}$$

In particular, this implies that $\|f - v\|^2 \geq \|f - u\|^2$. Conversely, suppose that (8.2.1) is satisfied for $u \in K$. Then for $v \in K$ and $t \in \mathbb{R}$, as K is a linear subspace of H, we have that $u + tv \in K$ and so $\|f - u\|^2 \leq \|f - (u + tv)\|^2$. Consider,

$$0 \leq \|f - (u + tv)\|^2 - \|f - u\|^2 = 2t(u - f, v) + t^2\|v\|^2 =: g(t).$$

If $(u - f, v) \neq 0$, then $g(t)$ is minimised at $t = -\frac{(u-f,v)}{\|v\|^2}$, giving a minimum value

$$\begin{aligned}
g\left(-\frac{(u - f, v)}{\|v\|^2}\right) &= -2\frac{(u - f, v)^2}{\|v\|^2} + \frac{(f - u, v)^2}{\|v\|^2} \\
&= -\frac{(u - f, v)^2}{\|v\|^2}
\end{aligned}$$

which is strictly negative as we are assuming $(u - f, v) \neq 0$. This is a contradiction and so it must be the case that $(f - u, v) = 0$. □

Remark 8.2.3.

1. *Suppose that M is a closed linear subspace. Then $P : H \to M$ given by $f \mapsto u$, as in Theorem 8.2.1, is a linear operator. It is characterised by the property that $Pf \in M$ and*

$$\|f - Pf\| = \min_{v \in M} \|f - v\|.$$

Equivalently, it can be characterised by the property that $Pf \in M$ and

$$(f - Pf, v) = 0$$

for all $v \in M$. In particular, $(f - Pf, Pf) = 0$, and so we recover a Pythagoras type relation

$$\|f\|^2 = \|f - Pf\|^2 + \|Pf\|^2.$$

2. *Convexity is necessary for the uniqueness statement of Theorem 8.2.1. Consider $H = \mathbb{R}^2$, $f = (0,0)$ and K the annulus with centre $(0,0)$. Although the distance from f to K is well-defined, the projection of f to K is not unique.*

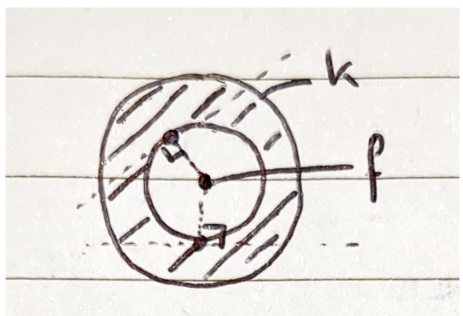


Figure 13: A non-convex set that does not satisfy the uniqueness statement of Theorem 8.2.1. Note that the angle between v and f is obtuse.

For a linear subspace F of a Hilbert space H, the orthogonal complement of F in H is the set

$$F^\perp = \{y \in H : (y, x) = 0 \text{ for all } x \in F\}.$$

Proposition 8.2.4. *Let F be a closed subspace of a Hilbert space H. Then $H = F \oplus F^\perp$. In particular, for $v \in H$ we have that $v = Pv + P^\perp v$, where Pv is the projection of v onto F, and $P^\perp v$ is the projection of v onto $F^\perp$.*

Proof. Suppose that $y \in F \cap F^\perp$ then $(y, x) = 0$ for all $x \in F$. In particular, $(y, y) = 0$ which implies that $y = 0$, hence, $F \cap F^\perp \subseteq \{0\}$. As F and $F^\perp$ are linear subspaces we have $0 \in F \cap F^\perp$ and so $\{0\} \subseteq F \cap F^\perp$, meaning $F \cap F^\perp = \{0\}$. Now let $v \in H$. Then for $\tilde{v} \in F^\perp$ we have that

$$\begin{aligned}(v - (v - Pv), \tilde{v} - (v - Pv)) &= (Pv, \tilde{v} - v + Pv) \\ &= (Pv, Pv - v) \\ &= (v - Pv, 0 - Pv).\end{aligned}$$

As $0 \in F$ we use the fact that Pv is the projection of v onto F to note that $(v - Pv, 0 - Pv) \le 0$. Hence, $(v - (v - Pv), \tilde{v} - (v - Pv)) \le 0$ which implies that $v - Pv = P^\perp v$ and so $v = Pv + P^\perp v$. $\square$

Corollary 8.2.5. *Let F be a closed subspace of a Hilbert space H. Then for $v \in H$ it follows that*

$$\|v\|_H^2 = \|Pv\|_H^2 + \left\|P^{\perp}v\right\|_H^2,$$

where Pv is the projection of v onto F and $P^{\perp}v$ is the projection of v onto $F^{\perp}$.

Proof. Note that $\left(Pv, P^{\perp}v\right) = 0$ as $Pv \in F$ and $P^{\perp}v \in F^{\perp}$. Hence,

$$\begin{aligned}
\|v\|_H^2 &= (v, v) \\
&= \left(Pv + P^{\perp}v, Pv + P^{\perp}v\right) \\
&= (Pv, Pv) + 2\left(Pv, P^{\perp}v\right) + \left(P^{\perp}v, P^{\perp}v\right) \\
&= \|Pv\|_H^2 + \left\|Pv^{\perp}\right\|_H^2.
\end{aligned}$$

□

Corollary 8.2.6. *For every closed and non-empty subspace F of a Hilbert space H, there exists a unique linear map $\pi : H \to F$ such that*

1. $\|\pi\|_{H\to H} = 1$,
2. $\pi^2 = \pi$, *and*
3. $\ker(\pi) = F^{\perp}$.

Proof. For $v \in H$, let $\pi(v) = Pv$.

1. using Corollary 8.2.5 it is clear that $\|v\|_H \geq \|\pi(v)\|_H$. Hence,

$$\|\pi\|_{H\to H} = \sup_{v\in H\setminus\{0\}} \frac{\|\pi(v)\|_H}{\|v\|_H} \leq 1.$$

However, as for $v \in F \setminus \{0\}$ we have $\|v\|_H = \|\pi(v)\|_H$ it follows that $\|\pi\|_{H\to H} = 1$.

2. As $Pv \in F$ it is clear that $P(Pv) = Pv$ and so $\pi^2 = \pi$.
3. If $v \in F^{\perp}$, then $\pi(v) = 0$ and so $v \in \ker(\pi)$. On the other hand, if $\pi(v) = 0$, then $v \in F^{\perp}$ by Proposition 8.2.4.

□

Exercise 8.2.7. *Let F and G be linear subspaces of a Hilbert space H. Prove the following statements.*

1. $H^{\perp} = \{0\}$ *and* $\{0\}^{\perp} = H$.
2. $F^{\perp}$ *is a closed linear subspace of H.*
3. *If $F \subseteq G$ then $G^{\perp} \subseteq F^{\perp}$.*
4. $\left(F^{\perp}\right)^{\perp} = \bar{F}$.
5. *If F and G are closed, show that the following hold.*
   (a) $F \cap G = \left(F^{\perp} + G^{\perp}\right)^{\perp}$.
   (b) $F^{\perp} \cap G^{\perp} = (F + G)^{\perp}$.
   (c) $(F \cap G)^{\perp} = \overline{F^{\perp} + G^{\perp}}$.

(d) $\left(F^{\perp} \cap G^{\perp}\right)^{\perp} = \overline{F + G}$.

Example 8.2.8. *Consider*

$$E = \left\{g \in L^2(0,1) : g \geq 0 \text{ almost everywhere}\right\}.$$

Then for $g_1, g_2 \in E$ and $t \in [0,1]$ we have

$$t g_1(x) + (1-t) g_2(x) \geq 0$$

almost everywhere. Moreover, for $(g_n)_{n\in\mathbb{N}} \subseteq E$ converging to g, there exists a subsequence $(g_{n_k})_{k\in\mathbb{N}}$ which converges pointwise almost everywhere to g. Therefore, $g(x) \geq 0$ almost everywhere as $g_{n_k}(x) \geq 0$ almost everywhere for each $k \in \mathbb{N}$. By Theorem 8.2.1, for $f \in L^2(0,1)$ there exists a unique projection onto E. More specifically, considering $\tilde{f} \in E$ given by

$$\tilde{f}(x) = \begin{cases} f(x) & f(x) \geq 0 \\ 0 & f(x) < 0, \end{cases}$$

we note that for $g \in E$ we have

$$\begin{aligned} (\tilde{f} - f, \tilde{f} - g) &= \left(\int_{\{x : f(x) \geq 0\}} + \int_{\{x : f(x) < 0\}}\right) (\tilde{f} - f, \tilde{f} - g) \\ &= \int_{\{x : f(x) < 0\}} fg \\ &\leq 0, \end{aligned}$$

where the inequality follows as $f(x) < 0$ and $g(x) > 0$ on the specified domain. Therefore, by the uniqueness of the projection we deduce that $Pf = \tilde{f}$. Furthermore, by Proposition 8.2.4 we deduce that

$$P^{\perp} f(x) = \begin{cases} 0 & f(x) \geq 0 \\ f(x) & f(x) < 0. \end{cases}$$

8.3 The Dual Space

Observe that for any $u \in H$, the map $\varphi_u : H \to \mathbb{R}$ given by $v \mapsto (u, v)$ is in the dual space of H, denoted H^*. Moreover, using the Cauchy-Schwartz inequality we can show that the map $H \to H^*$ given by $u \mapsto \varphi_u$ is an isometry. If $\dim(H) < \infty$, then it follows by arguments involving linear algebra, that any element of H^* is of the form φ_u for some $u \in H$.

Theorem 8.3.1 (Riesz-Frechet Representation Theorem)**.** *For any $\varphi \in H^*$, there exists a $u \in H$ such that $\varphi = \varphi_u$ and $\|\varphi\|_{H^*} = \|u\|_H$.*

Proof. For $\varphi \in H^*$, let $M = \varphi^{-1}(\{0\})$. By the continuity of φ we know that M is a closed subspace. If $\varphi = 0$ then $M = H$, so we assume instead that there exists a $g_0 \in H \setminus M$. Let P_M be the projection on M, and let $g_1 = P_M g_0$ and $g = \frac{g_0 - g_1}{\|g_0 - g_1\|}$. Then g is such that $\|g\| = 1$ and $(g, v) = 0$ for all $v \in M$. In particular, this means that $g \notin M$ which implies that $\varphi(g) \neq 0$. For $u \in H$ we have $\varphi(u - \lambda g) = 0$ for $\lambda = \frac{\varphi(u)}{\varphi(g)}$. Thus, $(g, u - \lambda g) = 0$ which implies that

$$(g, u) = \frac{\varphi(u)}{\varphi(g)},$$

so that $\varphi(u) = \varphi(g)(g, u)$. Therefore, $\varphi = \varphi_{\varphi(g)g}$. $\square$

Remark 8.3.2. *As $u \mapsto \varphi_u$ is an isometry it is injective. As Theorem 8.3.1 shows that $u \mapsto \varphi_u$ is surjective, we have that $H = H^*$ for H a Hilbert space. As $(L^p)' = L^{p'}$ by Theorem 6.4.2, it follows that L^p can only be a Hilbert space if $p = p'$, which is only true for $p = 2$.*

Theorem 8.3.3 (Lax-Milgram)**.** *Let H be a real Hilbert space. Assume $a : H \times H \to \mathbb{R}$ is such that the following hold.*

1. *It is bilinear, that is $a(x, \cdot)$ and $a(\cdot, y)$ are linear for all $x, y \in H$.*
2. *It is continuous, that is $|a(x, y)| \leq C\|x\|\|y\|$ for all $x, y \in H$.*
3. *It is coercive, that is $|a(x, x)| \geq c\|x\|^2$ for all $x \in H$.*

Then for $f \in H$ there exists a unique u such that

$$a(u, v) = \langle f, v \rangle$$

for all $v \in H$.

Proof. Step 1: The linear operator associated with a.
For fixed u, we look at $v \mapsto a(u, v) \in H^*$. By Theorem 8.3.1 there exists $A(u) \in H$ such that

$$a(u, v) = \langle A(u), v \rangle$$

for every $v \in H$. Observe that $A : H \to H$ is linear. Moreover, A is bounded as

$$|\langle A(u), v \rangle| = |a(u, v)| \leq C\|u\|\|v\|$$

and so continuous. Furthermore, A is non-degenerate as

$$\|u\|\|Au\| \geq \langle Au, u \rangle = a(u, u) \geq c\|u\|^2$$

and so $\|Au\| \geq c\|u\|$.
Step 2: Solving $Au = f$.

1. A is injective as $\|Au\| \geq c\|u\|$.
2. Let $(g_n)_{n\in\mathbb{N}} \subseteq \mathrm{Ran}(A)$ such that $g_n \to g$ in H. We know that there exists a $u_n \in H$ such that $A(u_n) = g_n$. In particular, $A(u_n - u_m) = g_n - g_m$. Hence, by coercivity it follows that

   $$\|u_n - u_m\| \leq \frac{1}{c}\|g_n - g_m\|.$$

   Therefore, as $(g_n)_{n\in\mathbb{N}}$ converges it is Cauchy and so $(u_n)_{n\in\mathbb{N}} \subseteq H$ is Cauchy. Using completeness it follows that $u_n \to u$ in H. Passing to the limit we deduce that $A(u_n) = g_n \to g = A(u)$ where $A(u) \in \mathrm{Ran}(A)$. Thus we conclude that $\mathrm{Ran}(A)$ is closed.
3. Suppose that $\mathrm{Ran}(A)$ is not dense. Then its orthogonal complement is non-zero. That is, there exists a $v \neq 0$ such that $\langle A(u), v \rangle = 0$ for all $u \in H$. In particular, choosing $u = v$ we obtain

   $$0 = \langle Av, v \rangle \geq c\|v\|^2$$

   which is a contradiction. Therefore, $\mathrm{Ran}(A)$ is dense.

Using statements 2 and 3 it follows that $\mathrm{Ran}(A) = H$, meaning A is surjective. Combining this with statement 1 we deduce that A is bijective and so a unique solution $u \in H$ to $A(u) = f$ exists. $\square$

Remark 8.3.4.

1. *Note that $\langle f, u\rangle = \varphi(u)$ for some $\varphi \in H^*$. So taking $a(u, v) = \langle u, v\rangle$ the problem solved by Theorem 8.3.3 is equivalent to the problem solved by Theorem 8.3.1. Hence, one can view Theorem 8.3.3 as an extension of Theorem 8.3.1.*

2. *Note that a is not symmetric and so in general not an inner product.*

Theorem 8.3.3 has applications in partial differential equations. For a domain $\Omega \subseteq \mathbb{R}^d$ and $f \in \mathcal{C}^0_\infty$, the Dirichlet problem is to solve

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

Taking the inner product of the first equation with $\varphi \in \mathcal{C}_c^\infty(\Omega)$ yields

$$-\int_\Omega (\Delta u) \cdot \varphi \, \mathrm{d}x = \int_\Omega f \cdot \varphi \, \mathrm{d}x.$$

Integrating by parts gives

$$\int_\Omega \nabla u \cdot \nabla \varphi \, \mathrm{d}x = \int_\Omega f \cdot \varphi \, \mathrm{d}x \tag{8.3.1}$$

as φ vanishes on $\partial\Omega$. Note that the right-hand of (8.3.1) is the inner product of f and φ on $L^2(\Omega)$ and the left-hand side is of the form $a(u, \varphi)$. The idea now is to use Theorem 8.3.3 to solve the Dirichlet problem. To do this H needs to be chosen such that a satisfies the conditions of Theorem 8.3.3.

8.4 Hilbert Sums and Orthonormal Bases

If H is a finite-dimensional Hilbert space, there exists a bases $(e_n)_{n=1}^d \subseteq H$ such that for any $x \in H$ we can write

$$x = \sum_{n=1}^d x_n e_n$$

for some $x_n \in \mathbb{R}$. In particular, if $(e_n)_{n=1}^d$ is an orthonormal basis it follows that

$$\|x\|^2 = \sum_{n=1}^d \|x_n\|^2. \tag{8.4.1}$$

We would like to generalise the idea of a basis for infinite dimensional Hilbert spaces. Using the relation (8.4.1), which holds for orthonormal bases, this generalisation amounts to understanding the convergence of sums.

Definition 8.4.1. *Let $(E_n)_{n\in\mathbb{N}}$ be a sequence of closed subspaces of a Hilbert space H. Then H is a Hilbert sum of the $(E_n)_{n\in\mathbb{N}}$, written $H = \bigoplus_{n=1}^\infty E_n$, if the following hold.*

1. *The E_n are mutually orthogonal. Namely, $(x, y) = 0$ if $x \in E_n$ and $y \in E_m$ for $n \neq m$.*

2. *The subspace $\operatorname{span}\left(\bigcup_{n=1}^\infty E_n\right)$ is dense in H.*

Remark 8.4.2. *The span of a set of vectors refers to all finite linear combinations of the vectors.*

Lemma 8.4.3. *Let $(v_n)_{n\in\mathbb{N}} \subseteq H$ be such that $(v_n, v_m) = 0$ for $n \neq m$ and $\sum_{n=1}^\infty \|v_n\|^2 < \infty$. Then*

$S_n = \sum_{k=1}^{n} v_k$ converges, to S say. Furthermore,

$$\|S\|^2 = \sum_{k=1}^{\infty} \|v_k\|^2.$$

Proof. For $n < m$, using (8.4.1) we have that

$$\|S_n - S_m\|^2 = \sum_{k=n+1}^{m} \|v_k\|^2. \tag{8.4.2}$$

Since, $\sum_{k=1}^{\infty} \|v_k\|^2 < \infty$, using (8.4.2) it follows that $(S_n)_{n\in\mathbb{N}} \subseteq H$ is Cauchy. Therefore, by completeness $(S_n)_{n\in\mathbb{N}}$ has a limit, say S. Furthermore, using (8.4.1) we know that $\|S_n\|^2 = \sum_{k=1}^{n} \|v_k\|^2$ and so passing to the limit we deduce that

$$\|S\|^2 = \sum_{k=1}^{\infty} \|v_k\|^2.$$

□

Theorem 8.4.4. *Assume that $H = \bigoplus_{n=1}^{\infty} E_n$ is a Hilbert sum of the closed subspaces $(E_n)_{n\in\mathbb{N}}$. For $u \in H$, let $u_n = P_{E_n} u$ and $S_n = \sum_{k=1}^{n} u_k$. Then $S_n \to u$ as $n \to \infty$ and*

$$\sum_{n=1}^{\infty} \|u_n\|^2 = \|u\|^2. \tag{8.4.3}$$

Proof. Step 1: Show that the limit exists.
On the one hand,

$$\|S_n\|^2 = \sum_{k=1}^{n} \|u_k\|^2$$

using (8.4.1). On the other hand, as $u_n = P_{E_n} u$ we have that

$$(u, u_n) = \|u_n\|^2$$

which implies that $(u, S_n) = \sum_{k=1}^{n} \|u_k\|^2$ using the orthogonality of the $E_1, \ldots, E_n$. Therefore, using the Cauchy-Schwartz inequality it follows that

$$\|S_n\|^2 = (u, S_n) \leq \|u\|\|S_n\|,$$

which implies that

$$\left(\sum_{k=1}^{n} \|u_k\|^2 \right)^{\frac{1}{2}} = \|S_n\| \leq \|u\|.$$

Passing to the limit it follows that

$$\sum_{k=1}^{\infty} \|u_k\|^2 \leq \|u\|^2 < \infty.$$

Hence, the conditions of Lemma 8.4.3 are satisfied and thus we deduce that S_n converges to S with

$$\|S\|^2 = \sum_{k=1}^{\infty} \|u_k\|^2.$$

Step 2: Identification of the limit.
Note that $(u - S_n, v) = 0$ for all $v \in E_m$ where $m \leq n$, by the characterisation of the projection. Letting $n \to \infty$ it follows that $(u - S, v) = 0$ for all $v \in E_m$ where $m \in \mathbb{N}$. By linearity it follows that $(u - S, v) = 0$ for all $v \in \operatorname{span}(\bigcup_{m\in\mathbb{N}} E_m)$. Moreover, by the density of $\operatorname{span}\left(\bigcup_{m\in\mathbb{N}} E_m\right)$ it follows that $(u - S, v) = 0$ for all $v \in H$. Therefore, $u = S$. □

Remark 8.4.5.

1. *Equation* (8.4.3) *is often referred to as the Bessel-Parseval identity.*
2. *The vector S_n in Theorem 8.4.4 is the projection of u onto $\operatorname{span}\left(\bigcup_{k=1}^n E_n\right)$ and so the convergence $S_n \to u$ is expected from statement 2 of Definition 8.4.1. Moreover,* (8.4.1) *is reasonable due to the orthogonality assumptions we impose on the $(E_n)_{n\in\mathbb{N}}$.*
3. *Henceforth, we write $\sum_{n=1}^{\infty} u_n = u$ to mean $\lim_{n\to\infty} S_n = u$.*

Definition 8.4.6. *A sequence $(e_n)_{n\in\mathbb{N}} \subseteq H$ is an orthonormal basis if the following hold.*

1. $(e_n, e_m) = \delta_{nm}$.
2. $\overline{\operatorname{span}((e_n)_{n\in\mathbb{N}})} = H$.

Remark 8.4.7. *An orthonormal basis of a Hilbert is sometimes referred to as a Hilbert basis.*

Exercise 8.4.8. *Let H be a Hilbert space and let $V := \operatorname{span}(v)$ for $v \in H \setminus \{0\}$. Show that V is a closed linear subspace of H. Moreover, for $u \in H$ show that $P_V u = \frac{(u,v)}{\|v\|^2} v$.*

Corollary 8.4.9. *If $(e_n)_{n\in\mathbb{N}} \subseteq H$ is an orthonormal basis, then for all $u \in H$ we have*

$$u = \sum_{n=1}^{\infty} (u, e_n) e_n$$

and

$$\|u\|^2 = \sum_{n=1}^{\infty} |(u, e_n)|^2 .$$

Proof. Consider the subspaces $(E_n)_{n\in\mathbb{N}}$ of H given by $E_n = \operatorname{span}(e_n)$. By Exercise 8.4.8 the subspace E_n is closed and $u_n := P_{E_n} u = (u, e_n)e_n$. Moreover, if $x \in E_n$ and $y \in E_m$, for $n \neq m$, then $x = \lambda e_n$ and $y = \mu e_m$. Using the orthogonality of $(e_n)_{n\in\mathbb{N}}$ it follows that that

$$\langle x, y\rangle = \lambda\mu\langle e_n, e_m\rangle = 0.$$

Similarly, as $(e_n)_{n\in\mathbb{N}} \subseteq \bigcup_{n\in\mathbb{N}} E_n$ we have that

$$H = \overline{\operatorname{span}\left((e_n)_{n\in\mathbb{N}}\right)} \subseteq \overline{\operatorname{span}\left(\bigcup_{n\in\mathbb{N}} E_n\right)} \subseteq H,$$

which implies that $\overline{\operatorname{span}\left(\bigcup_{n\in\mathbb{N}} E_n\right)} = H$. Therefore, we can apply Theorem 8.4.4 to conclude that

$$u = \sum_{n=1}^{\infty} (u, e_n) e_n$$

and

$$\|u\|^2 = \sum_{n=1}^{\infty} \|(u, e_n)e_n\|^2 = \sum_{n=1}^{\infty} |(u, e_n)|^2.$$

□

Definition 8.4.10. *A Hilbert space H is separable if it admits a countably dense subset.*

Theorem 8.4.11. *A Hilbert space H is separable if and only if H has an orthonormal basis.*

Proof. ($\Leftarrow$). Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of H and consider the subset

$$F = \left\{ \sum_{k=1}^{n} r_k e_k : r_k \in \mathbb{Q},\, n \in \mathbb{N} \right\} \subseteq H.$$

Let $u \in H$ and $\epsilon > 0$. By Corollary 8.4.9 we know that $u = \sum_{k=1}^{\infty}(u, e_k)e_k$ and

$$\sum_{k=1}^{\infty} |(u, e_k)|^2 = \|u\|^2 < \infty.$$

Hence, we can find an $N \in \mathbb{N}$ such that

$$\sum_{k=N+1}^{\infty} |(u, e_k)|^2 < \frac{\epsilon}{2}.$$

Moreover, for $k \leq N$ we can find $r_k \in \mathbb{Q}$ such that $|(u, e_n) - r_k|^2 < \frac{\epsilon}{2N}$. Let

$$\tilde{u} = \sum_{k=1}^{N} r_k e_k \in F,$$

then

$$\begin{aligned}
\|u - \tilde{u}\|^2 &= \left\| \sum_{k=1}^{\infty}(u, e_k)e_k - \sum_{k=1}^{N} r_k e_k \right\|^2 \\
&= \left\| \sum_{k=1}^{N} \left((u, e_k) - r_k\right) e_k + \sum_{k=N+1}^{\infty} (u, e_k)e_k \right\|^2 \\
&\overset{\text{Cor 8.4.9}}{=} \sum_{k=1}^{N} |(u, e_k) - r_k|^2 + \sum_{k=N+1}^{\infty} |(u, e_k)|^2 \\
&< \sum_{k=1}^{N} \frac{\epsilon}{2N} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}$$

Therefore, F is a countable dense subset of H.
($\Rightarrow$). Let $(u_n)_{n\in\mathbb{N}} \subseteq H$ be a countably dense subset. Construct the sequence $(e_n)_{n\in\mathbb{N}}$ in the following way.

1. $E_1 := \operatorname{span}(u_1)$, and let $e_1 = \frac{u_1}{\|u_1\|}$.
2. $E_2 := \operatorname{span}(u_1, u_2)$ and choose e_2 such that $\{e_1, e_2\}$ is an orthonormal basis for E_2.
   - Note that we assume that u_1 and u_2 are not aligned. We label the subset $(u_n)_{n\in\mathbb{N}}$ in this way as the subset is countably dense.
3. For general $k \in \mathbb{N}$, let $E_k := \operatorname{span}(u_1, \ldots, u_k)$ and choose e_k such that $\{e_1, \ldots, e_k\}$ is an orthonormal basis for E_k.
   - Again we can assume that the $u_1, \ldots, u_k$ are not aligned by the fact that $(u_n)_{n\in\mathbb{N}}$ is countably dense.

The sequence $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of H. $\square$

Remark 8.4.12. *Let H and H' be separable real Hilbert spaces. Then orthonormal bases $(e_n)_{n\in\mathbb{N}} \subseteq H$ and $(e'_n)_{n\in\mathbb{N}} \subseteq H'$ exist. Hence, we can consider the map $J : H \to H'$ given by*

$$\sum_{n=1}^{\infty} x_n e_n \mapsto \sum_{n=1}^{\infty} x_n e'_n.$$

This is an isometric isomorphism. In particular, fix $H = \ell^2$ and consider the orthonormal basis $(e_n)_{n\in\mathbb{N}}$ where

$$e_n = (\underbrace{0, \dots, 1}_{n}, 0, \dots).$$

Then the above arguments imply that any separable real Hilbert space has the same structure as ℓ^2. One may think then that we can characterise all properties of general Hilbert spaces by investigating ℓ^2. After all the isometric isomorphism captures all the structural information regarding the inner product and norm. However, certain interesting Hilbert spaces have additional structures that are not captured within this isometric isomorphism.

Example 8.4.13. *Let $H = L^2(0, 2\pi)$ be a complex Hilbert space and consider $e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$ for $n \in \mathbb{N}$. Then*

$$(e_n, e_m) = \frac{1}{2\pi} \int_0^{2\pi} \overline{e^{inx} e^{imx}} \, \mathrm{d}x = \delta_{nm}.$$

With additional computations one can show that $\overline{\operatorname{span}((e_n)_{n\in\mathbb{N}})} = H$. With this it follows that $(e_n)_{n\in\mathbb{N}} \subseteq H$ is an orthonormal basis of H.

8.5 Linear Operators

8.5.1 Adjoint Operators

Consider the finite-dimensional real Hilbert space $H = \mathbb{R}^d$. Let $x, y \in \mathbb{R}^d$ and $M \in \mathbb{R}^{d\times d}$. Then

$$\langle Mx, y \rangle = \langle x, M^\top y \rangle .$$

For $H = L^2\left(\mathbb{R}^d\right)$ consider

$$(Lu)(x) = \int K(x, y) u(y) \, \mathrm{d}y,$$

where $K(x, y)$ is sufficiently smooth and decays fast enough such that the map $u \mapsto Lu$ is well-defined. Then under sufficient assumptions, we can write

$$\begin{aligned} \langle Lu, v \rangle &= \int \left(\int K(x, y) u(y) \, \mathrm{d}y \right) u(x) \, \mathrm{d}x \\ &= \int \left(\int K(x, y) u(x) \, \mathrm{d}x \right) u(y) \, \mathrm{d}y \\ &= \langle u, L^* v \rangle , \end{aligned}$$

where

$$(L^* u)(x) = \int K(y, x) u(y) \, \mathrm{d}y.$$

Proposition 8.5.1. *Let H be a real Hilbert space and consider $T \in \mathcal{L}(H)$. Then there exists a unique $T^* \in \mathcal{L}(H)$ such that*

$$\langle Tx, y \rangle = \langle x, T^* y \rangle$$

for all $x, y \in H$ with $\|T\|_{\mathcal{L}(H)} = \|T^*\|_{\mathcal{L}(H)}$

Proof. For fixed $y \in H$ let $\varphi_y : H \to \mathbb{R}$ be given by $x \mapsto \langle Tx, y\rangle$. Note that $\varphi_y \in H^*$, and so by Theorem 8.3.1 there exists a $u_y \in H$ such that $\langle Tx, y\rangle = \langle x, u_y\rangle$ for all $x \in H$ with $\|\varphi_y\|_{H^*} = \|u_y\|_H$. Let $T^* : H \to H$ be given by $y \mapsto u_y$, then

$$\langle Tx, y\rangle = \langle x, u_y\rangle = \langle x, T^*y\rangle$$

for all $x, y \in H$. For $y_1, y_2 \in H$, $\lambda \in \mathbb{R}$ and any $x \in H$ we have that

$$\begin{aligned}
\langle x, T^*(y_1 + \lambda y_2)\rangle &= \langle Tx, y_1 + \lambda y_2\rangle \\
&= \langle Tx, y_1\rangle + \lambda\langle Tx, y_2\rangle \\
&= \langle x, \lambda T^* y_1\rangle + \langle x, \lambda T^* y_2\rangle \\
&= \langle x, T^* y_1 + \lambda T^* y_2\rangle.
\end{aligned}$$

As this holds for all $x \in H$ it follows that $T^*(y_1 + y_2) = T^*y_1 + \lambda T^* y_2$ meaning the operator T^* is linear. Recall $\|\varphi_y\|_{H^*} = \|T^*y\|_H$, where

$$\begin{aligned}
\|\varphi_y\|_{H^*} &= \sup_{x \in H\setminus\{0\}} \frac{|\langle Tx, y\rangle|}{\|x\|_H} \\
&\leq \sup_{x \in H\setminus\{0\}} \frac{\|Tx\|_H \|y\|_H}{\|x\|_H}
\end{aligned}$$

Hence,

$$\begin{aligned}
\|T^*\|_{\mathcal{L}(H)} &= \sup_{y \in H\setminus\{0\}} \frac{\|T^*y\|_H}{\|y\|_H} \\
&= \sup_{y \in H\setminus\{0\}} \frac{\|\varphi_y\|_{H^*}}{\|y\|_H} \\
&\leq \sup_{y \in H\setminus\{0\}} \sup_{x \in H\setminus\{0\}} \frac{\|Tx\|_H}{\|x\|_H} \\
&= \sup_{x \in H\setminus\{0\}} \frac{\|Tx\|_H}{\|x\|_H} \\
&= \|T\|_{\mathcal{L}(H)}.
\end{aligned}$$

As $T \in \mathcal{L}(H)$ it is bounded and so as T^* is linear it follows that $T^* \in \mathcal{L}(H)$. Through similar computations one deduces that $\|T\|_{\mathcal{L}(H)} \leq \|T^*\|_{\mathcal{L}(H)}$ to conclude that $\|T\|_{\mathcal{L}(H)} = \|T^*\|_{\mathcal{L}(H)}$. ☐

Remark 8.5.2. *The operator T^* of Proposition 8.5.1 is known as the adjoint of T.*

Definition 8.5.3. *An operator T is self-adjoint if $T^* = T$.*

From our previous discussions, it follows that operators in finite-dimensional real Hilbert spaces are self-adjoint if the corresponding matrices are symmetric. Similarly, a kernel operator of the form

$$(Tf)(x) = \int K(x, y) f(y)\,\mathrm{d}y$$

is self-adjoint if $K(x, y) = K(y, x)$.

Theorem 8.5.4 (Schauder)**.** *An operator $T \in \mathcal{L}(H)$ is compact if and only if its adjoint, T^*, is compact.*

Proposition 8.5.5. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator. Then*

$$\|T\|_{\mathcal{L}(H)} = \sup_{\|x\|=1} |\langle Tx, x\rangle| .$$

Proof. Let $M = \sup_{\|x\|=1} |\langle Tx, x\rangle|$. Then by the Cauchy-Schwartz inequality it follows that

$$M \leq \sup_{\|x\|=1} \|Tx\|_H \|x\|_H = \sup_{\|x\|=1} \|Tx\|_H = \|T\|_{\mathcal{L}(H)}.$$

Now consider $x, y \in H$ with $\|x\| = \|y\| = 1$. Using the self-adjoint property of T note that

$$\begin{aligned}\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle &= (\langle Tx, x\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle + \langle Ty, y\rangle) \\ &\quad - (\langle Tx, x\rangle - \langle Tx, y\rangle - \langle Ty, x\rangle + \langle Ty, y\rangle) \\ &= 4\langle Tx, y\rangle.\end{aligned}$$

Therefore,

$$\begin{aligned}|\langle Tx, y\rangle| &\leq \frac{M\|x+y\|^2 + M\|x-y\|^2}{4} \\ &= \frac{M}{4}\left(\|x+y\|^2 + \|x-y\|^2\right) \\ &\stackrel{(1)}{=} \frac{M}{2}\left(\|x\|^2 + \|y\|^2\right) \\ &= M,\end{aligned}$$

where (1) in application of the parallelogram law. Setting $y = \frac{Tx}{\|Tx\|}$ it follows that $\|Tx\| \leq M$, and so

$$\|T\|_{\mathcal{L}(H)} = \sup_{\|x\|=1} \|Tx\| \leq M.$$

Therefore, we conclude that

$$\|T\|_{\mathcal{L}(H)} = M = \sup_{\|x\|=1} |\langle Tx, x\rangle|.$$

□

8.5.2 Fredholm's Theory

Fredholm's theory aims to solve problems of the form

$$f(x) - \int T(x,y) f(y) \, \mathrm{d}y = h(x),$$

where f is unknown, h is given and T is an operator. The term h is often referred to as the inhomogeneous component of the problem. To make progress we focus on the case when T is compact and reduce the problem to one of the form

$$(\mathrm{Id} - T) f = h. \tag{8.5.1}$$

Finding solutions to (8.5.1) is equivalent to determining $\mathrm{Ran}(\mathrm{Id} - T)$.

Theorem 8.5.6 (Fredholm's Alternative)**.** *Let $T \in \mathcal{K}(H)$. Then the following hold.*

1. $\ker(\mathrm{Id} - T)$ *is finite-dimensional.*
2. $\mathrm{Ran}(\mathrm{Id} - T)$ *is closed, in particular,* $\mathrm{Ran}(\mathrm{Id} - T) = \ker\left(\mathrm{Id} - T^*\right)^{\perp}$.
3. $\ker(\mathrm{Id} - T) = \{0\}$ *if and only if* $\mathrm{Ran}(\mathrm{Id} - T) = H$.
4. $\dim\left(\ker(\mathrm{Id} - T)\right) = \dim\left(\ker\left(\mathrm{Id} - T^*\right)\right)$.

Proof.

1. Let $E = \ker(\mathrm{Id} - T) = \{z : Tz = z\}$. Note that E is closed by the continuity of T. Furthermore,

$$\overline{T\left(\overline{B_H}\right)} \overset{(1)}{\supseteq} \overline{T\left(\overline{B_E}\right)} \overset{(2)}{=} \overline{B_E},$$

where (1) follows as $\overline{B_H} \supseteq \overline{B_E}$ and (2) follows as T is the identity of E. By assumption T is compact and so $\overline{T\left(\overline{B_H}\right)}$ is compact. So as $\overline{B_E}$ is closed it must also be compact. However, by Theorem 7.1.3 we know that $\overline{B_E}$ is only compact if it is of finite dimension.

2. Let $(f_n)_{n\in\mathbb{N}} \subseteq \mathrm{Ran}(\mathrm{Id} - T)$, with $f_n = (\mathrm{Id} - T)u_n$, be a sequence converging to f in H.
Step 1: Project u_n onto $\ker(\mathrm{Id} - T)^{\perp}$.
Let $d_n = \mathrm{dist}(u_n, \ker(\mathrm{Id} - T))$. By statement 1 we know that $\ker(\mathrm{Id} - T)$ is finite-dimensional and thus closed, moreover, it is a subspace and hence convex. Therefore, by Theorem 8.2.1 we can write

$$u_n = v_n + (u_n - v_n)$$

where $v_n \in \ker(\mathrm{Id} - T)$ and $u_n - v_n \in \ker(\mathrm{Id} - T)^{\perp}$. Note that $\|u_n - v_n\| = d_n$ by Corollary 8.2.5.
Step 2: Show that $(d_n)_{n\in\mathbb{N}}$ is bounded.
For contradiction suppose that, up to subsequences, we have $d_n \to \infty$. Let

$$w_n = \frac{u_n - v_n}{\|u_n - v_n\|} = \frac{u_n - v_n}{d_n}$$

so that $\|w_n\| = 1$. As $v_n \in \ker(\mathrm{Id} - T)$ it follows that $(\mathrm{Id} - T)(u_n - v_n) = f_n$ so we deduce that

$$(\mathrm{Id} - T)w_n = \frac{f_n}{d_n} \overset{n\to\infty}{\longrightarrow} 0 \tag{8.5.2}$$

By compactness of T we can assume that $Tw_n \to z$, up to subsequences. Hence, by (8.5.2) it follows that $w_n \to z$ with $z \in \ker(\mathrm{Id} - T)$. However, this is a contradiction as $w_n \in \ker(\mathrm{Id} - T)^{\perp}$ which is a closed subspace by statement 2 of Exercise 8.2.7.
Step 3: Show that $f \in \mathrm{Ran}(\mathrm{Id} - T)$, meaning that $\mathrm{Ran}(\mathrm{Id} - T)$ closed.
From step 2 we know that $(u_n - v_n)_{n\in\mathbb{N}}$ is a bounded sequence, and so using the compactness of T we can assume that $(T(u_n - v_n))_{n\in\mathbb{N}}$ converges to l, up to subsequences. Hence,

$$u_n - v_n = (\mathrm{Id} - T)(u_n - v_n) + T(u_n - v_n) \to f + l := g.$$

Consequently, we can use the continuity of $\mathrm{Id} - T$ to deduce that

$$(\mathrm{Id} - T)(u_n - v_n) \to (\mathrm{Id} - T)g.$$

On the other hand, we know that

$$(\mathrm{Id} - T)(u_n - v_n) = (\mathrm{Id} - T)u_n \to f,$$

and so we see that $(\mathrm{Id} - T)g = f$.
Step 4: Show that $\mathrm{Ran}(\mathrm{Id} - T) = \ker\left(\mathrm{Id} - T^*\right)^{\perp}$.
($\subseteq$). Let $y = (\mathrm{Id} - T)x \in \mathrm{Ran}(\mathrm{Id} - T)$. As $(\mathrm{Id} - T)^* = \mathrm{Id} - T^*$, for $z \in \ker(\mathrm{Id} - T^*)$ we have

$$\langle y, z\rangle = \langle (\mathrm{Id} - T)x, z\rangle = \langle x, \left(\mathrm{Id} - T^*\right) z\rangle = 0.$$

Therefore, $\mathrm{Ran}(\mathrm{Id} - T) \subseteq \ker\left(\mathrm{Id} - T^*\right)^{\perp}$.
($\supseteq$). Assume that $\ker\left(\mathrm{Id} - T^*\right)^{\perp} \setminus \mathrm{Ran}(\mathrm{Id} - T) \neq \{0\}$. Let $x \in \ker\left(\mathrm{Id} - T^*\right)^{\perp} \setminus \mathrm{Ran}(\mathrm{Id} - T)$ be non-zero. Since $\mathrm{Ran}(\mathrm{Id} - T)$ is closed, the orthogonal projection on $\mathrm{Ran}(\mathrm{Id} - T)$ is well-defined as it is also a subspace and so convex. Let

$$x = Px + (x - Px)$$

where $x \in \mathrm{Ran}(\mathrm{Id} - T)$ and $(x - Px) \in \mathrm{Ran}(\mathrm{Id} - T)^{\perp}$. By assumption we know that $x \in \mathrm{Ker}(\mathrm{Id} - T^*)^{\perp}$, and as $\mathrm{Ran}(\mathrm{Id} - T) \subseteq (\ker(\mathrm{Id} - T^*)^{\perp}$ we know that $Px \in \ker(\mathrm{Id} - T^*)^{\perp}$. Therefore, $x - Px \in \ker(\mathrm{Id} - T^*)^{\perp}$ as it is a linear subspace. It follows that $y := x - Px \in \mathrm{Ran}(\mathrm{Id} - T)^{\perp} \cap \ker(\mathrm{Id} - T^*)^{\perp}$. Where we also note that $y \neq 0$ as $x \notin \mathrm{Ran}(\mathrm{Id} - T)$ by assumption. Using that $y \in \mathrm{Ran}(\mathrm{Id} - T)^{\perp}$ it follows that for all $c \in H$ we have $\langle y, (\mathrm{Id} - T)c \rangle = 0$ which happens if and only if $\langle (\mathrm{Id} - T^*)y, c \rangle = 0$ for all $c \in H$. Consequently, $T^*y = 0$ and $y \in \ker(\mathrm{Id} - T^*)$, but we know $y \in \ker(\mathrm{Id} - T^*)^{\perp}$. Thus $y = 0$, which is a contradiction.

3. Suppose that $\ker(\mathrm{Id} - T) = \{0\}$ but $\mathrm{Ran}(\mathrm{Id} - T) \neq H$. Let $Y_n = \mathrm{Ran}\left((\mathrm{Id} - T)^n\right)$ for $n \in \mathbb{N}$. Note that the set of inclusions
$$H = Y_0 \supseteq Y_1 \supseteq Y_2 \supseteq \dots \tag{8.5.3}$$
are proper due to our assumption that $\mathrm{Id} - T$ is injective but not surjective. Moreover, note that
$$(\mathrm{Id} - T)^n = \mathrm{Id} - \sum_{k=1}^{n} \binom{n}{k} T^k =: \mathrm{Id} - S$$
where S is a compact operator as T is compact. Therefore, applying statement 2 to $\mathrm{Id} - S$ it follows that Y_n is closed for every $n \in \mathbb{N}$. Applying Theorem 6.3.3 to $Y_{n+1} \subseteq Y_n$, we find an element $\varphi_{f_n} \in \mathcal{L}(Y_n, \mathbb{R})$ given by $\varphi_{f_n}(x) = \langle f_n, x \rangle$ such that $\varphi_{f_n}(x) = 0$ for all Y_{n+1} and $\|\varphi_{f_n}\|_{\mathcal{L}(Y_n, \mathbb{R})} = 1$. Consequently, $f_n \in Y_n^{\perp}$ and $\|f_n\|_H = 1$. By Theorem 6.2.5 we can extend φ_{f_n} to $\mathcal{L}(H, \mathbb{R})$. For $n > m$ observe that
$$\begin{aligned} \|T^* f_n - T^* f_m\| &= \|T^*(f_n - f_m)\| \\ &= \|(\mathrm{Id} - T^*)(f_n - f_m) + (f_m - f_n)\| \\ &\geq \sup_{x \in \bar{B}^{Y_n}} |\langle (\mathrm{Id} - T^*)(f_n - f_m) + (f_m - f_n), x \rangle| \\ &= \sup_{x \in \bar{B}^{Y_n}} |\langle f_n - f_m, (\mathrm{Id} - T)x \rangle + \langle f_m - f_n, x \rangle| \,. \end{aligned}$$
As $f_n - f_m \in Y_{n+1}^{\perp}$ and $(\mathrm{Id} - T)x \in Y_{n+1}$ it follows that $\langle f_n - f_m, (\mathrm{Id} - T)x \rangle$=0. Similarly, as $n > m$ we have $f_m \in Y_n^{\perp}$ so that $\langle f_m, x \rangle = 0$ as $x \in Y_n$. Hence,
$$\|T^* f_n - T^* f_m\| \geq \sup_{x \in \bar{B}^{Y_n}} |\langle f_n, x \rangle| = 1.$$
Therefore, $(T^* f_n)$ contains no convergent subsequences and so cannot be compact. This contradicts Theorem 8.5.4 as T is compact. So it must be the case that the inclusions (8.5.3) are not proper meaning $\mathrm{Id} - T$ is surjective. Conversely, if we assume that $\mathrm{Ran}(\mathrm{Id} - T) = H$, then using statement 2 it follows that $\ker(\mathrm{Id} - T^*) = \{0\}$. So from the arguments we have just made it follows that $\mathrm{Ran}(\mathrm{Id} - T^*) = H$, and so we can apply statement 2 again to conclude that $\ker(\mathrm{Id} - T) = \{0\}$.

4. Consider the following quantities.
   - $\alpha = \dim(\ker(\mathrm{Id} - T))$.
   - $\beta = \dim\left(H / \mathrm{Ran}(\mathrm{Id} - T)\right)$.
   - $\alpha^* = \dim(\ker(\mathrm{Id} - T^*))$.
   - $\beta^* = \dim\left(H / \mathrm{Ran}(\mathrm{Id} - T^*)\right)$

   By statement 1 we know that $\alpha, \alpha^* < \infty$. Also note that by statement 2 we have that $\mathrm{Ran}(\mathrm{Id} - T)$ is closed with $\mathrm{Ran}(\mathrm{Id} - T) = \ker(\mathrm{Id} - T^*)^{\perp}$. As $\ker(\mathrm{Id} - T^*)$ is finite-dimensional and thus closed it follows that $\mathrm{Ran}(\mathrm{Id} - T)^{\perp} = \ker(\mathrm{Id} - T^*)$. Therefore, $H = \mathrm{Ran}(\mathrm{Id} - T) \oplus \ker(\mathrm{Id} - T^*)$. Consequently, one can show that $H / \mathrm{Ran}(\mathrm{Id} - T) \subseteq \ker(\mathrm{Id} - T^*)$ which implies that $\beta \leq \alpha^*$. Similarly, $\beta^* \leq \alpha$. Now suppose that $\alpha > \beta$. Then we can write
$$H = \ker(\mathrm{Id} - T) \oplus E = \mathrm{Ran}(\mathrm{Id} - T) \oplus F$$
for E and F closed subspaces of H with $\dim(F) = \beta$. For $x \in H$ write $x = x_1 + x_2$ for $x_1 \in \ker(\mathrm{Id} - T)$ and $x_2 \in E$. Let $\pi : H \to \ker(\mathrm{Id} - T)$ be the continuous map given by $\pi x = x_1$. As we assume $\alpha > \beta$,

it follows that there is a surjective linear map $\phi : \ker(\mathrm{Id} - T) \to F$ such that there exists $x_0 \neq 0$ with $\phi x_0 = 0$. Note that ϕ is a finite range operator and so compact. Hence, the operator $\Phi = T + \phi \circ \pi$ is also compact as it is bounded. Moreover,

$$(\Phi - \mathrm{Id})(E) = \mathrm{Ran}(\mathrm{Id} - T)$$

and

$$(\Phi - \mathrm{Id})(\ker(\mathrm{Id} - T)) = \phi(\ker(\mathrm{Id} - T)) = F.$$

Therefore,

$$\mathrm{Ran}(\Phi - \mathrm{Id}) \supseteq \mathrm{Ran}(\mathrm{Id} - T) + F = H.$$

However, this contradicts statement 3 as $\ker(\Phi - \mathrm{Id}) \neq \{0\}$. Therefore, $\alpha \leq \beta$ which implies that $\alpha \leq \alpha^*$. Similarly, one shows that $\alpha^* \leq \alpha$ to deduce that $\alpha = \alpha^*$.

□

Remark 8.5.7.

1. *We can explore each of the components of Theorem 8.5.6 in the context of finite-dimensional real Hilbert spaces. Statement 1 is meaningless in finite dimensions. Similarly, statement 2 is meaningless as any finite-dimensional vector space is closed. The equality of statement 2 follows from standard manipulations in linear algebra. Let* $y = (\mathrm{Id} - T)x \in \mathrm{Ran}(\mathrm{Id} - T)$ *and* $z \in \ker(\mathrm{Id} - T^*)$*. Then as,*

   $$(\mathrm{Id} - T)^* = (\mathrm{Id} - T)^\top = \mathrm{Id} - T^\top = \mathrm{Id} - T^*$$

   we have

   $$\langle y, z\rangle = \langle (\mathrm{Id} - T)x, z\rangle = \langle x, (\mathrm{Id} - T^*)\, z\rangle = 0.$$

   Therefore, $y \in \ker(\mathrm{Id} - T^*)^\perp$ *meaning* $\mathrm{Ran}(\mathrm{Id} - T) \subseteq \ker(\mathrm{Id} - T^*)^\perp$*. To argue for equality one uses the rank-nullity theorem. In our setting, statement 3 says that the operator* $\mathrm{Id} - T$ *is injective if and only if it is surjective. Statement 4 is the fundamental theorem of linear algebra. Consequently, we see that Theorem 8.5.6 establishes conditions for when standard properties familiar from finite-dimensional operators also hold for infinite-dimensional operators.*

2. *As Theorem 8.5.6 gives a correspondence between* $\mathrm{Ran}(\mathrm{Id} - T) = \ker(\mathrm{Id} - T^*)^\perp$*, it reduces the problem of determining* $\mathrm{Ran}(\mathrm{Id} - T)$ *to a finite set of orthogonality conditions as Theorem 8.5.6 also tells us that* $\ker(\mathrm{Id} - T^*)$ *finite-dimensional.*

3. *The alternative nature of Theorem 8.5.6 refers to the fact that either* $\ker(\mathrm{Id} - T) \neq \{0\}$ *so the homogeneous formulation of* (8.5.1) *has a non-zero solution. Or,* $\ker(\mathrm{Id} - T) = \{0\}$ *so that* $\mathrm{Ran}(\mathrm{Id} - T) = H$ *meaning the inhomogeneous variation of* (8.5.1) *always has a solution.*

8.5.3 Spectral Theory

In finite dimensions linear operators are represented by matrices and there exists a concise understanding of the properties of this matrix when the operator is self-adjoint, that is when the matrix is symmetric. We will now try and generalise such a result to the infinite-dimensional case. However, we will only be able to consider compact self-adjoint operators. Removing the compactness assumptions leads to a result attributed to von Neumann that is beyond our scope.

Theorem 8.5.8 (The Spectral Theorem in Finite Dimensions)**.** *Let* H *be a finite-dimensional real Hilbert space. Consider* $M \in \mathcal{L}(H)$ *a symmetric matrix, then there exists an orthonormal basis* $(e_n)_{n=1}^d \subseteq H$ *such that*

$$Mx = \sum_{n=1}^{d} \lambda_n \langle x, e_n\rangle e_n,$$

where the λ_n are the eigenvalues of M and e_n are the corresponding eigenvectors.

When we transition to infinite dimensions terms become more nuanced. In the finite-dimensional case let $M : H \to H$ be an operator. Then the spectrum of M can be formulated in different ways.

1. The union of the eigenvalues, which are the λ such that $\ker(M - \lambda \mathrm{Id}) \neq \{0\}$.

2. The union of λ such that $M - \lambda \mathrm{Id}$ is not invertible.

In infinite dimensions, these notions are no longer equivalent and require use to make a distinction.

Definition 8.5.9. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator.*

1. *λ is an eigenvalue if there exists an $x \in H \setminus \{0\}$ such that $(T - \lambda \mathrm{Id})x = 0$.*

2. *λ is in the spectrum if $T - \lambda \mathrm{Id} : H \to H$ is not invertible.*

3. *The resolvent set is the complement of the spectrum.*

Example 8.5.10. *To see why we require Definition 8.5.9 to distinguish these notions in infinite dimensions consider the following. Let $T : L^2(0,1) \to L^2(0,1)$ be given by $f \mapsto mf$. Let $m \in L^\infty(0,1)$ so that $T \in \mathcal{L}\left(L^2(0,1)\right)$. Moreover, suppose the measure of $m^{-1}(\{y\})$ is zero for any y.*

- *T has no eigenvalues. Suppose T has an eigenvalue λ then $(m(x) - \lambda)f(x) = 0$ which cannot be the case for $f \in L^2(0,1)$ unless $f = 0$.*

- *The spectrum of f is $m([0,1])$. To see this observe that*

$$(T - \lambda \mathrm{Id})(x) = (m - \lambda)(x)$$

  is invertible with inverse $f \mapsto \frac{f}{m-\lambda}$ which is in $\mathcal{L}\left(L^2(0,1)\right)$ if and only if $\lambda \notin m([0,1])$.

Thus in infinite dimensions, there exists operators whose eigenvalues and spectrum do not coincide.

Proposition 8.5.11. *Let H be an infinite-dimensional Hilbert space. Let $T \in \mathcal{L}(H)$ be a self-adjoint compact operator. Then either $\pm\|T\|_{\mathcal{L}(H)}$ is an eigenvalue of T.*

Proof. Let $\lambda = \pm\|T\|_{\mathcal{L}(H)}$. Then using Proposition 8.5.5, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq H$ with $\|x_n\|_H = 1$ such that $\langle Tx_n, x_n\rangle \to \lambda$. Hence,

$$\begin{aligned}
0 &\leq \|Tx_n - \lambda x_n\|_H^2 \\
&= \|Tx_n\|_H^2 + \lambda^2\|x_n\|^2 - 2\lambda\langle Tx_n, x_n\rangle \\
&\leq \lambda^2 + \lambda^2 - 2\lambda\langle Tx_n, x_n\rangle \\
&\overset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Using the compactness of T we also know that there is a subsequence $(x_{n_k})_{k\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$ such that $Tx_{n_k} \overset{k\to\infty}{\longrightarrow} x$ for some $x \in \bar{B}^H$. We note $x \neq 0$ as $\|x_{n_k}\|_H = 1$. Therefore, as $Tx_{n_k} - \lambda x_{n_k} \to 0$ it follows that $\lambda x_{n_k} \to x$, and so

$$Tx = \lim_{k\to\infty} T(\lambda x_{n_k}) = \lambda \lim_{k\to\infty} Tx_{n_k} = \lambda x.$$

We conclude that λ is an eigenvalue of T. $\square$

Theorem 8.5.12. *For an infinite-dimensional separable Hilbert space H, let $T \in \mathcal{L}(H)$ be a compact self-adjoint operator. Then the following hold.*

1. *Zero is in the spectrum of T.*
2. *If λ is in the spectrum and non-zero then λ is an eigenvalue.*
3. *The eigenvalues can be ordered as a sequence $\lambda_n \to 0$.*
4. *The eigenspaces $\ker(T - \lambda_n \mathrm{Id}) = E_n$ are finite-dimensional.*
5. $\bigoplus_{\lambda_n \neq 0} E_n \oplus \ker(T) = H$.

Proof.

1. Suppose that T were invertible with inverse T^{-1}. As $T^{-1} \in \mathcal{L}(H)$ it is bounded and so $T^{-1}\left(\bar{B}^H\right) \subseteq K\bar{B}^H$ for some $K > 0$. Therefore,
$$\bar{B}^H = T\left(T^{-1}\left(\bar{B}^H\right)\right) = T\left(K\bar{B}^H\right) = KT\left(\bar{B}^H\right),$$
which implies that
$$\bar{B}^H \subseteq K\overline{T\left(\bar{B}^H\right)} \tag{8.5.4}$$
As T is compact we know that $\overline{T\left(\bar{B}^H\right)}$ is compact. Thus, as $\bar{B}^H$ is closed, (8.5.4) implies that $\bar{B}^H$ is compact. However, this contradicts Theorem 7.1.3 as H is infinite-dimensional. Therefore, T is not invertible and so zero is in the spectrum of T.

2. For $\lambda \neq 0$ the operator $\frac{1}{\lambda}T$ is compact. Hence, using Theorem 8.5.6 we have that the operator $T - \lambda \mathrm{Id} = \lambda\left(\frac{1}{\lambda}T - \mathrm{Id}\right)$ is invertible if and only if it is injective. Hence, if λ is in the spectrum it follows that $\ker(T - \lambda\mathrm{Id}) \neq \{0\}$. Which implies that there exists an $x \in H \setminus \{0\}$ such that $(T - \lambda\mathrm{Id})x = 0$, meaning λ is an eigenvalue.

3. Suppose that $\lambda_n \not\to 0$. Then we can extract a subsequence $(\lambda_{n_k})_{k\in\mathbb{N}}$ such that $|\lambda_{n_k}| \geq \epsilon$ for some $\epsilon > 0$. Moreover, as $(Te_{n_k})_{k\in\mathbb{N}} \subseteq T\left(\bar{B}^H\right)$, and T is compact, it follows that $(Te_{n_k})_{k\in\mathbb{N}}$ admits a convergent subsequence. In particular, the subsequence is Cauchy. For simplicity, we will also denote this subsequence $(Te_{n_k})_{k\in\mathbb{N}}$. However,
$$\begin{aligned}\left\|Te_{n_k} - Te_{n_{k'}}\right\| &= \left\|\lambda_{n_k}e_{n_k} - \lambda_{n_{k'}}e_{n_{k'}}\right\| \\ &\overset{(1)}{=} \sqrt{|\lambda_{n_k}|^2 + \left|\lambda_{n_{k'}}\right|^2} \\ &\geq \sqrt{2\epsilon^2} \\ &= \sqrt{2}\epsilon,\end{aligned}$$
where (1) is an application of Parseval's identity. This contradicts the $(Te_{n_k})_{k\in\mathbb{N}}$ being Cauchy, and so it must be the case that $\lambda_n \to 0$.

4. Note that $\dim(\ker(T - \lambda_n\mathrm{Id})) = \dim\left(\ker\left(\frac{1}{\lambda_n}T - \mathrm{Id}\right)\right)$. As $\frac{1}{\lambda_n}T$ is also compact it follows from statement 1 of Theorem 8.5.6 that $\dim(\ker(T - \lambda_n\mathrm{Id}))$ is finite-dimensional.

5. Let $x_n \in E_n \setminus \{0\}$ and $x_m \in E_m \setminus \{0\}$ for $n \neq m$, so that $\lambda_n \neq \lambda_m$. Then using the self-adjoint property of T it follows that
$$\lambda_n\langle x_n, x_m\rangle = \langle Tx_n, x_m\rangle = \langle x_n, Tx_m\rangle = \lambda_m\langle x_n, x_m\rangle.$$
Hence,
$$(\lambda_n - \lambda_m)\langle x_n, x_m\rangle = 0$$
which implies that $\langle x_n, x_m\rangle = 0$ as $\lambda_n - \lambda_m \neq 0$. Similarly, for $x \in \ker(T) \setminus \{0\}$ and $x_n \in E_n \setminus \{0\}$ we have
$$0 = \langle Tx, x_n\rangle = \langle x, Tx_n\rangle = \lambda_n\langle x, x_n\rangle.$$
So that $\langle x, x_n\rangle = 0$ as $\lambda_n \neq 0$. Now let $x \in H$. As E_1 is a closed linear subspace, the projection of x onto E_1 is well-defined. In particular, we write $x = x_1 + \tilde{x}_1$ for $x_1 \in E_1$ and $\tilde{x}_1 \in E_1^\perp$. By Theorem 8.5.6 we

know that $E_1^\perp = \operatorname{Ran}(T - \lambda_1 \mathrm{Id})$ is a closed linear subspace, which means that it is also a Hilbert space that is separable as H is separable. Let $T_2 = T|_{\operatorname{Ran}(T-\lambda_1\mathrm{Id})}$. As $\ker(T - \lambda_n \mathrm{Id}) \subseteq \operatorname{Ran}(T - \lambda_1 \mathrm{Id})$ for all $n \geq 2$ it follows that λ_n for $n \geq 2$ is in the spectrum of T_2. Moreover, λ_1 is not in the spectrum of T_2. From Proposition 8.5.11 we know that $\|T_2\|_{\mathcal{L}(H)} = |\lambda_2|$, as we have assumed the ordering of statement 3. Similarly to before, we can consider the projection of $\tilde{x}_1 \in \operatorname{Ran}(T-\lambda_1\mathrm{Id})$ onto E_2 and write $x = x_1 + x_2 + \tilde{x}_2$ where $x_2 \in E_2$ and $\tilde{x}_2 \in E_2^\perp$. Then we can let $T_3 = T_2|_{\operatorname{Ran}(T-\lambda_2\mathrm{Id})}$, noting that $\|T_3\|_{\mathcal{L}(H)} = |\lambda_3|$ by Proposition 8.5.11.

- If $\|T_{n+1}\|_{\mathcal{L}(H)} = 0$ for some $n \in \mathbb{N}$, it follows that $x = \sum_{k=1}^n x_k + \tilde{x}_n$ where $\tilde{x}_n \in \ker(T)$. Therefore, $x \in \operatorname{span}\left(\ker(T) \cup \bigcup_{n\in\mathbb{N}} E_n\right)$.
- If $\|T_{n+1}\|_{\mathcal{L}(H)} > 0$ for all $n \in \mathbb{N}$ we have

$$\begin{aligned}\left\| Tx - T\left(\sum_{k=1}^n x_n\right)\right\|_H &= \|T\tilde{x}_n\|_H \\ &= \|T_{n+1}\tilde{x}_n\|_H \\ &\leq \|T_{n+1}\|_{\mathcal{L}(H)}\|\tilde{x}_n\| \\ &\leq |\lambda_{n+1}|\|x\|_H,\end{aligned}$$

  where the right-hand side tends to zero as $n \to \infty$. It follows that $\operatorname{span}\left(\ker(T) \cup \bigcup_{n\in\mathbb{N}} E_n\right)$ is dense in H.

Therefore, the conditions of Definition 8.4.1 are satisfied and so $H = \bigoplus_{\lambda_n \neq 0} E_n \oplus \ker(T)$.

□

Remark 8.5.13.

1. *The sequence in statement 3 of Theorem 8.5.12 may be set to be eventually zero if there are only finitely many eigenvalues of T.*
2. *By Theorem 8.5.12, if T is compact then we can represent*

$$Tx = \sum_{\lambda_n \neq 0} \lambda_n \langle x, e_n \rangle e_n$$

for $e_n \in E_n$.

8.6 Solution to Exercises

Exercise 8.2.7

Solution.

1. If $y \in H^\perp$ then $(y,y) = 0$ which implies $y = 0$. On the other hand, for $y \in H$ it follows that $(y,0) = (y, 0+0) = 2(y,0)$ which implies that $(y,0) = 0$ and so $y \in \{0\}^\perp$.
2. By the bilinearity of $(\cdot,\cdot)$, we have that $F^\perp$ is a linear subspace. Let $(y_n)_{n\in\mathbb{N}} \subseteq F^\perp$ converge to y in H. Then for any $x \in F$ it follows that

$$\begin{aligned}|(y,x) - (y_n,x)| &= |(y - y_n, x)| \\ &\overset{\text{C.S}}{\leq} \|y - y_n\|\|x\|\end{aligned}$$

   where the right-hand side converges to zero by the assumption that $\|y - y_n\| \to 0$ and $x \in F$ is fixed. Therefore, $y \in F^\perp$ which implies that $F^\perp$ is closed.
3. Let $y \in G^\perp$. Then for $x \in F$ it follows that $x \in G$ which implies $(y,x) = 0$ and so $y \in F^\perp$.

4. Suppose F is closed. Then for $x \in F$ it follows that $(x, y) = 0$ for all $y \in F^\perp$ which implies that $x \in \left(F^\perp\right)^\perp$. Hence, we have $F \subseteq \left(F^\perp\right)^\perp$. Let $x \in \left(F^\perp\right)^\perp$, then we can consider $\tilde{x} = P_F x \in F \subseteq \left(F^\perp\right)^\perp$. As F is closed we know that $x = P_F x + P_{F^\perp} x$ and so $x - \tilde{x} \in F^\perp$. As $\left(F^\perp\right)^\perp$ is a linear space we also have that $x - \tilde{x} \in \left(F^\perp\right)^\perp$. Therefore, as $F^\perp \cap \left(F^\perp\right)^\perp = \{0\}$ we deduce that $x - \tilde{x} = 0$ which implies that $x = \tilde{x} \in F$. Hence, $F = \left(F^\perp\right)^\perp$. For general F, we know by the continuity of $(\cdot, \cdot)$ that $\bar{F}^\perp = F^\perp$. Therefore, $\left(\bar{F}^\perp\right)^\perp = \left(F^\perp\right)^\perp$. Using the fact that $\bar{F}$ is closed we deduce that $\bar{F} = \left(F^\perp\right)^\perp$.

5. Let F and G be closed.

   (a) If $x \in F \cap G$ then for $y_1 + y_2 \in F^\perp + G^\perp$ we have that $(x, y_1 + y_2) = (x, y_1) + (x, y_2) = 0 + 0 = 0$ and so $F \cap G \subseteq \left(F^\perp + G^\perp\right)^\perp$. On the other hand, if $x \in \left(F^\perp + G^\perp\right)^\perp$, then $(x, y_1 + y_2) = 0$ for all $y_1 \in F^\perp$ and $y_2 \in G^\perp$. In particular, for $y_1 = 0$ we get that $x \in \left(G^\perp\right)^\perp = G$ and for $y_2 = 0$ we get that $x \in \left(F^\perp\right)^\perp = F$. Therefore, $x \in F \cap G$.

   (b) Replacing F with $F^\perp$ and G with $G^\perp$ in statement 5(a) gives

   $$F^\perp \cap G^\perp = \left(\left(F^\perp\right)^\perp \cap \left(G^\perp\right)^\perp\right)^\perp = (F + G)^\perp$$

   as F and G are closed.

   (c) Note that

   $$(F \cap G)^\perp \overset{\text{Stat 5(a)}}{=} \left(\left(F^\perp + G^\perp\right)^\perp\right)^\perp \overset{\text{Stat 4}}{=} \overline{F^\perp + G^\perp}.$$

   (d) Follows by similar arguments as statement 5(c) where instead we use statement 5(b) and statement 4.

□

Exercise 8.4.8

Proof. Let $(\lambda_k v)_{k \in \mathbb{N}} \subseteq V$ be a sequence converging to $u \in H$. Note that there is a bijection between V and $\mathbb{R}$, namely $\lambda v \mapsto \lambda$. As metrics are equivalent in finite dimensions it follows that $\lambda_k \to \lambda \in \mathbb{R}$, and so $\lambda_k v \to \lambda v \in V$. Hence, V is closed. Consequently, we can write $H = V \oplus V^\perp$ using Proposition 8.2.4. In particular, for $u \in H$ we have that $u = \lambda v + P_{V^\perp} u$, where $P_V u = \lambda v \in V$ and $P_{V^\perp} u \in V^\perp$. Therefore, $(u, v) = \lambda (v, v)$ which implies that $\lambda = \frac{(u,v)}{\|v\|^2}$. □

9 Integral Operators

The theory of bounded linear operators can be harnessed to solve linear differential equations and integral equations.

9.1 Kernel Functions

Definition 9.1.1. *Let $X, Y \subseteq \mathbb{R}$ be interval. Then an operator A is an integral operator if there exists a function $k : X \to Y \to \mathbb{R}$ such that*

$$(Af)(t) = \left(A_{[k]}f\right)(t) := \int_Y k(t,s)f(s)\,\mathrm{d}s$$

for all $t \in X$ and functions f for which A is defined.

Remark 9.1.2. *The function k of Definition 9.1.1 is referred to as the integral kernel or the kernel function of A.*

Example 9.1.3. *Consider $[a,b] \subseteq \mathbb{R}$ and the initial value problem*

$$u' = f$$

with $u(a) = 0$, where $f \in L^2(a,b)$. The solution to the problem is given by $u = Jf$ where J is the operator

$$Jf(t) = \int_a^t f(s)\,\mathrm{d}s = \int_a^b \mathbf{1}_{[a,t]}(s)f(s)\,\mathrm{d}s$$

for $t \in [a,b]$. Letting

$$k(t,s) := \mathbf{1}_{[a,t]}(s)$$

we have that

$$Jf(t) = \int_a^b k(t,s)f(s)\,\mathrm{d}s$$

for $f \in L^2(a,b)$. Therefore, J is an integral operator and k is the kernel function as per Definition 9.1.1.

For f a measurable function defined on an interval X, and g a measurable function defined on an interval Y, let $f \otimes g : X \times Y \to \mathbb{R}$ be given by

$$(f \otimes g)(x,y) := f(x)g(y).$$

Theorem 9.1.4. *Let $p \in [1,\infty)$. If $f \in L^p(X)$ and $g \in L^p(Y)$ then $f \otimes g \in L^p(X \times Y)$ with*

$$\|f \otimes g\|_{L^p(X\times Y)} = \|f\|_{L^p(X)}\|g\|_{L^p(Y)}.$$

Moreover,

$$\overline{\operatorname{span}\{f \otimes g : f \in L^p(X),\, g \in L^p(Y)\}} = L^p(X \times Y).$$

Proof. Note that $f \otimes g$ is measurable. Moreover,

$$\begin{aligned}
\int_{X\times Y} |f\otimes g|^p \,\mathrm{d}\lambda^2 &= \int_X \int_Y |f(x)g(y)|^p \,\mathrm{d}y\,\mathrm{d}x \\
&= \int_X |f(x)|^p \int_Y |g(y)|^p \,\mathrm{d}y\,\mathrm{d}x \\
&= \int_X |f(x)|^p \|g\|_{L^p(Y)}^p \,\mathrm{d}x \\
&= \|f\|_{L^p(X)} \|g\|_{L^p(Y)}.
\end{aligned}$$

Therefore,

$$\|f\otimes g\|_{L^p(X\times Y)} = \|f\|_{L^p(X)}\|g\|_{L^p(Y)}.$$

□

Lemma 9.1.5. *Let $f \in L^1(a,b)$ and*

$$(Jf)(t) = \int_a^b \mathbf{1}_{[a,t]}(s) f(s)\,\mathrm{d}s.$$

Then for $n \geq 1$ we have

$$\left(J^n f\right)(t) = \frac{1}{(n-1)!}\int_a^t (t-s)^{n-1} f(s)\,\mathrm{d}s \tag{9.1.1}$$

for all $t \in [a,b]$. In particular, J^n is an integral operator with kernel function

$$k_n(t,s) = \frac{1}{(n-1)!}\mathbf{1}_{[a,t]}(s)(t-s)^{n-1}$$

for $s,t \in [a,b]$.

Proof. For $n = 1$ it is clear that (9.1.1) holds. Assume (9.1.1) holds for $n \leq k$. Then

$$\begin{aligned}
\left(J^{k+1} f\right)(t) &= J^k(Jf)(t) \\
&\overset{(9.1.1)}{=} \frac{1}{(k-1)!}\int_a^t (t-s)^{k-1} Jf(s)\,\mathrm{d}s \\
&= \frac{1}{(k-1)!}\int_a^t (t-s)^{k-1} \int_a^s f(u)\,\mathrm{d}u\,\mathrm{d}s \\
&\overset{\text{Fubini}}{=} \frac{1}{(k-1)!}\int_a^t \int_u^t (t-s)^{k-1} f(u)\,\mathrm{d}s\,\mathrm{d}u \\
&= \frac{1}{(k-1)!}\int_a^t f(u)\left[-\frac{1}{k}(t-s)^k\right]_u^t \mathrm{d}u \\
&= \frac{1}{k!}\int_a^t (t-u)^k f(u)\,\mathrm{d}u.
\end{aligned}$$

Therefore, by induction the proof is complete. □

9.2 The Hilbert-Schmidt Kernel Function

Definition 9.2.1. *Let $X, Y \subseteq \mathbb{R}$ be intervals and let $k : X\times Y \to \mathbb{R}$ be measurable with respect to the product measure. If $k \in L^2(X\times Y)$, then k is referred to as a Hilbert-Schmidt kernel function.*

Theorem 9.2.2. *Let $k \in L^2(X \times Y)$ be a Hilbert-Schmidt kernel function. Then the corresponding integral operator, $A_{[k]}$, satisfies*

$$\|A_{[k]}f\|_{L^2(X)} \leq \|k\|_{L^2(X\times Y)}\|f\|_{L^2(Y)}$$

for all $f \in L^2(Y)$. Moreover, the kernel function k is uniquely determined, almost everywhere, by $A_{[k]}$.

Proof. Using Cauchy-Schwartz we have

$$\left|\int_Y k(x,y)f(y)\,\mathrm{d}y\right| \leq \int_Y |k(x,y)f(y)|\,\mathrm{d}y \leq \left(\int_Y |k(x,y)|^2\,\mathrm{d}y\right)^{\frac{1}{2}} \|f\|_{L^2(Y)}$$

for $x \in [a,b]$. Therefore,

$$\begin{aligned}\|A_{[k]}f\|^2_{L^2(X)} &= \int_X \left|\int_Y k(x,y)f(y)\,\mathrm{d}y\right|^2 \mathrm{d}x \\ &\leq \left(\int_X\int_Y |k(x,y)|^2\,\mathrm{d}y\,\mathrm{d}x\right)\|f\|^2 L^2(Y)\end{aligned}$$

as required. For the uniqueness statement, suppose that k_1 and k_2 are kernel functions of $A_{[k]}$. Then,

$$\begin{aligned}\langle k_1 - k_2, f\otimes g\rangle_{L^2(X\times Y)} &= \int_X\int_Y (k_1(x,y) - k_2(x,y))\,f(x)g(y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_X \left(A_{[k]}g - A_{[k]}g\right)(x)f(x)\,\mathrm{d}x \\ &= 0\end{aligned}$$

for all $f \in L^2(X)$ and $g \in L^2(Y)$. Using Theorem 9.1.4 we know that

$$k_1 - k_2 \in \overline{\operatorname{span}\{f\otimes g : f \in L^p(X),\, g \in L^p(Y)\}}$$

and so it follows that $\langle k_1 - k_2, k_1 - k_2\rangle_{L^2(X,Y)} = 0$ which implies that $k_1 = k_2$ almost everywhere. □

Remark 9.2.3. *The integral operator $A_{[k]} : L^2(Y) \to L^2(X)$ associated with a Hilbert-Schmidt kernel $k \in L^2(X \times Y)$ is referred to as a Hilbert-Schmidt integral operator.*

The Hilbert-Schmidt norm for $A_{[k]}$ is

$$\|A_{[k]}\|_{\mathrm{HS}} := \|k\|_{L^2(X\times Y)} = \left(\int_X\int_Y |k(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y\right).$$

The Hilbert-Schmidt norm for $A_{[k]}$ is well-defined as k is almost surely unique by Theorem 9.2.2. Moreover, from Theorem 9.2.2 we deduce that

$$\|A_{[k]}\|_{\mathcal{L}} \leq \|A_{[k]}\|_{\mathrm{HS}}.$$

9.2.1 Application to the Poisson Problem

The Dirichlet Laplacian on (a,b) is the operator $\Delta_D : \mathrm{H}^2(a,b) \cap H^1_0(a,b) \to L^2(a,b)$ given by

$$\Delta_D u := u''.$$

For $f \in L^2(a,b)$ the Poisson problem is

$$\Delta u = -f \tag{9.2.1}$$

with $u(a) = u(b) = 0$. Equation (9.2.1) is said to be well-posed if it has a unique solution for each $f \in L^2(X\times Y)$ which depends continuously on f. Equivalently, there exists a continuous operator that maps an input function to its unique solution.

Exercise 9.2.4. *Show that* $\mathrm{H}^2(a,b) \cap H_0^1(a,b)$ *is a closed subspace of* $\mathrm{H}^2(a,b)$ *with respect to* $\|\cdot\|_{\mathrm{H}^2}$.

Definition 9.2.5. *For normed spaces, a bounded linear operator* $T : E \to F$ *is invertible if* T *is bijective and* T^{-1} *is bounded.*

Proposition 9.2.6. *The Dirichlet Laplacian,* Δ_D*, has inverse given by*

$$(\Delta_D)^{-1} f(t) = -\int_a^b g(t,s) f(s)\,\mathrm{d}s$$

where

$$g(t,s) = \begin{cases} \frac{1}{b-a}(b-t)(s-a) & s \leq t \\ \frac{1}{b-a}(b-s)(t-a) & t \leq s, \end{cases}$$

is Green's function for the Poisson problem.

Proof. Integrating (9.2.1) twice yields

$$u(t) = -\left(J^2 f\right)(t) + (t-a) + d$$

for $t \in [a,b]$ and some scalars $c, d \in \mathbb{R}$. Using $u(a) = u(b) = 0$ it follows that $d = 0$ and

$$c = \frac{1}{b-a}\left(J^2 f\right)(b).$$

Therefore,

$$u(t) = -\left(J^2 f\right)(t) + \frac{\left(J^2 f\right)(b)}{b-a}(t-a).$$

Using Lemma 9.1.5 we conclude that

$$\begin{aligned} u(t) &= -\int_a^t (t-s) f(s)\,\mathrm{d}s + \frac{t-a}{b-a}\int_a^b (b-s) f(s)\,\mathrm{d}s \\ &= \int_a^b \left(-\mathbf{1}_{[a,t]}(s)(t-s) + \frac{(t-a)(b-s)}{b-a}\right) f(s)\,\mathrm{d}s \\ &= \int_a^b g(t,s) f(s)\,\mathrm{d}s. \end{aligned}$$

□

Note that Green's function for the Poisson problem satisfies

$$|g(t,s)| \leq b-a$$

for all $s, t \in [a,b]$ and so

$$\int_a^b \int_a^b |g(t,s)|^2\,\mathrm{d}s\,\mathrm{d}t < \infty.$$

Therefore, Green's function for the Poisson problem is a Hilbert-Schmidt Kernel function. Thus using Theorem 9.2.2 we have that Δ_D^{-1} is a Hilbert-Schmidt integral operator on $L^2(a,b)$.

Exercise 9.2.7. *Show that* $\Delta_D^{-1} : L^2(a,b) \to \mathrm{H}^2(a,b)$ *is bounded.*

Using Exercise 9.2.7, it follows that (9.2.1) is well-posed.

9.3 The Neumann Series

In more general settings than the Poisson problem, there is likely no closed form for the solution operator. Instead, a series of approximate solutions to simplified problems are constructed in what are known as approximation methods. This motivates the study of sequences of operators. In particular, recall that if E, F and G are normed spaces, then

$$\|ST\| \leq \|S\|\|T\|$$

for $T \in \mathcal{L}(E,F)$ and $S \in \mathcal{L}(F,G)$. Consequently,

$$\|T^n\| \leq \|T\|^n$$

for $n \in \mathbb{N}$.

Lemma 9.3.1. *Let E, F and G be normed space. Let $T, T_n \in \mathcal{L}(E,F)$ and $S, S_n \in \mathcal{L}(F,G)$ for $n \in \mathbb{N}$. Then $T_n \overset{n\to\infty}{\longrightarrow} T$ and $S_n \overset{n\to\infty}{\longrightarrow} S$ implies that $S_nT_n \overset{n\to\infty}{\longrightarrow} ST$. Moreover, if $f, f_n \in E$ for $n \in \mathbb{N}$, then $T_n \overset{n\to\infty}{\longrightarrow} T$ and $f_n \overset{n\to\infty}{\longrightarrow} f$ implies that $T_nf \overset{n\to\infty}{\longrightarrow} Tf$.*

Proof. Observe that

$$S_nT_n - ST = (S_n - S)(T_n - T) + S(T_n - T) + (S_n - S)T.$$

Taking norms it follows that

$$\begin{aligned}
\|S_nT_n - ST\| &\leq \|(S_n - S)(T_n - T)\| + \|S(T_n - T)\| + \|(S_n - S)T\| \\
&\leq \|S_n - S\|\|T_n - T\| + \|S\|\|T_n - T\| + \|S_n - S\|\|T\| \\
&\overset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

We similarly conclude that $T_nf \overset{n\to\infty}{\longrightarrow} Tf$ under the appropriate conditions. $\square$

Remark 9.3.2. *A sequence $(T_n)_{n\in\mathbb{N}} \subseteq \mathcal{L}(E,F)$ converges strongly to $T \in \mathcal{L}(E,F)$ if*

$$T_nf \overset{n\to\infty}{\longrightarrow} Tf$$

with respect to $\|\cdot\|_F$ for every $f \in E$.

With these, we study the Neumann series which is a useful approximation method for solving perturbed problems. In particular, we look at the Neumann series for solving for $u \in \mathrm{H}^2(a,b)$ satisfying

$$u'' - Tu = -f \tag{9.3.1}$$

with $u(a) = u(b) = 0$ for given $f \in L^2(a,b)$. Note that when $T = 0$ (9.3.1) becomes (9.2.1), and so T can be seen as a perturbation on the Poisson problem. As (9.2.1) is well-posed, we would desire that for a small perturbation, that is small T, (9.3.1) would be well-posed. Using the Dirichlet Laplacian operator, (9.3.1) is equivalent to

$$\left(I - T\Delta_D^{-1}\right)\Delta_D u$$

for $u \in \mathrm{H}^2(a,b) \cap \mathrm{H}^1(a,b)$. Hence, the well-posedness of (9.3.1) is reduced to understanding when $I - T\Delta_D^{-1} : L^2(a,b) \to L^2(a,b)$ is invertible. We study this problem more generally by considering $A \in \mathcal{L}(E)$ for E a Banach space and the conditions for unique solutions to the equation $(I - A)u = f$.

Lemma 9.3.3. *Let E be a Banach space and let $A \in \mathcal{L}(E)$. If $f \in E$ is such that $u := \sum_{n=0}^{\infty} A^n f$ converges in E, then $(I - A)u = f$.*

Proof. Let $u_0 := 0$ and $u_n := f + Au_n$ for $n \geq 1$. Observe that $u_{n+1} = Tu_n$ where $T(\cdot) = f + A(\cdot)$ is a continuous. Moreover,

$$\|Tx - Ty\| = \|A(x - y)\| = \|A\|\|x - y\|.$$

As $\sum_{n=0}^{\infty} A^n f$ it must be the case that $\|A\| < 1$, and so T is a strict contraction. Thus we can apply Theorem 1.2.25 to conclude that $(u_n)_{n\in\mathbb{N}} \subseteq E$ converges to the unique fixed point of T. So by the uniqueness of limits, we have that $Tu = u$ which implies that $(I - A)u = f$. □

Lemma 9.3.4. *Let E be a Banach space and let $(f_n)_{n\in\mathbb{N}} \subseteq E$ be such that $\sum_{n=1}^{\infty} \|f_n\| < \infty$. Then $\sum_{n=1}^{\infty} f_n$ converges in E. Moreover,*

$$\left\|\sum_{n=1}^{\infty} f_n\right\| \leq \sum_{n=1}^{\infty} \|f_n\|.$$

Proof. Let $(s_n)_{n\in\mathbb{N}}$ be given by $s_n := \sum_{j=1}^{n} f_j$. Let $m > n$ then

$$\begin{aligned}
\|s_m - s_n\| &= \left\|\sum_{j=n+1}^{m} f_j\right\| \\
&\leq \sum_{j=n+1}^{m} \|f_j\| \\
&\leq \sum_{j=n+1}^{\infty} \|f_j\| \\
&\overset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Moreover, using the continuity of $\|\cdot\|$ we deduce that

$$\begin{aligned}
\left\|\sum_{n=1}^{\infty} f_n\right\| &= \left\|\lim_{n\to\infty} \sum_{j=1}^{n} f_j\right\| \\
&= \lim_{n\to\infty} \left\|\sum_{j=1}^{n} f_j\right\| \\
&\leq \lim_{n\to\infty} \sum_{j=1}^{n} \|f_j\| \\
&= \sum_{j=1}^{\infty} \|f_j\|.
\end{aligned}$$

□

Theorem 9.3.5. *Let E be a Banach space and let $A \in \mathcal{L}(E)$ be such that*

$$\sum_{n=0}^{\infty} \|A^n\| < \infty.$$

Then $I - A$ is invertible, with its inverse given by

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n. \tag{9.3.2}$$

Proof. If $(I-A)f=0$, then $Af=f$ and so

$$\begin{aligned}\sum_{n=0}^{\infty}\|f\| &= \sum_{n=0}^{\infty}\|A^n f\| \\ &\le \sum_{n=0}^{\infty}\|A^n\|\|f\| \\ &< \infty.\end{aligned}$$

Therefore, $f=0$ which implies that $I-A$ is injective. As E is Banach we have that $\mathcal{L}(E)$ is Banach by Proposition 6.1.2. Hence, by Lemma 9.3.4 the limit $B := \sum_{n=0}^{\infty} A^n$ exists in $\mathcal{L}(E)$ with respect to the operator norm. Since convergence with respect to the operator norm implies strong convergence we have that

$$Bf = \left(\sum_{n=0}^{\infty} A^n\right) f = \sum_{n=0}^{\infty} A^n f$$

as a convergent series in E. Therefore, by Lemma 9.3.3 it follows that $(I-A)Bf=f$, meaning $(I-A)$ is bijective with B as its inverse. □

Remark 9.3.6.

1. *The series of* (9.3.2) *is known as the Neumann series of* A.
2. *A sufficient, but by no means necessary, condition for the conditions of Theorem 9.3.5 to hold is that* A *is a strict contraction. That is,* $\|A\|<1$.

Using Theorem 9.3.5 we see that $\left\|T\Delta_D^{-1}\right\|<1$ is sufficient for (9.3.1) to be well-posed.

9.3.1 Application to the Volterra Integral Equations

For $m : [a,b]\times[a,b]\to\mathbb{R}$, the Volterra integral equation is given by

$$u(t) - \int_a^t m(t,s)u(s)\,\mathrm{d}s = f(t) \tag{9.3.3}$$

for $t\in[a,b]$, and $f\in\mathcal{C}([a,b])$ given. Consequently, the Volterra operator $V:\mathcal{C}([a,b])\to\mathcal{C}([a,b])$ is the integral operator with kernel function

$$k(t,s) = \mathbf{1}_{[a,t]}(s)m(t,s).$$

Exercise 9.3.7. *Verify that the Volterra operator* $V:\mathcal{C}([a,b])\to\mathcal{C}([a,b])$ *is well-defined.*

Lemma 9.3.8. *Let* $m:[a,b]\times[a,b]\to\mathbb{R}$ *be continuous, with* V *its associated Volterra operator. Then*

$$|V^n f(t)| \le \frac{\|m\|_\infty^n\|f\|_\infty(t-a)^n}{n!} \tag{9.3.4}$$

for all $f\in\mathcal{C}([a,b])$, $t\in[a,b]$ *and* $n\in\mathbb{N}$. *Consequently,*

$$\|V^n\|_{\mathcal{L}(\mathcal{C}([a,b]))} \le \frac{\|m\|_\infty^n(b-a)^n}{n!}$$

for every $n\in\mathbb{N}$.

Proof. For $n = 1$ we have that

$$\begin{aligned}
|Vf(t)| &= \left| \int_a^t m(t,s) f(s) \, \mathrm{d}s \right| \\
&\le \int_a^t |m(t,s) f(s)| \, \mathrm{d}s \\
&\le \|m\|_\infty \|f\|_\infty \int_a^t \mathrm{d}s \\
&= \|m\|_\infty \|f\|_\infty (t-a).
\end{aligned}$$

Assume (9.3.4) holds for $n \le k$, then

$$\begin{aligned}
\left|V^{k+1} f(t)\right| &= \left|V\left(V^k f\right)(t)\right| \\
&= \left| \int_a^t m(t,s) V^k f(s) \, \mathrm{d}s \right| \\
&\le \int_a^t |m(t,s)| \left|V^k f(s)\right| \, \mathrm{d}s \\
&\overset{(9.3.4)}{\le} \|m\|_\infty \int_a^t \frac{\|m\|_\infty^k \|f\|_\infty (s-a)^k}{k!} \, \mathrm{d}s \\
&= \frac{\|m\|_\infty^{k+1} \|f\|_\infty (t-a)^{k+1}}{(k+1)!}.
\end{aligned}$$

Therefore, (9.3.4) holds for all $n \in \mathbb{N}$. Consequently,

$$\begin{aligned}
\|V^n\|_{\mathcal{L}(\mathcal{C}([a,b]))} &= \sup_{f \in \mathcal{C}([a,b]) \setminus \{0\}} \frac{\|V^n f\|_\infty}{\|f\|_\infty} \\
&= \sup_{f \in \mathcal{C}([a,b]) \setminus \{0\}} \frac{\sup_{t \in [a,b]} |V^n f(t)|}{\|f\|_\infty} \\
&\le \sup_{t \in [a,b]} \frac{\|m\|_\infty^n (t-a)^n}{n!} \\
&= \frac{\|m\|_\infty^n (b-a)^n}{n!}.
\end{aligned}$$

□

Corollary 9.3.9. *Let $m : [a,b] \times [a,b] \to \mathbb{R}$ be continuous, with V its associated Volterra operator. Then for every $f \in \mathcal{C}([a,b])$ the corresponding Volterra integral equation given by* (9.3.3) *has a unique solution $u \in \mathcal{C}([a,b])$.*

Proof. Using Lemma 9.3.8 it follows that

$$\sum_{n=0}^{\infty} \|V^n\| \le \sum_{n=0}^{\infty} \frac{\|m\|_\infty^n (b-a)^n}{n!} = \exp\left(\|m\|_\infty (b-a)\right) < \infty.$$

Therefore, V satisfies the conditions of Theorem 9.3.5, and so $I - V$ is invertible which implies that (9.3.1) has a unique solution. □

9.4 Solution to Exercises

Exercise 9.2.4

Solution. Let $(f_n)_{n\in\mathbb{N}} \subseteq \mathrm{H}^2(a,b) \cap \mathrm{H}^1_0(a,b)$ be convergent to f with respect to $\|\cdot\|_{\mathrm{H}^2}$. Then as $\left(\mathrm{H}^2(a,b), \|\cdot\|_{\mathrm{H}^2}\right)$ is a Hilbert space it follows that $f \in \mathrm{H}^2(a,b)$. Moreover, as

$$\|f\|_{\mathrm{H}^1} \leq \|f\|_{\mathrm{H}^2}$$

for $f \in \mathrm{H}^2(a,b)$, we have that $(f_n)_{n\in\mathbb{N}}$ converges to f in $\|\cdot\|_{\mathrm{H}^1}$. Using Theorem 10.5.23 it follows that $f \in \mathrm{H}^1_0(a,b)$ which implies that $f \in \mathrm{H}^2(a,b) \cap \mathrm{H}^1_0(a,b)$. Therefore, $\mathrm{H}^2(a,b) \cap \mathrm{H}^1_0(a,b)$ is closed with respect to $\|\cdot\|_{\mathrm{H}^2}$. □

Exercise 9.2.7

Solution. Using Theorem 9.2.2 and Proposition 9.2.6 it follows that

$$\left\|\Delta_D^{-1} f\right\|_{\mathrm{H}^2(a,b)} \leq \|g\|_{L^2([a,b]^2)} \|f\|_{L^2(a,b)}$$

for all $f \in L^2(Y)$. Therefore,

$$\left\|\Delta_D^{-1}\right\| = \sup_{f \in L^2(a,b)\setminus\{0\}} \frac{\left\|\Delta_D^{-1} f\right\|_{\mathrm{H}^2(a,b)}}{\|f\|_{L^2(a,b)}} \leq \|g\|_{L^2([a,b]^2)} < \infty.$$

□

Exercise 9.3.7

Solution. For $t_1, t_2 \in [a,b]$ it follows that

$$\begin{aligned} |(Vf)(t_1) - (Vf)(t_2)| &= \left| \int_{t_1}^{t_2} m(t,s) f(s) \, \mathrm{d}s \right| \\ &\leq |t_2 - t_1| \|m\|_\infty \|f\|_\infty. \end{aligned}$$

As m and f are continuous functions on a compact domain we know that $\|m\|_\infty, \|f\|_\infty < \infty$. Therefore, Vf is continuous, and thus $V : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ is well-defined. □

10 Appendix

10.1 Ordered Sets

Let P be a set. Then $\leq$ is said to define a partial order relation on P if it satisfies the following.

- Reflexivity, $a \leq a$ for all $a \in P$.
- Anti-symmetry, $a \leq b$ and $b \leq a$ implies that $a = b$ for all $a, b \in P$.
- Transitivity, $a \leq b$ and $b \leq c$ implies $a \leq c$ for all $a, b, c \in P$.

Definition 10.1.1. *A subset $S \subseteq P$ is totally ordered if $a \leq b$ or $b \leq a$ for any $a, b \in S$.*

Definition 10.1.2. *If $Q \subseteq P$, then $c \in P$ is an upper bound for Q if $a \leq c$ for all $a \in Q$.*

Definition 10.1.3. *An element $m \in S \subseteq P$ is maximal if $m \leq x$ for $x \in S$ implies that $m = x$.*

Definition 10.1.4. *A set P is inductive if any totally ordered subset Q has an upper bound.*

Lemma 10.1.5 (Zorn's Lemma). *Every non-empty ordered set that is inductive has a maximal element.*

10.2 Hardy's Inequality

Theorem 10.2.1 (Hardy's Inequality). *Let $1 < p \leq \infty$ and let $f \in L^p(0, \infty)$. Then there exists a $C_p > 0$ such that*

$$\left\| \frac{f(x)}{x} \right\|_{L^p} \leq C_p \left\| f'(x) \right\|_{L^p} .$$

Equivalently, if $F(x) = \int_0^x f(t) \, \mathrm{d}t$ then

$$\left\| \frac{F(x)}{x} \right\|_{L^p} \leq C_p \left\| f \right\|_{L^p} .$$

Proof. For a function f let $F(x) := \frac{1}{x} \int_0^x f(t) \, \mathrm{d}t$.
Step 1: Let $f \in \mathcal{C}_c^\infty(0, \infty)$ be non-negative. Show that $F \in \mathcal{C}^1(0, \infty)$ and $xF' = f - F$.
Note that by the fundamental theorem of calculus

$$F'(x) = \frac{1}{x} f(x) - \frac{1}{x^2} F(x)$$

and so $xF' = f - F$. It is clear that F and F' are continuous. We now show that F and F' are bounded to complete the step. As f is a bounded function the only concerns of unboundedness arise for the $\frac{1}{x}$ terms as $x \to 0$. Recall, that $f \in \mathcal{C}_c^\infty(0, \infty)$. Hence, $\operatorname{supp}(f) = K$ is a compact set of $(0, \infty)$. Suppose that for every $\epsilon > 0$ the set $[0, \epsilon] \cap K \neq \emptyset$. Then there exists a sequence $(x_n)_{n \in \mathbb{N}} \subseteq K$ such that $x_n \to 0$ as $n \to \infty$. As K is closed this would imply that $0 \in K$ which contradicts $K \subseteq (0, \infty)$. Therefore, there exists an $\epsilon > 0$ such that $[0, \epsilon] \cap K = \emptyset$. Consequently, $f(x) = 0$ for all $x \in [0, \epsilon]$. Therefore, $\int_0^x f(x) \, \mathrm{d}x = 0$ for all $x \in [0, \epsilon]$. Hence, $\frac{1}{x} \int_0^x f(x) \, \mathrm{d}x = 0$ for $x \in [0, \epsilon]$. One carries out a similar argument to show that F' is bounded near zero. Thus, F and F' are continuous and bounded which implies that $F \in \mathcal{C}^1(0, \infty)$.
Step 2: Show that $\int_0^\infty F(x)^p \, \mathrm{d}x = -p \int_0^\infty x F(x)^{p-1} F'(x) \, \mathrm{d}x$.

Consider $I_R = \int_0^R F(x)^p \,\mathrm{d}x$. Performing integration by parts with $u = F(x)^p$ and $\frac{\mathrm{d}v}{\mathrm{d}x} = 1$ we deduce that

$$\int_0^R F(x)^p \,\mathrm{d}x = [xF(x)^p]_0^R - \int_0^R pxF(x)^{p-1}F'(x)\,\mathrm{d}x.$$

Letting K be the compact support of f we know that K is bounded and so for sufficiently large R it follows that

$$\int_K f(x)\,\mathrm{d}x = \int_0^R f(x)\,\mathrm{d}x = \int_0^\infty f(x)\,\mathrm{d}x.$$

As f is bounded on K it follows that

$$\int_0^\infty f(x)\,\mathrm{d}x \leq M$$

for some $M > 0$ which implies that $xF(x)^p \leq \frac{M^p}{x^{p-1}}$. Hence,

$$[xF(x)^p]_0^R \overset{R\to\infty}{\longrightarrow} 0.$$

Therefore,

$$\int_0^\infty F(x)^p \,\mathrm{d}x = -p\int_0^\infty xF(x)^{p-1}F'(x)\,\mathrm{d}x,$$

which is well-defined as the functions F and F' are bounded.
<u>Step 3: Deduce that $\|F\|_{L^p}^p \leq C_p\|f\|_{L^p}$.</u>
Combining steps 1 and 2 we deduce that

$$\begin{aligned}\int_0^\infty F(x)^p \,\mathrm{d}x &= -p\int_0^\infty xF(x)^{p-1}F'(x)\,\mathrm{d}x\\ &= -p\int_0^\infty F(x)^{p-1}(f(x)-F(x))\,\mathrm{d}x\\ &= p\int_0^\infty F(x)^p\,\mathrm{d}x - p\int_0^\infty F(x)^{p-1}f(x)\,\mathrm{d}x.\end{aligned}$$

Therefore,

$$\int_0^\infty F(x)^p\,\mathrm{d}x = \frac{p}{p-1}\int_0^\infty F(x)^{p-1}f(x)\,\mathrm{d}x.$$

As $f(x) \geq 0$ for all $x \in (0,\infty)$ it follows that $F(x) \geq 0$ for all $x \in (0,\infty)$. Therefore,

$$\begin{aligned}\|F\|_{L^p}^p &= \int_0^\infty |F(x)|^p\,\mathrm{d}x\\ &= \int_0^\infty F(x)^p\,\mathrm{d}x\\ &= \frac{p}{p-1}\int_0^\infty F(x)^{p-1}f(x)\,\mathrm{d}x.\end{aligned}$$

Let p' be such that $1 = \frac{1}{p} + \frac{1}{p'}$ so that $p' = \frac{p}{p-1}$. Then by applying Hölder's inequality, we deduce that

$$\begin{aligned}\|F\|_{L^p}^p &= \int_0^\infty F(x)^p\,\mathrm{d}x\\ &\leq \frac{p}{p-1}\|f\|_{L^p}\left\|F^{p-1}\right\|_{L^{p'}}\\ &= \frac{p}{p-1}\|f\|_{L^p}\left(\int_0^\infty \left(F(x)^{p-1}\right)^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}\\ &= \frac{p}{p-1}\|f\|_{L^p}\left(\int_0^\infty F(x)^p\,\mathrm{d}x\right)^{\frac{p-1}{p}}\\ &= \frac{p}{p-1}\|f\|_{L^p}\|F\|_{L^p}^{p-1}.\end{aligned}$$

Therefore,

$$\|F\|_{L^p} \le \frac{p}{p-1}\|f\|_{L^p}.$$

Step 4: Extend the result to general $g \in \mathcal{C}_c^\infty(0,\infty)$.
For $g \in \mathcal{C}_c^\infty(0,\infty)$, note that $|g|$ is still a continuous function with compact support. As the continuous differentiability of f in the previous steps is not used the claims still hold true for $|g|$ as $|g(x)| \ge 0$ for all $x \in (0,\infty)$. Therefore,

$$\left\|\frac{1}{x}\int_0^x |g(x)|\,\mathrm{d}x\right\|_{L^p} \le \frac{p}{p-1}\||g|\|_{L^p}.$$

As $\||g|\|_{L^p} = \|g\|_{L^p}$ and $\frac{1}{x}\int_0^x g(t)\,\mathrm{d}t \le \frac{1}{x}\int_0^x |g(t)|\,\mathrm{d}t$ for all $t \in (0,\infty)$ we deduce that

$$\|G\|_{L^p} \le \frac{p}{p-1}\|g\|_{L^p}$$

where $G(x) := \frac{1}{x}\int_0^x g(t)\,\mathrm{d}t$.
Step 5: Extend the result to $f \in L^p(0,\infty)$.
Recall that $\mathcal{C}_c^\infty(0,\infty)$ is dense in $L^p(0,\infty)$. Therefore, given $f \in L^p(0,\infty)$ there exists a sequence $(f_n) \subseteq \mathcal{C}_c^\infty(0,\infty)$ such that $f_n \xrightarrow{L^p} f$. Letting $F_n(x) = \frac{1}{x}\int_0^x f_n(t)\,\mathrm{d}t$ we observe that

$$\begin{aligned}
\|F_n(x) - F(x)\|_{L^p_x} &= \left(\int_0^\infty \left|\int_0^x \frac{1}{x}f_n(t) - \frac{1}{x}f(t)\,\mathrm{d}t\right|^p \mathrm{d}x\right)^{\frac{1}{p}} \\
&= \left(\int_0^\infty \left|\int_0^1 f_n(xt) - f(xt)\,\mathrm{d}t\right|^p \mathrm{d}x\right)^{\frac{1}{p}} \\
&\overset{(1)}{\le} \int_0^1 \left(\int_0^\infty |f_n(xt) - f(xt)|\,\mathrm{d}x\right)^{\frac{1}{p}} \mathrm{d}t \\
&= \int_0^1 \frac{1}{t^{\frac{1}{p}}}\|f_n - f\|_{L^p}\,\mathrm{d}x \\
&\overset{(2)}{=} M\|f_n - f\|_{L^p},
\end{aligned}$$

where (1) follows from Minkowski's integral inequality[1], and (2) follows from the fact that $p > 1$ and so the integral is finite. Therefore, $F_n \xrightarrow{L^p} F$. As $f_n \in \mathcal{C}_c^\infty(0,\infty)$ we know that the inequality $\|F_n\|_{L^p} \le C_p\|f_n\|_{L^p}$ holds. Sending $n \to \infty$ it follows that $\|F\|_{L^p} \le C_p\|f\|_{L^p}$. $\square$

10.3 Hölder Spaces

Definition 10.3.1. *For an open set $\Omega \subseteq \mathbb{R}^d$, the $\alpha \in (0,1)$ Hölder space denoted $\mathcal{C}^\alpha\left(\bar{\Omega}\right)$ is the set of continuous functions $f \in \mathcal{C}^0\left(\bar{\Omega}\right)$ such that*

$$\sup_{x\neq y,(x,y)\in\Omega^2} \frac{|f(x)-f(y)|}{|x-y|^\alpha} < \infty.$$

The norm on $\mathcal{C}^\alpha\left(\bar{\Omega}\right)$ is given by

$$\|f\|_{\mathcal{C}^\alpha(\bar{\Omega})} = \|f\|_\infty + \sup_{x\neq y,(x,y)\in\Omega^2} \frac{|f(x)-f(y)|}{|x-y|^\alpha}.$$

[1] https://en.wikipedia.org/wiki/Minkowski_inequality

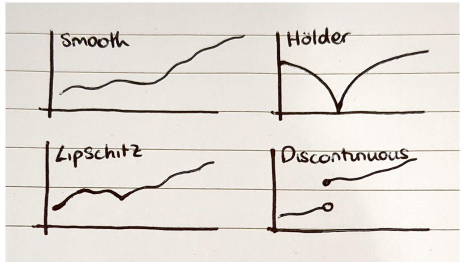


Figure 14: Smooth functions are the strongest class of continuous functions. Lipschitz continuous functions have joins where the gradients at the joins are finite. Lipschitz continuous functions can be thought of as Hölder continuous with $\alpha = 1$. Hölder continuous functions for $\alpha \in (0,1)$ can have cusps where the gradient at the cusp is potentially unbounded. Discontinuous functions contains jumps that do not satisfy the conditions of the previous spaces.

Theorem 10.3.2. *The space* $\left(\mathcal{C}^\alpha\left(\bar{\Omega}\right), \|\cdot\|_{\mathcal{C}^\alpha\left(\bar{\Omega}\right)}\right)$ *is a Banach space.*

Proof. Let $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}^\alpha\left(\bar{\Omega}\right)$ be Cauchy sequence. Then $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}^0\left(\bar{\Omega}\right)$ is a Cauchy sequence with respect to $\|\cdot\|_\infty$. As $\left(\mathcal{C}^0\left(\bar{\Omega}\right), \|\cdot\|_\infty\right)$ is a Banach space we know that $f_n \to f \in \mathcal{C}^0\left(\bar{\Omega}\right)$. It remains to show that $f \in \mathcal{C}^\alpha\left(\bar{\Omega}\right)$ and $f_n \to f$ in $\mathcal{C}^\alpha\left(\bar{\Omega}\right)$. For any $(x,y) \in \Omega^2$ with $x \neq y$, let $\delta = |x-y|$. Then as $f_n \to f$ in $\|\cdot\|_\infty$ it follows that there exists an $N \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| < \frac{\delta^\alpha}{2}$$

for all $x \in \Omega$. Therefore, for $n \geq N$ it follows that

$$\begin{aligned}
\frac{|f(x)-f(y)|}{|x-y|^\alpha} &\leq \frac{|f(x)-f_n(x)| + |f_n(x)-f_n(y)| + |f_n(y)-f(y)|}{|x-y|^\alpha} \\
&= \frac{|f(x)-f_n(x)| + |f_n(y)-f(y)|}{\delta^\alpha} + \frac{|f_n(x)-f_n(y)|}{|x-y|^\alpha} \\
&\leq \frac{\frac{\delta^\alpha}{2} + \frac{\delta^\alpha}{2}}{\delta^\alpha} + \frac{|f_n(x)-f_n(y)|}{|x-y|^\alpha} \\
&= 1 + \frac{|f_n(x)-f_n(y)|}{|x-y|^\alpha}.
\end{aligned}$$

As $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}^\alpha\left(\bar{\Omega}\right)$ is Cauchy we know that the sequence $(f_n)_{n\in\mathbb{N}}$ is bounded and so $\frac{|f_n(x)-f_n(y)|}{|x-y|} \leq C$ for all $n \in \mathbb{N}$ and $(x,y) \in \Omega^2$. Therefore,

$$\sup_{x\neq y,(x,y)\in\Omega^2} \frac{|f(x)-f(y)|}{|x-y|^\alpha} \leq 1 + C$$

and so $f \in \mathcal{C}^\alpha\left(\bar{\Omega}\right)$. By similar arguments we show that given an $\epsilon > 0$ and $(x,y) \in \Omega^2$ there exits a $N \in \mathbb{N}$ such that for $n \geq N$ we have that

$$\frac{|f(x)-f_n(x)-(f(y)-f_n(y))|}{|x-y|^\alpha} \leq \frac{\epsilon}{2}.$$

Therefore,

$$\sup_{x\neq y(x,y)\in\Omega^2} \frac{|f(x)-f_n(x)-(f(y)-f_n(y))|}{|x-y|^\alpha} \leq \frac{\epsilon}{2}.$$

Moreover, there exists a $M \in \mathbb{N}$ such that for $n \geq M$ we have that $\|f - f_n\|_\infty \leq \frac{\epsilon}{2}$ by the fact that $f_n \to f$ in $\|\cdot\|_\infty$. Therefore,

$$\|f - f_n\|_{\mathcal{C}^\alpha\left(\bar{\Omega}\right)} = \|f - f_n\|_\infty + \sup_{x\neq y(x,y)\in\Omega^2} \frac{|f(x)-f_n(x)-(f(y)-f_n(y))|}{|x-y|^\alpha} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence, $f_n \to f$ in $\mathcal{C}^\alpha\left(\bar{\Omega}\right)$. □

Example 10.3.3. *Let $p \in (1,\infty]$ and consider the operator $T : L^p(0,1) \to \mathcal{C}^{1-\frac{1}{p}}(0,1)$ be given by*

$$Tf(x) = \int_0^x f(z)\,\mathrm{d}z$$

for $x \in [0,1]$. For $x < y$ we have that,

$$\begin{aligned}
|Tf(x) - Tf(y)| &= \left|\int_0^x f(z)\,\mathrm{d}z - \int_0^y f(z)\,\mathrm{d}z\right| \\
&= \left|\int_y^x f(z)\,\mathrm{d}z\right| \\
&\overset{\text{T.I}}{\leq} \int_0^1 \mathbf{1}_{[x,y]}|f(z)|\,\mathrm{d}z \\
&\overset{\text{Hölders}}{\leq} \left\|\mathbf{1}_{[x,y]}\right\|_{L^{p'}(0,1)} \|f\|_{L^p(0,1)} \\
&= |x-y|^{1-\frac{1}{p}} \|f\|_{L^p(0,1)}.
\end{aligned}$$

Hence, for $1 - \frac{1}{p} > 0$ we have that $Tf \in \mathcal{C}^0(0,1)$. Moreover, we have that

$$\begin{aligned}
\|Tf\|_{\mathcal{C}^0(0,1)} &= \sup_{x\in(0,1)} \left|\int_0^x f(z)\,\mathrm{d}z\right| \\
&\leq \int_0^1 |f(z)|\,\mathrm{d}z \\
&\leq \|\mathbf{1}\|_{L^{p'}(0,1)} \|f\|_{L^p(0,1)} \\
&= \|f\|_{L^p(0,1)}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|Tf\|_{\mathcal{C}^{1-\frac{1}{p}}(0,1)} &= \|Tf\|_{\mathcal{C}^0(0,1)} + \sup_{x\neq y,(x,y)\in(0,1)^2} \frac{|Tf(x) - Tf(y)|}{|x-y|^{1-\frac{1}{p}}} \\
&\leq \|f\|_{L^p(0,1)} + \|f\|_{L^p(0,1)} < \infty.
\end{aligned}$$

Thus $Tf \in \mathcal{C}^{1-\frac{1}{p}}(0,1)$ and the operator T is well-defined. Moreover, this show that

$$\|T\|_{L^p(0,1)\to\mathcal{C}^{1-\frac{1}{p}}(0,1)} \leq 2.$$

Therefore, as T is a linear map we also deduce that T is continuous. Note that for all $f \in \bar{B}^{L^p(0,1)}$ we have that

$$|Tf(x) - Tf(y)| \leq |x-y|^{1-\frac{1}{p}},$$

hence, $T\left(\bar{B}^{L^p(0,1)}\right) \subseteq \mathcal{C}^0(0,1)$. Moreover, it follows that any sequence $(Tf_n)_{n\in\mathbb{N}} \subseteq T\left(\bar{B}^{L^p(0,1)}\right) \subseteq \mathcal{C}^0(0,1)$ is bounded and equicontinuous. Therefore, by Theorem 7.1.7 any sequence $(Tf_n)_{n\in\mathbb{N}} \subseteq T\left(\bar{B}^{L^p(0,1)}\right)$ admits a convergent subsequence. Thus, $T\left(\bar{B}^{L^p(0,1)}\right)$ is pre-compact, implying that $T : L^p(0,1) \to L^p(0,1)$ is a compact operator.

10.4 Weak Convergence in Hilbert Spaces

Definition 10.4.1. *Let H be a Hilbert space. A sequence $(x_n)_{n\in\mathbb{N}} \subseteq H$ weakly converges to $x \in H$ if*

$$(x_n, y) \to (x, y)$$

for all $y \in H$.

Remark 10.4.2.

1. *Symbolically one writes $x_n \rightharpoonup x$ to say that the sequence $(x_n)_{n\in\mathbb{N}} \subseteq H$ converges weakly to $x \in H$.*
2. *If $x_n \to x$ in the usual sense, then as*

$$|(x_n, y) - (x, y)| \leq \|x - x_n\|\|y\|$$

by Cauchy-Schwarz, it follows that $x_n \rightharpoonup x$.

Example 10.4.3. *In a finite-dimensional Euclidean space, the notions of strong and weak convergence are equivalent. In Remark 10.4.2 2. we saw that strong convergence implies weak convergence using the Cauchy-Schwarz inequality. Conversely, consider the finite-dimensional Euclidean space $\mathbb{R}^d$ and suppose that $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$ converges weakly to $x \in \mathbb{R}^d$. Then it follows that $(x_n, e_i) \overset{n\to\infty}{\longrightarrow} (x, e_i)$ where e_i is the $i^{th} \in \mathbb{R}^d$ is the i^{th} coordinate vector. This implies that $x_n^{(i)} \overset{n\to\infty}{\longrightarrow} x^{(i)}$ for each $i \in \{1, \dots, d\}$. Consequently,*

$$\|x_n - x\| \leq \sum_{i=1}^{d} \left| x_n^{(i)} - x^{(i)} \right| \overset{n\to\infty}{\longrightarrow} 0,$$

and so $x_n \to x$ strongly.

Theorem 10.4.4. *Let H be a Hilbert space. Then every bounded sequence $(x_n)_{n\in\mathbb{N}} \subseteq H$ has a weakly convergent subsequence.*

Proof. Let $M > 0$ be such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$. It follows by Cauchy-Schwarz that for fixed $m \in \mathbb{N}$ the sequence $(x_n, x_m)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ is bounded. Therefore, it has a convergent subsequence. By Cantor's diagonal argument we can find a subsequence $(x_{n_k})_{k\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$ such that $(x_{n_k}, x_m)_{k\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$ converges for every $m \in \mathbb{N}$ as $k \to \infty$. Consequently, for $y' \in \operatorname{span}(\{x_n\}_{n\in\mathbb{N}}) =: S$ it follows that $(x_{n_k}, y')_{k\in\mathbb{N}}$ converges as $k \to \infty$. Now consider $y \in \bar{S}$. For $y' \in S$ it follows that

$$\begin{aligned}\left|(x_{n_j} - x_{n_k}, y)\right| &\leq \left|(x_{n_j}, y - y')\right| + \left|(x_{n_j} - x_{n_k}, y')\right| + |(x_{n_k}, y' - y)| \\ &\leq 2M\|y - y'\| + \left|(x_{n_j} - x_{n_k}, y')\right|.\end{aligned}$$

Hence, given $\epsilon > 0$, let $y' \in S$ be such that $\|y' - y\| < \frac{\epsilon}{4M}$, and let j, k be large enough such that $\left|(x_{n_j} - x_{n_k}, j)\right|\| < \frac{\epsilon}{2}$. It follows that

$$\left|(x_{n_j} - x_{n_k}, y)\right| < \epsilon,$$

and so $\left|(x_{n_j} - x_{n_k}, y)\right| \to 0$ as $j, k \to \infty$. This implies that for $y \in \bar{S}$ the sequence (x_{n_k}, y) is Cauchy, and so has a limit. Let $Ly := \lim_{k\to\infty}(x_{n_k}, y)$. It is clear that $L : \bar{S} \to \mathbb{R}$ is linear. We also note that L is bounded using Cauchy-Schwarz and the fact that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$. Therefore, by Theorem 8.3.1 there exists an $x \in \bar{S}$ such that $(x, y) = Ly$ for all $y \in \bar{S}$. Now as $\bar{S}$ is closed we can write $H = \bar{S} \oplus \bar{S}^{\perp}$ by Proposition 8.2.4. Hence, for any $y \in H$ we can write $y = y_1 + y_2$, where $y_1 \in \bar{S}$ and $y_2 \in \bar{S}^{\perp}$. It follows that $(x_n, y) = (x_n, y_1)$ for all $n \in \mathbb{N}$. In particular, we have shown that $(x_{n_k}, y_1)_{k\in\mathbb{N}}$ converges for any $y_1 \in \bar{S}$ and so it follows that $(x_{n_k}, y)_{k\in\mathbb{N}}$ converges for any $y \in H$. Thus we have that the subsequence $(x_{n_k})_{k\in\mathbb{N}}$ converges weakly. ☐

Corollary 10.4.5. *Let H be a Hilbert space. If $(x_n)_{n\in\mathbb{N}} \subseteq H$ converges weakly to x, then*

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|.$$

Moreover, $\lim_{n\to\infty} \|x_n\| = \|x\|$ if and only if $x_n \to x$ strongly in H.

Proof. As

$$0 \leq (x_n - x, x_n - x) = \|x_n\|^2 - 2(x_n, x) + \|x\|^2 \tag{10.4.1}$$

and $(x_n, x) \to (x, x)$ as $n \to \infty$, it follows that

$$0 \leq \liminf \|x_n\|^2 - \|x\|^2.$$

Moreover, it is clear from (10.4.1) that if $\lim_{n\to\infty} \|x_n\| = \|x\|$ then $(x_n - x, x_n - x) \to 0$ which implies strong convergence. Conversely, by the triangle inequality, we know that $\|x_n - x\| \geq |\|x_n\| - \|x\||$, and so strong convergence implies $\lim_{n\to\infty} \|x_n\| = \|x\|$. $\square$

Definition 10.4.6. *Let H be a Hilbert space. A family $(e_n)_{n\in\mathbb{N}} \subseteq H$ is orthonormal if*

$$(e_n, e_m) = \delta_{nm}$$

for every $n, m \in \mathbb{N}$. If additionally,

$$x = \sum_{n\in\mathbb{N}} (x, e_n) e_n$$

for every $x \in H$, then the family is complete.

Example 10.4.7. *Consider the Hilbert space $L^2((-\pi, \pi))$ and the family $E = (e_n)_{n\in\mathbb{N}}$*

1. $e_1 = \frac{1}{\sqrt{2\pi}}$,
2. $e_{2n} = \frac{1}{\sqrt{\pi}} \sin(nx)$, *and*
3. $e_{2n+1} \frac{1}{\sqrt{\pi}} \cos(nx)$

for $n \geq 1$. One can show that E is an orthonormal family. Moreover, one can consider E as an orthonormal sequence in the infinite-dimensional Hilbert space $H = L^2((-\pi, \pi))$. Suppose that $(e_n)_{n\mathbb{N}}$ did not converge weakly to zero. Then we can choose a subsequence and an $x \in H$ such that

$$|(x, e_n)| \geq \epsilon \tag{10.4.2}$$

for all $n \in \mathbb{N}$ and some $\epsilon > 0$. Consider $E_m = \operatorname{span}(e_m)$, which is a closed subspace of H as it is finite-dimensional. Hence, by Proposition 8.2.4 $x = \lambda e_m + y$ for unique $\lambda \in \mathbb{R}$ and $y \in E_m^\perp$, where in particular λe_m is the projection of x onto E_m. Considering (x, e_m) we see that $\lambda = (x, e_m)$, and so $(x, e_m)e_m$ is the projection of x onto E_m. Similarly,

$$\sum_{n=1}^{N} (x, e_n) e_n$$

is the projection of x onto $E_{1,\ldots,N} := \operatorname{span}(e_1, \ldots, e_N)$. Thus using (10.4.2) it follows that

$$\|x\|^2 = \left\| x - \sum_{n=1}^{N} (x, e_n) e_n \right\|^2 + \left\| \sum_{n=1}^{N} (x, e_n) e_n \right\|^2 \geq \sum_{n=1}^{N} (x, e_n)^2 \geq N\epsilon^2$$

which is contradicts $\|x\|^2 < \infty$. Thus we conclude that $e_n \rightharpoonup 0$. In particular, we have shown that in the setting of Corollary 10.4.5 we cannot ask for equality. Moreover, $(e_n)_{n\in\mathbb{N}}$ is an example of a sequence that converges weakly, but whose norm does not converge to the norm of the limit, and so we do not have strong convergence.

Corollary 10.4.8 (Banach-Saks). *Let H be a Hilbert Space. Let $(x_n)_{n\in\mathbb{N}}$ be such that $\|x_n\| \leq K$ for all $n \in \mathbb{N}$. Then there exists a subsequence $\left(x_{n_j}\right)_{j\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$ and $x \in H$ such that*

$$\frac{1}{k}\sum_{j=1}^{k} x_{n_j} \overset{k\to\infty}{\longrightarrow} x$$

in H.

Proof. Let x be the weak limit of a subsequence $(x_{n_i})_{i\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$ as given by Theorem 10.4.4. Now consider the sequence $(y_i)_{i\in\mathbb{N}}$ given by $y_i := x_{n_i} - x$. It is clear that $y_i \rightharpoonup 0$ and $\|y_i\| \leq K'$ for some fixed K'. Consequently, one can choose a subsequence $\left(y_{i_j}\right)$ successively such that

$$\left|\left(y_{i_l}, y_{i_j}\right)\right| \leq \frac{1}{j}$$

for $l < j$. This is because for $j \in \mathbb{N}$ we have that $(y_{i_l}, y_i) \overset{i\to\infty}{\longrightarrow} 0$ for each $l < j-1$. Hence, there exists an I such that

$$\left|(y_{i_l}, y_i)\right| \leq \frac{1}{j}$$

for all $l < j$ and $i \geq I$. Thus, we can let $i_j = \max\left(I, i_{j-1}\right)$. Therefore,

$$\begin{aligned}
\left\|\frac{1}{k}\sum_{j=1}^{k} y_{i_j}\right\|^2 &= \frac{1}{k^2}\sum_{l,j=1}^{k}\left(y_{i_l}, y_{i_j}\right) \\
&= \frac{1}{k^2}\left(\sum_{j=1}^{k}\left(\left(y_{i_j}, y_{i_j}\right) + 2\sum_{l=1}^{j-1}\left(y_{i_l}, y_{i_j}\right)\right)\right) \\
&\leq \frac{1}{k^2}\left(k\left(K'\right)^2 + 2\sum_{j=1}^{k} j\frac{1}{j}\right) \\
&\leq \frac{\left(K'\right)^2 + 2}{k} \\
&\overset{k\to\infty}{\longrightarrow} 0.
\end{aligned}$$

$\square$

Lemma 10.4.9. *Let H be a Hilbert space. Then every weakly convergent sequence $(x_n)_{n\in\mathbb{N}} \subseteq H$ is bounded.*

Proof. Consider the sequence of linear functions $(L_n)_{n\in\mathbb{N}}$ given by $L_n y := (x_n, y)$. Now suppose that $(L_n)_{n\in\mathbb{N}}$ is not bounded on any closed ball of H. Then there exists a sequence $(K_i)_{i\in\mathbb{N}}$ of closed balls such that

1. $K_i := \{y : |y - y_i| \leq r_i\}$,
2. $K_{i+1} \subseteq K_i$, and
3. $r_i \to 0$.

Moreover, there exists a subsequence $(x_{n_i})_{i\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$ with $|L_{n_i} y| > i$ for all $y \in K_i$. Note that the $(y_i)_{i\in\mathbb{N}}$ form a Cauchy sequence and so have a limit $y_0 \in H$. As $y_0 \in \bigcap_{i=1}^{\infty} K_i$ it follows that $|L_{n_i} y_0| > i$ for all $i \in \mathbb{N}$. This contradicts the weak convergence of $(x_{n_i})_{i\in\mathbb{N}}$, and so there must exist a closed ball on which the linear functions $(L_n)_{n\in\mathbb{N}}$ are bounded. It follows by the linearity of the L_n that the set of linear functions $(L_n)_{n\in\mathbb{N}}$ is bounded on the closed unit ball, that is $\|L_n y\| = \|(x_n, y)\| \leq M$ for some $M > 0$ and for all $n \in \mathbb{N}$. In particular, letting $y = \frac{x_n}{\|x_n\|}$ it follows that

$$\|x_n\| = \left(x_n, \frac{x_n}{\|x_n\|}\right) \leq M$$

for all $n \in \mathbb{N}$, hence, the sequence $(x_n)_{n\in\mathbb{N}}$ is bounded. □

Corollary 10.4.10. *Let H be a Hilbert space. If $K \subseteq H$ is closed and convex, then K is closed with respect to weak convergence.*

Proof. Let $(x_n)_{\mathbb{N}} \subseteq K$ be weakly convergent to $x \in H$. Then by Lemma 10.4.9 the sequence $(x_n)_{n\in\mathbb{N}}$ is bounded, and by Corollary 10.4.8 there exists a subsequence $\left(x_{n_j}\right)_{j\in\mathbb{N}}$ such that

$$\frac{1}{k}\sum_{j=1}^{k} x_{n_j} \to x.$$

As K is convex we know that $\frac{1}{k}\sum_{j=1}^{k} x_{n_j} \in K$ for all j, so because K is closed it follows that $x \in K$. □

10.5 Sobolev Spaces

10.5.1 Weak Derivatives

Definition 10.5.1. *Let $[a,b] \subseteq \mathbb{R}$ be finite. Then a function $\psi \in \mathcal{C}^1([a,b])$ with $\psi(a) = \psi(b) = 0$ is referred to as a test function. The space of all such test functions on $[a,b]$ is denoted $\mathcal{C}_0^1([a,b])$.*

Lemma 10.5.2. *For $[a,b] \subseteq \mathbb{R}$ finite we have that*

$$\overline{\mathcal{C}_0^1([a,b])} = L^2(a,b).$$

In particular, if $g, h \in L^2(a,b)$ are such that

$$\int_a^b g(s)\psi(s)\,\mathrm{d}s = \int_a^b h(s)\psi(s)\,\mathrm{d}s$$

for all $\psi \in \mathcal{C}_0^1([a,b])$, then $g = h$ almost everywhere.

Definition 10.5.3. *Let $f \in L^2(a,b)$. A function $g \in L^2(a,b)$ is said to be the weak derivative of f if*

$$\int_a^b g(s)\psi(s)\,\mathrm{d}s = -\int_a^b f(s)\psi'(s)\,\mathrm{d}s$$

for all $\psi \in \mathcal{C}_0^1([a,b])$.

Remark 10.5.4. *From Lemma 10.5.2 it follows that if f has a weak derivative, then it is unique.*

10.5.2 The Fundamental Theorem of Calculus

For $f \in L^2(a,b)$ let

$$(Jf)(t) := \int_a^t f(x)\,\mathrm{d}x \tag{10.5.1}$$

for $t \in [a,b]$.

Lemma 10.5.5. *The operator $J : L^2(a,b) \to \mathcal{C}([a,b])$ is linear and bounded. Moreover, $(Jf)' = f$ in the weak sense for all $f \in L^2(a,b)$.*

Proof. For $t \in [a,b]$ let $(t_n)_{n\in\mathbb{N}} \subseteq [a,b]$ be such that $t_n \to t$. Clearly, $\mathbf{1}_{[a,t_n]}(x)f(x) \to \mathbf{1}_{[a,t]}(x)f(x)$ almost everywhere. Moreover, by Proposition 4.2.2 we have $L^2(a,b) \subseteq L^1(a,b)$ and so

$$\int_a^b |f(x)|\,\mathrm{d}x < \infty.$$

Therefore, as $|\mathbf{1}_{[a,t_n]}(x)f(x)| \leq |f(x)|$ almost everywhere it follows by the dominated convergence theorem that

$$(Jf)(t_n) = \int_a^b \mathbf{1}_{[a,t_n]}(x)f(x)\,\mathrm{d}x \overset{n\to\infty}{\longrightarrow} \int_a^b \mathbf{1}_{[a,t]}(x)f(x)\,\mathrm{d}x = (Jf)(t).$$

Therefore, $Jf \in \mathcal{C}([a,b])$ and J is well-defined. Moreover,

$$\begin{aligned}\|Jf\|_\infty &= \sup_{t\in[a,b]} \left|\int_a^t f(x)\,\mathrm{d}x\right| \\ &\leq \sup_{t\in[a,b]} \int_a^t |f(x)|\,\mathrm{d}x \\ &\leq \int_a^b |f(x)|\,\mathrm{d}x.\end{aligned}$$

In particular, using the Cauchy-Schwartz inequality it follows that

$$\|Jf\|_\infty \leq \sqrt{b-a}\,\|f\|_{L^2}. \tag{10.5.2}$$

Thus, J is bounded. Fix $\psi \in \mathcal{C}_0^1([a,b])$ and consider $T : L^2(a,b) \to \mathbb{R}$ given by

$$f \mapsto \langle Jf, \psi'\rangle + \langle f, \psi\rangle.$$

Then T is bounded and using integration by parts we have $Tf = 0$ for all $f \in \mathcal{C}^1([a,b])$. Since $\mathcal{C}^1([a,b])$ is dense in $L^2(a,b)$ we have $Tf = 0$ for all $f \in L^2(a,b)$. From this, we deduce that $\langle Jf, \psi'\rangle = -\langle f, \psi\rangle$ which is equivalent to saying that $(Jf)' = f$. ☐

Lemma 10.5.6. *For $F := \left\{\psi' : \mathcal{C}_0^1([a,b])\right\} \subseteq L^2(a,b)$ and $G := \mathbf{1}^\perp \subseteq L^2(a,b)$, show that $G = \bar{F}$.*

Proof. As $\psi \in \mathcal{C}_0^1([a,b])$ is such that $\psi(a) = \psi(b) = 0$, it is clear that $\mathbf{1} \in F^\perp$. Therefore, $F \subseteq \mathbf{1}^\perp = G$ and moreover $\bar{F} \subseteq \bar{G} = G$. On the other hand, by Proposition 8.2.4 we have

$$P_G f = (I - P_{\mathbb{R}\mathbf{1}})\,f \overset{\text{Ex } 8.4.8}{=} f - \frac{\langle f, \mathbf{1}\rangle}{b-a}\mathbf{1}$$

for $f \in L^2(a,b)$. In particular, for $f \in \mathcal{C}([a,b])$, then

$$JP_G f = Jf - \frac{\langle f, \mathbf{1}\rangle}{b-a}J\mathbf{1} \in \mathcal{C}([a,b]).$$

Note that $JP_G f(a) = 0$ by definition of J and similarly,

$$\begin{aligned}(JP_G f)(b) &= (Jf)(b) - \frac{\langle f, \mathbf{1}\rangle}{b-a}(J\mathbf{1})(b) \\ &= \int_a^b f(s)\,\mathrm{d}s - \frac{\langle f, \mathbf{1}\rangle}{b-a}\int_a^b \mathrm{d}s \\ &= \langle f, \mathbf{1}\rangle - \langle f, \mathbf{1}\rangle \\ &= 0.\end{aligned}$$

Therefore, $JP_G f \in \mathcal{C}_0^1([a,b])$ which implies that $P_G f \in F$ by Lemma 10.5.5. Hence as $\mathcal{C}([a,b])$ is dense in $L^2(a,b)$, there exists a sequence $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}([a,b])$ such that $f_n \to f$ in $L^2(a,b)$. As P_G is bounded, and thus continuous, we have $P_G f_n \to P_G f$ and so $P_G f \in \bar{F}$. Thus as $\operatorname{Ran}(P_G) = G$ we deduce that $G \subseteq \bar{F}$. □

Proposition 10.5.7. *The decomposition*

$$L^2(a,b) = \mathbb{R}\mathbf{1} \oplus \overline{\{\psi' : \psi \in \mathcal{C}_0^1([a,b])\}},$$

where closure is with respect to $\|\cdot\|_{L^2(a,b)}$.

Proof. Applying Proposition 8.2.4 with $\bar{F}$ and using Lemma 10.5.6, the result follows. □

Corollary 10.5.8. *Let $f \in L^2(a,b)$ be such that $f' = 0$ in the weak sense. Then f is constant almost everywhere.*

Proof. Let $f \in L^2(a,b)$ with $f' = 0$ in the weak sense. Then

$$\int_a^b f(s)\psi'(s)\,\mathrm{d}s = 0$$

for every $\psi \in \mathcal{C}_0^1([a,b])$. In other words, $f \in F^\perp$ where $F = \left\{\psi' : \psi \in \mathcal{C}_0^1([a,b])\right\}$ and in particular $f \in \bar{F}^\perp$. Using Lemma 10.5.6 we know that $\bar{F}^\perp = \left((\mathbb{R}\mathbf{1})^\perp\right)^\perp = \mathbb{R}\mathbf{1}$. Therefore $f \in \mathbb{R}\mathbf{1}$ and is thus constant almost everywhere. □

Corollary 10.5.9. *The inclusion $\mathrm{H}^1(a,b) \subseteq \mathcal{C}^0([a,b])$ holds. More specifically, $f \in \mathrm{H}^1(a,b)$ if and only if*

$$f = Jg + c\mathbf{1}$$

for $g \in L^2(a,b)$ and $c \in \mathbb{R}$ are uniquely given by

$$g = f'$$

and

$$c = \frac{\langle f - Jf', \mathbf{1}\rangle}{b-a}.$$

Moreover,

$$\int_c^d f'(s)\,\mathrm{d}s = f(d) - f(c) \tag{10.5.3}$$

for every $[c,d] \subseteq [a,b]$.

Proof. $(\Rightarrow)$. Let $f = Jg + c\mathbf{1}$ for $g \in L^2(a,b)$ and $c \in \mathbb{R}$. Then

$$f' = (Jg)' + 0 = g,$$

and so $f \in \mathrm{H}^1(a,b)$.
($\Leftarrow$). Let $f \in \mathrm{H}^1(a,b)$, and set $g := f'$, then

$$(f - Jg) = f' - (Jg)' \overset{\text{Lem 10.5.5}}{=} f' - g = 0.$$

Therefore, by Corollary 10.5.8 there exists a $c \in \mathbb{R}$ such that $f - Jg = c\mathbf{1}$.
For (10.5.3) note that

$$\begin{aligned} \int_c^d f'(s)\,\mathrm{d}s &= (Jg)(d) - (Jg)(c) \\ &= f(d) - \tilde{c}\mathbf{1} - f(c) + \tilde{c}\mathbf{1} \\ &= f(d) - f(c). \end{aligned}$$

□

10.5.3 Sobolev Spaces

Definition 10.5.10. *The first-order Sobolev space on $[a,b] \subseteq \mathbb{R}$ is*

$$\mathrm{H}^1(a,b) := \left\{ f \in L^2(a,b) : f \textit{ has a weak derivative} \right\}.$$

Lemma 10.5.11. *The map $\langle \cdot, \cdot \rangle_{\mathrm{H}^1} : \mathrm{H}^1(a,b) \to \mathbb{R}$ given by*

$$\langle f, g \rangle_{\mathrm{H}^1} = \langle f, g \rangle_{L^2} + \langle f', g' \rangle_{L^2},$$

where derivatives are in the weak sense, is an inner product on $\mathrm{H}^1(a,b)$.

Proof. The map $\langle \cdot, \cdot \rangle_{\mathrm{H}^1}$ is clearly symmetric as $\langle \cdot, \cdot \rangle_{L^2}$ is symmetric. By the linearity of the integral, if $f_1, f_2 \in \mathrm{H}^1(a,b)$ have weak derivatives g_1 and g_2 respectively, then $g_1 + \lambda g_2$ is the weak derivative of $f_1 + \lambda f_2$ for $\lambda \in \mathbb{R}$. Therefore, $f \mapsto f'$ is linear and thus $\langle \cdot, \cdot \rangle_{\mathrm{H}^1}$ is symmetric as $\langle \cdot, \cdot \rangle_{L^2}$ are symmetric. Similarly,

$$\langle f, f \rangle_{\mathrm{H}^1} = \langle f, f \rangle_{L^2} + \langle f', f' \rangle_{L^2} \geq 0$$

for all $f \in \mathrm{H}^1(a,b)$. Moreover, $\langle f, f \rangle_{\mathrm{H}^1} = 0$ if and only if $\langle f, f, \rangle_{L^2} = 0$ which happens if and only if $f = 0$. Therefore, $\langle \cdot, \cdot \rangle_{\mathrm{H}^1}$ is an inner product. □

Corollary 10.5.12. *The map $\| \cdot \|_{\mathrm{H}^1} : \mathbb{H}^1(a,b) \to \mathbb{R}$ given by*

$$\|f\|_{\mathrm{H}^1} := \langle f, f \rangle_{\mathrm{H}^1}$$

Proof. This follows from Lemma 10.5.11 and Proposition 8.1.5. □

Lemma 10.5.13. *The map $H^1(a,b) \to L^2(a,b)$ given by $f \mapsto f'$ is linear and bounded.*

Proof. Linearity was noted in the proof of 10.5.11. Moreover, as

$$\|f'\|_{L^2}^2 \leq \|f\|_{L^2}^2 + \|f'\|_{L^2}^2 = \|f\|_{\mathrm{H}^1}^2,$$

it follows that $f \mapsto f'$ is bounded. □

Theorem 10.5.14. *For $[a,b] \subseteq \mathbb{R}$ the following statements hold.*

1. *The space $\left(\mathrm{H}^1(a,b), \| \cdot \|_{\mathrm{H}^1}\right)$ is a Hilbert space.*

2. The inclusion $\left(\mathrm{H}^1(a,b), \|\cdot\|_{\mathrm{H}^1}\right) \subseteq \left(\mathcal{C}^0([a,b]), \|\cdot\|_\infty\right)$ *is continuous.*

Proof.

1. Consider $(f_n)_{n\in\mathbb{N}} \subseteq \mathrm{H}^1(a,b)$ a Cauchy sequence. That is,

$$\|f_n - f_m\|_{L^2}^2 + \|f' - f_m'\|_{L^2}^2 = \|f_n - f_m\|_{\mathrm{H}^1}^2 \overset{n,m\to\infty}{\longrightarrow} 0.$$

Therefore, $(f_n)_{n\in\mathbb{N}}, (f_n')_{n\in\mathbb{N}} \subseteq L^2(a,b)$ are Cauchy sequence and thus convergent. Let $f, \in L^2(a,b)$ be such that $f_n \to f$ and $f_n' \to g$ in $L^2(a,b)$. Then for $\psi \in \mathcal{C}_0^1([a.b])$, by using the fact that f_n' is the weak derivative of f_n it follows that

$$\langle g, \psi\rangle_{L^2} = \lim_{n\to\infty} \langle f_n', \psi\rangle_{L^2} = \lim_{n\to\infty} \left(-\langle f_n, \psi'\rangle\right) = -\langle f, \psi'\rangle .$$

In other words, g is the weak derivative of f. Therefore, $f \in \mathrm{H}^1(a,b)$ with f_n converging to f in $\mathrm{H}^1(a,b)$. Therefore, $\mathrm{H}^1(a,b)$ is complete and thus a Hilbert space.

2. For $f \in \mathrm{H}^1(a,b)$, using Corollary 10.5.9, we can write

$$f = Jg + c\mathbf{1}$$

for $g \in L^2(a,b)$ and $c \in \mathbb{R}$, where $c = \frac{\langle f - Jf' \rangle, \mathbf{1}\rangle}{b-a}$ and $g = f'$. Therefore,

$$\|f\|_\infty = \|Jg + c\mathbf{1}\|_\infty \le \|Jg\|_\infty + |c|.$$

Using (10.5.2) we have that

$$\|Jg\|_\infty \le \sqrt{b-a}\|g\|_{L^2} = \sqrt{b-a}\,\|f'\|_{L^2} .$$

Similarly,

$$\begin{aligned}
|c| &= \frac{1}{b-a}\left|\int_a^b f - Jf' \,\mathrm{d}x\right| \\
&\le \frac{1}{b-a}\left(\|f\|_{L^1} + \|Jf'\|_{L^1}\right) \\
&\overset{\text{C.S}}{\le} \frac{1}{b-a}\|\mathbf{1}\|_{L^2}\left(\|f\|_{L^2} + \|Jf'\|_{L^2}\right) \\
&\overset{(10.5.2)}{\le} \frac{1}{\sqrt{b-a}}\left(\|f\|_{L^2} + \sqrt{b-a}\,\|f'\|_{L^2}\right).
\end{aligned}$$

It follows that

$$\|f\|_\infty \le c\|f\|_{\mathrm{H}^1}$$

where

$$c := \max\left(\sqrt{b-a} + 1, \frac{1}{\sqrt{b-a}}\right)$$

is only dependent on $b-a$. As this holds for all $f \in \mathrm{H}^1(a,b)$, the inclusion map $\left(\mathrm{H}^1(a,b), \|\cdot\|_{\mathrm{H}^1}\right) \to \left(\mathcal{C}^0([a,b]), \|\cdot\|_\infty\right)$ is bounded and thus continuous.

□

Lemma 10.5.15. *The map* $J : L^2(a,b) \to \mathrm{H}^1(a,b)$*, given by* (10.5.1)*, is bounded.*

Proof. Note that

$$\begin{aligned}
\|J\|_{L^2(a,b)\to \mathrm{H}^1(a,b)} &= \sup_{f\in L^2(a,b)\setminus\{0\}} \frac{\|Jf\|_{\mathrm{H}^1}}{\|f\|_{L^2}} \\
&= \sup_{f\in L^2(a,b)\setminus\{0\}} \frac{\sqrt{\|Jf\|_{L^2}^2 + \left\|(Jf)'\right\|_{L^2}^2}}{\|f\|_{L^2}} \\
&\overset{\text{Lem 10.5.5}}{=} \sup_{f\in L^2(a,b)\setminus\{0\}} \frac{\sqrt{\|Jf\|_{L^2}^2 + \|f\|_{L^2}^2}}{\|f\|_{L^2}}.
\end{aligned}$$

Using (10.5.2) and Proposition 4.1.6 we have

$$\|Jf\|_{L^2} \le \|Jf\|_\infty \|\mathbf{1}\|_{L^2} \le (b-a)\|f\|_{L^2}.$$

Therefore,

$$\|J\|_{L^2(a,b)\to \mathrm{H}^1(a,b)} \le \sup_{f\in L^2(a,b)\setminus\{0\}} \frac{\sqrt{(b-a)^2+1}\|f\|_{L^2}}{\|f\|_{L^2}} = \sqrt{(b-a)^2+1} < \infty.$$

□

Corollary 10.5.16. *For $[a,b]\subseteq\mathbb{R}$, we have $\overline{\mathcal{C}^1([a,b])} = H^1(a,b)$.*

Proof. Let $f\in \mathrm{H}^1(a,b)$ with $g := f'$. Then by Corollary 10.5.9 we have $f = Jg + c\mathbf{1}$ for some $c\in\mathbb{R}$. Since, $\overline{\mathcal{C}^0([a,b])} = L^2(a,b)$ there exists a sequence $(g_n)_{n\in\mathbb{N}}\subseteq \mathcal{C}^0([a,b])$ such that $\|g_n - g\|_2 \overset{n\to\infty}{\longrightarrow} 0$. Then using Lemma 10.5.15 we have that $Jg_n \overset{n\to\infty}{\longrightarrow} Jg$ in $\mathrm{H}^1(a,b)$. Therefore,

$$f_n := Jg_n + c\mathbf{1} \overset{n\to\infty}{\longrightarrow} Jg + c\mathbf{1} = f$$

in $\mathrm{H}^1(a,b)$. As $(f_n)_{n\in\mathbb{N}}\subseteq \mathcal{C}^1([a,b])$ we conclude that $\mathcal{C}^1([a,b])$ is dense in $\mathrm{H}^1(a,b)$. □

Definition 10.5.17. *The n^{th} order Sobolev space, for $n\ge 2$, on $[a,b]\subseteq\mathbb{R}$ is*

$$\mathrm{H}^n(a,b) := \left\{ f\in \mathrm{H}^1(a,b) : f'\in \mathrm{H}^{n-1}(a,b)\right\}.$$

Proceeding in the same way as Lemma 10.5.11 we have that

$$\langle f,g\rangle_{\mathrm{H}^n} := \sum_{k=0}^{n} \left\langle f^{(k)}, g^{(k)}\right\rangle_{L^2}$$

is an inner product on $\mathrm{H}^n(a,b)$, and thus

$$\|f\|_{\mathrm{H}^n} = \sqrt{\|f\|_{L^2}^2 + \cdots + \left\|f^{(n)}\right\|_{L^2}^2}$$

is a norm on $\mathrm{H}^n(a,b)$. Moreover, we have that $(\mathrm{H}^n(a,b), \|\cdot\|_{\mathrm{H}^n})$ is a Hilbert space.

Definition 10.5.18. *For $[a,b]\subseteq\mathbb{R}$, let*

$$\mathrm{H}_0^1(a,b) := \mathrm{H}^1(a,b)\cap \mathcal{C}_0([a,b]),$$

where

$$\mathcal{C}_0^0([a,b]) := \{\varphi : \varphi(a) = \varphi(b)\} \cap \mathcal{C}^0([a,b]).$$

Lemma 10.5.19. *The map* $\langle\cdot,\cdot\rangle_{\mathrm{H}_0^1}:\mathrm{H}_0^1(a,b)\to\mathbb{R}$ *given by*

$$\langle u,v\rangle_{\mathrm{H}_0^1}=\langle u',v'\rangle_{L^2}$$

is an inner-product on $\mathrm{H}_0^1(a,b)$*. Moreover,* $\|\cdot\|_{\mathrm{H}_0^1}:\mathrm{H}_0^1(a,b)\to\mathbb{R}$ *given by*

$$\|u\|_{\mathrm{H}_0^1}=\|u'\|_{L^2}$$

is a norm on $\mathrm{H}_0^1(a,b)$*.*

Proof. As J is linear, from Lemma 10.5.13, and $\langle\cdot,\cdot\rangle_{L^2}$ is an inner product it is clear that $\langle\cdot,\cdot\rangle_{\mathrm{H}_0^1}$ is bilinear and symmetric. Moreover, $\langle u,u\rangle_{\mathrm{H}_0^1}=0$ if and only if $u'=0$ which happens if and only if $u=c$ for some constant c by Corollary 10.5.8. As $u\in\mathrm{H}_0^1$ we have that $u(a)=0$ which implies that $c=0$. Hence, $\langle u,u\rangle_{\mathrm{H}_0^1}\geq 0$ with equality if and only if $u=0$. Therefore, $\langle\cdot,\cdot\rangle_{\mathrm{H}_0^1}$ is an inner product and so by Proposition 8.1.5 we have that $\|\cdot\|_{\mathrm{H}_0^1}$ is a norm. □

Remark 10.5.20. *The norm of* H_0^1 *is often referred to as the energy norm.*

Lemma 10.5.21. *There exists a* $C\geq 0$*, dependent on* $b-a$*, such that for all* $u\in\mathrm{H}_0^1(a,b)$ *we have*

$$\|u\|_{L^2(a,b)}\leq C\,\|u'\|_{L^2(a,b)}\,.$$

Proof. Let $u\in\mathrm{H}_0^1(a,b)$. Then by Lemma 10.5.5 we have that $(Ju')'=u'$ and so by Corollary 10.5.8 it follows that $Ju'-u=c$ almost everywhere. As $Ju'-u$ vanishes at a it follows that $c=0$. Therefore, $Ju'=u$ and so

$$\|u\|_{L^2}=\|Ju'\|_{L^2}\leq C\,\|u'\|_{L^2}\,,$$

where C is given by the boundedness of J shown in Lemma 10.5.5. □

Remark 10.5.22. *From Lemma 10.5.21 it follows that on* $\mathrm{H}_0^1(a,b)$*, the norms* $\|\cdot\|_{\mathrm{H}^1}$ *and* $\|\cdot\|_{\mathrm{H}_0^1}$ *are equivalent. Indeed,*

$$\|u\|_{\mathrm{H}_0^1}\leq\|u\|_{\mathrm{H}^1}$$

and

$$\|u\|_{H_1}^2=\|u\|_2^2+\|u'\|_2^2\leq\left(C^2+1\right)\|u'\|_2^2=\left(C^2+1\right)\|u\|_{\mathrm{H}_0^1}.$$

Theorem 10.5.23. *The space* $\mathrm{H}_0^1(a,b)$ *is* $\|\cdot\|_{\mathrm{H}^1}$*-closed in* $\mathrm{H}^1(a,b)$ *and a Hilbert space with respect to* $\|\cdot\|_{\mathrm{H}_0^1}$*.*

Proof. Let $(f_n)_{n\in\mathbb{N}}\subseteq\mathrm{H}_0^1(a,b)$ be convergent to f with respect to $\|\cdot\|_{\mathrm{H}^1}$. As $\left(\mathrm{H}^1(a,b),\|\cdot\|_{\mathrm{H}^1}\right)$ is a Hilbert space and $(f_n)_{n\in\mathbb{N}}\subseteq\mathrm{H}^1(a,b)$ it follows that $f\in\mathrm{H}^1(a,b)$. Similarly, by statement 2. of Theorem 10.5.14 we have that

$$\|f\|_\infty\leq c\|f\|_{\mathrm{H}^1}$$

for all $f\in\mathrm{H}^1(a,b)$ and some $c>0$. Therefore, $(f_n)_{n\in\mathbb{N}}\subseteq\mathrm{H}_0^1(a,b)\subseteq\mathcal{C}_0^0([a,b])$ converges to f with respect to $\|\cdot\|_\infty$. As $\left(\mathcal{C}_0^0([a,b]),\|\cdot\|_\infty\right)$ is a Hilbert space we must have $f\in\mathcal{C}_0^0([a,b])$ and so $f\in\mathrm{H}_0^1(a,b)$. Showing that $\mathrm{H}_0^1(a,b)$ is closed. Consequently, as $\|\cdot\|_{\mathrm{H}_0^1}$ and $\|\cdot\|_{\mathrm{H}^1}$ coincide on H_0^1 we can conclude that $\mathrm{H}_0^1(a,b)$ is a Hilbert space. □

References

[1] Markus Haase. *Functional Analysis: An Elementary Introduction*. 2014.