Some notes on functional analysis

Brian Chan

May 22, 2024

Contents

0.1 Purpose

(DUE FOR A REVAMP) This document serves two purposes. Firstly, it serves as a summary of what I have learnt from functional analysis. Secondly, typesetting this document allows me to learn LaTeX. Essentially, we are "killing two birds with one stone" with this endeavour. The main reference for my studies are "Lecture Notes on Functional Analysis and Linear Partial Differential Equations" by Alberto Bressan (see [AB10]).

Chapter 1

Basic Definitions

1.1 Normed vector spaces

Generally speaking, a norm is a function which allows us to measure magnitude. A norm could refer to the norm of a Euclidean domain or the norm induced by the inner product of a Hilbert space. For now, we will focus on equipping a vector space V with a norm, giving rise to the definition of a normed vector space.

We will assume that the underlying field of the vector space, denoted by $\mathbb{K}$, is either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$.

Definition 1.1.1. A **normed vector space** is a vector space V, equipped with a function $\|.\| : V \to \mathbb{R}_{\geq 0}$ (called a *norm*) such that the following axioms are satisfied:

1. $\forall x \in V$ and $\forall \lambda \in \mathbb{K}$, $\|\lambda x\| = |\lambda| \, \|x\|$ (Scalar Multiplication).
2. $\forall x \in V$, $\|x\| = 0$ if and only if $x = 0$ (Positive definiteness).
3. $\forall x, y \in V$, $\|x + y\| \leq \|x\| + \|y\|$ (Triangle Inequality).

We will also like to quantify the notion of "distance" between two vectors. Fortunately, we can use the norm to do this. A natural definition of distance is

$$d(x, y) = \|x - y\| \qquad \forall x, y \in V \tag{1.1}$$

From the axioms satisfied by the norm, we can determine the axioms $d(x, y)$ satisfies. This leads to the important definition of a metric space

Definition 1.1.2. A **metric space** (X, d) is a set X, equipped with a function $d : X \times X \to \mathbb{R}_{\geq 0}$ (called a *metric*) which satisfies the following axioms:

1. $\forall x, y \in X, d(x, y) = d(y, x)$ (Symmetry).
2. $\forall x, y \in X$, $d(x, y) = 0$ if and only if $x = y$ (Positive definiteness).
3. $\forall x, y, z \in X$, $d(x, z) \leq d(x, y) + d(y, z)$ (Triangle Inequality).

Note that in the definition of a metric space, we are equipping a *set* with a metric, whereas for normed vector spaces, we are equipping a *vector space* with a norm. Now, we will focus on normed vector spaces and give illustrative examples.

Example 1.1.1. The vector space $\mathbb{R}^n$ with the *Euclidean norm* $\|(x_1, x_2, \ldots, x_n)\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ is a normed vector space.

Example 1.1.2. Once again, consider the vector space $\mathbb{R}^n$. Let $1 \leq p < \infty$. Then, the norms $\|(x_1, x_2, \ldots, x_n)\|_p = (x_1^p + x_2^p + \cdots + x_n^p)^{\frac{1}{p}}$ and $\|(x_1, x_2, \ldots, x_n)\|_\infty = \max(|x_1|, |x_2|, \ldots, |x_n|)$ also turn $\mathbb{R}^n$ into normed vector spaces.

Example 1.1.3. Consider the vector space of real number sequences $l^p = \{(x_1, x_2, \ldots, x_n, \ldots) | \, x_i \in \mathbb{R}, \sum_{i=1}^{\infty} |x_i|^p < \infty\}$. This is a normed vector space with the norm defined as $\|x\| = (\sum_{i=1}^{\infty} |x_i|^p)^{\frac{1}{p}}$.

Example 1.1.4. Consider the vector space of bounded real number sequences $l^\infty = \{(x_1, x_2, \ldots, x_n, \ldots) | \, x_i \in \mathbb{R}, \forall x_i \in \mathbb{R} \, x_i < \infty\}$. This is also a normed vector space with the norm defined as $\|x\| = \sup\{|x_1|, |x_2|, \ldots, |x_n|, \ldots\}$.

Example 1.1.5. Let X be an open subset of $\mathbb{R}$. Then, consider the vector space $L^p = \{f : X \to \mathbb{R} \mid |f|^p$ is Lebesgue integrable and $\int_X |f|^p \, dx = 0$ if and only if $f = 0\}$. This is a normed vector space with norm defined as $\|f\| = (\int_X |f|^p \, dx)^{\frac{1}{p}}$

It is worth observing that (1.1) provides a natural, intuitive way of defining a metric for vector spaces. This allows us to utilise topological concepts in order to analyse normed vector spaces. We will begin with basic definitions first.

Let X be a normed vector space. Let $x \in X$. We define the **open ball** centred at x with radius r as follows:

$$B(x,r) = \{y \in X | \ \|y - x\| < r\} \tag{1.2}$$

Similarly, the **closed ball** centred at x with radius r is defined as

$$\overline{B}(x,r) = \{y \in X | \ \|y - x\| \leq r\} \tag{1.3}$$

Definition 1.1.3. Let X be a set and $x \in X$. If for some $r \in \mathbb{R}_{>0}, B(x,r) \subset V$, then V is called a **neighbourhood** of the point x.

Definition 1.1.4. Let $V \subseteq X$ be a set. We say that V is an **open set** if for all $x \in V$, there exists a $r \in \mathbb{R}_{>0}$ such that $B(x,r) \subset V$.

Definition 1.1.5. Let $V \subseteq X$ be a set. If for some $x \in V$, there exists a $r \in \mathbb{R}_{>0}$ such that $B(x,r) \subset V$, then x is called an **interior point** of V.

Definition 1.1.6. Let $V \subseteq X$ be a set. The **interior** V° of V is the set of all interior points of V.

An alternative definition of an open set uses the definitions above. A set $V \subseteq X$ is open if $V = V^\circ$.

Theorem 1.1.1. *Let X be a set and $V \subseteq X$ be a subset of V. Then, V° is the largest open set that is contained within V.*

Proof. Assume that X is a set. Assume that $V \subset X$ is an open subset of X.

To show: (a) $V \subset X^\circ$.

(b) X° is an open set.

(a) Assume that $v \in V$. Since V is an open subset, there exists $\epsilon \in \mathbb{R}_{>0}$ such that the open ball $B(v,\epsilon) \subset V$. Since $V \subset X$, $B(v,\epsilon) \subset X$. So, there exists $\epsilon \in \mathbb{R}_{>0}$ such that $B(v,\epsilon) \subset X$. Hence, $v \in X^\circ$. So, $V \subset X^\circ$

(b) Assume that $x \in X^\circ$. Then, there exists a $\epsilon \in \mathbb{R}_{>0}$ such that the open ball $B(x,\epsilon) \subset X$. Assume that $y \in B(x,\epsilon)$. We can use part (a) to deduce that $y \in X^\circ$. Hence, $B(x,\epsilon) \subset X^\circ$ and subsequently, X° is an open subset of X.

Consequently X° is the largest open set contained in X. □

Definition 1.1.7. Let $V \subseteq X$ be a set. We say that V is a **closed set** if its complement $X \backslash V$ is open.

Definition 1.1.8. Let $x \in X$. We say that x is an **adherent point** of $V \subseteq X$ if for all $r \in \mathbb{R}_{>0}, B(x,r) \cap V \neq \emptyset$.

Definition 1.1.9. Once again, let $V \subseteq X$ be a subset of X. The set of all adherent points of V is called the **closure** of V, denoted by $\overline{V}$.

Similarly to open sets, we can use the closure of a set to give an alternative definition of a closed set. We say that $V \subseteq X$ is a closed set if $V = \overline{V}$. Adherent points are also referred to as *accumulation points* or *cluster points*.

Theorem 1.1.2. *Let X be a set and $V \subseteq X$ be a subset of V. Then, $\overline{V}$ is the smallest closed set containing V.*

Proof. Assume that X is a set. Assume that $X \subset V$, where V is a closed set.

To show: (a) $\overline{X} \subset V$.

(b) $\overline{X}$ is a closed set.

(a) Assume that $x \in \overline{X}$. Then, for all $r \in \mathbb{R}_{>0}$, $B(x,r) \cap X \neq \emptyset$. Since X is a subset of V, we can further conclude that $B(x,r) \cap V \neq \emptyset$. This means that x is an adherent point of V and so, $x \in \overline{V}$. Finally, since V is a closed set, $V = \overline{V}$. So, $x \in V$ and therefore, $\overline{X} \subset V$.

(b) Assume that $y \in \overline{X}$. Then, for all $r \in \mathbb{R}_{>0}$, $B(y,r) \cap X \neq \emptyset$. Assume that the point $z \in B(y,r) \cap X$. Then, $z \in X \subset \overline{X}$. Additionally, $z \in B(y,r)$. So, $B(y,r) \cap \overline{X} \neq \emptyset$. Hence, y is an adherent point of $\overline{X}$ and subsequently, $\overline{X}$ is a closed set.

So, $\overline{X}$ is the smallest closed set containing X. □

Note that if a given set is not open, it is not necessarily the case that it is closed and vice versa. For example, $\mathbb{C}$ is both open and closed, whereas the interval $(0,1] \subset \mathbb{R}$ is neither open nor closed.

We will end this section by proving that L^p is indeed a normed vector space.

Proof. Before we delve into verifying the axioms of a normed vector space, we will prove the following identity first.

To show: (a) For all $a, b \in \mathbb{R}_{>0}$ and $p \in [1, \infty)$,

$$\inf_{0<t<1} [t^{1-p}a^p + (1-t)^{1-p}b^p] = (a+b)^p.$$

(a) Assume that $a, b \in \mathbb{R}_{>0}$ and $p \in [1, \infty)$. Define the function

$$g(t) = t^{1-p}a^p + (1-t)^{1-p}b^p.$$

Then, we differentiate to obtain

$$g'(t) = (1-p)t^{-p}a^p - (1-p)(1-t)^{-p}b^p.$$

$g'(t)$ vanishes if and only if $t = t_1 = a/(a+b)$. To determine the nature of this stationary point, we compute $g''(t_1)$ to be

$$g''(t_1) = -[(1-p)p{t_1}^{-p-1}a^p + (1-p)p(1-t_1)^{-p-1}b^p] > 0.$$

Hence, $g(t)$ has a local minimum at t_1, based off the convexity of the stationary point. Note that $g(t_1) = (a+b)^p$. So, we have

$$\inf_{0<t<1} [t^{1-p}a^p + (1-t)^{1-p}b^p] = (a+b)^p.$$

Now we are ready to prove that L^p satisfies the axioms of a normed vector space.

To show: (b) For all $f \in L^p, \|f\|_p = 0$ if and only if $f = 0$.

(c) For all $f \in L^p$ and $\lambda \in \mathbb{R}$, $\|\lambda f\|_p = |\lambda| \|f\|_p$.

(d) For all $f, g \in L^p$, $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

(b) From the definition of the integral, we have that for all $f \in L^p$, $\|f\|_p \geq 0$. When $\|f\|_p = 0$, this means that

$$\int_X |f|^p \, \mathrm{d}x = 0.$$

Once again, by the positive definiteness of the integral (integral pair), $|f|^p = 0$. Therefore, $f = 0$. Conversely, if $f = 0$, then $|f|^p = 0$ and consequently, $\|f\|_p = 0$.

(c) Assume that $\lambda \in \mathbb{R}$. Then, we proceed as follows

$$\begin{aligned}
\|\lambda f\|_p &= (\int_X |\lambda f|^p \mathrm{d}x)^{1/p} \\
&= (|\lambda|^p \int_X |f|^p \mathrm{d}x)^{1/p} \\
&= |\lambda| (\int_X |f|^p \mathrm{d}x)^{1/p} \\
&= |\lambda| \|f\|_p.
\end{aligned}$$

Hence, $\|\lambda f\|_p = |\lambda| \|f\|_p$.

(d) Assume that $f, g \in L^p$. Then we use the identity proven in part (a) to argue as follows:

$$\begin{aligned}
\|f+g\|_p^p &= \int_X |f+g|^p \mathrm{d}x \\
&\leq \int_X (|f| + |g|)^p \mathrm{d}x \\
&\leq \int_X [t^{1-p}|f|^p + (1-t)^{1-p}|g|^p] \mathrm{d}x \quad (t \in (0,1)) \\
&= t^{1-p} \int_X |f|^p \mathrm{d}x + (1-t)^{1-p} \int_X |g|^p \mathrm{d}x. \\
&= t^{1-p} \|f\|_p^p + (1-t)^{1-p} \|g\|_p^p.
\end{aligned}$$

Note that this holds for all $t \in (0,1)$. Hence, we can deduce that

$$\|f+g\|_p^p \leq \inf_{t \in (0,1)} [t^{1-p} \|f\|_p^p + (1-t)^{1-p} \|g\|_p^p] = (\|f\|_p + \|g\|_p)^p.$$

Taking the p^{th} root of both sides gives us the triangle inequality we are after.

Therefore, L^p is a normed vector space.

□

1.2 Banach Spaces and Completeness

The notions of convergence, continuity and completeness are critically important ideas in analysis and topology. Using the norm of a normed vector space, we will define the convergence of sequences, continuous functions, Cauchy sequences and completeness. Then, we will investigate an important subset of normed vector spaces - **Banach spaces**.

Definition 1.2.1. Let X be a normed vector space and let $\{x_n\}_{n\geq 1}$ be a sequence of points in X. We say that $\{x_n\}$ **converges** to the point x if for all $\epsilon > 0$, there exists a $N \in \mathbb{R}_{>0}$ such that for all $n > N$,

$$\|x_n - x\| < \epsilon.$$

We also write $\lim_{n\to\infty}\|x_n - x\| = 0$.

Closely related to the notion of convergence is the notion of a *Cauchy sequence*.

Definition 1.2.2. Let X be a normed vector space and let $\{x_n\}_{n\geq 1}$ be a sequence of points in X. We say that $\{x_n\}$ is a **Cauchy sequence** if for all $\epsilon > 0$, there exists a $N \in \mathbb{R}_{>0}$ such that for all $m, n > N$,

$$\|x_n - x_m\| < \epsilon.$$

Theorem 1.2.1. *Let X be a normed vector space and let $\{x_n\}$ be a sequence of points in X, converging to a point $x \in X$. Then, $\{x_n\}$ is a Cauchy sequence.*

Proof. Assume X is a normed vector space and $\{x_n\}$ is a convergent sequence of points in X, converging to $x \in X$. Then, for all $\epsilon > 0$, there exists a $N \in \mathbb{R}_{>0}$ such that for all $n > N$,

$$\|x_n - x\| < \frac{\epsilon}{2}.$$

Similarly, for all $\epsilon > 0$, there exists a $M \in \mathbb{R}_{>0}$ such that for all $m > M$

$$\|x_m - x\| < \frac{\epsilon}{2}.$$

Hence, we define $N_2 = \min\{M, N\}$. So, for all $\epsilon > 0$ and $m, n > N_2$,

$$\begin{aligned}
\|x_n - x\| &= \|x_n - x + x - x_m\| \\
&\leq \|x_n - x\| + \|x_m - x\| \quad \text{(Triangle Inequality)} \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}$$

Therefore, the sequence is Cauchy. □

Note that these definitions also apply to metric spaces. It is natural to ask whether the converse of the previous theorem is true. In general, Cauchy sequences are not necessarily convergent. However, satisfying the converse statement leads to a very special property normed vector spaces/metric spaces can have.

Definition 1.2.3. Let X be a normed vector space. X is **complete** if every Cauchy sequence in X converges (same definition works for metric spaces).

Definition 1.2.4. A **Banach space** is a normed vector space which is also complete with respect to the norm.

Each of the normed vector spaces given as examples on page 3 are all Banach spaces. We will now list some important theorems with proofs pertaining to completeness.

Theorem 1.2.2. *Let X be a metric space and let M be a non-empty subset of X. Then, $x \in \overline{M}$ if and only if there exists a sequence $\{x_n\}$ in M such that $x_n \to x$.*

Proof. To show: (a) If $x \in \overline{M}$, then there exists a sequence $\{x_n\}$ in M such that $x_n \to x$.

(b) If we have a sequence $\{x_n\}$ in M such that $x_n \to x$, then $x \in \overline{M}$.

(a) Assume X is a metric space with metric d. Assume M is a non-empty subset of X and that $x \in \overline{M}$. Then, this divides into two cases. Firstly, consider the case where $x \in M$. Then, we can define the sequence $\{x_n\} = (x, x, x, \dots)$. Since $x \in M$, the sequence $\{x_n\} \in M$ and it converges to x.

Next, consider the case where $x \notin M$. Observe that since $x \in \overline{M}$, x must be an adherent point of M. Now consider the set of open balls $B(x, \frac{1}{n})$, where n is a positive integer. We choose the points x_n such that for all $n \in \mathbb{Z}_{>0}, x_n \in B(x, \frac{1}{n})$. Define the sequence $\{x_n\} = (x_1, x_2, x_3, \dots)$. Due to our construction, this sequence converges to x because $d(x_n, x) \leq \frac{1}{n} \to 0$ as $n \to \infty$.

(b) Assume that we have a sequence $\{x_n\}$ in M such that $x_n \to x$. Then, either $x \in M$ or x is an adherent point because every neighbourhood of x contains at least one point x_n in the sequence $\{x_n\}$. In either case, $x \in \overline{M}$. □

Furthermore, note that if $x \in M$ in part (b), then M must be a closed subset of X.

Theorem 1.2.3. *Let (X, d) be a complete metric space and $M \subset X$. Then, the subspace (M, d) is complete if and only if M is closed.*

Proof. To show: (a) If M is a complete subspace with d as the metric, then M is closed.

(b) If M is closed, then (M, d) is complete.

(a) Assume (X, d) is a complete metric space. Assume (M, d) is a complete subspace. Then, every Cauchy sequence in M converges. We already know that $M \subseteq \overline{M}$.

To show: (aa) $\overline{M} \subseteq M$.

(aa) Assume that $x \in \overline{M}$. Utilising the previous theorem, we know that there exists a sequence $\{x_n\}$ in M such that $x_n \to x$. Since $\{x_n\}$ is convergent, $\{x_n\}$ is a Cauchy sequence. Since (M, d) is a complete subspace by assumption, $\{x_n\}$ converges inside M. Therefore, $x \in M$ and consequently, $\overline{M} \subseteq M$.

(a) Since $\overline{M} \subseteq M$ and $M \subseteq \overline{M}$, $M = \overline{M}$ and so M is closed.

(b) For the converse, assume that M is a closed subset of X. Let $\{x_n\}$ be a Cauchy sequence in M. Due to the completeness of X, $\{x_n\}$ must converge to (say) x. From the previous theorem, $x \in \overline{M}$. Since M is closed, $x \in M$. Since our choice of Cauchy sequence was arbitrary, we deduce that every Cauchy sequence in M converges to a point in M. Therefore, (M, d) is a complete metric subspace. □

Example 1.2.1. Consider $\mathbb{R}$ as a *metric space*, equipped with the usual Euclidean metric $d(x, y) = |x - y|$. Consider the closed interval $[0, 1] \in \mathbb{R}$. From the theorem above, we know that the subspace $([0, 1], d)$ must be complete because $(\mathbb{R}, d)$ is a complete metric space. On the other hand, the subspace $((0, 1), d)$ is not complete because the Cauchy sequence $(1, \frac{1}{2}, \frac{1}{3}, \dots)$ does not converge in the open interval $(0, 1)$.

Theorem 1.2.4. *Let $\{x_n\}_{n \in \mathbb{Z}_{>0}}$ be a Cauchy sequence in a normed vector space X. Then, $\{x_n\}_{n \in \mathbb{Z}_{>0}}$ is bounded.*

Proof. Assume that X is a normed vector space and $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in X. Then there exists $N \in \mathbb{Z}_{>0}$ such that if $m, n > N$ then

$$\|x_n - x_m\| < 1.$$

Now fix $n \in \mathbb{Z}_{>N}$. If $m \in \mathbb{Z}_{>N}$ then

$$\|x_m\| \leq \|x_m - x_n\| + \|x_n\| < \|x_n\| + 1.$$

We conclude that if $j \in \mathbb{Z}_{>0}$ then

$$\|x_j\| < \max(\|x_1\|, \|x_2\|, \ldots, \|x_N\|, \|x_n\| + 1) < \infty.$$

So the Cauchy sequence $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ is bounded. □

Observe that this theorem also works perfectly for metric spaces.

Theorem 1.2.5. *Let $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ be a bounded sequence in $\mathbb{R}$. Then, $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ has a convergent subsequence.*

Proof. Assume that $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ is a bounded sequence in $\mathbb{R}$. Without loss of generality, assume that if $n \in \mathbb{Z}_{>0}$ then $x_n \in [0, 1]$. We can do this because $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ is bounded and changing the signs of each x_n does not affect the value of $|x_n|$.

We proceed by using a *diagonalisation* argument. Divide the interval $[0, 1]$ into two halves — $[0, 1/2]$ and $[1/2, 1]$. Either one of these intervals must contain infinitely many terms of the sequence $\{x_n\}_{n\in\mathbb{Z}_{>0}}$. Denote the interval which satisfies this property by $I_1 = [a_1, b_1]$.

Next divide the interval I_1 into its two halves. Since I_1 contains infinitely many terms of the sequence $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ then either one of the half-intervals of I_1 must contain infinitely many terms of $\{x_n\}_{n\in\mathbb{Z}_{>0}}$. We denote the interval which satisfies this property by $I_2 = [a_2, b_2]$.

Continuing in this fashion, we construct a sequence of nested intervals

$$\cdots \subsetneq I_n \subsetneq \cdots \subsetneq I_2 \subsetneq I_1 \subsetneq [0, 1]$$

such that if $j \in \mathbb{Z}_{>0}$ then $I_j = [a_j, b_j]$ contains infinitely many terms of the sequence $\{x_n\}_{n\in\mathbb{Z}_{>0}}$. Note that by construction, if $j \in \mathbb{Z}_{>0}$ then $b_j - a_j = 2^{-j}$ (this is the Lebesgue measure of the interval I_j). Now pick x_{n_j} from the sequence $\{x_n\}$ such that $x_{n_j} \in I_j$ and $x_{n_j} \notin I_{j+1}$. This gives us a subsequence $\{x_{n_j}\}_{j\in\mathbb{Z}_{>0}}$ of $\{x_n\}$.

Now observe that the sequence $\{a_j\}_{j\in\mathbb{Z}_{>0}}$ is increasing and bounded above by 1. So it must converge to some $x \in [0,1]$. We claim that the subsequence $\{x_{n_j}\}_{j\in\mathbb{Z}_{>0}}$ converges to x.

By assumption if $j \in \mathbb{Z}_{>0}$ then $x_{n_j} \in I_j = [a_j, b_j]$. So

$$|x_{n_j} - a_j| \leq |b_j - a_j| = \frac{1}{2^j}.$$

Assume that $\epsilon \in \mathbb{R}_{>0}$. Since $\{a_j\}_{j\in\mathbb{Z}_{>0}}$ converges to x then there exists $N \in \mathbb{Z}_{>0}$ such that if $k > N$ then

$$|a_k - x| < \frac{\epsilon}{2} \qquad \text{and} \qquad |x_{n_k} - a_k| \leq \frac{1}{2^k} < \frac{\epsilon}{2}.$$

By the triangle inequality, we obtain

$$|x_{n_k} - x| \leq |x_{n_k} - a_k| + |a_k - x| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence the subsequence $\{x_{n_j}\}_{j\in\mathbb{Z}_{>0}}$ converges to x. □

It is not too difficult to use 1.2.5 to show that every bounded sequence in $\mathbb{R}^n, \mathbb{C}$ and $\mathbb{C}^n$ has a convergent subsequence.

We will now use 1.2.5 to establish a very important result - the completeness of $\mathbb{R}$.

Theorem 1.2.6 (Completeness of $\mathbb{R}$)**.** *Let $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\mathbb{R}$. Then, $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ converges. In particular, the pair $(\mathbb{R}, |-|)$ is a Banach space where $|-| : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is the absolute value.*

Proof. The fact that $(\mathbb{R}, |-|)$ is a normed vector space follows from the properties of the absolute value $|-| : \mathbb{R} \to \mathbb{R}_{\geq 0}$. Assume that $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathbb{R}$. Then it is bounded by (REFERENCE).

By the Bolzano-Weierstrass theorem in Theorem 1.2.5 to obtain a subsequence $\{x_{n_k}\}_{k\in\mathbb{Z}_{>0}}$ which converges to some $x \in \mathbb{R}$. Assume that $\epsilon \in \mathbb{R}_{>0}$. Then there exists $M \in \mathbb{Z}_{>0}$ such that if $k > M$ then

$$|x_{n_k} - x| < \frac{\epsilon}{2}.$$

There also exists $P \in \mathbb{Z}_{>0}$ such that if $n_k, j > P$ then

$$|x_{n_k} - x_j| < \frac{\epsilon}{2}.$$

Fix $k \in \mathbb{Z}_{>M}$ such that $n_k > P$. If $j > \max(M, P)$ then

$$|x_j - x| \leq |x_j - x_{n_k}| + |x_{n_k} - x| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Therefore the Cauchy sequence $\{x_n\}_{n \in \mathbb{Z}_{>0}}$ converges to $x \in \mathbb{R}$. So the pair $(\mathbb{R}, |-|)$ is a Banach space. □

We can use the completeness of $\mathbb{R}$ in order to prove the completeness of other metric spaces/normed vector spaces.

Theorem 1.2.7 (Completeness of $\mathbb{R}^n$)**.** *Let $\mathbb{R}^n$ be a normed vector space, equipped with the usual Euclidean norm. Then, $(\mathbb{R}^n, \|.\|)$ is a Banach space.*

Proof. Assume $\{x_m\} = ({y_1}^{(m)}, {y_2}^{(m)}, \dots, {y_n}^{(m)})$ is a Cauchy sequence in $\mathbb{R}^n$.

To show: (a) $\{x_m\}$ is convergent.

(a) If $\{x_m\}$ is a Cauchy sequence, then for all $\epsilon > 0$, there exists a $N \in \mathbb{Z}_{>0}$ such that for all $l, m > N$

$$\|x_l - x_m\| = (\sum_{i=1}^{n} ({y_i}^{(l)} - {y_i}^{(m)})^2)^{\frac{1}{2}} < \epsilon$$

Utilising a bounding argument, we find that for all $j \in \{1, 2, \dots, n\}$,

$$\begin{aligned}({y_j}^{(l)} - {y_j}^{(m)})^2 &\leq \sum_{i=1}^{n} ({y_i}^{(l)} - {y_i}^{(m)})^2 \\ &\leq \epsilon^2\end{aligned}$$

Therefore,

$$|{y_j}^{(l)} - {y_j}^{(m)}| \leq \epsilon$$

This tells us that for all $j \in \{1, 2, \dots, n\}$, the sequence $({y_j}^{(1)}, {y_j}^{(2)}, \dots)$ in $\mathbb{R}$ is a Cauchy sequence. Due to the completeness of $\mathbb{R}$ (see 1.2.6), this sequence converges to (say) y_j for all $j \in \{1, 2, \dots, n\}$. This tells us that for all $j \in \{1, 2, \dots, n\}$ and $\epsilon > 0$, there exists a $N \in \mathbb{Z}_{>0}$ such that for all $m > N$

$$|{y_j}^{(m)} - y_j| \leq \frac{\epsilon}{\sqrt{n}}$$

Returning to $\mathbb{R}^n$, we define $x = (y_1, y_2, \ldots, y_n)$. Clearly, $x \in \mathbb{R}^n$ and for all $i \in \{1, 2, \ldots, n\}$ and $l > N$,

$$\sum_{i=1}^{n} (y_i{}^{(l)} - y_i)^2 \leq \frac{(n \times \epsilon^2)}{n}$$
$$= \epsilon^2$$

Subsequently,

$$\|x_l - x\| = (\sum_{i=1}^{n} (y_i{}^{(l)} - y_i)^2)^{\frac{1}{2}} \leq \epsilon$$

Therefore, the sequence $\{x_m\}$ converges to $x \in \mathbb{R}^n$. This proves the completeness of $\mathbb{R}^n$ □

Theorem 1.2.8 (Completeness of l^∞)**.** *The normed vector space l^∞ is a Banach space.*

Proof. Assume $\{x_m\} = (x_1{}^{(m)}, x_2{}^{(m)}, \ldots, x_n{}^{(m)}, \ldots)$ is a Cauchy sequence in l^∞.

To show: (a) $\{x_m\}$ is convergent.

(a) If $\{x_m\}$ is a Cauchy sequence, then for all $\epsilon > 0$, there exists a $N \in \mathbb{Z}_{>0}$ such that for all $l, m > N$

$$\|x_l - x_m\| = \sup\{|x_1{}^{(l)} - x_1{}^{(m)}|, |x_2{}^{(l)} - x_2{}^{(m)}|, \ldots, |x_n{}^{(l)} - x_n{}^{(m)}|, \ldots\} < \epsilon$$

Then, for all $j \in \{1, 2, ..., n\}$,

$$|x_j{}^{(l)} - x_j{}^{(m)}| < \epsilon.$$

Once again, this tells us that the real sequence $(x_j{}^{(1)}, x_j{}^{(2)}, \ldots)$ is a Cauchy sequence. Due to the completeness of $\mathbb{R}$, the sequence converges to (say) x_j. Now, we define

$$x = (x_1, x_2, \ldots, x_n, \ldots)$$

For clarity, x is a sequence of real numbers. To see that $x \in l^\infty$, we note that for all $j \in \{1, 2, ..., n\}$ and $l > N$,

$$\begin{aligned}|x_j| &= |x_j - {x_j}^{(l)} + {x_j}^{(l)}| \\ &\leq |{x_j}^{(l)} - x_j| + |{x_j}^{(l)}| \quad \text{(Triangle Inequality)} \\ &\leq \epsilon + |{x_j}^{(l)}| \\ &< \infty\end{aligned}$$

Thus, $x \in l^\infty$. Finally, we note that

$$\begin{aligned}\|x_l - x\| &= \sup\{|{x_1}^{(l)} - x_1|, |{x_2}^{(l)} - x_2|, \ldots, |{x_n}^{(l)} - x_n|, \ldots\} \\ &< \epsilon.\end{aligned}$$

This is because for all $j \in \{1, 2, ..., n\}$, the sequence $({x_j}^{(1)}, {x_j}^{(2)}, \ldots)$ in $\mathbb{R}$ converges to x_j as established previously. Hence, the sequence $\{x_m\}$ in l^∞ is convergent. Consequently, l^∞ is a Banach space. □

Finally, we will end this section with the definition of a continuous function.

Definition 1.2.5. Let $(X, \|.\|_x)$ and $(Y, \|.\|_y)$ be normed vector spaces. Then, a **continuous function** $f : X \to Y$ has the property that for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $a, b \in X$,

$$\|a - b\|_x < \delta \;\Rightarrow\; \|f(a) - f(b)\|_y < \epsilon.$$

1.3 Linear Operators

When one learns about a new space, it is natural to ask about the functions/mappings between them. Similarly to how we have linear transformations between vector spaces, normed vector spaces are connected via *linear operators*. We will make this more precise later.

Definition 1.3.1. Let X and Y be normed vector spaces over the same field $\mathbb{F}$. A **linear operator** is a mapping $\Lambda : X \to Y$ from $Dom(\Lambda) \subseteq X$ (the domain of Λ) to Y such that for all $c_1, c_2 \in \mathbb{F}$ and $x_1, x_2 \in X$, $\Lambda(c_1x_1 + c_2x_2) = c_1\Lambda(x_1) + c_2\Lambda(x_2)$.

Note the striking similarity to the definition of a linear transformation between two vector spaces. Again, the notions of kernel and image remain the same.

Definition 1.3.2. The **kernel** of Λ is the subspace of X defined by

$$ker(\Lambda) = \{x \in X \mid \Lambda(x) = 0\}$$

Definition 1.3.3. The **image** of Λ is the subspace of Y defined by

$$im(\Lambda) = \{\Lambda(x) \mid x \in Dom(\Lambda)\}$$

Definition 1.3.4. The linear operator Λ is said to be one-to-one or **injective** if and only if $ker(\Lambda) = \{0\}$.

Definition 1.3.5. The linear operator Λ is said to be onto or **surjective** if and only if $im(\Lambda) = Y$.

The main difference between linear transformations on vector spaces and linear operators on normed vector spaces is the fact that due to the extra structure on normed vector spaces induced by the norm, linear operators also have a topological structure. The next few definitions represent this.

Definition 1.3.6. The linear operator Λ is said to be **densely defined** if and only if $\overline{Dom(\Lambda)} = X$.

Definition 1.3.7. Let X and Y be normed vector spaces. The mapping $\phi : X \to Y$ is said to be **bounded** if the image of a bounded subset of X is a bounded subset of Y.

We will return to bounded maps/linear operators later. As an aside, we will shift our focus to the aforementioned similarity between linear transformations between vector spaces and linear operators between normed vector spaces. A major aspect of pure maths is to take the properties of something familiar (such as the integers for instance) and create and study abstractions with these properties. Consider vector spaces over the same field and their associated linear transformations. It is clear that the similar properties between linear transformations and linear operators is something that can be abstracted. The result is the notion of a **category**.

Definition 1.3.8. A **category** $\mathscr{C}$ is a triple, consisting of:

1. A class of **objects** ob($\mathscr{C}$)
2. A class of **morphisms** (or arrows) between the objects $Hom(\mathscr{C})$. We say that the morphism $f : A \to B$ is an element of $Hom(A, B)$, which denotes the class of all morphisms from A to B. A is deemed the **source object** and B is the **target object** in this case.

3. A binary operation called **composition of morphisms**, defined by $\circ : Hom(B, C) \times Hom(A, B) \to Hom(A, C)$, $\circ(g, f) = g \circ f$.

Additionally, the composition of morphisms must satisfy the following two properties:

1. **Associativity**: $(f \circ g) \circ h = f \circ (g \circ h)$
2. **Identity**: For all objects $A \in ob(\mathscr{C})$, there exists a morphism $1_A : A \to A$ such that for all morphisms $f \in Hom(A, B)$, $f \circ 1_A = f$ and for all morphisms $g \in Hom(B, A)$, $1_A \circ g = g$.

Theorem 1.3.1. *The collection of normed vector spaces over a field* $\mathbb{K}$*, together with linear operators, is a category* (we will call this category **K-Norm Vect**).

Proof. To show: (a) For all linear operators f, g, h, $(f \circ g) \circ h = f \circ (g \circ h)$

(b) There exists a linear operator $1_A : A \to A$ such that for all linear operators $f \in Hom(A, B)$, $f \circ 1_A = f$ and for all linear operators $g \in Hom(B, A)$, $1_A \circ g = g$

(a) Assume $f \in Hom(C, D)$, $g \in Hom(B, C)$ and $h \in Hom(A, B)$. Assume $ax + by \in A$ where $a, b \in \mathbb{K}$. Then,

$$\begin{aligned}
(f \circ g) \circ h(ax + by) &= (f \circ g) \circ (ah(x) + bh(y)) \\
&= a(f \circ g) \circ h(x) + b(f \circ g) \circ h(y) \\
&= f(a(g \circ h)(x) + b(g \circ h)(y)) \\
&= f \circ (g \circ h)(ax + by).
\end{aligned}$$

(b) Define the mapping $1_A : A \to A$ such that $1_A(a) = a$ for all $a \in A$.

To show: (ba) 1_A is a linear operator.

(bb) For all linear operators $f \in Hom(A, B)$, $f \circ 1_A = f$.

(bc) For all linear operators $g \in Hom(B, A)$, $1_A \circ g = g$.

(ba) Again, assume that $ax + by \in A$ where $a, b \in \mathbb{K}$. Then, from the definition of 1_A, we have

$$\begin{aligned} 1_A(ax+by) &= ax+by \\ &= a1_A(x)+b1_A(y) \end{aligned}$$

Hence, 1_A is a linear operator.

(bb) Assume that $f \in Hom(A,B)$ and that $ax+by \in A$ where $a,b \in \mathbb{K}$. Then,

$$f \circ 1_A(ax+by) = f(ax+by)$$

So, $f \circ 1_A = f$.

(bc) Assume that $g \in Hom(B,A)$ and that $ax+by \in A$ where $a,b \in \mathbb{K}$. Then,

$$1_A \circ g(ax+by) = 1_A(g(ax+by)) = g(ax+by)$$

So, $1_A \circ g = g$. Therefore, **K-Norm Vect** is a category. □

There are a multitude of categories that one might be already familiar with. Here are some examples:

Example 1.3.1. **Grp** is the category of groups with groups as objects and group homomorphisms as the morphisms.

Example 1.3.2. **Ab** is the category of abelian groups with abelian groups as objects and group homomorphisms as the morphisms. In fact, **Ab** is a subcategory of **Grp**.

Example 1.3.3. **Ring** is the category of rings with rings as objects and ring homomorphisms as the morphisms.

Example 1.3.4. **Top** is the category of topological spaces with topological spaces as objects and continuous functions as the morphisms.

Example 1.3.5. **R-Mod** is the category of R-modules with R-modules as objects and module homomorphisms as the morphisms. Here, R refers to an arbitrary ring.

Now we will return to our study of linear operators.

Definition 1.3.9. Let $(V_1, \|-\|_1)$ and $(V_2, \|-\|_1)$ be normed vector spaces and $\Lambda : V_1 \to V_2$ be a linear operator. We say that Λ is **bounded** if there exists $C \in \mathbb{R}_{>0}$ such that if $v_1 \in V_1$ then

$$\|\Lambda v_1\|_2 \leq C\|v_1\|_1.$$

One of the most foundational concepts of functional analysis is that a linear operator between normed vector spaces is continuous if and only if it is bounded.

Theorem 1.3.2. *Let $(V_1, \|-\|_1)$ and $(V_2, \|-\|_2)$ be normed vector spaces and $\Lambda : V_1 \to V_2$ be a linear operator. The following are equivalent:*

1. *Λ is continuous at $0 \in V_1$.*
2. *Λ is bounded.*
3. *Λ is continuous.*

Proof. Assume that $\Lambda : V_1 \to V_2$ is a linear operator.

(a) Assume that Λ is continuous at $0 \in V_1$. This means that for all $\epsilon \in \mathbb{R}_{>0}$, there exists $\delta \in \mathbb{R}_{>0}$ such that if $\|v - 0\|_1 = \|v\|_1 < \delta$ then $\|\Lambda v\|_2 < \epsilon$.

Take $w \in V_1 - \{0\}$ and set

$$v' = \frac{\delta}{2\|w\|_1} w.$$

Then, $\|v'\|_1 = \delta/2 < \delta$ and

$$\begin{aligned}\|\Lambda v'\|_2 &= \|\Lambda(\frac{\delta}{2\|w\|_1} w)\|_2 \\ &= \frac{\delta}{2\|w\|_1}\|\Lambda w\|_2 < \epsilon.\end{aligned}$$

So, if $w \in V_1$ then

$$\|\Lambda w\|_2 \leq \frac{2\epsilon}{\delta}\|w\|_1.$$

Thus, Λ must be bounded.

(b) Assume that Λ is bounded — there exists $C \in \mathbb{R}_{>0}$ such that if $v_1 \in V_1$ then $\|\Lambda v_1\|_2 \leq C\|v_1\|_1$. Assume that $\epsilon \in \mathbb{R}_{>0}$ and set $\delta = \epsilon/C$. If $\|v_1 - v_2\|_1 < \delta$ then

$$\begin{aligned}\|\Lambda(v_1) - \Lambda(v_2)\|_2 &= \|\Lambda(v_1 - v_2)\|_2 \\ &\leq C\|v_1 - v_2\|_1 \\ &< C \cdot \frac{\epsilon}{C} = \epsilon.\end{aligned}$$

So, Λ must be continuous.

(c) If Λ is continuous then it is continuous at all points $v \in V_1$. So, it must be continuous at $0 \in V_1$. □

Definition 1.3.10. We will define the **norm of a linear operator** to be

$$\|\Lambda\| = \sup_{\|x\|=1} \|\Lambda x\| = \sup_{x \neq 0} \frac{\|\Lambda x\|}{\|x\|}.$$

To understand why we defined the norm of a linear operator in this manner, consider the following characterisation of a bounded linear operator below, which states that a linear operator $\Lambda : X \to Y$ is bounded if there exists $c \in \mathbb{R}_{>0}$ such that if $x \in X$ then

$$\|\Lambda x\| \leq c\|x\|.$$

Dividing both sides by $\|x\|$, we obtain for $x \neq 0$

$$\frac{\|\Lambda x\|}{\|x\|} \leq c.$$

Hence, c is an upper bound of the left hand side. The supremum (least upper bound) of the left hand side over all $x \in X - \{0\}$ is the smallest such c and this is what we take to be the definition of the norm of a linear operator. Furthermore,

$$\|\Lambda\| = \sup_{x \neq 0} \frac{\|\Lambda x\|}{\|x\|} = \sup_{x \neq 0} \|\Lambda(\frac{x}{\|x\|})\| = \sup_{\|x\|=1} \|\Lambda x\|.$$

A natural question that remains is whether our definition of the norm satisfies the properties of a norm.

Theorem 1.3.3. *The norm of a linear operator as defined above is actually a norm*

Proof. Assume that $\Lambda : V \to W$ is a linear operator and that V and W are normed vector spaces over the field $\mathbb{K}$.

To show: (a) $\forall \alpha \in \mathbb{K}$, $\|\alpha\Lambda\| = |\alpha|\,\|\Lambda\|$.

(b) $\|\Lambda\| = 0$ if and only if $\Lambda = 0$.

(c) For all linear operators Λ_1, Λ_2, $\|\Lambda_1 + \Lambda_2\| \leq \|\Lambda_1\| + \|\Lambda_2\|$.

(a) Assume that $\alpha \in \mathbb{K}$. Then,

$$\begin{aligned} \|\alpha\Lambda\| &= \sup_{\|x\|=1} \|\alpha\Lambda x\| \\ &= |\alpha| \sup_{\|x\|=1} \|\Lambda x\| \\ &= |\alpha|\|\Lambda\|. \end{aligned}$$

(b) First, assume that $\|\Lambda\| = 0$. Then, it must be the case that

$$\sup_{x\neq 0} \frac{\|\Lambda x\|}{\|x\|} = 0.$$

This tells us that whenever $x \neq 0$, $\|\Lambda x\| = 0$ and consequently, $\Lambda = 0$. For the case where $x = 0$, we observe that for all linear operators Ω and for all $v \in V$, $\Omega(v) = \Omega(0 + v) = \Omega(0) + \Omega(v)$. So, $\Omega(0) = 0$. Consequently, $\Lambda = 0$ for all $x \in V$ because $\Lambda(0) = 0$.

For the converse, assume that $\Lambda = 0$. Then,

$$\|\Lambda\| = \sup_{\|x\|=1} \|\Lambda x\| = 0.$$

(c) Assume that Λ_1 and Λ_2 are linear operators. Then,

$$\begin{aligned} \|\Lambda_1 + \Lambda_2\| &= \sup_{\|x\|=1} \|(\Lambda_1 + \Lambda_2)x\| \\ &= \sup_{\|x\|=1} \|\Lambda_1 x + \Lambda_2 x\| \\ &\leq \sup_{\|x\|=1} \|\Lambda_1 x\| + \|\Lambda_2 x\| \\ &\leq \sup_{\|x\|=1} \|\Lambda_1 x\| + \sup_{\|x\|=1} \|\Lambda_2 x\| \\ &= \|\Lambda_1\| + \|\Lambda_2\|. \end{aligned}$$

Therefore, the norm of a linear operator does indeed satisfy the required properties of a norm. □

Let us quickly prove that a linear operator $\Lambda : X \to Y$ is bounded if and only if its norm $\|\Lambda\|$ is finite.

Theorem 1.3.4. *Let $(X, \|-\|_X)$ and $(Y, \|-\|_Y)$ be normed vector spaces and $\Lambda : X \to Y$ be a linear operator. Then, Λ is bounded if and only if $\|\Lambda\|$ is finite.*

Proof. Assume that X and Y are normed vector space and that $\Lambda : X \to Y$ is a linear operator.

Assume that $\|\Lambda\| < \infty$. If $x \in X - \{0\}$ then

$$\begin{aligned} \|\Lambda x\|_Y &= \frac{\|\Lambda x\|_Y}{\|x\|_X} \|x\|_X \\ &\leq \big(\sup_{x \neq 0} \frac{\|\Lambda x\|_Y}{\|x\|_X} \big) \|x\|_X = \|\Lambda\| \|x\|_X. \end{aligned}$$

Hence, if $x \in X$ then $\|\Lambda x\|_Y \leq \|\Lambda\| \|x\|_X$. Since $\|\Lambda\| \in \mathbb{R}_{>0}$, Λ is bounded in the sense of Definition 1.3.9.

Next, assume that $\Lambda : X \to Y$ is bounded. Then, there exists $C \in \mathbb{R}_{>0}$ such that if $x \in X$ then $\|\Lambda x\|_Y \leq C\|x\|_X$. Taking the supremum over all $x \in X$ which satisfy $\|x\|_X = 1$, we find that

$$\|\Lambda\| = \sup_{\|x\|_X = 1} \|\Lambda x\|_Y \leq C.$$

Thus, $\|\Lambda\|$ is finite. □

Definition 1.3.11. A bounded linear operator $\Lambda : X \to Y$ is said to be **compact** if for all bounded sequences $\{x_n\}$, there exists a subsequence $\{x_{n_k}\}$ such that Λx_{n_k} converges to some point $y \in Y$.

In order to make the concept of a linear operator more palatable, we will list some examples of linear operators.

Example 1.3.6. Let $C((0, 1), \mathbb{R})$ be the normed vector space of bounded, real valued, continuous functions with domain $(0, 1)$, with norm given by

$$\|f\| = \max_{x \in (0,1)} |f(x)|.$$

Define the differentiation operator Λ as $\Lambda f = f'$. From the properties of differentiation, we know that this is a linear operator. However, this operator is unbounded. To see this, consider the sequence of functions $f_k = \sin(kx)$ where $k \in \mathbb{Z}_{>0}$. Then, $\Lambda f_k = k\cos(kx)$ and

$$\|\Lambda\| = \sup_{\|f_k\|=1} \|\Lambda f_k\| = k.$$

Notably, k is unbounded because $k \in \mathbb{Z}_{>0}$. Hence, the differentiation operator is unbounded.

Example 1.3.7. This time, the normed vector space in question is $L^p(\mathbb{R})$, which has norm

$$\|f\| = (\int_{\mathbb{R}} |f|^p \, \mathrm{d}x)^{\frac{1}{p}}.$$

Let $a \in \mathbb{R}$. Define the shift operator $\Lambda_a(f(x)) = f(x-a)$. A noteworthy observation about this operator stems from an integral substitution:

$$\begin{aligned} \|\Lambda_a f\| &= (\int_{\mathbb{R}} |f(x-a)|^p \, \mathrm{d}x)^{\frac{1}{p}} \\ &= (\int_{\mathbb{R}} |f(y)|^p \, \mathrm{d}y)^{\frac{1}{p}} \\ &= \|f\|. \end{aligned}$$

Hence, it is clear that

$$\|\Lambda_a\| = \sup_{\|f\|=1} \|\Lambda_a f\| = 1.$$

So, Λ_a is a bounded linear operator.

Example 1.3.8. The normed vector space we will focus on this time is l^p, which has norm

$$\|x\| = (\sum_{i=1}^{\infty} |x_i|^p)^{\frac{1}{p}}.$$

Define the two shift operators as $\Lambda_+(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$ and $\Lambda_-(x_1, x_2, \dots) = (x_2, x_3, x_4, \dots)$. In a similar manner to the previous example, $\|\Lambda_+\| = \|\Lambda_-\| = 1$. So, Λ_+ and Λ_- are both bounded linear operators.

Example 1.3.9. In this example, our normed vector space is $C((a,b),\mathbb{R})$, which is the normed vector space of bounded, real valued, continuous functions with domain (a,b). Consider the integral operator, which is defined by

$$\Lambda_x f = \int_a^x f(y)\ \mathrm{d}y.$$

Here, $x \in [a,b]$. We know from the properties of integrals that this operator is definitely linear. In order to demonstrate that the integral is bounded, we proceed as follows

$$\begin{aligned}
\|\Lambda_x\| &= \sup_{\|f\|=1} \|\Lambda_x f\| \\
&= \sup_{\|f\|=1} \max_{x\in(a,b)} |\int_a^x f(y)\ \mathrm{d}y| \\
&= \sup_{\|f\|=1} |\int_a^b f(y)\ \mathrm{d}y| \\
&\le \sup_{\|f\|=1} \int_a^b |f(y)|\ \mathrm{d}y \\
&\le (b-a) \times \sup_{\|f\|=1} |f(y)| \\
&= b-a.
\end{aligned}$$

So, $\|\Lambda_x\| \le b-a < \infty$, thus demonstrating that the integral operator is bounded.

Finally, we end this section with an important theorem.

Theorem 1.3.5. *Let V and W be normed vector spaces over the field $\mathbb{K}$. Let $B(V;W)$ represent the space of bounded linear operators from V to W. If W is a Banach space, then $B(V;W)$ is also a Banach space, with norm defined in 1.3.10.*

Proof. To show: (a) $B(V;W)$ is a normed vector space

(b) A Cauchy sequence in $B(V;W)$ converges.

(a) We have already established that the norm of a linear operator satisfies the axioms required of a norm. This demonstrates that $B(V;W)$ is a normed vector space.

(b) Assume that $\{\Lambda_n\}$ is a Cauchy sequence of bounded linear operators. Assume W is a Banach space. Assume $f \in V$. Then, for all $\epsilon > 0$, there exists a $N \in \mathbb{Z}_{>0}$ such that for all $m, n > N$,

$$\|\Lambda_m - \Lambda_n\| < \epsilon.$$

As a result, $\|(\Lambda_m - \Lambda_n)(f)\| < \epsilon$ where $\|f\| = 1$. So, $\|\Lambda_m f - \Lambda_n f\| < \epsilon$. This tells us that the sequence $\{\Lambda_n f\}$ is Cauchy. Since this is a sequence in W, which is a Banach space, $\{\Lambda_n f\}$ must be convergent. Let $\{\Lambda_n f\}$ converge to Λf.

To show: (ba) Λ is a bounded operator.

(bb) Λ is a linear operator

(ba) Using the convergence of $\{\Lambda_n f\}$, we choose $N \in \mathbb{Z}_{>0}$ such that for all $n > N$,

$$\|\Lambda_n f - \Lambda f\| < 1.$$

Utilising this, we note that

$$\begin{aligned} \|\Lambda\| &= \sup_{\|f\|=1} \|\Lambda f\| \\ &\leq \sup_{\|f\|=1} (\|\Lambda f - \Lambda_n f\| + \|\Lambda_n f\|) \quad \text{(Triangle Inequality)} \\ &\leq 1 + \|\Lambda_n f\| \\ &< \infty \end{aligned}$$

Hence, Λ is a bounded operator from V to W.

(bb) Assume $a, b \in \mathbb{K}$ and $f, g \in V$.

To show: (bba) For all $\epsilon > 0$, $\|\Lambda(af + bg) - a\Lambda(f) - b\Lambda(g)\| < \epsilon$.

(bba) Assume that $\epsilon > 0$. Once again, we utilise the triangle inequality on the expression $\|\Lambda(af + bg) - a\Lambda(f) - b\Lambda(g)\|$. We know that the sequence $\{\Lambda_n f\}$ converges to Λf and that f is arbitrary. So, we choose $N_1 \in \mathbb{Z}_{>0}$ such that for all $n > N_1$,

$$\|\Lambda(af + bg) - \Lambda_n(af + bg)\| < \frac{\epsilon}{3}.$$

Then, we choose $N_2 \in \mathbb{Z}_{>0}$ such that for all $n > N_2$,

$$\|\Lambda_n(f) - \Lambda(f)\| < \frac{\epsilon}{3|a|}.$$

Finally, we choose $N_3 \in \mathbb{Z}_{>0}$ such that for all $n > N_3$,

$$\|\Lambda_n(g) - \Lambda(g)\| < \frac{\epsilon}{3|b|}.$$

Now, select $M = \max\{N_1, N_2, N_3\}$. Then, for all $n > M$ (and using the fact that Λ_n is a linear operator),

$$\begin{aligned}
&\|\Lambda(af + bg) - a\Lambda(f) - b\Lambda(g)\| \\
&= \|\Lambda(af + bg) - \Lambda_n(af + bg) + a\Lambda_n(f) + b\Lambda_n(g) - a\Lambda(f) - b\Lambda(g)\| \\
&\leq \|\Lambda(af + bg) - \Lambda_n(af + bg)\| + \|a\Lambda_n(f) - a\Lambda(f)\| + \|b\Lambda_n(g) - b\Lambda(g)\| \\
&< \frac{\epsilon}{3} + (|a| \times \frac{\epsilon}{3|a|}) + (|b| \times \frac{\epsilon}{3|b|}) \\
&= \epsilon.
\end{aligned}$$

Therefore, for all $\epsilon > 0$, $\|\Lambda(af + bg) - a\Lambda(f) - b\Lambda(g)\| < \epsilon$. Consequently, $\|\Lambda(af + bg) - a\Lambda(f) - b\Lambda(g)\| = 0$ and so, $\Lambda(af + bg) = a\Lambda(f) + b\Lambda(g)$.

(bb) Hence, Λ is a linear operator.

(b) Now consider $\|\Lambda_n - \Lambda\|$. Using the fact that $\{\Lambda_n f\}$ is a Cauchy sequence in W and hence convergent, we obtain

$$\|\Lambda_n - \Lambda\| = \sup_{\|f\|=1} \|\Lambda_n f - \Lambda f\| \leq \epsilon.$$

Therefore, the Cauchy sequence of operators $\{\Lambda_n\}$ in $B(V; W)$ converges. □

Chapter 2

Dual Spaces and the Hahn-Banach Extension Theorem

2.1 Dual Spaces

Definition 2.1.1. Let X be a normed vector space over a field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We define the **dual space of X** as the set of all continuous linear functionals $\phi : X \to \mathbb{K}$. The dual space of X is often denoted by X^*.

Theorem 2.1.1. *Let X be a normed vector space and X^* be its associated dual space. Then, X^* is a Banach space, with norm given by*

$$\|\phi\| = \sup_{x \neq 0} \frac{|\phi(x)|}{\|x\|} = \sup_{\|x\|=1} |\phi(x)|$$

Proof. An application of 1.3.5 demonstrates this. Luckily for us, $\mathbb{R}$ and $\mathbb{C}$ are both complete with respect to the standard Euclidean norm. □

The Hahn-Banach theorem is particularly powerful because it is one of the main reasons why dual spaces are worth studying. It ensures that there will always be enough continuous linear functionals for the theory of dual spaces to remain rich. We will prove three versions of the Hahn-Banach theorem.

Theorem 2.1.2 (Hahn-Banach extension theorem V1)**.** *Let X be a $\mathbb{R}$-vector space and $p : X \to \mathbb{R}$ be a sub-linear function, which satisfies the following two properties:*

1. *If $x, y \in X$ then $p(x + y) \leq p(x) + p(y)$*

2. If $x \in X$ and $t \in [0, \infty)$, $p(tx) = tp(x)$

Let Y be a subspace of X and $\lambda : Y \to \mathbb{R}$ be a functional with $\lambda(x) \leq p(x)$ for all $x \in Y$. Then, there exists a functional $\Lambda : X \to \mathbb{R}$ such that the restriction $\Lambda|_Y = \lambda$ and if $x \in X$ then

$$-p(-x) \leq \Lambda(x) \leq p(x).$$

Proof. Assume that X is a $\mathbb{R}$-vector space and p, λ are the functions defined as above. Let $z \in X \backslash Y$. We would like to extend λ first to the subspace $Y + span(z)$.

Assume that $x, y \in Y$. Then,

$$\begin{aligned} f(x) + f(y) &= f(x+y) \\ &\leq p(x+y) = p(x - z + z + y) \\ &\leq p(x-z) + p(z+y) \end{aligned}$$

where in the last line, we used the triangle inequality. Consequently, $f(x) - p(x-z) \leq p(z+y) - f(y)$. Now define

$$\beta = \sup_{x \in V} (f(x) - p(x-z))$$

so that

$$f(x) - p(x-z) \leq \beta \leq p(z+y) - f(y).$$

Now let $t \in \mathbb{R}$ and define for $x \in Y$, the extension $\tilde{f}(x + tz) = f(x) + \beta t$. Since we have extended f linearly to $Y + span(z)$, $\tilde{f}$ is still a linear functional on $Y + span(z)$. We will now show that

$$-p(-x - tz) \leq \tilde{f}(x + tz) \leq p(x + tz).$$

Notice that when $t = 0$, the result follows from the assumption that $f(x) \leq p(x)$ for all $x \in X$. Without loss of generality, assume that $t > 0$. Recalling the inequality $f(x) - p(x-z) \leq \beta \leq p(z+y) - f(y)$, we replace x with $-x/t$ and y with x/t which yields

$$f(-\frac{x}{t}) - p(-\frac{x}{t} - z) \leq \beta \leq p(z + \frac{x}{t}) - f(\frac{x}{t})$$

Now, we can multiply by t to obtain the inequality

$$-f(x) - p(-x - tz) \leq \beta t \leq p(x + tz) - f(x)$$

Adding $f(x)$, we deduce that

$$-p(-x - tz) \leq f(x) + \beta t = \tilde{f}(x + tz) \leq p(x + tz).$$

as required.

We have successfully extended f to $Y + span(z)$. Let C be the family of pairs (W, ϕ) such that $W \subseteq X$ is a subspace of X and $\phi : W \to \mathbb{R}$ is a linear functional such that $\phi(w) \leq p(w)$.

We define a partial ordering on C by saying that $(W, \phi) < (W', \phi')$ if and only if $W \subset W'$ and ϕ is the restriction of ϕ' to W. The pair $(C, <)$ is now a poset.

With this partial ordering, we can invoke Zorn's lemma to deduce that C has a maximal element, which we will denote by $(\widehat{W}, F)$.

To show: (a) $\widehat{W} = X$.

(a) Suppose for the sake of contradiction that $\widehat{W} \neq X$. Then, take some $v \in X \backslash \widehat{W}$. By the previous argument, we can extend the functional $F : \widehat{W} \to \mathbb{R}$ to another linear functional $F' : \widehat{W} + span(v) \to \mathbb{R}$ such that $F'|_{\widehat{W}} = F$. However, this means that in C,

$$(\widehat{W}, F) < (\widehat{W} + span(v), F')$$

which contradicts the maximality of $(\widehat{W}, F)$ in C. Therefore, $\widehat{W} = X$.

Hence, $F : X \to \mathbb{R}$ is a linear functional such that if $x \in X$ then $F(x) \leq p(x)$. By the linearity of F, we finally have

$$-p(-x) \leq -F(-x) = F(x) \leq p(x).$$

□

The second version of the Hahn-Banach theorem applies to a situation where the linear functional is bounded above by a convex function.

Definition 2.1.2. Let X be a Banach space. A function $p : X \to \mathbb{R}$ is **convex** if for all $t \in (0, 1)$ and $x, y \in X$,

$$p(tx + (1-t)y) \leq tp(x) + (1-t)p(y).$$

Theorem 2.1.3 (Hahn-Banach extension theorem V2)**.** *Let X be a $\mathbb{R}$-vector space and $p : X \to \mathbb{R}$ be a* ***convex*** *function. Let Y be a subspace of X and $f : Y \to \mathbb{R}$ be a functional with $f(x) \leq p(x)$ for all $x \in Y$. Then, there exists a functional $\tilde{f} : X \to \mathbb{R}$ such that the restriction $\tilde{f}|_Y = f$ and if $x \in X$ then $\tilde{f}(x) \leq p(x)$.*

Proof. Assume that X is a $\mathbb{R}$-vector space and p, λ are the functions defined as above. The idea is to use the Hahn-Banach extension theorem for sub-linear functions (see Theorem 2.1.2) and extend it to account for convex functions.

Define the function

$$\begin{array}{rccc} Q: & X & \to & \mathbb{R} \\ & x & \mapsto & \inf_{t>0} \frac{1}{t} p(tx). \end{array}$$

We claim that Q is a sub-linear function.

To show: (a) If $x, y \in X$ then $Q(x+y) \leq Q(x) + Q(y)$.

(b) If $x \in X$ and $\lambda \in \mathbb{R}_{\geq 0}$, $Q(\lambda x) = \lambda Q(x)$.

(a) Assume that $x, y \in X$. Pick $u, v \in \mathbb{R}_{>0}$ so that the fraction

$$\frac{uv}{u+v} > 0.$$

By using the convexity of p, we have

$$\begin{aligned} Q(x+y) &= \inf_{t>0} \frac{1}{t} p(t(x+y)) \\ &\leq \frac{u+v}{uv} p(\frac{uv}{u+v}(x+y)) \\ &= \frac{u+v}{uv} p(\frac{u}{u+v}(vx) + \frac{v}{u+v}(uy)) \\ &\leq \frac{u+v}{uv} \big(\frac{u}{u+v} p(vx) + \frac{v}{u+v} p(uy)\big) \quad (p \text{ is convex}) \\ &= \frac{1}{v} p(vx) + \frac{1}{u} p(uy). \end{aligned}$$

Now, we can take the infimum over all $v, u \in \mathbb{R}_{>0}$ to obtain $Q(x+y) \leq Q(x) + Q(y)$.

(b) Assume that $\lambda \in \mathbb{R}_{\geq 0}$. Then,

$$\begin{aligned} Q(\lambda x) &= \inf_{t>0} \frac{1}{t} p(t\lambda x) \\ &= \inf_{t>0} \frac{\lambda}{t\lambda} p(t\lambda x) \\ &= \inf_{s>0} \frac{\lambda}{s} p(sx) = \lambda Q(x). \end{aligned}$$

Combining parts (a) and (b), we deduce that Q is sub-linear.

We will now prove two critical properties of Q.

To show: (c) If $f \leq p$ on Y then $f \leq Q$ on Y.

(d) If $x \in X$ then $Q(x) \leq p(x)$.

(c) Assume that $f(v) \leq p(v)$ for $v \in Y$. Using the linearity of f, we have for $t > 0$

$$f(v) = tf(\frac{1}{t}v) \leq p(v)$$

So, $f(\frac{1}{t}v) \leq \frac{1}{t}p(v)$. By replacing v with tv, we deduce that $f(v) \leq \frac{1}{t}p(tv)$ and by taking the infimum over all $t > 0$, we obtain $f(v) \leq Q(v)$.

(d) Assume that $x \in X$. Then,

$$Q(x) = \inf_{t>0} \frac{1}{t} p(tx) \leq p(x)$$

where the last inequality follows from setting $t = 1$.

Part (c) tells us that if $v \in Y$ then $f(v) \leq Q(v) \leq p(v)$. Since Q is sub-linear, we can use Theorem 2.1.2 to obtain an extension $\tilde{f} : X \to \mathbb{R}$ such that $\tilde{f}|_Y = f$ and if $x \in X$ then $\tilde{f}(x) = Q(x)$ and

$$-Q(-x) \leq \tilde{f}(x) \leq Q(x).$$

But, by part (d), $Q(x) \leq p(x)$. So, if $x \in X$ then $\tilde{f}(x) \leq p(x)$. So, $\tilde{f}$ is the desired extension of f. □

The final version of the Hahn-Banach theorem deals with the extension of complex linear functionals.

Theorem 2.1.4 (Hahn-Banach extension theorem V3)**.** *Let X be a $\mathbb{C}$ vector space and $p : X \to \mathbb{R}$ be a function which satisfies*

$$p(\alpha x + \beta y) \leq |\alpha|p(x) + |\beta|p(y)$$

for $x, y \in X$ and $\alpha, \beta \in \mathbb{C}$ with $|\alpha| + |\beta| = 1$. Let $Y \subseteq X$ be a subspace of X and $f : Y \to \mathbb{C}$ be a complex linear functional such that if $y \in Y$ then $|f(y)| \leq p(y)$. Then, there exists a complex linear functional $F : X \to \mathbb{C}$ such that the restriction $F|_Y = f$ and if $x \in X$ then $|F(x)| \leq p(x)$.

Proof. Assume that X is a $\mathbb{C}$-vector space and $f : Y \to \mathbb{C}$ and $p : X \to \mathbb{R}$ are defined as above. Let

$$\ell(y) = Re(f(y)).$$

Then, ℓ defines a real-valued linear functional on Y. Observe also that

$$\ell(iy) = Re(f(iy)) = Re(if(y)) = -Im(f(y))$$

and consequently, $\lambda(y) = \ell(y) - i\ell(iy)$. Now, if $y \in Y$ then

$$|\ell(y)| = |Re(f(y))| \leq |f(y)| \leq p(y).$$

So, we can apply Theorem 2.1.3 to extend ℓ to a real-valued $\mathbb{R}$-linear functional $L : X \to \mathbb{R}$ such that $L(x) \leq p(x)$ for all $x \in X$. Now define for $x \in X$

$$F(x) = L(x) - iL(ix).$$

Then, the restriction $F|_Y = f$ and F is $\mathbb{R}$-linear. Also,

$$F(ix) = L(ix) - iL(-x) = iF(x).$$

So, F must be a $\mathbb{C}$-linear functional.

To see that $|F(x)| \leq p(x)$ for $x \in X$, observe that by the condition on p, if $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ then $p(\alpha x) = |\alpha|p(x) = p(x)$. If we set $\theta = \arg(F(x))$ then $F(x) = |F(x)|e^{i\theta}$ and if $x \in X$ then

$$\begin{aligned} |F(x)| &= e^{-i\theta}F(x) \\ &= F(e^{-i\theta}x) = Re(F(e^{-i\theta}x)) \\ &= L(e^{-i\theta}x) \leq p(e^{-i\theta}x) = p(x). \end{aligned}$$

□

An immediate application of the Hahn-Banach theorem is to the case where $p(x) = \|x\|$ (the norm function).

Theorem 2.1.5 (Extension with constant norm)**.** *Let X be a Banach space and $Y \subseteq X$ be a subspace of X. Let $\lambda : Y \to \mathbb{C}$ be a linear functional. Then, there exists a linear functional $\Lambda : X \to \mathbb{C}$ extending λ such that $\|\Lambda\| = \|\lambda\|$.*

Proof. Assume that X is a Banach space (over $\mathbb{C}$). Assume that $Y \subseteq X$ is a subspace and $\lambda \in Y^*$. Define the map $p : X \to \mathbb{R}$ by

$$\begin{array}{rccc} p: & X & \to & \mathbb{R} \\ & x & \mapsto & \|\lambda\|\|x\|. \end{array}$$

If $y \in Y$ then $|\lambda(y)| \leq \|\lambda\|\|y\| = p(y)$. If $\alpha, \beta \in \mathbb{C}$ and $x_1, x_2 \in X$ then

$$p(\alpha x_1 + \beta x_2) = \|\lambda\|\|\alpha x_1 + \beta x_2\| \leq \|\lambda\|(|\alpha|\|x_1\| + |\beta|\|x_2\|).$$

So, $p(\alpha x_1 + \beta x_2) \leq |\alpha| p(x_1) + |\beta| p(x_2)$.

Now we can apply Theorem 2.1.4 to find a complex linear functional $\Lambda : X \to \mathbb{C}$ such that $\Lambda|_Y = \lambda$ and if $x \in X$ then $|\Lambda(x)| \leq \|\lambda\|\|x\|$.

Taking the supremum over all $x \in X$ with $\|x\| = 1$, we find that $\|\Lambda\| \leq \|\lambda\|$. But, $\|\lambda\| \leq \|\Lambda\|$ because the norm cannot decrease when we extend λ from Y to X. Therefore, $\|\Lambda\| = \|\lambda\|$ as required. □

Theorem 2.1.5 tells us that if we have a bounded, linear, real-valued functional on a subspace of X, then we can extend it to the whole of X, while keeping the norm intact.

We will now discuss two consequences of Theorem 2.1.5 that reinforce the notion that there are enough continuous linear functionals in order to make the study of dual spaces worthwhile.

Theorem 2.1.6. *Let X be a Banach space and $x \in X$. Then, there exists a continuous linear functional $\phi \in X^*$ such that $\phi(x) = \|x\|$ and $\|\phi\| = 1$.*

Proof. Assume that X is a Banach space and $x \in X$. Let $X' = span(x)$ and define the map

$$\begin{array}{rccc} \psi: & X' & \to & \mathbb{C} \\ & w = kx & \mapsto & |k|\|x\|. \end{array}$$

Then, ψ is a $\mathbb{C}$-linear functional on X'. By Theorem 2.1.5, there exists an extension $\phi : X \to \mathbb{C}$ of ψ such that

$$\|\phi\| = \|\psi\| = \sup_{\|y\|=1, y \in X'} |\psi(y)| = |\psi(\frac{x}{\|x\|})| = 1.$$

Since $\phi|_{X'} = \psi$, $\phi(x) = \|x\|$. □

Theorem 2.1.7 (Separating points with functionals)**.** *Let X be a Banach space and $x, y \in X$ with $x \neq y$. Then, there exists a continuous linear functional $\phi \in X^*$ such that $\phi(x) \neq \phi(y)$.*

Proof. Assume that X is a Banach space and $x, y \in X$ with $x \neq y$. From Theorem 2.1.6, there exists a continuous linear functional $\phi \in X^*$ such that $\|\phi\| = 1$ and $\phi(x - y) = \|x - y\| \neq 0$. By linearity of ϕ, we deduce that since $\phi(x - y) \neq 0$, $\phi(x) \neq \phi(y)$ as required. □

Theorem 2.1.7 tells us that there are enough continuous linear functionals defined on the dual space X^* in order to distinguish the vectors in X.

Our next goal is to look at some examples of dual spaces. Before this, we will briefly discuss the idea of an isomorphism.

Definition 2.1.3. Let $\mathscr{C}$ be a category. Let $a, b \in ob(\mathscr{C})$ and $f \in Hom(a, b)$. Then, f is said to be an **isomorphism** if there exists another morphism $f^{-1} \in Hom(b, a)$ such that $f^{-1} \circ f = 1_a$ and $f \circ f^{-1} = 1_b$.

From the perspective of category theory, we can see why the notion of an isomorphism is ubiquitous in many areas of mathematics.

Example 2.1.1. Consider the category of groups **Grp**. From the definition of an isomorphism, a group isomorphism is a bijective group homomorphism.

Example 2.1.2. Consider the category of rings **Ring**. From the definition of an isomorphism, a ring isomorphism is a bijective ring homomorphism.

Example 2.1.3. Consider the category of vector spaces over a field $\mathbb{K}$ **K-Vect**. From the definition of an isomorphism, a vector space isomorphism is a bijective linear transformation.

Now consider the category **K-Norm Vect** of normed vector spaces. It makes sense for an isomorphism in **K-Norm Vect** to be similar to an isomorphism in **K-Vect** because the objects in both categories are vector spaces on the most basic level. However, we have the extra structure of a norm to deal with in **K-Norm Vect**. If we want two normed vector spaces to be isomorphic, then their norms have to behave in exactly the same manner as each other. This motivates the following definition.

Definition 2.1.4. Let X and Y be normed vector spaces. Let $\Lambda : X \to Y$ be a linear operator. Then, Λ is said to be an isomorphism if Λ is bijective and $\|\Lambda x\|_Y = \|x\|_X$. Additionally, if there exists an isomorphism between X and Y, they are said to be isomorphic.

In the following examples of dual spaces, we will be proving that the dual spaces of certain Banach spaces are isomorphic to other familiar Banach spaces. For our first example, we will need Hölder's inequality, which we will prove below.

Theorem 2.1.8 (Young's inequality)**.** *Let $q, p \in (1, \infty)$. Assume p and q are related by the following identity:*

$$\frac{1}{p} + \frac{1}{q} = 1$$

Let $a, b \in \mathbb{R}_{>0}$. Then, the following inequality holds

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{a}{p} + \frac{b}{q}.$$

Proof. Assume that $q, p \in (1, \infty)$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Assume that $a, b \in \mathbb{R}_{>0}$. Assume that $f(x)$ is the function defined as above. By differentiating $f(x)$, we obtain

$$f'(x) = \alpha x^{\alpha-1} - \alpha = \alpha(\frac{1}{x^{1-\alpha}} - 1).$$

By solving $f'(x) = 0$, we deduce that $f(x)$ has a maximum at $x = 1$. So, for $x \in \mathbb{R}_{>0}$,

$$x^{\alpha} - \alpha x \leq f(1) = 1 - \alpha.$$

If we substitute $\alpha = 1/p$ and $x = a/b$, we deduce that

$$(\frac{a}{b})^{\frac{1}{p}} - \frac{a}{bp} \leq= 1 - \frac{1}{p} = \frac{1}{q}.$$

Multiplying both sides by b, we deduce that

$$a^{\frac{1}{p}} b^{1-\frac{1}{p}} - \frac{a}{p} = a^{\frac{1}{p}} b^{\frac{1}{q}} - \frac{a}{p} \leq \frac{b}{q}.$$

□

Theorem 2.1.9 (Hölder's inequality)**.** *Let* $p, q \in (1, \infty)$, $x = (x_1, x_2, \dots) \in l^p$, $y = (y_1, y_2, \dots) \in l^q$ *and* $\langle x, y\rangle = x_1y_1 + x_2y_2 + \dots$. *Assume p and q are related by the following identity:*

$$\frac{1}{p} + \frac{1}{q} = 1$$

Then,

$$|\langle x, y\rangle| \le \|x\|_p\|y\|_q.$$

Proof. Assume $x = (x_1, x_2, \dots) \in l^p$ and $y = (y_1, y_2, \dots) \in l^q$. Assume $q \in \mathbb{R}_{\ge 1}$, $p \in \mathbb{R}_{>1} \cup \{\infty\}$ and $\frac{1}{p} + \frac{1}{q} = 1$. Using Young's inequality with $a = (|x_i|/\|x\|_p)^p$ and $b = (|y_i|/\|y\|_q)^q$ $(i \in \mathbb{Z}_{>0})$,

$$\frac{|x_i|}{\|x\|_p}\frac{|y_i|}{\|y\|_q} \le \frac{1}{p}(\frac{|x_i|}{\|x\|_p})^p + \frac{1}{q}(\frac{|y_i|}{\|y\|_q})^q.$$

Summing over both sides from $i = 1$ to ∞, we obtain

$$\sum_{i=1}^{\infty}\frac{|x_iy_i|}{\|x\|_p\|y\|_q} \le \sum_{i=1}^{\infty}\frac{1}{p}(\frac{|x_i|}{\|x\|_p})^p + \frac{1}{q}(\frac{|y_i|}{\|y\|_q})^q.$$

Working on the RHS, we find that

$$\begin{aligned}\sum_{i=1}^{\infty}\frac{1}{p}(\frac{|x_i|}{\|x\|_p})^p + \frac{1}{q}(\frac{|y_i|}{\|y\|_q})^q &= \frac{1}{p(\|x\|_p)^p}\sum_{i=1}^{\infty}|x_i|^p + \frac{1}{q(\|y\|_q)^q}\sum_{i=1}^{\infty}|y_i|^q.\\ &= \frac{1}{p(\|x\|_p)^p}(\|x\|_p)^p + \frac{1}{q(\|y\|_q)^q}(\|y\|_q)^q.\\ &= \frac{1}{p} + \frac{1}{q}\\ &= 1.\end{aligned}$$

So,

$$\sum_{i=1}^{\infty}\frac{|x_iy_i|}{\|x\|_p\|y\|_q} \le 1.$$

As a result,

$$\sum_{i=1}^{\infty}|x_iy_i| \le \|x\|_p\|y\|_q.$$

Now consider $|\langle x, y\rangle|$. We argue as follows

$$
\begin{aligned}
|\langle x, y\rangle| &= |\sum_{i=1}^{\infty} x_i y_i| \\
&\le \sum_{i=1}^{\infty} |x_i y_i| \\
&\le \|x\|_p \|y\|_q.
\end{aligned}
$$

Thus, we have proved Hölder's inequality. □

Now, we can analyse our first example of a dual space.

Example 2.1.4. Assume $p, q \in \mathbb{R}_{>1}$ and $\frac{1}{p} + \frac{1}{q} = 1$. Consider the Banach space l^p. Then, the dual space of l^p is l^q. In other words, $(l^p)^* = l^q = B(l^p, \mathbb{R})$.

Proof. Define the map $\phi : l^q \to (l^p)^*$ by

$$\phi(y) = \psi_y$$

In turn, the functional $\psi_y : l^p \to \mathbb{R}$ is defined by

$$\psi_y(x) = \langle y, x\rangle = \sum_{i=1}^{\infty} y_i x_i.$$

Our aim is to show that the map ϕ is an isomorphism between $(l^p)^*$ and l^q.

To show: (a) ϕ is a linear operator.

(b) ϕ is invertible.

(c) If $y \in l^q$ then $\|\phi(y)\| = \|y\|_q$.

(a) Assume $k \in \mathbb{R}$, $x \in l^p$ and $a, b \in l^q$. Assume $a = (a_1, a_2, \dots)$, $b = (b_1, b_2, \dots)$ and $x = (x_1, x_2, \dots)$.

To show: (aa) $\phi(a + b) = \phi(a) + \phi(b)$.

(ab) $\phi(ka) = k\phi(a)$.

(aa) Using the definition of ϕ, we compute directly

$$
\begin{aligned}
\phi(a+b)(x) &= \psi_{a+b}(x) = \langle a+b, x\rangle \\
&= \sum_{i=1}^{\infty}(a_i+b_i)x_i = \sum_{i=1}^{\infty} a_i x_i + \sum_{i=1}^{\infty} b_i x_i \\
&= \langle a, x\rangle + \langle b, x\rangle = \psi_a(x) + \psi_b(x) = \phi(a)(x) + \phi(b)(x).
\end{aligned}
$$

Hence, $\phi(a+b) = \phi(a) + \phi(b)$.

(ab) Once again, we compute directly

$$
\begin{aligned}
\phi(ka)(x) &= \psi_{ka}(x) = \langle ka, x\rangle \\
&= \sum_{i=1}^{\infty} ka_i x_i = k\sum_{i=1}^{\infty} a_i x_i \\
&= k\langle a, x\rangle = k\psi_a(x) = k\phi(a)(x).
\end{aligned}
$$

Hence, $\phi(ka) = k\phi(a)$.

(b) Define the map $\alpha : (l^p)^* \to l^q$ by

$$
\alpha(\gamma) = (\gamma(e_1), \gamma(e_2), \dots)
$$

Here, $e_i \in l^p$ denotes the sequence with a 1 in the i^{th} position and zeros elsewhere.

To show: (ba) If $\gamma \in (l^p)^*$ then $\alpha(\gamma) \in l^q$.

(bb) $\phi(\alpha(\gamma))(x) = \gamma(x)$.

(bc) $\alpha(\phi(y)) = y$.

(ba) For $n \in \mathbb{Z}_{>0}$, let $x_n = (\xi_k^{(n)})$, where

$$
\xi_k^{(n)} = \begin{cases} |\gamma(e_k)|^q/\gamma(e_k), \text{ if } k \le n \text{ and } \gamma(e_k) \ne 0, \\ 0, \text{ if } k > n \text{ or } \gamma(e_k) = 0. \end{cases} \in \mathbb{C}
$$

We can write $x_n = \sum_{k=1}^{\infty} \xi_k^{(n)} e_k$ so that $\gamma(x_n) = \sum_{k=1}^{\infty} \xi_k^{(n)} \gamma(e_k)$. From the definition of $\xi_k^{(n)}$,

$$
\sum_{k=1}^{\infty} \xi_k^{(n)} \gamma(e_k) = \sum_{k=1}^{n} |\gamma(e_k)|^q.
$$

We bound $\gamma(x_n)$ above by

$$\begin{aligned}\gamma(x_n) &\leq \|\gamma\|\|x_n\| \\ &= \|\gamma\|(\sum_{k=1}^{\infty}|\xi_k^{(n)}|^p)^{\frac{1}{p}} \\ &= \|\gamma\|(\sum_{k=1}^{n}|\gamma(e_k)|^{(q-1)p})^{\frac{1}{p}} \\ &= \|\gamma\|(\sum_{k=1}^{n}|\gamma(e_k)|^q)^{\frac{1}{p}}.\end{aligned}$$

In the above working, we used the fact that $\frac{1}{p}+\frac{1}{q}=1$ and $(q-1)p=q$. Thus,

$$\sum_{k=1}^{n}|\gamma(e_k)|^q \leq \|\gamma\|(\sum_{k=1}^{n}|\gamma(e_k)|^q)^{\frac{1}{p}}.$$

Consequently,

$$(\sum_{k=1}^{n}|\gamma(e_k)|^q)^{\frac{1}{q}} \leq \|\gamma\|.$$

Since this holds for arbitrary $n \in \mathbb{Z}_{>0}$, we can take the limit as $n \to \infty$ to obtain

$$\begin{aligned}\|\alpha(\gamma)\|_q &= \|(\gamma(e_1), \gamma(e_2), \dots)\|_q \\ &= (\sum_{i=1}^{\infty}|\gamma(e_i)|^q)^{\frac{1}{q}} \\ &= \lim_{n\to\infty}(\sum_{i=1}^{n}|\gamma(e_i)|^q)^{\frac{1}{q}} \leq \|\gamma\| < \infty.\end{aligned}$$

Hence, $\alpha(\gamma) \in l^q$. This means that the map $\alpha : (l^p)^* \to l^q$ is well-defined.

(bb) Assume that $x = (x_1, x_2, \dots) \in l^p$ and $y = (y_1, y_2, \dots) \in l^q$. Expanding the LHS yields

$$\begin{aligned}\phi(\alpha(\gamma))(x) &= \phi((\gamma(e_1), \gamma(e_2), \dots))(x) \\ &= \psi_{(\gamma(e_1),\gamma(e_2),\dots)}(x) \\ &= \langle(\gamma(e_1), \gamma(e_2), \dots), (x_1, x_2, \dots)\rangle \\ &= \sum_{i=1}^{\infty} \gamma(e_i)x_i = \sum_{i=1}^{\infty} \gamma(x_i e_i) \\ &= \gamma(\sum_{i=1}^{\infty} x_i e_i) = \gamma(x).\end{aligned}$$

Hence, $\phi \circ \alpha = 1_{(l^p)^*}$.

(bc) Once again, we compute directly

$$\begin{aligned}\alpha(\phi(y)) &= \alpha(\psi_y) \\ &= (\psi_y(e_1), \psi_y(e_2), \dots) = (\langle y, e_1\rangle, \langle y, e_2\rangle, \dots) \\ &= (\sum_{i=1}^{\infty} y_i(e_1)_i, \sum_{i=1}^{\infty} y_i(e_2)_i, \dots) \\ &= (y_1, y_2, \dots) = y.\end{aligned}$$

Subsequently, $\alpha \circ \phi = 1_{l^q}$.

(b) From parts (bb) and (bc), ϕ is invertible.

(c) Assume $y \in l^q$.

To show: (ca) $\|\phi(y)\| \geq \|y\|_q$.

(cb) $\|\phi(y)\| \leq \|y\|_q$.

(ca) We will show that there exists $x \in l^p$ such that $|\psi_y(x)| \geq \|x\|_p \|y\|_q$

Assume $y = (y_1, y_2, \dots)$. Define $x \in l^p$ by

$$x = (sgn(y_1)|y_1|^{q-1}, sgn(y_2)|y_2|^{q-1}, \dots).$$

Calculating $\|x\|_p$ yields

$$
\begin{aligned}
\|x\|_p &= (\sum_{i=1}^{\infty}|x_i|^p)^{\frac{1}{p}} \\
&= (\sum_{i=1}^{\infty}|sgn(y_i)|y_i|^{q-1}|^p)^{\frac{1}{p}} \\
&= (\sum_{i=1}^{\infty}|y_i|^{p(q-1)})^{\frac{1}{p}} = (\sum_{i=1}^{\infty}|y_i|^{pq(1-1/q)})^{\frac{1}{p}} \\
&= (\sum_{i=1}^{\infty}|y_i|^q)^{\frac{1}{p}} = (\sum_{i=1}^{\infty}|y_i|^q)^{\frac{1}{q}\times\frac{q}{p}} \\
&= (\|y\|_q)^{\frac{q}{p}} = (\|y\|_q)^{q(1-1/q)} = (\|y\|_q)^{q-1}.
\end{aligned}
$$

Using the above, we now find $|\psi_y(x)|$ to be

$$
\begin{aligned}
|\psi_y(x)| &= |\sum_{i=1}^{\infty} y_i x_i| \\
&= |\sum_{i=1}^{\infty} sgn(y_i)y_i \times |y_i|^{q-1}| \\
&= |\sum_{i=1}^{\infty}|y_i|^q| = (\|y\|_q)^q \\
&= (\|y\|_q)^{q-1} \times \|y\|_q = \|x\|_p\|y\|_q.
\end{aligned}
$$

So, $\frac{|\psi_y(x)|}{\|x\|_p} = \|y\|_q$, assuming $x \neq 0$. Taking the supremum of the LHS gives the following inequality

$$
\|y\|_q = \frac{|\psi_y(x)|}{\|x\|_p} \leq \sup_{x\neq 0}\frac{|\psi_y(x)|}{\|x\|_p} = \|\phi(y)\|.
$$

Hence in this case, $\|y\|_q \leq \|\phi(y)\|$.

(cb) Using the norm of a functional and the definition of ϕ, we obtain

$$
\|\phi(y)\| = \|\psi_y\| = \sup_{x\neq 0}\frac{|\psi_y(x)|}{\|x\|_p}.
$$

Assume $x \neq 0$. From Hölder's inequality, we obtain

$$\begin{aligned} |\langle y, x\rangle| &\leq \|x\|_p \|y\|_q \\ |\psi_y(x)| &\leq \|x\|_p \|y\|_q \\ \frac{|\psi_y(x)|}{\|x\|_p} &\leq \|y\|_q \end{aligned}$$

Taking the supremum of both sides gives $\|\phi(y)\| \leq \|y\|_q$.

(c) Combining (ca) and (cb), we can conclude that $\|\phi(y)\| = \|y\|_q$.

So, ϕ is an isomorphism and $(l^p)^* = l^q$.

□

Example 2.1.5. Consider $\mathbb{R}^n$ with the standard Euclidean norm. The dual space of $\mathbb{R}^n$ is $\mathbb{R}^n$ itself. In other words, $(\mathbb{R}^n)^* = \mathbb{R}^n$.

Proof. Assume that $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. In a similar fashion to the previous example, define the map $\beta : \mathbb{R}^n \to (\mathbb{R}^n)^*$ by

$$\beta(x) = \psi_x.$$

In turn, the functional $\psi_x : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\psi_x((y_1, y_2, \dots, y_n)) = \sum_{i=1}^{n} x_i y_i.$$

Let $y = (y_1, \dots, y_n)$. We will also use the notation $\psi_x(y) = \langle x, y\rangle$.

To show: (a) β is a linear functional.

(b) β is invertible.

(c) If $x \in \mathbb{R}^n$, then $\|\beta(x)\| = \|x\|$.

(a) Assume $k \in \mathbb{R}$ and $a, b, x \in \mathbb{R}^n$. Assume $a = (a_1, a_2, \dots)$, $b = (b_1, b_2, \dots)$ and $x = (x_1, x_2, \dots)$.

To show: (aa) $\beta(a + b) = \beta(a) + \beta(b)$.

(ab) $\beta(ka) = k\beta(a)$.

(aa) Using the definition of β, we compute directly

$$
\begin{aligned}
\beta(a+b)(x) &= \psi_{a+b}(x) \\
&= \langle a+b, x\rangle \\
&= \sum_{i=1}^{n}(a_i+b_i)x_i \\
&= \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} b_i x_i \\
&= \langle a, x\rangle + \langle b, x\rangle \\
&= \psi_a(x) + \psi_b(x) \\
&= \beta(a)(x) + \beta(b)(x).
\end{aligned}
$$

Hence, $\beta(a+b) = \beta(a) + \beta(b)$.

(ab) Once again, we compute directly

$$
\begin{aligned}
\beta(ka)(x) &= \psi_{ka}(x) \\
&= \langle ka, x\rangle \\
&= \sum_{i=1}^{n} ka_i x_i \\
&= k\sum_{i=1}^{n} a_i x_i \\
&= k\langle a, x\rangle \\
&= k\psi_a(x) \\
&= k\beta(a)(x).
\end{aligned}
$$

Therefore, $\beta(ka) = k\beta(a)$.

(b) Define the map $\zeta : (\mathbb{R}^n)^* \to \mathbb{R}^n$ by

$$\zeta(\gamma) = (\gamma(e_1), \gamma(e_2), \dots, \gamma(e_n))$$

Here, $\{e_1, e_2, \dots, e_n\}$ denotes the standard basis for $\mathbb{R}^n$.

To show: (ba) $\beta(\zeta(\gamma))(x) = \gamma(x)$.

(bb) $\zeta(\beta(y)) = y$.

(ba) Assume that $x = (x_1, x_2, \dots) \in \mathbb{R}^n$ and $y = (y_1, y_2, \dots) \in \mathbb{R}^n$. Expanding the LHS yields

$$\begin{aligned}
\beta(\zeta(\gamma))(x) &= \beta((\gamma(e_1), \gamma(e_2), \dots, \gamma(e_n)))(x) \\
&= \psi_{(\gamma(e_1),\dots,\gamma(e_n))}(x) \\
&= \langle(\gamma(e_1), \dots, \gamma(e_n)), (x_1, x_2, \dots, x_n)\rangle \\
&= \sum_{i=1}^{n} \gamma(e_i)x_i \\
&= \sum_{i=1}^{n} \gamma(x_i e_i) \\
&= \gamma(\sum_{i=1}^{n} x_i e_i) \\
&= \gamma(x).
\end{aligned}$$

Hence, $\beta \circ \zeta = 1_{(\mathbb{R}^n)^*}$.

(bb) Once again, we compute directly

$$\begin{aligned}
\zeta(\beta(y)) &= \zeta(\psi_y) \\
&= (\psi_y(e_1), \psi_y(e_2), \dots, \psi_y(e_n)) \\
&= (\langle y, e_1\rangle, \dots, \langle y, e_n\rangle) \\
&= (\sum_{i=1}^{n} y_i(e_1)_i, \dots, \sum_{i=1}^{n} y_i(e_n)_i) \\
&= (y_1, y_2, \dots, y_n) \\
&= y.
\end{aligned}$$

Subsequently, $\zeta \circ \beta = 1_{\mathbb{R}^n}$.

(b) Therefore, β is invertible.

(c) Assume $y \in \mathbb{R}^n$.

To show: (ca) $\|\beta(y)\| \leq \|y\|$.

(cb) $\|\beta(y)\| \geq \|y\|$.

(ca) Using the norm of a functional and the definition of β, we obtain

$$\|\beta(y)\| = \|\psi_y\| = \sup_{x\neq 0} \frac{|\psi_y(x)|}{\|x\|}.$$

Assume $x \neq 0$. From Hölder's inequality with $p = q = 2$ (Cauchy-Schwarz Inequality), we obtain

$$\begin{aligned} |\langle y, x\rangle| &\leq \|x\|\|y\| \\ |\psi_y(x)| &\leq \|x\|\|y\| \\ \frac{|\psi_y(x)|}{\|x\|} &\leq \|y\| \end{aligned}$$

Taking the supremum of both sides gives us the desired inequality. Hence, $\|\beta(y)\| \leq \|y\|$.

(cb) To show: There exists $x \in \mathbb{R}^n$ such that $|\psi_y(x)| \geq \|x\|\|y\|$

Assume $y = (y_1, y_2, \ldots, y_n)$. Set $x = y$. A calculation similar to part (ca) in the previous theorem shows that

$$\|y\| = \frac{|\psi_y(x)|}{\|x\|} \leq \sup_{x\neq 0} \frac{|\psi_y(x)|}{\|x\|} = \|\beta(y)\|.$$

Therefore, in this particular case, $\|\beta(y)\| \geq \|y\|$.

(c) Consequently, $\|\beta(y)\| = \|y\|$.

So, β is an isomorphism and $(\mathbb{R}^n)^* = \mathbb{R}^n$. □

The above example is particularly important and will reoccur in the context of the Riesz representation theorem for Hilbert spaces. Due to Theorem 1.2.7, $\mathbb{R}^n$ with the Euclidean inner product is the archetypal example of a Hilbert space and every Hilbert space (over $\mathbb{R}$ or $\mathbb{C}$) satisfies the Riesz representation theorem — they are *self-dual* with respect to a isomorphism similar in definition to β in the above example.

2.2 Embedding into the double dual

A particularly powerful application of the Hahn-Banach theorem is the fact that a Banach space X embeds into its double dual $(X^*)^*$. The precise statement is given below.

Theorem 2.2.1 (Embedding into the double dual)**.** *Let X be a Banach space. Define the map*

$$\begin{array}{rrcl} E: & X & \to & (X^*)^* \\ & x & \mapsto & E(x) \end{array}$$

where $E(x)$ is the map $\lambda \mapsto \lambda(x)$. Then, E is an injective isometry.

Proof. Assume that X is a Banach space and E is the map defined as above.

To show: (a) If $x \in X$ then $E(x) \in (X^*)^*$.

(b) E is injective.

(c) E is an isometry.

(a) Assume that $x \in X$. Then, the norm of $E(x)$ is bounded above by

$$\|E(x)\| = \sup_{\|\lambda\|=1} |\lambda(x)| \leq \sup_{\|\lambda\|=1} \|\lambda\|\|x\| = \|x\|.$$

Therefore, $E(x)$ is bounded. It is easy to check that $E(x)$ is a linear functional on X^*. So, $E(x) \in (X^*)^*$.

(b) Assume that $x \in \ker E$. Then, $E(x) = 0$ in X^*. This means that if $\lambda \in X^*$ then $\lambda(x) = 0$. Hence, $x = 0$ and E must be injective as a result.

(c) We already know from part (a) that if $x \in X$ then $\|E(x)\| \leq \|x\|$.

To show: (ca) $\|x\| \leq \|E(x)\|$.

(ca) By Theorem 2.1.6, there exists a functional $\phi \in X^*$ such that if $x \in X$ then $\phi(x) = \|x\|$ and $\|\phi\| = 1$. Therefore,

$$\|E(x)\| = \sup_{\|\lambda\|=1} |\lambda(x)| \geq |\phi(x)| = \|x\|.$$

From part (ca), we deduce that $\|E(x)\| = \|x\|$. This shows that E is an injective isometry from X to the double dual $(X^*)^*$. □

Theorem 2.2.1 leads to an important definition.

Definition 2.2.1. Let X be a Banach space. We say that X is **reflexive** if the injective isometry $X \hookrightarrow (X^*)^*$ is surjective and hence, bijective.

As we will see later, the Riesz representation theorem (see Theorem 3.3.1) provides us with examples of reflexive Banach spaces. We claim that the sequence space ℓ^1 is not reflexive. The argument we will make is based on the notion of separability.

Definition 2.2.2. Let X be a Banach space. We say that X is **separable** if there exists a countable dense subset D of X.

First, we will show that the dual of ℓ^1 is $(\ell^1)^* = \ell^\infty$. We know that the set $\{e_i \mid i \in \mathbb{Z}_{>0}\}$ forms a basis for ℓ^1, where e_i is the sequence with a 1 in the i^{th} position and zeros elsewhere.

Define the map

$$\begin{array}{cccc} \phi : & \ell^\infty & \to & (\ell^1)^* \\ & v = (v_1, v_2, \dots) & \mapsto & \phi(v) \end{array}$$

In turn, $\phi(v)$ is defined by

$$\begin{array}{cccc} \phi(v) : & \ell^1 & \to & \mathbb{C} \\ & (x_1, x_2, \dots) & \mapsto & \sum_{i=1}^\infty v_i x_i. \end{array}$$

To show: (a) If $v = (v_1, v_2, \dots) \in \ell^\infty$ then $\phi(v) \in (\ell^1)^*$.

(b) ϕ is a linear map.

(a) Assume that $v = (v_1, v_2, \dots) \in \ell^\infty$ and $x = (x_1, x_2, \dots) \in \ell^1$. We will first show that $\phi(v)$ is linear. Assume that $y = (y_1, y_2, \dots) \in \ell^1$ and $\lambda \in \mathbb{C}$. Then,

$$\begin{aligned} \phi(v)(x+y) &= \sum_{i=1}^\infty v_i(x_i + y_i) \\ &= \sum_{i=1}^\infty v_i x_i + \sum_{i=1}^\infty v_i y_i = \phi(v)(x) + \phi(v)(y) \end{aligned}$$

and

$$
\begin{aligned}
\phi(v)(\lambda x) &= \sum_{i=1}^{\infty} v_i(\lambda x_i) \\
&= \lambda \sum_{i=1}^{\infty} v_i x_i = \lambda \phi(v)(x).
\end{aligned}
$$

Hence, $\phi(v)$ is a linear functional on ℓ^1. To see that $\phi(v)$ is bounded, we compute its operator norm directly as

$$
\begin{aligned}
\|\phi(v)\| &= \sup_{\|x\|_1=1} |\phi(v)(x)| \\
&= \sup_{\|x\|_1=1} |\sum_{i=1}^{\infty} v_i x_i| \\
&\leq \sup_{\|x\|_1=1} \sum_{i=1}^{\infty} |v_i x_i| \\
&\leq \sup_{\|x\|_1=1} \sup_{j\in\mathbb{Z}_{>0}} |v_j| \sum_{i=1}^{\infty} |x_i| \\
&= \sup_{\|x\|_1=1} \|v\|_\infty \|x\|_1 = \|v\|_\infty.
\end{aligned}
$$

Since $v \in \ell^\infty$, $\|v\|_\infty$ is finite. Hence, $\|\phi(v)\| \leq \|v\|_\infty < \infty$ and so, $\phi(v)$ is bounded. Therefore, we have proven that if $v \in \ell^\infty$ then $\phi(v) \in (\ell^1)^*$.

(b) Assume that $w = (w_1, w_2, \dots) \in \ell^\infty$ and $\mu \in \mathbb{C}$. We compute directly that if $x = (x_1, x_2, \dots) \in \ell^1$ then

$$
\begin{aligned}
\phi(v + w)(x) &= \sum_{i=1}^{\infty} (v_i + w_i) x_i \\
&= \sum_{i=1}^{\infty} v_i x_i + \sum_{i=1}^{\infty} w_i x_i = \phi(v)(x) + \phi(w)(x)
\end{aligned}
$$

and

$$
\begin{aligned}
\phi(\mu v)(x) &= \sum_{i=1}^{\infty} (\mu v_i) x_i \\
&= \mu \sum_{i=1}^{\infty} v_i x_i = \mu \phi(v)(x).
\end{aligned}
$$

Therefore, ϕ must be a linear map.

In order to see that ϕ is invertible, we will define an explicit inverse for ϕ. Define

$$\begin{array}{rccc} \psi: & (\ell^1)^* & \to & \ell^\infty \\ & F & \mapsto & (F(e_1), F(e_2), \dots) \end{array}$$

To show: (c) If $F \in (\ell^1)^*$ then $\psi(F) \in \ell^\infty$.

(d) F is linear.

(e) $\psi \circ \phi = id_{\ell^\infty}$, where id_{ℓ^∞} is the identity operator on ℓ^∞.

(f) $\phi \circ \psi = id_{(\ell^1)^*}$, where $id_{(\ell^1)^*}$ is the identity operator on $(\ell^1)^*$.

(c) Assume that $F \in (\ell^1)^*$. By direct computation, we have

$$\begin{aligned} \|\psi(F)\|_\infty &= \sup_{i \in \mathbb{Z}_{>0}} |F(e_i)| \\ &\leq \sup_{i \in \mathbb{Z}_{>0}} \|F\| \|e_i\|_1 \\ &= \|F\| \sup_{i \in \mathbb{Z}_{>0}} \|e_i\|_1 = \|F\|. \end{aligned}$$

Since $\|F\| < \infty$ because F is a bounded linear functional on ℓ^1, we deduce that $\|\psi(F)\|_\infty < \infty$. So, $\psi(F) \in \ell^\infty$.

(d) Assume that $G \in (\ell^1)^*$ and $\lambda \in \mathbb{C}$. We compute directly that

$$\psi(F+G) = (F(e_1), F(e_2), \dots) + (G(e_1), G(e_2), \dots) = \psi(F) + \psi(G).$$

and

$$\psi(\lambda F) = (\lambda F(e_1), \lambda F(e_2), \dots) = \lambda \psi(F).$$

Consequently, ψ is linear.

(e) If $v = (v_1, v_2, \dots) \in \ell^\infty$ then

$$\psi(\phi(v)) = (\phi(v)(e_1), \phi(v)(e_2), \dots) = (v_1, v_2, \dots) = v.$$

Hence, $\psi \circ \phi = id_{\ell^\infty}$.

(f) Assume that $F \in (\ell^1)^*$. If $x = (x_1, x_2, \dots) \in \ell^1$ then

$$\begin{aligned}\phi(\psi(F))(x) &= \phi((F(e_1), F(e_2), \dots))(x) \\ &= \sum_{i=1}^{\infty} F(e_i)x_i = F(\sum_{i=1}^{\infty} x_i e_i) = F(x).\end{aligned}$$

Therefore, $\phi \circ \psi = id_{(\ell^1)^*}$.

By combining parts (c), (d), (e) and (f), we deduce that ψ is the inverse of ϕ. Finally, we will show that if $v = (v_1, v_2, \dots) \in \ell^\infty$ then $|\phi(v)| = \|v\|_\infty$. We already know from part (a) that $\|\phi(v)\| \leq \|v\|_\infty$.

To show: (g) $\|v\|_\infty \leq \|\phi(v)\|$.

(g) We know from part (c) that if $F \in (\ell^1)^*$ then $\|\psi(F)\|_\infty \leq \|F\|$. Since $\phi(v) \in (\ell^1)^*$, we have from part (e)

$$\|v\|_\infty = \|\psi(\phi(v))\|_\infty \leq \|\phi(v)\|.$$

Consequently, we have $\|\phi(v)\| = \|v\|_\infty$. So, ϕ is an isometric isomorphism between $(\ell^1)^*$ and ℓ^∞, establishing that the dual space of ℓ^1 is $(\ell^1)^* = \ell^\infty$.

In order to show that ℓ^1 is not reflexive, it suffices to prove that $(\ell^\infty)^* \neq \ell^1$. We require two more results for this purpose.

Theorem 2.2.2 (Mapping to the distance)**.** *Let X be a normed vector space and $Z \subseteq X$ be a non-zero subspace of X. Let $y \in X$ and*

$$d = \inf_{z \in Z} \|y - z\|.$$

Then, there exists a functional $\Lambda \in X^*$ such that $\|\Lambda\| \leq 1$, $\Lambda(y) = d$ and $\Lambda(z) = 0$ for $z \in Z$.

Proof. Assume that X is a normed vector space and $Z \subseteq X$ be a non-zero subspace of X. Assume that $y \in X \backslash Z$ is non-zero and $d \in \mathbb{R}_{\geq 0}$ is defined as above.

We will first define a linear functional Λ' on the subspace $span(y) + Z$. Define

$$\begin{array}{rccc} \Lambda' : & span(y) + Z & \to & \mathbb{C} \\ & \lambda y + z & \mapsto & |\lambda| \|y\|. \end{array}$$

Then, Λ' is a linear functional which satisfies $\Lambda'(z) = 0$ for all $z \in Z$ and $\Lambda'(y) = \|y\|$. It is bounded because if $\lambda y + z \in span(y) + Z$ for $\lambda \in \mathbb{C} - \{0\}$ then

$$\begin{aligned} \Lambda'(\frac{\lambda y + z}{\|\lambda y + z\|}) &= \frac{|\lambda| \|y\|}{\|\lambda y + z\|} \\ &= \frac{|\lambda| \|y\|}{|\lambda| \|y + \frac{1}{\lambda} z\|} \\ &= \frac{\|y\|}{\|y - (\frac{-1}{\lambda} z)\|} \leq \frac{\|y\|}{d}. \end{aligned}$$

Therefore, $\|\Lambda'\| \leq \frac{\|y\|}{d}$. Now we can apply Theorem 2.1.5 to obtain a functional $\tilde{\Lambda} \in X^*$ such that $\tilde{\Lambda}|_{span(y)+Z} = \Lambda'$ and $\|\tilde{\Lambda}\| = \|\Lambda'\|$.

Now define the linear functional $\Lambda \in X^*$ by $\Lambda(x) = \frac{d}{\|y\|} \tilde{\Lambda}(x)$. By the construction of $\tilde{\Lambda}$, we have

$$\|\Lambda\| = \frac{d}{\|y\|} \|\tilde{\Lambda}\| \leq \frac{d}{\|y\|} \frac{\|y\|}{d} = 1.$$

On the subspace $span(y) + Z$, we have

$$\Lambda(y) = \frac{d}{\|y\|} \tilde{\Lambda}(y) = \frac{d}{\|y\|} \Lambda'(y) = d$$

and if $z \in Z$ then

$$\Lambda(z) = \frac{d}{\|y\|} \tilde{\Lambda}(z) = \frac{d}{\|y\|} \Lambda'(z) = 0.$$

□

Recall that our proof that $\ell_\infty^* \neq \ell_1$ is based on the notion of separability. Specifically, we have the following theorem:

Theorem 2.2.3 (Separable dual space)**.** *Let X be a Banach space. If the dual space X^* is separable then X is separable.*

Proof. Assume that X is a Banach space and that X^* is separable. Then, there exists a countable dense subset $\{\lambda_n\}_{n\in\mathbb{Z}_{>0}}$ of X^*.

Define a sequence $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ such that $\|x_n\| = 1$ for all $n \in \mathbb{Z}_{>0}$ and

$$|\lambda_n(x_n)| \geq \frac{\|\lambda_n\|}{2}.$$

We can do this due to the definition of the operator norm.

Now define the set

$$\mathcal{D} = \{\sum_{i=1}^{k} \mu_i x_i \mid \mu_i \in \mathbb{Q}, k \in \mathbb{Z}_{>0}\}.$$

Since $\mathcal{D}$ is countable, it suffices to show that $\mathcal{D}$ is dense in X.

Suppose for the sake of contradiction that $\mathcal{D}$ is not dense in X. Then, there exists $y \in X - \mathcal{D}$ and a linear functional $\lambda \in X^*$ such that $\lambda(y) \neq 0$, but $\lambda(x) = 0$ for all $x \in \mathcal{D}$. The existence of λ uses Theorem 2.2.2.

Since $\{\lambda_n\}$ is a dense subset of X^*, there exists a subsequence $\{\lambda_{n_k}\}$ such that $\lambda_{n_k} \to \lambda$ as $k \to \infty$. By construction of the sequence $\{x_n\}$ in X, we have

$$\begin{aligned}\|\lambda - \lambda_{n_k}\| &= \sup_{\|x\|=1} |(\lambda - \lambda_{n_k})(x)| \\ &\geq |(\lambda - \lambda_{n_k})(x_{n_k})| = |\lambda_{n_k}(x_{n_k})| \geq \frac{\|\lambda_{n_k}\|}{2}.\end{aligned}$$

where the second last equality follows from the fact that $x_{n_k} \in \mathcal{D}$. By taking the limit as $k \to \infty$, we deduce that since $\lambda_{n_k} \to \lambda$, $\|\lambda_{n_k}\| \to 0$ as $k \to \infty$. Therefore, $\lambda = 0$, which contradicts the fact that $\lambda(y) \neq 0$. Hence, $\mathcal{D}$ is a countable dense subset of X and X must be separable as required. □

Using Theorem 2.2.3, we will now prove that $(\ell^\infty)^* \neq \ell^1$. Suppose for the sake of contradiction that $\ell^1 = (\ell^\infty)^*$.

To show: (a) ℓ^1 is separable.

(b) ℓ^∞ is not separable.

(a) Define the set

$$\mathcal{D} = \{x = (x_1, x_2, \dots) \mid x_i \in \mathbb{Q}, \text{there exists } j \in \mathbb{Z}_{>0} \text{ such that if } i > j \text{ then } x_i = 0\}.$$

Observe that the set $\mathcal{D}$ must also be countable. To see why this is the case, write

$$\mathcal{D} = \bigcup_{j \in \mathbb{Z}_{>0}} \mathcal{D}_j$$

where

$$\mathcal{D}_j = \{(x_1, x_2, \dots, x_j, 0, 0, \dots) \mid x_i \in \mathbb{Q} \text{ for } i \in \{1, 2, \dots, j\}\}.$$

Since $\mathbb{Q}$ is countable, $\mathcal{D}_j$ must be countable for $j \in \mathbb{Z}_{>0}$. Since $\mathcal{D}$ is a countable union of countable sets, it must be countable as a result.

We will show that $\mathcal{D}$ is a dense subset of

$$\ell^1(\mathbb{R}) = \{x = (x_1, x_2, \dots) \in \ell^1 \mid x_i \in \mathbb{R}\} \subseteq \ell^1.$$

Assume that $x = (x_1, x_2, \dots) \in \ell^1(\mathbb{R})$. For $i \in \mathbb{Z}_{>0}$, $x_i \in \mathbb{R}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists a sequence $\{q_{i,j}\}_{j \in \mathbb{Z}_{>0}}$ in $\mathbb{Q}$ such that $q_{i,j} \to x_i$ as $j \to \infty$.

This propels us to define the sequences

$$Q_i = (q_{1,i}, q_{2,i}, \dots, q_{i-1,i}, q_{i,i}, 0, 0, \dots)$$

for $i \in \mathbb{Z}_{>0}$. Then, $\{Q_i\}_{i \in \mathbb{Z}_{>0}}$ is a sequence in the set $\mathcal{D}$. To see that it converges to x, assume that $\epsilon \in \mathbb{R}_{>0}$. Since the sequence of partial sums $\{\sum_{j=1}^{k} |x_j|\}_{k \in \mathbb{Z}_{>0}}$ converges to $\|x\|_1 < \infty$, there exists $N' \in \mathbb{Z}_{>0}$ such that if $i > N'$ then

$$\sum_{j=i+1}^{\infty} |x_j| < \frac{\epsilon}{2}.$$

Now fix $\alpha \in \mathbb{Z}_{>0}$. For each $j \in \{1, 2, \dots, \alpha\}$ there exists $N_j \in \mathbb{Z}_{>0}$ such that if $i' > N_j$ then

$$|x_j - q_{j,i'}| < \frac{\epsilon}{2\alpha}.$$

Now let $N = \max(N_1, \dots, N_\alpha)$. If $i' > N$ then we have from the previous inequality that

$$|x_j - q_{j,i'}| < \frac{\epsilon}{2\alpha} \qquad \text{for all} \qquad j \in \{1, 2, \dots, \alpha\}.$$

Let $M = \max(N, N')$. If $i' > M$ then we choose $\alpha > i'$ so that

$$\begin{aligned} \|x - Q_{i'}\|_1 &= \sum_{j=1}^{\infty} |x_j - q_{j,i'}| \\ &= \sum_{j=1}^{i'} |x_j - q_{j,i'}| + \sum_{j=i'+1}^{\infty} |x_j| \\ &< \sum_{j=1}^{i'} \frac{\epsilon}{2\alpha} + \frac{\epsilon}{2} = \frac{\epsilon i'}{2\alpha} + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Therefore, the sequence $\{Q_i\}_{i \in \mathbb{Z}_{>0}}$ in $\mathcal{D}$ converges to $x \in \ell^1(\mathbb{R})$. This shows that $\mathcal{D}$ is a countable dense subset of $\ell^1(\mathbb{R})$.

Notice that we can write

$$\ell^1 = \ell^1(\mathbb{R}) + i\ell^1(\mathbb{R}).$$

where the sequences in $i\ell^1(\mathbb{R})$ are of the form $(iy_1, iy_2, \dots)$, where $y_j \in \mathbb{R}$ for $j \in \mathbb{Z}_{>0}$. Since $\mathcal{D}$ is dense in $\ell^1(\mathbb{R})$, $i\mathcal{D}$ is a countable dense subset of $i\ell^1(\mathbb{R})$. So, the set

$$\mathcal{D} + i\mathcal{D}$$

is a countable dense subset of ℓ^1. Therefore, ℓ^1 is separable.

(b) Let $I \subseteq \mathbb{Z}_{>0}$. Define the sequence e_I by

$$(e_I)_i = \begin{cases} 1, & \text{if } i \in I \\ 0, & \text{if } i \notin I. \end{cases}$$

The sequence $e_I \in \ell^\infty$ because

$$\|e_I\|_\infty = \sup_{i \in \mathbb{Z}_{>0}} |(e_I)_i| = 1.$$

If I and J are distinct subsets of $\mathbb{Z}_{>0}$ then

$$\|e_I - e_J\|_\infty = \sup_{i \in \mathbb{Z}_{>0}} |(e_I)_i - (e_J)_i| = 1.$$

This means that the open balls $B(e_I, \frac{1}{2})$ and $B(e_J, \frac{1}{2})$ are disjoint whenever $I \neq J$. Here,

$$B(e_I, \frac{1}{2}) = \{x \in \ell^\infty \mid \|x - e_I\| < \frac{1}{2}\}.$$

Now define $\mathcal{B} = \{B(e_I, \frac{1}{2}) \mid I \subseteq \mathbb{Z}_{>0}\}$.

To show: (ba) $\mathcal{B}$ is uncountable.

(ba) For every $I \subseteq \mathbb{Z}_{>0}$, e_I is a sequence of zeros and ones. Each sequence of zeros and ones corresponds to the binary representation of some real number in $(0, 1]$. For instance, the sequence $(1, 0, 0, \dots)$ corresponds to $\frac{1}{2} \in (0, 1]$ because the binary representation of $\frac{1}{2}$ is 0.1. The bijection is given explicitly by

$$\begin{array}{rccc} \phi: & (0,1] & \to & \mathcal{B} \\ & r & \mapsto & B(e_I, \frac{1}{2}) \end{array}$$

where e_I is the sequence of zeros and ones corresponding to the binary representation of r. Since ϕ defines a bijection between the sets $(0, 1]$ and $\mathcal{B}$ and $(0, 1]$ is uncountable, we deduce that $\mathcal{B}$ is also an uncountable set.

(b) Now suppose that $\mathcal{S}$ is a dense subset of ℓ^∞. If $\lambda \in \mathcal{S}$ then there exists $I \subseteq \mathbb{Z}_{>0}$ such that $\lambda \in B(e_I, \frac{1}{2})$ because $\mathcal{S}$ is dense in ℓ^∞. Since the elements of $\mathcal{B}$ are all disjoint, every element of $\mathcal{S}$ must be contained in exactly one of the open balls in $\mathcal{B}$. Since $\mathcal{B}$ is uncountable, $\mathcal{S}$ must also be uncountable and hence, ℓ^∞ is not separable.

Since $\ell^1 = (\ell^\infty)^*$ by assumption and ℓ^1 is separable by part (a), we can apply Theorem 2.2.3 to deduce that ℓ^∞ is separable. However, this contradicts the finding that ℓ^∞ is not separable in part (b). Therefore, $\ell^1 \neq (\ell^\infty)^*$ and ℓ^1 is not a reflexive Banach space as required.

2.3 Weak Convergence

Let X be a normed vector space. In this section, we will introduce the notion of weak convergence. This concept of convergence utilises bounded linear functionals on X, reinforcing the need to study dual spaces alongside normed vector spaces.

In order to distinguish weak convergence from our usual definition of convergence, we will call the latter **strong convergence**.

Definition 2.3.1. Let X be a normed vector space and let $\{x_n\}_{n\geq 1}$ be a sequence of points in X. We say that $\{x_n\}$ is **strongly convergent** if there exists a $x \in X$ such that

$$\lim_{n\to\infty} \|x_n - x\| = 0.$$

Now, contrast this definition with weak convergence below:

Definition 2.3.2. Let X be a normed vector space and let $\{x_n\}_{n\geq 1}$ be a sequence of points in X. We say that $\{x_n\}$ is **weakly convergent** if there exists a $x \in X$ such that for all $f \in X^*$,

$$\lim_{n\to\infty} f(x_n) = f(x).$$

For weak convergence, we write $x_n \rightharpoonup x$. We call x the *weak limit* of $\{x_n\}$ and that $\{x_n\}$ *converges weakly* to x.

Upon seeing both of these definitions, a sensible inquiry would be about the relationship between strong convergence and weak convergence. Our first step towards analysing this question is the following theorem:

Theorem 2.3.1. *Let X be a normed vector space and let $\{x_n\}$ be a sequence of points in X. If $\{x_n\}$ is strongly convergent, then it is weakly convergent.*

Proof. Assume that $\{x_n\}$ is a sequence in X which is strongly convergent. Then, there exists a $x \in X$ such that

$$\lim_{n\to\infty} \|x_n - x\| = 0.$$

To show: (a) For all $f \in X^*$, $\lim_{n\to\infty} f(x_n) = f(x)$.

(a) Assume that $f \in X^*$. Then,

$$\begin{aligned}
\lim_{n\to\infty} |f(x_n) - f(x)| &\leq \lim_{n\to\infty} \sup_{x_n \neq x} |f(x_n) - f(x)| \\
&= \lim_{n\to\infty} \sup_{x_n \neq x} \frac{|f(x_n) - f(x)|}{\|x_n - x\|} \times \|x_n - x\| \\
&= \lim_{n\to\infty} \|f\| \|x_n - x\| \\
&= 0.
\end{aligned}$$

Hence, for all $f \in X^*$, $\lim_{n\to\infty}|f(x_n) - f(x)| = 0$ since the LHS is greater than or equal to zero. As a result, $\lim_{n\to\infty} f(x_n) = f(x)$. Therefore, $\{x_n\}$ is weakly convergent.

□

In general, weak convergence does not imply strong convergence. However, there is one particularly important case where weak convergence and strong convergence are equivalent. The next theorem also explains why the concept of weak convergence does not appear in real-valued calculus.

Theorem 2.3.2. *Let X be a* ***finite dimensional*** *normed vector space and let $\{x_n\}$ be a sequence of points in X. Then, if $\{x_n\}$ is weakly convergent, then $\{x_n\}$ is also strongly convergent.*

Proof. Assume that X is a finite dimensional normed vector space, $\dim X = n$ and $\{x_n\}$ is a sequence in X which is weakly convergent. Then, there exists a $x \in X$ such that for all $f \in X^*$,

$$\lim_{n\to\infty} f(x_n) = f(x).$$

To show: (a) $\lim_{n\to\infty}\|x_n - x\| = 0$.

(a) Let $\{e_1, \ldots, e_n\}$ be a basis for X. Then, x_n and x can be expressed as a linear combination of these basis vectors. We write $x_n = \sum_{i=1}^n a_i^{(n)} e_i$ and $x = \sum_{i=1}^n a_i e_i$. Assume $i \in \{1, \ldots, n\}$. Define $g_i \in X^*$ by

$$g_i(e_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Then, for $j \leq n$, $g_j(x_n) = a_j^{(n)}$ and $g_j(x) = a_j$. Since, $\{x_n\}$ is weakly convergent, $\lim_{n\to\infty} g_j(x_n) = g_j(x)$. So, $\lim_{n\to\infty} a_j^{(n)} = a_j$. Using this piece of information, we now calculate $\|x_n - x\|$.

$$\begin{aligned} \|x_n - x\| &= \|\sum_{i=1}^n (a_i^{(n)} - a_i)e_i\| \\ &\leq \sum_{i=1}^n |a_i^{(n)} - a_i| \|e_i\| \quad \text{(Triangle Inequality)} \\ &\to 0 \qquad \text{as } n \to \infty. \end{aligned}$$

So, $\lim_{n\to\infty}\|x_n - x\| = 0$. Hence, $\{x_n\}$ is strongly convergent. □

These two results demonstrate that weak convergence is a more general form of strong convergence; the idea of convergence that we are most used to. Note that in the Banach space l^1, strong convergence and weak convergence are equivalent, despite the fact that l^1 is infinite dimensional (see [EK78, Page 260]).

Theorem 2.3.3. *Let X be a normed vector space and let $\{x_n\}$ be a sequence of points in X. Assume $x_n \rightharpoonup x$. Then, the weak limit x is unique.*

Proof. Assume X is a normed vector space and $\{x_n\}$ is a sequence of points in X. Suppose for the sake of contradiction that the weak limit x of the sequence $\{x_n\}$ is not unique. Then, there exists $x, y \in X$ such that $x \neq y$, $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$. So, for all $f \in X^*$, $\lim_{n\to\infty} f(x_n) = f(x)$ and $\lim_{n\to\infty} f(x_n) = f(y)$. Therefore, $f(x) = f(y)$ for all $f \in X^*$ because limits are unique in $\mathbb{R}$.

However, we also know that there exists a linear functional $\phi \in X^*$ such that $\phi(x) \neq \phi(y)$ because $x \neq y$ (see Theorem 2.1.7). This contradicts the fact that $f(x) = f(y)$ for all $f \in X^*$. So, the weak limit x must be unique. □

Theorem 2.3.4. *Let X be a normed vector space and let $\{x_n\}$ be a sequence of points in X. Assume $x_n \rightharpoonup x$. Then, every subsequence $\{x_{n_k}\}$ of $\{x_n\}$ converges weakly to x.*

Proof. Assume that $\{x_n\}$ is a sequence of points in the normed vector space X. Assume that $\{x_n\}$ converges weakly to $x \in X$. Then, for all $f \in X^*$, $\lim_{n\to\infty} f(x_n) = f(x)$. This means that the real-valued sequence $\{f(x_n)\}$ is convergent. Since it is convergent, it must also be a Cauchy sequence and hence, bounded. In turn, since $\{f(x_n)\}$ was established to be bounded, it must have a subsequence $\{f(x_{n_k})\}$, which converges to $f(x)$, by the Bolzano-Weierstrass theorem. Hence, for all $f \in X^*$,

$$\lim_{n\to\infty} f(x_{n_k}) = f(x).$$

So, every subsequence $\{x_{n_k}\}$ of $\{x_n\}$ converges weakly to x. □

Theorem 2.3.5. *Let X be a normed vector space and let $\{x_n\}$ be a sequence of points in X. Assume $x_n \rightharpoonup x$. Then, the sequence $\{\|x_n\|\}$ is bounded.*

Proof. Assume that $\{x_n\}$ is a sequence of points in the normed vector space X. Assume that $\{x_n\}$ converges weakly to $x \in X$. Due to the completeness of $\mathbb{R}$, it suffices to show that the real-valued sequence $\{\|x_n\|\}$ is convergent because a convergent sequence in $\mathbb{R}$ is necessarily Cauchy and as a consequence of this, a Cauchy sequence must be bounded.

To show: (a) $\{\|x_n\|\}$ is convergent.

(a) Since $\{x_n\}$ converges weakly to $x \in X$, we know that for all $f \in X^*$, $\lim_{n\to\infty} f(x_n) = f(x)$. However, we also know that there exists a linear functional $\phi \in X^*$ such that $\|\phi\| = 1$ and $\phi(x) = \|x\|$. So, $\lim_{n\to\infty} \phi(x_n) = \phi(x)$, which simply means that

$$\lim_{n\to\infty} \|x_n\| = \|x\|.$$

Therefore from the above, the real-valued sequence $\{\|x_n\|\}$ is convergent. □

It turns out that the concept of weak convergence has important applications to the calculus of variations and the theory of partial differential equations. In this section, we have covered some of the basic properties of weak convergence. However, we will have to develop some more theory before we are able to learn deeper results about weak convergence and dual spaces. The main purpose of this chapter is to accentuate why it is important for one to study dual spaces alongside normed vector spaces. Weak convergence will reappear in the context of Hilbert spaces, the main subject of the next chapter.

Chapter 3

Hilbert Spaces

3.1 Definition and Examples

An appropriate analogy pertaining to Hilbert spaces is: "Banach space is to normed vector space as Hilbert space is to inner product space". Hilbert spaces are extremely important to certain areas of study, such as quantum physics and Fourier analysis. This is partly due to the numerous deep theorems associated with Hilbert spaces, some of which we shall explore in this chapter. Of course, we will first begin with the definition of a Hilbert space.

Definition 3.1.1. An **inner product space** H is a vector space over a field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, equipped with a bilinear map $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{K}$ which satisfies the following properties:

1. For all $x, y \in H$, $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (Skew Symmetry)
2. For all $x, y \in H$ and $a \in \mathbb{K}$, $\langle ax, y \rangle = a\langle x, y \rangle$ (Scalar Multiplication)
3. For all $x, y, z \in H$, $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ (Linearity)
4. For all $x \in H$, $\langle x, x \rangle \geq 0$, with equality holding if and only if $x = 0$. (Positive Definiteness)

A short application of these axioms reveal these two particular properties of the inner product. For all $x, y, z \in H$ and $a \in \mathbb{K}$, $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$ and $\langle x, ay \rangle = \overline{a}\langle x, y \rangle$.

Definition 3.1.2. Assume that H is an inner product space and that $x \in H$. We define the **norm** of an inner product space H to be

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Definition 3.1.3. Assume that H is an inner product space. Then, H is called a **Hilbert space** if it is complete with respect to the norm induced by the inner product.

Before we launch into some examples of Hilbert spaces, we will prove some important results relating to the norm of a Hilbert space.

Theorem 3.1.1 (Cauchy-Schwarz Inequality)**.** *Assume H is an inner product space and $x, y \in H$. Then,*

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

Proof. Assume H is an inner product space and that $x, y \in H$.

To show: (a) $\|x\|^2\|y\|^2 - |\langle x, y\rangle|^2 \geq 0$.

(a) Consider $\langle x + \alpha y, x + \alpha y\rangle$ for some scalar $\alpha \in \mathbb{K}$. Expanding this expression yields

$$\langle x + \alpha y, x + \alpha y\rangle = \|x\|^2 + \overline{\alpha}\langle x, y\rangle + \alpha\langle y, x\rangle + |\alpha|^2\|y\|^2 \geq 0.$$

Select $\alpha = -\frac{\langle x,y\rangle}{\|y\|^2}$. Substituting this into the above, one obtains the inequality

$$\|x\|^2 - 2|\langle x, y\rangle|^2/\|y\|^2 + |\langle x, y\rangle|^2/\|y\|^2 \geq 0.$$

So, $\|x\|^2 - |\langle x, y\rangle|^2/\|y\|^2 \geq 0$. Finally, multiplying both sides of the inequality by $\|y\|^2$ gives us the desired inequality. □

Theorem 3.1.2 (Minkowski Inequality)**.** *Assume H is an inner product space and $x, y \in H$. Then,*

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

Proof. Assume H is an inner product space and that $x, y \in H$.

To show: (a) $\|x + y\|^2 \leq (\|x\| + \|y\|)^2$.

(a) Starting with the LHS, we compute

$$
\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \|x\|^2 + \|y\|^2 + (\langle x,y\rangle + \langle y,x\rangle) \\
&= \|x\|^2 + \|y\|^2 + (\langle x,y\rangle + \overline{\langle x,y\rangle}) \\
&= \|x\|^2 + \|y\|^2 + 2Re(\langle x,y\rangle) \\
&\leq \|x\|^2 + \|y\|^2 + 2|\langle x,y\rangle| \\
&\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \quad \text{(Cauchy-Schwarz Inequality)} \\
&= (\|x\| + \|y\|)^2.
\end{aligned}
$$

□

Theorem 3.1.3 (Parallelogram Identity)**.** *Assume H is an inner product space and $x, y \in H$. Then,*

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

Proof. Assume H is an inner product space and $x, y \in H$. Expanding the LHS yields

$$
\begin{aligned}
\|x+y\|^2 + \|x-y\|^2 &= \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\
&= 2(\|x\|^2 + \|y\|^2) + \langle x,y\rangle + \langle y,x\rangle - \langle x,y\rangle - \langle y,x\rangle \\
&= 2(\|x\|^2 + \|y\|^2).
\end{aligned}
$$

□

Definition 3.1.4. Assume H is an inner product space and $x, y \in H$. Then, x and y are said to be **orthogonal** if $\langle x, y\rangle = 0$.

Theorem 3.1.4 (Pythagoras Theorem)**.** *Assume H is an inner product space and $x, y \in H$. Assume x and y are orthogonal. Then,*

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

Proof. Assume H is an inner product space and $x, y \in H$. Assume x and y are orthogonal. Once again, we expand the LHS to get

$$
\begin{aligned}
\|x\|^2 + \|y\|^2 &= \langle x,x\rangle + \langle y,y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \langle x+y, x+y\rangle \\
&= \|x+y\|^2.
\end{aligned}
$$

□

Here are some examples of Hilbert spaces.

Example 3.1.1. Let $H = \mathbb{R}^n$. Equip H with the standard dot product

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \sum_{i=1}^{n} x_i y_i.$$

Then, $\mathbb{R}^n$ is a Hilbert space.

Example 3.1.2. Let $H = \mathbb{C}^n$. Equip H with the standard complex dot product

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Then, $\mathbb{C}^n$ is a Hilbert space.

Example 3.1.3. Let $H = l^2$. We equip H with the inner product

$$\langle (x_1, x_2, \dots), (y_1, y_2, \dots) \rangle = \sum_{i=1}^{\infty} x_i y_i.$$

Then, l^2 is a Hilbert space.

Example 3.1.4. Let $H = L^2(\Omega)$, where Ω is an open subset of $\mathbb{R}$. We equip H with the inner product

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

Then, $L^2(\Omega)$ is a Hilbert space.

For the last two examples, it is reasonable for one to ask whether l^p and $L^p(\Omega)$ are Hilbert spaces when $p \neq 2$. We do know that l^p and $L^p(\Omega)$ are Banach spaces where $1 \leq p \leq \infty$. Does this necessarily mean that they are all Hilbert spaces? This will be addressed in the next two theorems.

Theorem 3.1.5. *If $p \neq 2$, then l^p is **not** a Hilbert space.*

Proof. Suppose for the sake of contradiction that l^p is a Hilbert space when $p \neq 2$. Then, l^p must satisfy the parallelogram identity (3.1.3).

To show: (a) There exists a $x, y \in l^p$ such that
$\|x+y\|^2 + \|x-y\|^2 \neq 2(\|x\|^2 + \|y\|^2)$.

(a) Define $x = (1, 1, 0, 0, 0, \dots)$ and $y = (-1, 1, 0, 0, 0, \dots)$. Clearly, $x, y \in l^p$. Utilising the norm of l^p, we calculate $\|x+y\| = \|x-y\| = 2$ and $\|x\| = \|y\| = 2^{1/p}$. Substituting all of this into the parallelogram identity, one obtains

$$4 + 4 = 2(2^{1/p} + 2^{1/p}).$$

This simplifies to $2 = 2^{2/p}$. This can only be true if $p = 2$. However, we have assumed that $p \neq 2$. This is a contradiction. Therefore, the parallelogram identity (3.1.3) does not holds for $p \neq 2$.

Hence, l^p is not a Hilbert space when $p \neq 2$.

□

Theorem 3.1.6. *Assume that Ω is an open subset of $\mathbb{R}$. If $p \neq 2$, then $L^p(\Omega)$ is **not** a Hilbert space.*

Proof. Assume that Ω is an open subset of $\mathbb{R}$. Suppose for the sake of contradiction that $L^p(\Omega)$ is a Hilbert space when $p \neq 2$. Then, $L^p(\Omega)$ must satisfy the parallelogram identity.

To show: (a) There exists a $f, g \in L^p(\Omega)$ such that
$\|f+g\|^2 + \|f-g\|^2 \neq 2(\|f\|^2 + \|g\|^2)$.

(a) Assume $\Omega = (a, b)$. Define the simple function $f(x)$ on (a, b) by

$$f(x) = \begin{cases} (\frac{3}{b-a})^{1/p} & \text{if } x \in (a, a + \frac{2(b-a)}{3}), \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, we define the simple function $g(x)$ on (a, b) by

$$g(x) = \begin{cases} -(\frac{3}{b-a})^{1/p} & \text{if } x \in (a, a + \frac{(b-a)}{3}), \\ (\frac{3}{b-a})^{1/p} & \text{if } x \in (a + \frac{(b-a)}{3}, a + \frac{2(b-a)}{3}), \\ 0 & \text{otherwise.} \end{cases}$$

Now, we compute the norms of $f, g, f+g$ and $f-g$ below.

$$
\begin{aligned}
\|f\| &= (\int_a^b |f|^p \,\mathrm{d}x)^{1/p} \\
&= (\int_a^{a+2(b-a)/3} |(\frac{3}{b-a})^{1/p}|^p \,\mathrm{d}x)^{1/p} \\
&= (\int_a^{a+2(b-a)/3} \frac{3}{b-a} \,\mathrm{d}x)^{1/p} \\
&= 2^{1/p}.
\end{aligned}
$$

$$
\begin{aligned}
\|g\| &= (\int_a^b |g|^p \,\mathrm{d}x)^{1/p} \\
&= (\int_a^{a+(b-a)/3} |-(\frac{3}{b-a})^{1/p}|^p \,\mathrm{d}x + \int_{a+(b-a)/3}^{a+2(b-a)/3} |(\frac{3}{b-a})^{1/p}|^p \,\mathrm{d}x)^{1/p} \\
&= (\int_a^{a+2(b-a)/3} \frac{3}{b-a} \,\mathrm{d}x)^{1/p} \\
&= 2^{1/p}.
\end{aligned}
$$

$$
\begin{aligned}
\|f+g\| &= (\int_a^b |f+g|^p \,\mathrm{d}x)^{1/p} \\
&= (\int_{a+(b-a)/3}^{a+2(b-a)/3} |2(\frac{3}{b-a})^{1/p}|^p \,\mathrm{d}x)^{1/p} \\
&= (\int_{a+(b-a)/3}^{a+2(b-a)/3} 2^p\frac{3}{b-a} \,\mathrm{d}x)^{1/p} \\
&= 2.
\end{aligned}
$$

$$
\begin{aligned}
\|f-g\| &= (\int_a^b |f-g|^p \,\mathrm{d}x)^{1/p} \\
&= (\int_a^{a+(b-a)/3} |2(\frac{3}{b-a})^{1/p}|^p \,\mathrm{d}x)^{1/p} \\
&= (\int_a^{a+(b-a)/3} 2^p\frac{3}{b-a} \,\mathrm{d}x)^{1/p} \\
&= 2.
\end{aligned}
$$

In a similar vein to the previous theorem, we find that the parallelogram identity is only satisfied if $p = 2$. However, we have assumed that $p \neq 2$.

This is a contradiction.

Hence, $L^p(\Omega)$ is not a Hilbert space for $p \neq 2$. $\square$

3.2 Orthogonality

Let V be a finite dimensional inner product space and $W \subset V$ be a subspace. We know that V is a direct sum of W and $W^\perp$, where $W^\perp$ is the *orthogonal complement* of W. The question that we will explore in this section is whether this decomposition holds more generally for infinite dimensional Hilbert spaces. First, we require some definitions.

Definition 3.2.1. Let X be a set of vectors from a Hilbert space H. Consider the subspace $span(X) \subset H$. The **space generated by** X, which we denote by V, is defined by $V = \overline{span(X)}$.

Definition 3.2.2. Let X be a set of vectors from a Hilbert space H. The set X is called **total** if for all $h \in H$, there exists a sequence $\{x_n\}$ with $x_n \in span(X)$ for all $n \in \mathbb{Z}_{>0}$, such that

$$\|x_n - h\| \to 0 \quad \text{as } n \to \infty.$$

Definition 3.2.3. Assume that X is a subset of the Hilbert space H. Then, the subset $X^\perp$ is called the **orthogonal subspace of** X and is defined by

$$X^\perp = \{h \in H \text{ such that } \langle h, x\rangle = 0 \text{ for all } x \in X\}.$$

This definition is no different to that of the orthogonal complement of a subspace of a finite dimensional inner product space.

Theorem 3.2.1. *Assume that X is a subset of the Hilbert space H. Then, the subspace $X^\perp$ is closed.*

Proof. Assume that X is a subset of the Hilbert space H.

To show: (a) $X^\perp \subseteq \overline{X^\perp}$.

(b) $\overline{X^\perp} \subseteq X^\perp$.

(a) Assume $w \in X^\perp$. Consider the constant sequence $\{w_n\} = (w, w, w, \dots)$. This is a sequence in $X^\perp$ which converges to w. As a consequence of 1.2.2, $w \in \overline{X^\perp}$. So, $X^\perp \subseteq \overline{X^\perp}$.

(b) Assume $w \in \overline{X^\perp}$. Then, from 1.2.2, there exists a sequence $\{w_n\}$ with $w_n \in X^\perp$ for all $n \in \mathbb{Z}_{>0}$ such that $w_n \to w$ as $n \to \infty$.

To show: (ba) For all $x \in X$, $\langle w, x\rangle = 0$

(ba) Starting with the expression $\langle w, x\rangle$ for all $x \in X$, we argue as follows

$$\begin{aligned}\langle w, x\rangle &= \langle w - w_n + w_n, x\rangle \\ &= \langle w - w_n, x\rangle + \langle w_n, x\rangle \\ &= \langle w - w_n, x\rangle \quad \text{since } w_n \in X^\perp.\end{aligned}$$

As $n \to \infty$, $\langle w, x\rangle = \langle 0, x\rangle = 0$. Therefore, $w \in X^\perp$ and so, $\overline{X^\perp} \subseteq X^\perp$.

Subsequently, $X^\perp = \overline{X^\perp}$. Hence, $X^\perp$ is closed. □

Now, we will investigate orthogonal projections onto closed subspaces of H.

Theorem 3.2.2. *Let H be a Hilbert space and V be a closed subspace of H. Then, $H = V \oplus V^\perp$. In other words, for all $x \in H$, $x = y + z$, where $y \in V$ is the unique point in V which has minimal distance from x. Similarly, $z \in V^\perp$ is the unique point in $V^\perp$ which has minimal distance from x.*

Proof. Assume that H is a Hilbert space and V is a closed subspace of H.

To show: (a) There exists a point $y \in V$ such that y has minimal distance from x.

(b) The point y is unique.

(c) There exists a point $z \in V^\perp$ such that z has minimal distance from x.

(d) The point z is unique.

(e) $x = y + z$.

(a) Assume that $x \in H$. Define

$$\alpha = d(x, V) = \inf\{\|x - v\| \mid v \in V\}.$$

Observe that there exists a sequence $\{y_n\}$ in V such that $\lim_{n\to\infty}\|x - y_n\| = \alpha$.

To show: (aa) The sequence $\{y_n\}$ is Cauchy.

(aa) We will utilise the parallelogram identity (3.1.3) for this. For all $u, v \in H$,

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Select $u = x - y_n$ and $v = x - y_m$ in order to obtain

$$\|2x - y_n - y_m\|^2 + \|y_m - y_n\|^2 = 2(\|x - y_n\|^2 + \|x - y_m\|^2).$$

This can be rearranged in order to get

$$\|y_m - y_n\|^2 = 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4\|x - (\frac{y_m + y_n}{2})\|^2.$$

From the definition of infimum and the fact that $(y_m + y_n)/2 \in V$, we know that

$$\|x - (\frac{y_m + y_n}{2})\|^2 \geq \alpha^2.$$

Using supremums, we have

$$\begin{aligned}
\sup \|y_m - y_n\|^2 &= 2(\sup \|x - y_n\|^2 + \sup \|x - y_m\|^2) - 4 \sup \|x - (\frac{y_m + y_n}{2})\|^2 \\
&\leq 2(\sup \|x - y_n\|^2 + \sup \|x - y_m\|^2) - 4 \inf \|x - (\frac{y_m + y_n}{2})\|^2 \\
&\leq 2(\alpha^2 + \alpha^2) - 4\alpha^2 \\
&= 0.
\end{aligned}$$

Taking limits of both sides as $m \to \infty$ and $n \to \infty$, we obtain $\lim_{m,n\to\infty} \|y_m - y_n\| = 0$. Hence, the sequence $\{y_n\}$ is Cauchy.

(a) Since V is a closed subspace of H, it is complete. Furthermore, the sequence $\{y_n\}$ is Cauchy. So, $\{y_n\}$ must converge to (say) $y \in V$. Hence, there exists a $y \in V$ such that $\|x - y\| = \alpha$.

(b) Assume that there exists $y' \in V$ such that $\|x - y'\| = \alpha$.

To show: (ba) $\|y - y'\| = 0$.

(ba) In the parallelogram identity, select $u = x - y$ and $v = x - y'$ in order to obtain

$$\|y-y'\|^2 = 2(\|x-y\|^2 + \|x-y'\|^2) - 4\|x-(\frac{y+y'}{2})\|^2.$$

Once again, since $(y+y')/2 \in V$,

$$\|x-(\frac{y+y'}{2})\|^2 \geq \alpha^2.$$

So,

$$\begin{aligned}\|y-y'\|^2 &= 2(\|x-y\|^2 + \|x-y'\|^2) - 4\|x-(\frac{y+y'}{2})\|^2 \\ &= 2(\alpha^2+\alpha^2) - 4\|x-(\frac{y+y'}{2})\|^2 \\ &\leq 4\alpha^2 - 4\alpha^2 \\ &= 0.\end{aligned}$$

Therefore, $\|y-y'\| = 0$ and so, $y = y'$.

(b) This demonstrates that y is unique.

(c) Define $z = x - y$.

To show: (ca) $z \in V^{\perp}$.

(cb) z has minimal distance from x.

(ca) Assume $v \in V$. Consider the expression $\|x-(y+\lambda v)\|^2$, where $\lambda \in \mathbb{R}$. Expanding this using the inner product, we get

$$\|x-(y+\lambda v)\|^2 = \|x-y\|^2 + |\lambda|^2\|v\|^2 + 2Re\langle x-y, \lambda v\rangle.$$

Due to the definition of y, the LHS attains a unique minimum when $\lambda = 0$. So, differentiating the RHS with respect to the variable λ yields

$$2|\lambda|\|v\|^2 + 2Re\langle x-y, v\rangle.$$

This is equal to zero when $\lambda = 0$. As a result, $Re\langle x-y, v\rangle = 0$ for all $v \in V$.

If H is a complex Hilbert space, then we observe that

$$Im\langle x-y, v\rangle = Re\langle x-y, iv\rangle.$$

Thus, $Im\langle x-y,v\rangle = 0$ because $Re\langle x-y,v\rangle = 0$ for all $v \in V$. So, in either case where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, we have $\langle x-y,v\rangle = 0$ for all $v \in V$. So, $z = x - y \in V^{\perp}$.

(cb) This follows from the fact that y is the point of minimal distance from x in V.

(d) Define $z' = x - y' \in V^{\perp}$.

To show: (da) $y = y'$.

(da) Similarly to before, we proceed as follows

$$\begin{aligned}\|y-y'\|^2 &= \langle y-y', y-y'\rangle \\ &= \langle y-y', x-y'\rangle - \langle y-y', x-y\rangle \\ &= \langle y-y', z'\rangle - \langle y-y', z\rangle \\ &= 0 \quad \text{since } y, y' \in V.\end{aligned}$$

Hence, $\|y-y'\| = 0$ and so, $y = y'$. Therefore, the point z is unique.

(e) $y + z = y + x - y = x$. Note that x is arbitrary. So, this decomposition works for all $x \in H$. □

We will use Theorem 3.2.2 to prove various properties about the orthogonal complement.

Theorem 3.2.3. *Let H be a Hilbert space.*

(a) If S is a subspace of H then $S \subseteq (S^{\perp})^{\perp}$.

(b) If S_1, S_2 are subspaces of H and $S_1 \subseteq S_2$ then $S_2^{\perp} \subseteq S_1^{\perp}$.

(c) If S is a closed subspace of H then $S = (S^{\perp})^{\perp}$.

(d) If S is a subspace of H, then $\overline{S} = (S^{\perp})^{\perp}$.

Proof. Assume that H is a Hilbert space.

(a) Assume that S is a subspace of H and that $s \in S$. If $t \in S^{\perp}$ then $\langle s,t\rangle = 0$. So, $s \in (S^{\perp})^{\perp}$ and $S \subseteq (S^{\perp})^{\perp}$.

(b) Assume that S_1, S_2 are subspaces of H such that $S_1 \subseteq S_2$. Then,

$$\begin{aligned} S_2^\perp &= \{x \in H \mid \langle x, v\rangle = 0 \text{ for } v \in S_2\} \\ &\subseteq \{x \in H \mid \langle x, v\rangle = 0 \text{ for } v \in S_1\} \\ &= S_1^\perp. \end{aligned}$$

(c) Assume that S is a closed subspace of H. Then, by Theorem 3.2.2, $H = S \oplus S^\perp$. We already have $S \subseteq (S^\perp)^\perp$ by part (a). It suffices to prove the reverse inclusion. Assume that $x \in (S^\perp)^\perp$. Then, $x = y + z$ where $y \in S$ and $z \in S^\perp$.

To show: (ca) $z = 0$.

(ca) Since $x \in (S^\perp)^\perp$ and $z \in S^\perp$, $\langle x, z\rangle = 0$. But,

$$\langle x, z\rangle = \langle y, z\rangle + \|z\|^2 = \|z\|^2 = 0.$$

Thus, $z = 0$ and $x = y \in S$. Hence, $(S^\perp)^\perp \subseteq S$ and $S = (S^\perp)^\perp$.

(d) Assume that S is a subspace of H. Then, $S \subseteq \overline{S}$. Applying part (b) twice, we deduce that $(S^\perp)^\perp \subseteq (\overline{S}^\perp)^\perp$. Since $\overline{S}$ is a closed subspace of H, $(\overline{S}^\perp)^\perp = \overline{S}$ from part (c). Thus, $(S^\perp)^\perp \subseteq \overline{S}$.

Now observe that from part (a), $\overline{S} \subseteq \overline{(S^\perp)^\perp} = (S^\perp)^\perp$. The last equality follows from the fact that $(S^\perp)^\perp$ is a closed subspace of H. So, $\overline{S} = (S^\perp)^\perp$ as required. □

In the context of Theorem 3.2.2, we define the projection operators $P_V : H \to V$ and $P_{V^\perp} : H \to V^\perp$ by

$$P_V(x) = y \text{ and } P_{V^\perp}(x) = z.$$

Theorem 3.2.4. *The projection operators are linear, continuous operators with norm less than or equal to* 1.

Proof. Assume $x, x' \in H$ and $a, b \in \mathbb{K}$.

To show: (a) $P_V(ax + bx') = aP_V(x) + bP_V(x')$.

(b) $P_{V^\perp}(ax + bx') = aP_{V^\perp}(x) + bP_{V^\perp}(x')$.

(c) $\|P_V\| \leq 1$.

(d) $\|P_{V^\perp}\| \leq 1$.

(a) Assume $P_V(x) = y$ and $P_V(x') = y'$. Then,

$$P_V(ax + bx') = ay + by' = aP_V(x) + bP_V(x').$$

Hence, P_V is linear.

(b) Assume $P_{V^\perp}(x) = z$ and $P_{V^\perp}(x') = z'$. Then,

$$P_{V^\perp}(ax + bx') = az + bz' = aP_{V^\perp}(x) + bP_{V^\perp}(x').$$

Hence, $P_{V^\perp}$ is linear.

(c) We know that

$$\|P_V\| = \sup_{\|x\|=1} \|P_V(x)\| = \sup_{\|x\|=1} \|y\|.$$

However, we also know from Pythagoras's theorem that

$$\|P_V(x)\|^2 + \|P_{V^\perp}(x)\|^2 = \|x\|^2.$$

Setting $\|x\| = 1$, we obtain the inequality $\|P_V\| \leq 1$.

(d) Setting $\|x\| = 1$, we also obtain the inequality $\|P_{V^\perp}\| \leq 1$.

Parts (c) and (d) prove that the projection operators are bounded. Hence, they are also continuous. □

We are able to characterise projection operators on a Hilbert space by its specific properties. The first property integral to this goal is that if H is a Hilbert space and S is a closed subspace of H, then the projection operator P_S onto S must satisfy $P_S^2 = P_S$. This is a consequence of 3.2.2.

The second defining property of projection operators is that they are self-dual with respect to the inner product on H.

Lemma 3.2.5. *Let H be a Hilbert space and S be a closed subspace of H. Let $P_S : H \to S$ be the projection operator onto S. Then, P_S is equal to its adjoint P_S^*. That is, for all $x, y \in H$, $\langle P_S(x), y\rangle = \langle x, P_S(y)\rangle$.*

Proof. Assume that H is a Hilbert space and S is a closed subspace of S. Assume that P_S is the projection operator from H onto S. Applying 3.3.1, the adjoint operator P_S^* admits the following definition:

$$P_S^* : S^* \to H^*$$

$$\langle -, s \rangle \mapsto \langle P_S(-), s \rangle$$

By definition of the adjoint and the decomposition $H = S \oplus S^\perp$ (see 3.2.2), we have for all $s \in S$,

$$\langle x, s \rangle = \langle P_S(x) + P_{S^\perp}(x), s \rangle = \langle P_S(x), s \rangle = \langle x, P_S^*(s) \rangle.$$

Therefore, $P_S^*(s) = s$ for all $s \in S$ since the above holds for all $x \in H$. Another useful observation one makes about the adjoint operator P_S^* is that its kernel is

$$\ker(P_S^*) = [\mathrm{im}(P_S)]^\perp = S^\perp.$$

One can check this directly from the definitions:

$$\begin{aligned}
\ker(P_S^*) &= \{x \in H \mid P_S^*(x) = 0\} \\
&= \{x \in H \mid \langle P_S^*(x), y \rangle = 0 \text{ for all } y \in H\} \\
&= \{x \in H \mid \langle x, P_S(y) \rangle = 0 \text{ for all } y \in H\} \\
&= S^\perp.
\end{aligned}$$

Using the above fact, we must have for all $x \in H$,

$$P_S^*(x) = P_S^*(s + s') = P_S^*(s) + P_S^*(s') = s$$

where $x = s + s'$ with $s \in S$ and $s' \in S^\perp$. This reveals that the projection operator is self-adjoint. □

Here is the promised characterisation of projection operators.

Theorem 3.2.6. *Let H be a Hilbert space and $P : H \to H$ be a bounded, linear operator which satisfies $P^2 = P$ and $P^* = P$. Then, P is the projection operator onto the closed subspace im(P).*

Proof. Assume that H is a Hilbert space and $P : H \to H$ is a bounded linear operator which satisfies $P^2 = P$ and $P^* = P$. The first part of the proof is to show that the image $\mathrm{im}(P)$ is a closed subspace of H.

To show: (a) The image $\operatorname{im}(P)$ is closed.

(a) Let $I : H \to H$ denote the identity operator on H. Consider the operator $I - P : H \to H$. We observe that $I - P$ satisfies the same properties as P. This is because

$$(I-P)^2 = I^2 - P - P + P^2 = I - P - P + P = I - P$$

and for all $x, y \in H$,

$$\begin{aligned}\langle (I-P)x, y\rangle &= \langle x, y\rangle - \langle P(x), y\rangle \\ &= \langle x, y\rangle - \langle x, P(y)\rangle \\ &= \langle x, (I-P)y\rangle.\end{aligned}$$

Furthermore, $\operatorname{im}(P) = \ker(I-P)$. To see why this is the case, suppose that $P(x) \in \operatorname{im}(P)$. Then, $(I-P)(Px) = Px - P^2x = 0$. So, $P(x) \in \ker(I-P)$ for all $x \in H$ and therefore, $\operatorname{im}(P) \subseteq \ker(I-P)$.

Conversely, suppose that $y \in \ker(I-P)$. Then, $(I-P)y = 0$ and consequently, $P(y) = y$. Hence, $y \in \operatorname{im}(P)$ and as a result, $\ker(I-P) \subseteq \operatorname{im}(P)$. Therefore, $\ker(I-P) = \operatorname{im}(P)$. Since $\ker(I-P)$ is a closed subspace of H, $\operatorname{im}(P)$ must also be closed as required.

We showed in part (a) that $I - P$ satisfies the same properties as P. Hence, we can apply the result of part (a), but to the operator $I - P$. So, $\operatorname{im}(I-P) = \ker P = [\operatorname{im}(P)]^{\perp}$. Now assume that $x \in H$. Then,

$$x = P(x) + (x - P(x))$$

where $P(x) \in \operatorname{im}(P)$ and $x - P(x) \in \operatorname{im}(I-P) = [\operatorname{im}(P)]^{\perp}$. Thus, P is the projection operator onto the closed subspace $\operatorname{im}(P)$, which completes the proof. □

3.3 Riesz Representation of Linear Functionals

The Riesz representation theorem establishes an important result about Hilbert spaces and their corresponding dual spaces. It roughly states that there is a type of isomorphism between Hilbert spaces and their duals.

Theorem 3.3.1. (Riesz Representation Theorem): *Let H be a Hilbert space and H^* be the corresponding dual space. Define the map $\psi : H \to H^*$ by*

$$\psi(x) = \phi^x.$$

In turn, the map $\phi^x : H \to \mathbb{K}$ is defined by

$$\phi^x(y) = \langle y, x\rangle.$$

Then, for all $x \in H$, the map ϕ^x is a continuous linear functional.

Additionally, let the map $y \mapsto Ay$ be a continuous linear functional. Then, there exists a unique element $c \in H$ such that $Ay = \langle y, c\rangle$ for all $y \in H$.

Proof. Assume that H is a Hilbert space and $x, y, z \in H$. Assume that $a, b \in \mathbb{K}$. Assume that H^* is the dual space of H. Assume that the map $y \mapsto Ay$ is a continuous linear functional.

To show: (a) The functional ϕ^x is linear.

(b) The functional ϕ^x is bounded.

(c) There exists a unique element $c \in H$ such that $Ay = \langle y, a\rangle$ for all $y \in H$.

(a) This is a straightforward application of the linearity of the inner product.

$$\begin{aligned}\phi^x(ay + bz) &= \langle ay + bz, x\rangle \\ &= a\langle y, x\rangle + b\langle z, x\rangle \\ &= a\phi^x(y) + b\phi^x(z).\end{aligned}$$

Hence, ϕ^x is linear.

(b) We argue as follows

$$\begin{aligned}\|\phi^x\| &= \sup_{\|y\|=1} |\phi^x(y)| \\ &= \sup_{\|y\|=1} |\langle y, x\rangle| \\ &\leq \sup_{\|y\|=1} \|x\|\|y\| \quad \text{(Cauchy-Schwarz Inequality)} \\ &= \|x\| \\ &< \infty.\end{aligned}$$

Hence, ϕ^x is bounded. By 1.3.2, it is continuous. Consequently, ϕ^x is a linear, continuous functional.

(c) Observe that if $A = 0$, then we can set $c = 0$ and the assertion follows. So, assume $A \neq 0$. Then, consider the closed subspace $ker(A) \neq H$ and its orthogonal complement

$$(ker(A))^{\perp} = \{h \in H \text{ such that } \langle h, x\rangle = 0 \text{ for all } x \in ker(A)\}.$$

Assume that $v_1, v_2 \in (ker(A))^{\perp}$.

To show: (ca) There exists a $\lambda \in \mathbb{K}$ such that $\lambda v_1 = v_2$.

(ca) Consider the scalars $Av_1, Av_2 \in \mathbb{K}$. Then, there exists a non-zero $\lambda \in \mathbb{K}$ such that $\lambda A v_1 = Av_2$. Using the linearity of the functional A, we obtain $A(\lambda v_1 - v_2) = 0$. So, $\lambda v_1 - v_2 \in ker(A)$. However, we also know that $\lambda v_1 - v_2 \in (ker(A))^{\perp}$. Therefore, $\lambda v_1 - v_2 \in ker(A) \cap (ker(A))^{\perp}$.

To show: (caa) If V and $V^{\perp}$ are subspaces of H, then $V \cap V^{\perp} = \{0\}$.

(caa) Assume that $x \in V$ and $x \in V^{\perp}$. Then, from the definition of $V^{\perp}$, we have

$$\langle x, x\rangle = 0$$

From positive definiteness, $x = 0$. Hence, $V \cap V^{\perp} = \{0\}$.

(ca) Since, $ker(A) \cap (ker(A))^{\perp} = \{0\}$, $\lambda v_1 - v_2 = 0$, which proves the assertion.

Part (ca) establishes that $(ker(A))^{\perp}$ is a one-dimensional closed subspace of H. Now we choose any non-zero $h \in (ker(A))^{\perp}$. Define $c = \overline{\kappa} h$, where

$$\kappa = \frac{Ah}{\langle h, h\rangle}.$$

Then, we have

$$Ah = \kappa\langle h, h\rangle = \langle h, \overline{\kappa} h\rangle = \langle h, c\rangle \neq 0.$$

We know from the previous theorem that $H = ker(A) \oplus (ker(A))^{\perp}$. Using this, we write y as

$$y = P_{ker(A)}(y) + \alpha h \quad \text{where } \alpha \in \mathbb{K}.$$

Finally, we consider Ay.

$$\begin{aligned}
Ay &= A(P_{ker(A)}(y) + \alpha h) \\
&= A(P_{ker(A)}(y)) + A(\alpha h) \\
&= A(P_{ker(A)}(y)) + \alpha Ah \\
&= A(P_{ker(A)}(y)) + \alpha \langle h, c \rangle \\
&= 0 + \alpha \langle h, c \rangle \\
&= \langle P_{ker(A)}(y), c \rangle + \alpha \langle h, c \rangle \quad \text{since } c \in (ker(A))^{\perp} \\
&= \langle y, c \rangle.
\end{aligned}$$

Therefore, there exists a unique element $c \in H$ such that $Ay = \langle y, a \rangle$ for all $y \in H$. □

What else can we say about the map ψ? During the proof of the Riesz representation theorem, we showed that $\|\psi(x)\| = \|\phi^x\| = \|x\|$. In other words, the map ψ is an isometry (distance preserving). As a consequence of this, we also know that $\|\psi\| = 1$ because

$$\|\psi\| = \sup_{\|x\|=1} \|\psi(x)\| = \sup_{\|x\|=1} \|\phi^x\| = \sup_{\|x\|=1} \|x\| = 1.$$

Theorem 3.3.2. *The map ψ is a bijection.*

Proof. To show: (a) ψ is injective.

(b) ψ is surjective.

(a) Assume that $\psi(x) = \psi(y)$. Assume that $z \neq 0$. Then, $\phi^x(z) = \phi^y(z)$. So, $\langle z, x \rangle = \langle z, y \rangle$. Using linearity, we obtain

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

Since z was assumed to not be zero, $x - y = 0$. Hence, ψ is injective.

(b) Assume that $\beta \in H^*$.

To show: (ba) There exists a $x \in H$ such that $\psi(x) = \beta$.

(ba) From the Riesz representation theorem, we know that there exists a $c \in H$ such that

$$\beta y = \langle y, c \rangle.$$

However, $\langle y, c \rangle = \phi^c(y) = \psi(c)(y)$. Hence, $\psi(c)(y) = \beta y$. Subsequently, $\beta = \psi(c)$. So, there exists a $x \in H$ such that $\psi(x) = \beta$.

(b) Therefore, ψ is surjective.

Finally, since ψ is both injective and surjective, it must be bijective. □

Definition 3.3.1. Let X and Y be normed vector spaces over a field $\mathbb{K}$. Let $\Lambda : X \to Y$ be an operator. We say that Λ is **skew-linear** if for all $x_1, x_2 \in X$ and $a, b \in \mathbb{K}$,

$$\Lambda(ax_1 + bx_2) = \overline{a}\Lambda(x_1) + \overline{b}\Lambda(x_2).$$

Theorem 3.3.3. *The map ψ is skew-linear.*

Proof. Assume that $a, b \in \mathbb{K}$ and $x, y, z \in H$.

To show: (a) $\psi(ax + by) = \overline{a}\psi(x) + \overline{b}\psi(y)$.

(a) Using the definition of ψ, we proceed as follows

$$\begin{aligned} \psi(ax + by)(z) &= \phi^{ax+by}(z) \\ &= \langle z, ax + by \rangle \\ &= \overline{a}\langle z, x \rangle + \overline{b}\langle z, y \rangle \\ &= \overline{a}\phi^x(z) + \overline{b}\phi^y(z) \\ &= \overline{a}\psi(x)(z) + \overline{b}\psi(y)(z). \\ &= (\overline{a}\psi(x) + \overline{b}\psi(y))(z) \end{aligned}$$

Hence, $\psi(ax + by) = \overline{a}\psi(x) + \overline{b}\psi(y)$. So, ψ is skew-linear.

□

Due to the skew-linearity of ψ, it links H and H^* in two different ways, depending on whether $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. First, we will consider the case when $\mathbb{K} = \mathbb{R}$.

When $\mathbb{K} = \mathbb{R}$, ψ is a bijective, linear map between H and H^* with $\|\psi(x)\| = \|x\|$. In this case, ψ defines an isometric isomorphism between the spaces H and H^*.

When $\mathbb{K} = \mathbb{C}$, ψ is a bijective, *skew-linear* map between H and H^* with $\|\psi(x)\| = \|x\|$. In this case, ψ is an isometry, but it is not an isomorphism. Instead, ψ is an anti-isomorphism, an isomorphism which also doubles as a skew-linear map, rather than a linear map.

Finally, we will end this section with one of the most important, if not the most important application of the Riesz representation theorem.

Example 3.3.1. Consider the vector space of continuous, complex-valued (and square-integrable) wavefunctions. We endow this space with the inner product defined below:

$$\langle \psi_a, \psi_b \rangle = \int (\psi_a(\vec{r}))^* \cdot \psi_b(\vec{r}) \, \mathrm{d}^3 r.$$

With this inner product, our vector space of wavefunctions becomes a Hilbert space. As we will see in the next section, there are many possible orthonormal bases for the vector space of wavefunctions. This motivates the need for a **ket**, a basis independent vector of the Hilbert space. With the wavefunction ψ_a, we write this as $|a\rangle$. A ket is usually thought of as a **state vector**, in a Hilbert space of state vectors with the same aforementioned inner product.

In turn, a **bra** is a linear functional f_a from the Hilbert space of state vectors to the complex numbers, defined by

$$f_a(|b\rangle) = \langle a|b\rangle = \langle \psi_a, \psi_b \rangle.$$

It is customary to write f_a as $\langle a|$. The Riesz representation theorem tells us that the map $\omega : \{\text{state vectors}\} \to \{\text{dual space of state vectors}\}$, which is defined by

$$\omega(|a\rangle) = \langle a|$$

is an isometric anti-isomorphism between the two spaces. This means that for a ket, there is always a bra and vice versa.

Finally, linear operators $\hat{A}$ map state vectors to other state vectors. This is written as $\hat{A}|a\rangle$.

Bra-ket notation (note the pun) was introduced by Paul Dirac in 1939 and enjoys universal usage in quantum physics ever since.

3.4 Gram-Schmidt Orthogonalization

This procedure should be very familiar from the theory of linear algebra and finite dimensional vector spaces. Nonetheless, it is worth discussing in this context because it works for Hilbert spaces as well. We will start off with some familiar definitions.

Definition 3.4.1. Let H be a Hilbert space and let $x \in H$. We say that x is **normalised**, if $\|x\| = 1$.

Definition 3.4.2. Let H be a Hilbert space and let $S = \{s_1, s_2, \dots\} \subset H$ be a subset of H. We say that S is an **orthonormal set**, if for all $i, j \in \mathbb{Z}_{>0}$,

$$\langle s_i, s_j \rangle = \delta_{ij}$$

where δ_{ij} denotes the Kronecker delta.

Assume that $S = \{s_1, s_2, \dots, s_n\}$ is a finite, linearly independent set of vectors. Assume that $x \in span(S)$, so

$$x = \sum_{i=1}^{n} \alpha_i s_i.$$

The question here is: how do we determine the coefficients α_i? We first note that for a fixed $j \in \{1, 2, \dots, n\}$,

$$\langle x, s_j \rangle = \sum_{i=1}^{n} \alpha_i \langle s_i, s_j \rangle.$$

This can be rewritten as the matrix equation

$$\begin{pmatrix} \langle s_1, s_1 \rangle & \langle s_2, s_1 \rangle & \dots & \langle s_n, s_1 \rangle \\ \langle s_1, s_2 \rangle & \langle s_2, s_2 \rangle & \dots & \langle s_n, s_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle s_1, s_n \rangle & \langle s_2, s_n \rangle & \dots & \langle s_n, s_n \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \langle x, s_1 \rangle \\ \langle x, s_2 \rangle \\ \vdots \\ \langle x, s_n \rangle \end{pmatrix}.$$

Before we continue on with the calculation, we will make an important definition.

Definition 3.4.3. Let $V = \{v_1, \dots, v_n\}$ be a set of vectors. The **Gram matrix** of these vectors is the $n \times n$ matrix with entries given by $G_{ij} = \langle v_i, v_j \rangle$. The **Gram determinant**, denoted by $G(v_1, v_2, \dots, v_n)$, is the determinant of the Gram matrix.

There is one important theorem about the Gram determinant.

Theorem 3.4.1. *Let $V = \{v_1, \dots, v_n\}$ be a set of vectors where $v_i \neq 0$ for all $i \in \{1, \dots, n\}$. These vectors are linearly independent if and only if $G(v_1, v_2, \dots, v_n) \neq 0$.*

Proof. To show: (a) If $G(v_1, v_2, \dots, v_n) = 0$, then the vectors in V are linearly dependent.

(b) If the vectors in V are linearly dependent, then $G(v_1, v_2, \dots, v_n) = 0$.

(a) Assume that $G(v_1, v_2, \dots, v_n) = 0$. Using cofactor expansion along the first row, we find that there exists non-zero $\beta_i \in \mathbb{K}$ such that

$$\sum_{i=1}^{n} \beta_i \langle v_1, v_i \rangle = 0.$$

Using the properties of the inner product, we simplify this to

$$\langle v_1, \sum_{i=1}^{n} \overline{\beta_i} v_i \rangle = 0.$$

Since $v_1 \neq 0$, $\sum_{i=1}^{n} \overline{\beta_i} v_i = 0$. Hence, there exists a non-trivial linear combination of vectors in V which sum to zero. Therefore, the vectors in V are linearly dependent.

(b) Assume that the vectors in V are linearly dependent. Then, there exists a non-trivial linear combination such that

$$\sum_{i=1}^{n} \gamma_i v_i = 0.$$

Here, $\gamma_i \in \mathbb{K} \backslash \{0\}$. We then write $v_1 = \sum_{i=1}^{n-1} -\frac{\gamma_{i+1}}{\gamma_1} \cdot v_{i+1}$. Now consider the Gram determinant $G(v_1, v_2, \dots, v_n)$ below:

$$\begin{vmatrix} \sum_{i=1}^{n-1} -\frac{\gamma_{i+1}}{\gamma_1} \cdot \langle v_{i+1}, v_1 \rangle & \sum_{i=1}^{n-1} -\frac{\gamma_{i+1}}{\gamma_1} \cdot \langle v_{i+1}, v_2 \rangle & \dots & \sum_{i=1}^{n-1} -\frac{\gamma_{i+1}}{\gamma_1} \cdot \langle v_{i+1}, v_n \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \dots & \langle v_2, v_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \dots & \langle v_n, v_n \rangle \end{vmatrix}$$

Define the matrix A as follows:

$$A = \begin{pmatrix} 1 & \frac{\gamma_2}{\gamma_1} & \dots & \frac{\gamma_n}{\gamma_1} \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

We also define G to be the Gram matrix corresponding to the Gram determinant $G(v_1, v_2, \dots, v_n)$. Computing the product AG yields:

$$\begin{aligned} AG &= \begin{pmatrix} 1 & \frac{\gamma_2}{\gamma_1} & \dots & \frac{\gamma_n}{\gamma_1} \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \times G \\ &= \begin{pmatrix} 0 & 0 & \dots & 0 \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \dots & \langle v_2, v_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \dots & \langle v_n, v_n \rangle \end{pmatrix}. \end{aligned}$$

Since $\det(AG) = \det(A)\det(G) = \det(A)\, G(v_1, v_2, \dots, v_n)$, we have $0 = 1 \times G(v_1, v_2, \dots, v_n)$. Therefore, $G(v_1, v_2, \dots, v_n) = 0$. This proves the assertion. □

Returning to the calculation we are working on, we observe that $(G(s_1, s_2, \dots, s_n))^T \neq 0$ because all the vectors in S are orthogonal to each other, since they are linearly independent. In fact, the corresponding Gram matrix is diagonal due to this. This results in the matrix equation

$$\begin{pmatrix} \langle s_1, s_1 \rangle & 0 & \dots & 0 \\ 0 & \langle s_2, s_2 \rangle & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \langle s_n, s_n \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \langle x, s_1 \rangle \\ \langle x, s_2 \rangle \\ \vdots \\ \langle x, s_n \rangle \end{pmatrix}.$$

So, we can solve this matrix equation in order to obtain

$$\alpha_i = \frac{\langle x, s_i \rangle}{\langle s_i, s_i \rangle}.$$

Note that if S is an orthonormal set, then $\alpha_i = \langle x, s_i \rangle$. Using this, we will now describe Gram-Schmidt orthogonalization, which takes any linearly independent set of vectors $\{v_1, \dots, v_n\}$ and provides an *orthonormal* set of vectors $\{e_1, \dots, e_n\}$ such that $span\{v_1, \dots, v_n\} = span\{e_1, \dots, e_n\}$ for all $n \geq 1$.

Example 3.4.1. (Gram-Schmidt Orthogonalization)

Step 1: Set $e_1 = \frac{v_1}{\|v_1\|}$.

Step 2: If we have already defined $e_1, e_2, \ldots, e_{n-1}$, we then define e_n to be

$$e_n = \frac{v_n - \sum_{i=1}^{n-1} \langle v_n, e_i \rangle e_i}{\|v_n - \sum_{i=1}^{n-1} \langle v_n, e_i \rangle e_i\|}$$

We definitely have $e_n \neq 0$ because $v_n \notin span\{e_1, \ldots, e_{n-1}\}$. We also have $\|e_n\| = 1$. We can also prove that e_n is perpendicular to all of $e_1, e_2, \ldots, e_{n-1}$ via induction.

3.5 Orthonormal sets in Hilbert spaces

In this section, we shall deal with results involving orthonormal sets in infinite-dimensional Hilbert spaces.

Theorem 3.5.1. *Let H be a Hilbert space and let S be a subset of H. Then, the set S is total if and only if $S^\perp = \{0\}$.*

Proof. Assume that H is a Hilbert space and $S \subset H$.

To show: (a) If S is total, then $S^\perp = \{0\}$.

(b) If $S^\perp = \{0\}$, then S is total.

(a) Assume that $S = \{s_1, s_2, \ldots\}$ is total. Then, $\overline{span(S)} = H$. For all $h \in H$, there exists a sequence $\{x_n\}$ with $x_n \in span(S)$ such that

$$\lim_{n \to \infty} \|x_n - h\| = 0.$$

Assume $x \in S^\perp$. Since $S^\perp$ is a closed subspace of H, there exists a sequence $\{y_n\}$ with $y_n \in span(S)$ such that $y_n \to x$ as $n \to \infty$. Since $y_n \in span(S)$, we write $y_n = \sum_{i=1}^{\infty} \alpha_i^{(n)} s_i$. Now consider $\langle x, x \rangle$.

$$\begin{aligned}\langle x, x\rangle &= \lim_{n\to\infty}\langle y_n, x\rangle \\ &= \lim_{n\to\infty}\langle\sum_{i=1}^{\infty}\alpha_i^{(n)}s_i, x\rangle \\ &= \lim_{n\to\infty}\sum_{i=1}^{\infty}\alpha_i^{(n)}\langle s_i, x\rangle \\ &= 0 \quad \text{since } x \in S^{\perp}.\end{aligned}$$

Therefore, $x = 0$ and so, $S^{\perp} = \{0\}$.

(b) Assume that $S^{\perp} = \{0\}$. Suppose for the sake of contradiction that $\overline{span(S)} \neq H$. Then, there exists a vector $v \in H$ such that $v \notin \overline{span(S)}$. Consider the vector v', defined by the relation

$$v' = v - P_{span(S)}(v).$$

From Theorem 3.2.2, $\langle v, v'\rangle = 0$ with $v' \in (span(S))^{\perp}$ because $v' = P_{(span(S))^{\perp}}(v)$. So, for all $s \in S$, $\langle v', s\rangle = 0$ because $S \subset span(S)$. Hence, v' is a non-zero vector in $S^{\perp}$, contradicting the assumption that $S^{\perp} = \{0\}$.

Therefore, $\overline{span(S)} = H$ and consequently, S is total. □

Example 3.5.1. In this example, let H be the Hilbert space of real numbers, equipped with the standard dot product (which is basically multiplication)

$$\langle x, y\rangle = xy.$$

Let $S = \mathbb{Q}$. We note that $\mathbb{Q}^{\perp} = \{0\}$ because 0 is the unique element in the ring $\mathbb{Q}$ such that $0x = 0$ for all $x \in \mathbb{Q}$. Therefore, from the previous theorem, $\mathbb{Q}$ must be a total set in $\mathbb{R}$. In other words, $\overline{span(\mathbb{Q})} = \mathbb{R}$, which makes sense because $\overline{\mathbb{Q}} = \mathbb{R}$ ($\mathbb{Q}$ is a dense set in $\mathbb{R}$).

Theorem 3.5.2. *Assume $S = \{e_1, e_2, \dots\}$ is an orthonormal subset of the Hilbert space H. Let V be the closed subspace generated by S. Define $P_V : H \to V$, the perpendicular projection onto V, by*

$$P_V(x) = \sum_{i=1}^{\infty}\langle x, e_i\rangle e_i.$$

Then, for all $x \in H$, Bessel's inequality holds

$$\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 = \|P_V(x)\|^2 \leq \|x\|^2.$$

Proof. Assume that H is a Hilbert space, $x \in H$ and that $S = \{e_1, e_2, \dots\}$ is an orthonormal subset of H. Assume V is the closed subspace generated by S. Define $V_n = span\{e_1, \dots, e_n\}$ for all $n \in \mathbb{Z}_{>0}$. Then, we consider the perpendicular projection of x onto V_n, given by

$$P_{V_n}(x) = \sum_{i=1}^{n} \langle x, e_i \rangle e_i.$$

Computing $\|P_{V_n}(x)\|^2$ with the definition of a norm, we obtain

$$\begin{aligned}
\|P_{V_n}(x)\|^2 &= \langle P_{V_n}(x), P_{V_n}(x) \rangle \\
&= \langle \sum_{i=1}^{n} \langle x, e_i \rangle e_i, \sum_{i=1}^{n} \langle x, e_i \rangle e_i \rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \langle x, e_i \rangle \overline{\langle x, e_j \rangle} \langle e_i, e_j \rangle. \\
&= \sum_{i=1}^{n} |\langle x, e_i \rangle|^2 \|e_i\|^2 \\
&= \sum_{i=1}^{n} |\langle x, e_i \rangle|^2.
\end{aligned}$$

Define the sequence $\{a_n\}$ by

$$a_n = \sum_{i=1}^{n} \langle x, e_i \rangle e_i.$$

To show: (a) The sequence $\{a_n\}$ is Cauchy.

(a) Assume $m < n$. Consider the expression $\|a_m - a_n\|^2$.

$$
\begin{aligned}
\|a_n - a_m\|^2 &= \|\sum_{k=1}^{n}\langle x, e_k\rangle e_k - \sum_{k=1}^{m}\langle x, e_k\rangle e_k\|^2 \\
&= \|\sum_{k=m+1}^{n}\langle x, e_k\rangle e_k\|^2 \\
&= \langle \sum_{k=m+1}^{n}\langle x, e_k\rangle e_k, \sum_{k=m+1}^{n}\langle x, e_k\rangle e_k\rangle \\
&= \sum_{k=m+1}^{n}|\langle x, e_k\rangle|^2 \quad (\{e_i\}_{i=1}^{n} \text{ is orthonormal}) \\
&\to 0 \quad \text{as } m, n \to \infty.
\end{aligned}
$$

Therefore, the sequence $\{a_n\}$ is Cauchy. Since H is complete, $\{a_n\}$ must be a convergent sequence. Assume that $\{a_n\}$ converges to (say) a. Then,

$$
a = \lim_{n\to\infty} a_n = \lim_{n\to\infty}\sum_{i=1}^{n}\langle x, e_i\rangle e_i = \sum_{i=1}^{\infty}\langle x, e_i\rangle e_i.
$$

To show: (b) $P_V(x) = a = \sum_{i=1}^{\infty}\langle x, e_i\rangle e_i$.

(b) First, we observe that for all $n \geq 1$, $V_n \subset V$. So, $a_n \in V$. Since $a_n \to a$, $a \in \overline{V}$. For a fixed $k \geq 1$,

$$
\begin{aligned}
\langle x - a, e_k\rangle &= \lim_{n\to\infty}\langle x - a_n, e_k\rangle \\
&= \lim_{n\to\infty}\langle x - \sum_{i=1}^{n}\langle x, e_i\rangle e_i, e_k\rangle \\
&= \lim_{n\to\infty}\langle x, e_k\rangle - \langle\sum_{i=1}^{n}\langle x, e_i\rangle e_i, e_k\rangle \\
&= \lim_{n\to\infty}\langle x, e_k\rangle - \sum_{i=1}^{n}\langle x, e_i\rangle\langle e_i, e_k\rangle. \\
&= \lim_{n\to\infty}\langle x, e_k\rangle - \langle x, e_k\rangle \\
&= 0.
\end{aligned}
$$

So, the vector $x - a$ is perpendicular to every vector in S. Consequently, $x - a$ is perpendicular to every linear combination of vectors in S. So, $x - a \in V^{\perp}$. Defining $P_{V^{\perp}}(x) = x - a$, we then obtain $P_V(x) = a$ by 3.2.2.

Finally, the inequality $\|P_V(x)\|^2 \leq \|x\|^2$ follows from a swift application of Pythagoras' theorem. Namely,

$$\|x\|^2 = \|P_V(x)\|^2 + \|P_{V^\perp}(x)\|^2 \geq \|P_V(x)\|^2.$$

So, $\|P_V(x)\| \leq \|x\|$, which gives us the required inequality because

$$\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 = \|a\|^2 = \|P_V(x)\|^2 \leq \|x\|^2.$$

□

For the special case where $\overline{span(S)} = H$, $P_V(x) = x$ and

$$x = \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i.$$

This holds for all $x \in H$.

3.6 Positive Definite Operators

Many of the foundational concepts of functional analysis stems from linear algebra and the analysis of finite dimensional vector spaces. For now, we will foray back into the world of linear algebra in order to discuss positive definite matrices and their significance in one of the fundamental questions of linear algebra - when does a linear system $Ax = b$ have a unique solution?

Definition 3.6.1. Let A be a $n \times n$ matrix with elements in $\mathbb{R}$ and let x be a $n \times 1$ column matrix with elements in $\mathbb{R}$. We say that A is **positive definite** if for all $x \neq 0$,

$$x^T Ax > 0.$$

Recall from linear algebra that if $rank(A) = n$, then $null(A) = 0$ by the rank-nullity theorem, indicating that the linear system $Ax = b$ has a unique solution. The next theorem establishes an important relation between positive definiteness and the existence of a unique solution.

Theorem 3.6.1. *Consider the linear system $Ax = b$, where A is a $n \times n$ matrix and x and b are $n \times 1$ column matrices, all with entries in $\mathbb{R}$. If A is positive definite, then $rank(A) = n$.*

Proof. Assume that $Ax = b$, where A is a $n \times n$ matrix and x and b are $n \times 1$ column matrices, all with entries in $\mathbb{R}$. Assume that A is positive definite. Suppose for the sake of contradiction that $rank(A) < n$. Then, there exists a column in A which is a linear combination of all other columns in A.

To show: (a) There exists a column matrix y such that $y^T Ay = 0$.

(a) Suppose that A is the matrix given by

$$A = \begin{pmatrix} a_1^{(1)} & a_1^{(2)} & \dots & a_1^{(j-1)} & \sum_{i=1,i\neq j}^{n} k_i a_1^{(i)} & a_1^{(j+1)} & \dots & a_1^{(n)} \\ a_2^{(1)} & a_2^{(2)} & \dots & a_2^{(j-1)} & \sum_{i=1,i\neq j}^{n} k_i a_2^{(i)} & a_2^{(j+1)} & \dots & a_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n^{(1)} & a_n^{(2)} & \dots & a_n^{(j-1)} & \sum_{i=1,i\neq j}^{n} k_i a_n^{(i)} & a_n^{(j+1)} & \dots & a_n^{(n)} \end{pmatrix}.$$

where $k_i \in \mathbb{R}$ and not all $k_i = 0$. Define the column matrix y by

$$y = \begin{pmatrix} -k_1 \\ -k_2 \\ \vdots \\ -k_{j-1} \\ 1 \\ -k_{j+1} \\ \vdots \\ -k_n \end{pmatrix}.$$

Then, we observe that $Ay = 0$. So, $y^T Ay = 0$. However, this contradicts the fact that A is a positive definite matrix. Therefore, $rank(A) \geq n$. Since $rank(A) \leq n$ by the definition of rank, we conclude that $rank(A) = n$. □

Hence, we have just shown that positive definite matrices are the ones that guarantee the existence of a unique solution. Analogously to this, we will now return to functional analysis in order to discuss the importance of *positive definite operators* on Hilbert spaces.

Definition 3.6.2. Let H be a Hilbert space over the field $\mathbb{R}$. Let $A : H \to H$ be a bounded, linear operator. A is said to be **positive definite** if there exists a $\beta > 0$ such that for all $h \in H$,

$$\langle Ah, h \rangle \geq \beta \|h\|^2.$$

Theorem 3.6.2. *Let H be a Hilbert space over the reals. Let $A : H \to H$ be a bounded, linear, positive definite operator. Then, for all $f \in H$, there exists a unique $u = A^{-1}f \in H$ such that $Au = f$ and*

$$\|A^{-1}\| \leq \frac{1}{\beta}.$$

Proof. Assume H is a Hilbert space over the reals. Assume A is a bounded, linear, positive definite operator.

To show: (a) A is injective.

(b) A is surjective.

(c) $\|A^{-1}\| \leq \frac{1}{\beta}$.

(a) Assume that $x \in H$ and $Ax = 0$. Assume that $\beta \in \mathbb{R}_{>0}$. From the positive definiteness of A and the Cauchy Schwarz inequality, we have

$$\beta\|x\|^2 \leq \langle Ax, x\rangle \leq \|Ax\|\|x\|.$$

This results in the inequality $\beta\|x\| \leq \|Ax\|$. Since $Ax = 0$, $\|Ax\| = 0$. Hence from the inequality, $\|x\| = 0$ and consequently $x = 0$. This demonstrates that $ker(A) = \{0\}$. So, A is injective.

(b) To show: (ba) $im(A)$ is closed.

(bb) $im(A) = H$.

(ba) Assume that $f \in \overline{im(A)}$. Then, there exists a sequence $\{f_n\}$ in $im(A)$ such that $f_n \to f$.

To show: (baa) $f \in im(A)$.

(baa) Since $f_n \in im(A)$ for all $n \in \mathbb{Z}_{>0}$, there exists $u_n \in H$ such that $Au_n = f_n$. Fix $N \in \mathbb{Z}_{>0}$. Then, for all $m, n > N$, we use the inequality $\beta\|x\| \leq \|Ax\|$ to obtain

$$\|u_n - u_m\| \leq \frac{1}{\beta}\|Au_n - Au_m\| = \frac{1}{\beta}\|f_n - f_m\|.$$

Since the sequence $\{f_n\}$ is convergent, it is Cauchy. Therefore, the sequence $\{u_n\}$ is Cauchy and thus, converges to (say) $u \in H$, by the completeness of

H. Observe that since A is a bounded operator, it is continuous. Utilising this, we find that $Au = f$. So, $f \in im(A)$.

(ba) Since $f \in im(A)$, $\overline{im(A)} \subset im(A)$. We already know that $im(A) \subset \overline{im(A)}$. Subsequently, $\overline{im(A)} = im(A)$ and so, $im(A)$ must be closed.

(bb) Suppose for the sake of contradiction that $im(A) \neq H$. Since $im(A)$ is a closed subspace of H, we can decompose H as the direct sum $H = im(A) \oplus (im(A))^{\perp}$. So, there exists a non-zero $y \in (im(A))^{\perp}$ such that

$$\langle Ay, y \rangle = 0.$$

Since $y \neq 0$, $Ay = 0$. However, this contradicts the injectivity of A. Therefore, $im(A) = H$ and consequently, A is surjective.

Furthermore, since A is both injective and surjective, it must be bijective. So, we now know that there exists an inverse operator $A^{-1} : H \to H$ such that $A \circ A^{-1} = 1_H$.

(c) From part (a), we have the inequality $\beta \|x\| \leq \|Ax\|$. With the inverse operator, we can rewrite this as

$$\beta \|A^{-1} f\| \leq \|f\|.$$

for some arbitrary $f \in H$. Using this, we find that

$$\begin{aligned} \|A^{-1}\| &= \sup_{\|f\|=1} \|A^{-1} f\| \\ &\leq \sup_{\|f\|=1} \frac{\|f\|}{\beta} \\ &= \frac{1}{\beta}. \end{aligned}$$

□

It turns out that this result can also be expressed with bilinear forms. This is an important theorem known as the *Lax-Milgram Theorem*. Of course, before proving the theorem, we will need to make important definitions involving bilinear forms.

Definition 3.6.3. Let H be a Hilbert space over the reals. A **bilinear functional** is a mapping $B : H \times H \to \mathbb{R}$ which satisfies the following properties

1. For all $a, b \in \mathbb{R}$ and $x, y, z \in H$, $B[ax + by, z] = aB[x, z] + bB[y, z]$.
2. For all $a, b \in \mathbb{R}$ and $x, y, z \in H$, $B[x, ay + bz] = aB[x, y] + bB[x, z]$.

Definition 3.6.4. Again, let H be a Hilbert space over the reals. Let $B : H \times H \to \mathbb{R}$ be a bilinear functional. B is said to be **continuous** if for all $u, v \in H$, there exists a constant $C \in \mathbb{R}$ such that

$$|B[u, v]| \leq C\|u\|\|v\|.$$

Definition 3.6.5. Let H be a Hilbert space over the reals. Let $B : H \times H \to \mathbb{R}$ be a bilinear functional. We say that B is **positive definite** if there exists a constant $\beta \in \mathbb{R}_{>0}$ such that for all $u \in H$,

$$B[u, u] \geq \beta\|u\|^2.$$

Theorem 3.6.3. (Lax-Milgram Theorem): *Let H be a Hilbert space over the reals and B be a continuous and positive definite bilinear functional. Then, for all $h \in H$, there exists a unique $x \in H$ such that for all $y \in H$*

$$B[x, y] = \langle h, y \rangle.$$

Moreover,

$$\|x\| \leq \beta^{-1}\|h\|.$$

Proof. Assume that H is a Hilbert space over the reals. Assume that B is a continuous and positive definite bilinear functional. Consider the map $y \mapsto B[x, y]$. From the definition of B, this is a continuous, linear functional. So, we can utilise the Riesz representation theorem to show that there exists a unique $Ax \in H$ such that for all $y \in H$

$$B[x, y] = \langle Ax, y \rangle.$$

To show: (a) The operator $A : H \to H$ is linear.

(b) The operator $A : H \to H$ is bounded.

(c) The operator $A : H \to H$ is positive definite.

(a) Assume $x, y \in H$ and $a, b \in \mathbb{R}$. We proceed as follows

$$\begin{aligned}\langle A(ax+by), z\rangle &= B[ax+by, z] \\ &= aB[x,z] + bB[y,z] \\ &= a\langle Ax, z\rangle + b\langle Ay, z\rangle \\ &= \langle aAx + bAy, z\rangle.\end{aligned}$$

Therefore, $A(ax+by) = aAx + bAy$. Hence, A is a linear operator.

(b) Using the continuity of B, we argue as follows

$$\begin{aligned}\|A\| &= \sup_{\|x\|=1} \|Ax\| \\ &= \sup_{\|x\|=1} \sup_{\|y\|=1} |\langle Ax, y\rangle| \quad \text{(Cauchy-Schwarz Inequality)} \\ &= \sup_{\|x\|=1} \sup_{\|y\|=1} |B[x,y]| \\ &\leq \sup_{\|x\|=1} \sup_{\|y\|=1} C\|x\|\|y\| \\ &= C.\end{aligned}$$

Hence, $\|A\| \leq C$. The result on the second line is found via the Cauchy-Schwarz inequality. We know that

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

Taking the supremum of both sides with $\|y\| = 1$ yields the result we are after.

This establishes the conclusion that A is bounded.

(c) Assume that $h \in H$. Then, we have

$$\langle Ah, h\rangle = B[h,h] \geq \beta\|h\|^2.$$

So, A is a positive definite operator.

Since A is a bounded, linear positive definite operator, we can apply the previous theorem to show that for all $f \in H$, there exists a unique $u = A^{-1}f \in H$ such that $Au = f$ and

$$\|A^{-1}\| \leq \frac{1}{\beta}$$

for some constant $\beta > 0$. Applying this to the fact that $B[x, y] = \langle Ax, y \rangle$, we find that for all $h \in H$, there exists a unique $x \in H$ such that for all $y \in H$, $B[x, y] = \langle h, y \rangle$. Finally, we observe that

$$\begin{aligned} \|x\| &= \|A^{-1}h\| \\ &\leq \frac{\|h\|}{\beta}. \end{aligned}$$

This follows from the inequality $\beta\|A^{-1}f\| \leq \|f\|$, which we proved in the previous theorem. □

The Lax-Milgram theorem is a powerful tool in the study of PDEs. It provides sufficient conditions required to invert a bilinear form and thus, deduce the existence and uniqueness of weak solutions to a boundary value problem. Interestingly, the Lax-Milgram theorem generalises even further to a result involving two Hilbert spaces rather than one. This was first proved by Babuška in [Bab71].

Definition 3.6.6. Let H_1 and H_2 be Hilbert spaces over $\mathbb{C}$. A **bilinear functional** is a mapping $B : H_1 \times H_2 \to \mathbb{C}$ which satisfies the following properties

1. For all $a, b \in \mathbb{C}$, $B[ax + by, z] = aB[x, z] + bB[y, z]$.
2. For all $a, b \in \mathbb{C}$, $B[x, ay + bz] = \overline{a}B[x, y] + \overline{b}B[x, z]$.

The above definition is from Kreyszig [EK78]. Note that the previous definition of a continuous bilinear functional still remains the same for complex-valued bilinear functionals.

Theorem 3.6.4 (Babuška Lax Milgram theorem)**.** *Let H_1 and H_2 be two Hilbert spaces over $\mathbb{C}$. Let $B : H_1 \times H_2 \to \mathbb{C}$ be a continuous bilinear form/functional, which is weakly coercive. That is, for all $u \in H_1$ and $v \in H_2$, there exists constants $C_2, C_3 \in \mathbb{R}_{>0}$ such that*

$$\sup_{\|u\|_{H_1}=1} |B(u, v)| \geq C_2\|v\|_{H_2} \text{ and } \sup_{\|v\|_{H_2}=1} |B(u, v)| \geq C_3\|u\|_{H_1}.$$

Furthermore, let $f : H_2 \to \mathbb{C}$ be a continuous linear functional on H_2. Then, there exists a unique element $u_1 \in H_1$ such that $B(u_1, v) = \overline{f(v)}$ for all $v \in H_2$. Moreover,

$$\|u_1\|_{H_1} \leq \frac{\|f\|}{C_3}.$$

Proof. Assume that H_1, H_2 are Hilbert spaces over $\mathbb{C}$ and that $B : H_1 \times H_2 \to \mathbb{C}$ is a continuous bilinear functional/form. Explicitly, continuous means that for all $u \in H_1$ and $v \in H_2$, there exists a constant $C_1 \in \mathbb{R}_{>0}$ such that

$$|B(u, v)| \leq C_1 \|u\|_{H_1} \|v\|_{H_2}.$$

Assume also that B is weakly coercive, as described in the statement of the theorem. We begin by defining for all $u \in H_1$,

$$\phi_u : H_2 \to \mathbb{C}$$

$$v \mapsto \overline{B(u, v)}$$

This is a linear functional on H_2 (remember that B is conjugate linear in the second input). It is also bounded because

$$\begin{aligned}
\|\phi_u\| &= \sup_{\|v\|_{H_2}=1} |\phi_u(v)| \\
&= \sup_{\|v\|_{H_2}=1} |\overline{B(u, v)}| \\
&= \sup_{\|v\|_{H_2}=1} |B(u, v)| \\
&\leq \sup_{\|v\|_{H_2}=1} C_1 \|u\|_{H_1} \|v\|_{H_2} \\
&= C_1 \|u\|_{H_1}.
\end{aligned}$$

Here, we have used the fact that B is continuous. Thus, by the Riesz representation theorem (see 3.3.1), there exists $z \in H_2$ such that for all $v \in H_2$,

$$B(u, v) = \langle z, v \rangle_{H_2}.$$

So, there exists a linear map $R : H_1 \to H_2$ such that $B(u, v) = \langle R(u), v \rangle_{H_2}$. Furthermore, we have the bound

$$\begin{aligned}\|R\| &= \sup_{\|u\|_{H_1}=1, \|v\|_{H_2}=1} |\langle R(u), v\rangle_{H_2}| \\ &= \sup_{\|u\|_{H_1}=1, \|v\|_{H_2}=1} |B(u,v)| \\ &\leq \sup_{\|u\|_{H_1}=1, \|v\|_{H_2}=1} C_1 \|u\|_{H_1} \|v\|_{H_2} \\ &= C_1.\end{aligned}$$

So, R is also a bounded and hence, a continuous linear operator.

Next, consider the image $R(H_1)$.

To show: (a) The image $R(H_1)$ is closed in H_2.

(a) We already have $R(H_1) \subseteq \overline{R(H_1)}$. To obtain the reverse inclusion, let $\{u_n\}$ be a sequence in $R(H_1)$ which converges to u. So, for all $n \in \mathbb{Z}_{>0}$, there exists $f_n \in H_1$ such that $Rf_n = u_n$.

To see that $u \in R(H_1)$, we note that for all $m, n > N$ for some $N \in \mathbb{Z}_{>0}$,

$$\begin{aligned}\|Rf_n - Rf_m\|_{H_2} &= \|R(f_n - f_m)\|_{H_2} \\ &= \sup_{\|v\|_{H_2}=1} |\langle R(f_n - f_m), v\rangle| \\ &= \sup_{\|v\|_{H_2}=1} |B(f_n - f_m, v)| \\ &\geq C_3 \|f_n - f_m\|_{H_1}.\end{aligned}$$

Thus, the sequence $\{f_n\}$ is a Cauchy sequence in H_1 since $\{Rf_n\}$ is a Cauchy sequence in $R(H_1)$. Since H_1 is complete, $\{f_n\}$ must converge to some element $f \in H_1$. To see that $Rf = u$, we argue with the continuity of R that

$$\begin{aligned}Rf &= R(\lim_{n\to\infty} f_n) \\ &= \lim_{n\to\infty} Rf_n \\ &= \lim_{n\to\infty} u_n = u.\end{aligned}$$

So, $u \in R(H_1)$ and thus, $R(H_1)$ is a closed subspace of H_2.

Now, we will argue that $R(H_1) = H_2$. Suppose that $R(H_1) \neq H_2$. Then, we can decompose H_2 as

$$H_2 = R(H_1) \oplus (R(H_1))^{\perp}.$$

Hence, there exists a non-zero $v_0 \in H_2$ such that for all $u \in H_1$,

$$0 = \langle R(u), v_0 \rangle_{H_2} = B(u, v_0).$$

However, recall that B is weakly coercive. In particular,

$$\sup_{\|u\|_{H_1}=1} |B(u, v)| \geq C_2 \|v\|_{H_2}.$$

So, there exists $u' \in H_1$ such that $|B(u', v_0)| \geq \frac{1}{2} C_2 \|v_0\|_{H_2}$. This contradicts the finding that $B(u, v_0) = 0$ for all $u \in H_1$.

Hence, $R(H_1) = H_2$, which demonstrates that R is surjective. To see that R is injective, assume that $Rx = 0$ for some $x \in H_1$. Then,

$$\begin{aligned} 0 &= \sup_{\|y\|_{H_2}=1} |\langle Rx, y \rangle_{H_2}| \\ &= \sup_{\|y\|_{H_2}=1} |B(x, y)| \\ &\geq C_3 \|x\|_{H_1}. \end{aligned}$$

Therefore, $x = 0$ and so, R is injective.

Thus, we find that R is bijective. Therefore, $R^{-1} : H_2 \to H_1$ is a well-defined linear operator. It is continuous because

$$\|R^{-1}\| \leq \frac{1}{C_3}.$$

Finally, let $f : H_2 \to \mathbb{C}$ be a linear functional on H_2. The Riesz representation theorem tells us that there exists $v_1 \in H_2$ such that

$$\overline{f(v)} = \langle v_1, v \rangle_{H_2}$$

with $\|v_1\|_{H_2} = \|f\|$. Using the fact that R is bijective, set $u_1 = R^{-1} v_1$. Then,

$$\overline{f(v)} = \langle R u_1, v \rangle_{H_2} = B(u_1, v)$$

for all $v \in H_2$. Note that $u_1 \in H_1$ is unique because R^{-1} is bijective. $\square$

Chapter 4

More on Linear Operators

4.1 Open Mapping Theorem

In this section of the notes, we will provide a more in-depth analysis on bounded linear operators on normed vector spaces. The first two results in this chapter rely on an important theorem pertaining to completeness - the *Baire category theorem.*

Theorem 4.1.1 (Baire Category Theorem)**.** *Let (X, d) be a complete metric space. Let $\{V_k\}$ be a collection of open, dense subsets of X, with $k \in \mathbb{Z}_{>0}$. Then, the intersection $V = \bigcap_{k=1}^{\infty} V_k$ is a non-empty, dense subset of X.*

Proof. Assume that (X, d) is a complete metric space. Assume that $\{V_k\}$ is a collection of open, dense subsets of X.

To show: (a) For every open ball $B(x, r)$, there exists a point y such that $y \in B(x, r) \cap \bigcap_{k=1}^{\infty} V_k$.

(a) We know that V_1 is an open and dense subset of X. Consider the set $B(x, r) \cap V_1$.

To show: (aa) If U and V are both open subsets of X with non-empty intersection, then $U \cap V$ is also open.

(ab) $B(x, r) \cap V_1$ is non-empty.

(aa) Assume that U and V are open subsets of X with non-empty intersection. Assume that $a \in U \cap V$. Since U is open, there exists a constant $r_1 \in \mathbb{R}_{>0}$ such that $B(a, r_1) \subset U$. Since V is open, there exists a constant $r_2 \in \mathbb{R}_{>0}$ such that $B(a, r_2) \subset V$. Define $r = \min\{r_1, r_2\}$. Then, $B(a, r) \subseteq B(a, r_1) \subset U$ and $B(a, r) \subseteq B(a, r_2) \subset V$. So, $B(a, r) \in U \cap V$. Hence, $U \cap V$ is open.

(ab) Suppose for the sake of contradiction that $B(x, r) \cap V_1 = \emptyset$. Since V_1 is dense, $\overline{V_1} = X$. So, for all points $p \in X$, there exists a sequence $\{v_i\}$ with $v_i \in V_1$ for all $i \in \{1, 2, \dots\}$ such that $v_i \to p$ as $i \to \infty$.

In particular, there exists a sequence $\{\zeta_i\}$ with $\zeta_i \in V_1$ for all $i \in \{1, 2, \dots\}$ such that $\zeta_i \to x$ as $i \to \infty$. Choose $N \in \mathbb{Z}_{>0}$ such that for all $i > N$,

$$d(x, \zeta_i) < r.$$

However, this means that for all $i > N$, $\zeta_i \in B(x, r)$. So, $\zeta_i \in B(x, r) \cap V_1$, which is a blatant contradiction of the assumption that $B(x, r) \cap V_1 = \emptyset$. So, $B(x, r) \cap V_1$ is non-empty.

(a) Combining parts (aa) and (ab), we conclude that the set $B(x, r) \cap V_1$ is open and non-empty. So, there exists a closed ball $\overline{B}(x_1, r_1)$ with $r_1 < 1$ such that $\overline{B}(x_1, r_1) \subset B(x, r) \cap V_1$. Since V_2 is also an open, dense subset of X, we can repeat this process to conclude that there exists a closed ball $\overline{B}(x_2, r_2)$ with $r_2 < \frac{1}{2}$ such that $\overline{B}(x_2, r_2) \subset \overline{B}(x_1, r_1) \cap V_2$.

Continuing this construction, we obtain a sequence of balls $\overline{B}(x_n, r_n)$ such that $\overline{B}(x_{n+1}, r_{n+1}) \subset \overline{B}(x_n, r_n) \cap V_{n+1}$ and $r_n < \frac{1}{n}$.

To show: (ac) The sequence $\{x_n\}$ is Cauchy.

(ac) From our construction, we note that for all $m \geq k$, $\overline{B}(x_m, r_m) \subset \overline{B}(x_k, r_k)$. Choose $\epsilon > 0$ such that $1/k < \epsilon/2$. So, for all $m, n \geq k$,

$$d(x_m, x_n) \leq d(x_m, x_k) + d(x_n, x_k) \leq r_k + r_k < \frac{2}{k} < \epsilon.$$

Therefore, the sequence $\{x_n\}$ is Cauchy.

(a) Since $\{x_n\}$ is a Cauchy sequence and X is complete, $\{x_n\}$ must converge to (say) y. Taking the limit as $n \to \infty$, we find that $y \in \overline{B}(x_k, r_k)$ for all $k \in \mathbb{Z}_{>0}$. Since $\overline{B}(x_1, r_1) \subset B(x, r) \cap V_1 \subset B(x, r)$, $y \in B(x, r)$. Moreover, since $\overline{B}(x_n, r_n) \subset \overline{B}(x_{n-1}, r_{n-1}) \cap V_n \subset V_n$, $y \in V_n$ for all

$n \in \mathbb{Z}_{>0}$. So, $y \in \bigcap_{n=1}^{\infty} V_n$.

Hence, $y \in B(x,r) \cap \bigcap_{n=1}^{\infty} V_n$. Consequently, $\bigcap_{n=1}^{\infty} V_n$ is non-empty. It is also dense because for all $x \in X$ and $r \in \mathbb{R}_{>0}$, $B(x,r) \cap \bigcap_{n=1}^{\infty} V_n \neq \emptyset$. $\square$

Here is one consequence of the Baire category theorem:

Theorem 4.1.2. *Let (X, d) be a complete metric space. Let $\{F_n\}$ be a sequence of nowhere dense sets. That is, $(\overline{F_n})^\circ = \emptyset$ for all $n \in \{1, 2, \dots\}$. Then, $\bigcup_{n=1}^{\infty} F_n$ has empty interior.*

Proof. Assume that (X, d) is a complete metric space and that $\{F_n\}$ is a sequence of nowhere dense sets. Suppose for the sake of contradiction that the set $\bigcup_{n=1}^{\infty} F_n$ has non-empty interior. Then, for some $x \in \bigcup_{n=1}^{\infty} F_n$, there exists a $r \in \mathbb{R}_{>0}$ such that $B(x,r) \subset \bigcup_{n=1}^{\infty} F_n$.

Define $U_n = X \backslash \overline{F_n}$.

To show: (a) U_n is open for all $n \in \{1, 2, \dots\}$.

(b) U_n is dense for all $n \in \{1, 2, \dots\}$.

(a) Suppose for the sake of contradiction that U_n is not open. Then, there exists a point $u \in U_n$ such that for all $r \in \mathbb{R}_{>0}$, $B(u,r) \cap X \backslash U_n \neq \emptyset$. From the construction of U_n, $B(u,r) \cap \overline{F_n} \neq \emptyset$ for all $r \in \mathbb{R}_{>0}$. Hence, $u \in \overline{F_n}$; u is an adherent point of F_n. However, we assumed that $u \in U_n = X \backslash \overline{F_n}$. From this assumption, we conclude that $u \notin \overline{F_n}$, which is a contradiction.

Therefore, U_n is open for all $n \in \{1, 2, \dots\}$.

(b) Suppose for the sake of contradiction that $\overline{U_n} \neq X$ for all $n \in \{1, 2, \dots\}$. Then, for some point $x \in X$, there exists a $s \in \mathbb{R}_{>0}$ such that $B(x,s) \cap U_n = \emptyset$. Then, it must be the case that $B(x,s) \subset X \backslash U_n = \overline{F_n}$. So, x is an interior point of $\overline{F_n}$. However, this contradicts the fact that the collection of sets $\{F_n\}$ are nowhere dense ($(\overline{F_n})^\circ = \emptyset$ for all $n \in \{1, 2, \dots\}$).

Consequently, U_n is dense for all $n \in \{1, 2, \dots\}$.

Since we now have a sequence $\{U_n\}$ of non-empty dense subsets of X, we can apply the Baire category theorem to deduce that the set $\bigcap_{n=1}^{\infty} U_n$ is non-empty and dense in X. In particular, since $\bigcap_{n=1}^{\infty} U_n$ is dense in X, for all $x \in X$ and for all $r \in \mathbb{R}_{>0}$, $B(x,r) \cap \bigcap_{n=1}^{\infty} U_n \neq \emptyset$. However from our assumption, $B(x,r) \subset \bigcup_{n=1}^{\infty} F_n \subset \bigcup_{n=1}^{\infty} \overline{F_n}$. So,

$$\begin{aligned}\emptyset &= B(x,r) \cap X \backslash \bigcup_{n=1}^{\infty} \overline{F_n} \\ &= B(x,r) \cap \bigcap_{n=1}^{\infty} X \backslash \overline{F_n} \\ &= B(x,r) \cap \bigcap_{n=1}^{\infty} U_n.\end{aligned}$$

But, this contradicts the fact that U_n is dense in X. Therefore, $\bigcup_{n=1}^{\infty} F_n$ must have empty interior. □

Roughly speaking, this theorem tells us that we cannot write X as the countable union of nowhere dense subsets of X, provided that the metric space (X, d) is complete.

Now we are ready to tackle the uniform boundedness principle.

Theorem 4.1.3 (Uniform Boundedness Principle)**.** *Let X and Y be Banach spaces. Let $F \subset B(X;Y)$ be an arbitrary family of bounded, linear operators. Then, it is either the case that F is uniformly bounded*

$$\sup_{T \in F} \|T\| < \infty$$

or there exists a dense subset $S \subset X$ such that for all $x \in S$,

$$\sup_{T \in F} \|Tx\| = \infty.$$

Proof. Assume that X and Y are Banach spaces. Assume that F is a family of bounded, linear operators. Define the collection of sets S_n given by

$$S_n = \{x \in X \text{ such that } \|Tx\| > n \text{ for some } T \in F\}.$$

Suppose that for a fixed m, the set S_m is not dense in X. Then, there exists a $x \in X$ and a constant $r \in \mathbb{R}_{>0}$ such that $B(x,r) \cap S_m = \emptyset$. Moreover, for all $s \in B(x,r)$ and for all $T \in F$,

$$\|Ts\| \leq m.$$

If $v \in B(0, r)$ (that is, $\|t\| < r$) then an application of the triangle inequality tells us that for all $T \in F$,

$$\|Tv\| = \|T(v + x - x)\| \leq \|T(v + x)\| + \|Tx\| \leq m + m = 2m$$

since $v + x \in B(x, r)$ and $\|T(v + x)\| \leq m$.

Now suppose that $z \in X - \{0\}$. By scaling z, we find that

$$\frac{r}{2\|z\|} z \in B(0, r)$$

and if $T \in F$ then

$$\|T(\frac{r}{2\|z\|} z)\| \leq 2m.$$

Since T is linear, we can rearrange the above inequality to obtain

$$\|Tz\| \leq \frac{4m}{r} \|z\|$$

Since $z \in X - \{0\}$ was arbitrary, an upper bound for $\|T\|$ is

$$\|T\| = \sup_{\|z\|=1} \|Tz\| \leq \sup_{\|z\|=1} \frac{4m}{r} \|z\| = \frac{4m}{r}.$$

Hence, $\|T\| \leq \frac{4m}{r}$ for all $T \in F$. Therefore, the set of bounded linear operators F is uniformly bounded in this case.

Now, suppose that all of the sets S_n are dense. Observe that they are all open as well. So, we can apply the Baire category theorem (4.1.1) in order to conclude that the set $S = \bigcap_{n=1}^{\infty} S_n$ is also dense in X. From the construction of the set S, we find that for all $x \in S$, there exists an operator $T \in F$ such that $\|Tx\| > n$ for all $n \in \mathbb{R}_{>0}$. □

Here is an important consequence of the uniform boundedness principle.

Theorem 4.1.4. *Let X and Y be Banach spaces. Let $\{T_n\}$ be a sequence of bounded, linear operators. Define the operator T by the relation $\lim_{n\to\infty} T_n x = Tx$ which holds for all $x \in X$. Then, T itself is a bounded linear operator.*

Proof. Assume that X and Y are Banach sequences and that $\{T_n\}$ is a sequence of bounded, linear operators. Assume that T is an operator, which is defined as specified above.

To show: (a) The operator T is linear.

(b) The operator T is bounded.

(a) Assume that $a, b \in \mathbb{K}$ and that $x, y \in X$. Then,

$$\begin{aligned} T(ax+by) &= \lim_{n\to\infty} T_n(ax+by) \\ &= \lim_{n\to\infty} aT_nx + bT_ny \\ &= a\lim_{n\to\infty} T_nx + b\lim_{n\to\infty} T_ny \\ &= aTx + bTy. \end{aligned}$$

Therefore, the operator T is linear.

(b) Observe that for all $n \in \mathbb{R}_{>0}$, each operator T_n is bounded for all $x \in X$. In particular, this rules out the possibility of finding a dense subset $S \subset X$ such that for all $s \in S$

$$\sup_{n\in\mathbb{Z}_{>0}} \|T_ns\| = \infty.$$

So, from the previous theorem, the sequence $\{T_n\}$ must be uniformly bounded. Knowing this, we calculate $\|T\|$ to be

$$\begin{aligned} \|T\| &= \sup_{\|x\|=1} \|Tx\| \\ &= \sup_{\|x\|=1} \lim_{n\to\infty} \|T_nx\| \\ &= \lim_{n\to\infty} \|T_n\| \\ &< \infty. \end{aligned}$$

Therefore, the operator T is bounded.

Combining parts (a) and (b), we conclude that $T \in B(X;Y)$. □

Before we progress to the open mapping theorem, we will introduce some more general definitions of a continuous function

Definition 4.1.1. Let X and Y be metric spaces and $f : X \to Y$ be a function. We say that f is a **continuous function** if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x \in X$

$$f(B(x,\delta)) \subset B(f(x),\epsilon).$$

In other words, for any $\epsilon > 0$, we can choose a suitable $\delta > 0$ such that the image of the open ball $B(x, \delta)$ is contained within the open ball $B(f(x), \epsilon)$. This is very similar to the usual epsilon-delta definition of continuity.

Our next definition of continuity is even more general and is frequently discussed in the context of *topological spaces*. Luckily, metric spaces can be turned into topological spaces. Consequently, the definition of continuity remains relevant for metric spaces.

Definition 4.1.2. Let X and Y be metric spaces and $f : X \to Y$ be a function. Let $U \subset Y$ be a subset of Y. Then, the **preimage** of U, denoted by $f^{-1}(U)$ is the set given by

$$f^{-1}(U) = \{x \in X \text{ such that } f(x) \in U\}.$$

Definition 4.1.3. Let X and Y be metric spaces and $f : X \to Y$ be a function. Then, f is a continuous function if for all open subsets $U \subset Y$, the preimage $f^{-1}(U)$ is open in X.

We will now show that these two definitions of continuity are equivalent.

Theorem 4.1.5. *The two definitions 4.1.1 and 4.1.3 of a continuous function are equivalent.*

Proof. Assume that X and Y are metric spaces. Assume that $f : X \to Y$ is a function. Assume that $U \subset Y$ is an open subset of Y.

To show: (a) If f is continuous in the sense of 4.1.1, then it is continuous in the sense of 4.1.3.

(b) If f is continuous in the sense of 4.1.3, then it is continuous in the sense of 4.1.1.

(a) Assume that f is continuous in the "open ball" sense (4.1.1). Then, for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x \in X$, $f(B(x, \delta)) \subset B(f(x), \epsilon)$. Suppose that the point $u \in f^{-1}(U)$. Then, $f(u) \in U$ from the definition of preimage. Since U is open, there exists a $\epsilon \in \mathbb{R}_{>0}$ such that $B(f(u), \epsilon) \in U$. However, we also know that f is continuous in the "open ball" sense (4.1.1). So, there exists a $\delta > 0$ such that $f(B(u, \delta)) \subset B(f(u), \epsilon)$. Taking preimages yields the following inclusions:

$$B(u, \delta) \subset f^{-1}(B(f(u), \epsilon)) \subset f^{-1}(U).$$

Hence, for all $u \in f^{-1}(U)$, there exists a $\delta > 0$ such that $B(u, \delta) \subset f^{-1}(U)$. Therefore, the preimage $f^{-1}(U)$ is an open subset of X. So, f must be continuous in the "open set" sense (4.1.3).

(b) Assume that f is continuous in the "open set" sense (4.1.3). Then, for all open subsets $W \subset Y$, the preimage $f^{-1}(W)$ is open in X. Once again, assume that $u \in f^{-1}(U)$, so that $f(u) \in U$. Observe that the open ball $B(f(u), \epsilon)$ is an open subset of Y for all $\epsilon > 0$. Since f is continuous in the "open set" sense of 4.1.3, the preimage of the open ball $f^{-1}(B(f(u), \epsilon))$ must also be open. Moreover, since $f(u) \in B(f(u), \epsilon)$, $u \in f^{-1}(B(f(u), \epsilon))$. Since $f^{-1}(B(f(u), \epsilon))$ is open, there exists a $\delta > 0$ such that $B(u, \delta) \subset f^{-1}(B(f(u), \epsilon))$. When mapped onto Y via f, we find that for all $\epsilon > 0$, there exists a $\delta > 0$ such that

$$f(B(u, \delta)) \subset B(f(u), \epsilon).$$

Therefore, f is continuous in the "open ball" sense of 4.1.1. Consequently, the two definitions are equivalent. □

Another interesting point to note is that the continuity of f in the "open set" sense can also be adapted for closed sets.

Definition 4.1.4. Let X and Y be metric spaces and $f : X \to Y$ be a function. Then, f is a continuous function if for all closed subsets $U \subset Y$, the preimage $f^{-1}(U)$ is closed in X.

Theorem 4.1.6. *The definition of a continuous function given in 4.1.4 is equivalent to the other definitions of continuity in 4.1.1 and 4.1.3.*

Proof. Assume that X and Y are metric spaces and that $f : X \to Y$ is a function.

To show: (a) If f is continuous in the sense of 4.1.3, then f is continuous in the sense of 4.1.4.

(b) If f is continuous in the sense of 4.1.4, then f is continuous in the sense of 4.1.3.

(a) Assume that f is continuous in the "open set" sense of 4.1.3. That is, for all open sets $U \subset Y$, the preimage $f^{-1}(U)$ is open in X. Assume that $W \subset Y$ is closed.

To show: (aa) The preimage $f^{-1}(W)$ is closed.

(aa) Since W is closed, its complement $Y \backslash W$ is open. As a result, the preimage $f^{-1}(Y \backslash W)$ is open in X. Observe that

$$f^{-1}(Y \backslash W) \cup f^{-1}(W) = \{x \in X \text{ such that } f(x) \in Y \backslash W \cup W\}.$$

Since $Y \backslash W \cup W = Y$, $f^{-1}(Y \backslash W) \cup f^{-1}(W) = X$. Subsequently, $f^{-1}(Y \backslash W) = X \backslash f^{-1}(W)$.

Now suppose for the sake of contradiction that $f^{-1}(W)$ is not closed. Then, there exists a point $w \in f^{-1}(W)$ and a constant $r \in \mathbb{R}_{>0}$ such that $B(w, r) \cap f^{-1}(W) = \emptyset$. So, $B(w, r) \subset X \backslash f^{-1}(W) = f^{-1}(Y \backslash W)$. Therefore, $w \in f^{-1}(Y \backslash W)$. Consequently, we conclude that w is an element of both $f^{-1}(W)$ and $X \backslash f^{-1}(W)$, leading to a contradiction.

Hence, the preimage $f^{-1}(W)$ is closed.

(b) Assume that f is continuous in the "closed set" sense of 4.1.4. Then, for all closed sets $U \subset Y$, the preimage $f^{-1}(U)$ is closed in X. Assume $Z \subset Y$ is an open set.

To show: (ba) The preimage $f^{-1}(Z)$ is open.

(ba) Since Z is an open set, its complement $Y \backslash Z$ must be a closed set. Subsequently, the preimage $f^{-1}(Y \backslash Z)$ is a closed subset of X by assumption. Similarly to part (aa), we note that $f^{-1}(Y \backslash Z) \cup f^{-1}(Z) = X$. So, $f^{-1}(Z) = X \backslash f^{-1}(Y \backslash Z)$ is an open set.

Therefore, the definition of a continuous function in the sense of 4.1.4 is equivalent to the other definitions of continuity purported in 4.1.1 and 4.1.3. $\square$

One preliminary definition required for the open mapping theorem is the concept of an open map.

Definition 4.1.5. Let X and Y be topological spaces and $f : X \to Y$ be a map between them. f is said to be **open** if for all open subsets $U \subseteq X$, the image $f(U)$ is open in Y.

Here is a nice proof for the purpose of getting acquainted with an open map.

Theorem 4.1.7. *Let X and Y be topological spaces. Let $f : X \to Y$ be a continuous map. Then, f is a homeomorphism if and only if f is bijective and open.*

Proof. Assume that X and Y are topological spaces. Assume that $f : X \to Y$ is a continuous function.

To show: (a) If f is a homeomorphism, then f is bijective and open.

(b) If f is bijective and open, then f is a homeomorphism.

(a) Assume that f is a homeomorphism. Then, by definition f is bijective. It remains to demonstrate that f is open. Assume that $U \subseteq X$ is an open subset of X. Since $f^{-1} : Y \to X$ is a continuous function, due to the assumption that f is a homeomorphism, we find that the preimage $(f^{-1})^{-1}(U)$ must be an open subset of Y. However, this simplifies as follows:

$$\begin{aligned}(f^{-1})^{-1}(U) &= \{y \in Y \mid f^{-1}(y) \in U\} \\ &= \{y \in Y \mid y \in f(U)\} \\ &= f(U).\end{aligned}$$

Hence, $f(U)$ is an open subset of Y. This proves the second property.

(b) Assume that f is bijective and the image $f(U) \subseteq Y$ is open in Y whenever $U \subseteq X$ is open in X. We already know that the map f is continuous by assumption. So, it suffices to show that f^{-1} is continuous.

To show: (ba) f^{-1} is continuous.

(ba) Assume that $V \subseteq X$ is an open subset of X. Then, the preimage $(f^{-1})^{-1}(V) = f(V)$. However, $f(V)$ is an open subset of Y because V is open in X. Therefore, $(f^{-1})^{-1}(V)$ is open and as a result, f^{-1} is continuous.

(b) Consequently, f must be a homeomorphism.

□

Now we finally proceed to the open mapping theorem.

Theorem 4.1.8 (Open Mapping Theorem)**.** *Let X and Y be Banach spaces. Let $T : X \to Y$ be a bounded and surjective linear operator. Then, T is open.*

Proof. Assume that X and Y are Banach spaces. Assume that $T : X \to Y$ is a bounded and surjective linear operator. Note that by linearity of T, the image of an open ball $B_X(x, r)$ can be expressed as

$$T(B_X(x, r)) = T(x) + rT(B_X(0, 1))$$

where $x \in X$ and $r \in \mathbb{R}_{>0}$. So, it suffices to show that $T(B_X(0, 1))$ is open in Y.

To show: (a) T is open.

(a) To show: (aa) The image $T(B_X(0, 1))$ is open in Y.

(aa) Firstly, we note that $X = \bigcup_{n \in \mathbb{Z}_{>0}} B_X(0, n)$. From this, we use the surjectivity of T to express Y as

$$Y = \bigcup_{n \in \mathbb{Z}_{>0}} T(B_X(0, n)).$$

Due to this, we can use 4.1.2 in order to deduce that there exists a $m \in \mathbb{Z}_{>0}$ such that $\overline{T(B_X(0, m))}$ has non-empty interior. However, we note that $B_X(0, m)$ and $B_X(0, 1)$ are homeomorphic to each other. Therefore, it must be the case that $\overline{T(B_X(0, 1))}$ has non-empty interior. So, there exists $y_0 \in Y$ and $s \in \mathbb{R}_{>0}$ such that

$$B_Y(y_0, s) \subseteq \overline{T(B_X(0, 1))}.$$

For the next part, we note that the unit ball is convex and symmetric. This also holds for $T(B_X(0, 1))$ and its closure. By symmetry, $B_Y(-y_0, s) \subseteq \overline{T(B_X(0, 1))}$ and by convexity,

$$B_Y(0, s) = \frac{1}{2}B_Y(y_0, s) + \frac{1}{2}B_Y(-y_0, s) \subseteq \overline{T(B_X(0, 1))}.$$

By linearity of T, we can rescale this equation by a factor of 2^{-n}. In turn, we find that for all $n \in \mathbb{Z}_{>0}$,

$$B_Y(0, 2^{-n}s) \subseteq \overline{T(B_X(0, 2^{-n}))}.$$

To show: (aaa) $B_Y(0, s/2) \subseteq T(B_X(0, 1))$.

(aaa) Assume that $y \in B_Y(0, s/2)$. Since $B_Y(0, s/2) \subseteq \overline{T(B_X(0, 1/2))}$ we can construct a sequence $\{y_{m,1}\}_{m\in\mathbb{Z}_{>0}}$ in $T(B_X(0, 1/2))$ such that $y_{m,1} \to y$ as $m \to \infty$. So, there exists $x_1 \in B_X(0, 1/2)$ such that $T(x_1) \in Y$ is a point in the sequence $\{y_{m,1}\}_{m\in\mathbb{Z}_{>0}}$ satisfying $\|y - Tx_1\| < 2^{-2}s$.

Now, $y - Tx_1 \in B_Y(0, 2^{-2}s)$. Since $B_Y(0, 2^{-2}s) \subseteq \overline{T(B_X(0, 1/4))}$, we can construct a sequence $\{y_{m,2}\}$ in $T(B_X(0, 1/4))$ such that $y_{m,2} \to y - Tx_1$ as $m \to \infty$. Similarly, there exists $x_2 \in B_X(0, 2^{-2}) \subseteq X$ such that Tx_2 is a point in the sequence $\{y_{m,2}\}$ which satisfies $\|y - Tx_1 - Tx_2\| < 2^{-3}s$.

Inductively, we construct a sequence $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ such that if $n \in \mathbb{Z}_{>0}$,

$$\|y - \sum_{i=1}^{n-1} Tx_i\| < 2^{-n}s.$$

Since $x_n \in B_X(0, 2^{-n})$ for all $n \in \mathbb{Z}_{>0}$, the sequence $\{x_n\}_{n\in\mathbb{Z}_{>0}}$ is Cauchy and hence, convergent to (say) $x \in X$. Observe that

$$\|x\| = \lim_{n\to\infty} \|x_n\| \leq \sum_{n=1}^{\infty} \|x_n\| < \sum_{n=1}^{\infty} 2^{-n} = 1.$$

We also have

$$\begin{aligned}
\|y - Tx\| &= \|y - \lim_{n\to\infty} \sum_{j=1}^{n} Tx_j\| \\
&= \lim_{n\to\infty} \|y - \sum_{j=1}^{n} Tx_j\| \\
&< \lim_{n\to\infty} 2^{-n-1}s = 0.
\end{aligned}$$

So, $y = Tx$. Since $x \in B_X(0, 1)$, we deduce that $y \in T(B_X(0, 1))$ so that $B_Y(0, s/2) \subseteq T(B_X(0, 1))$.

(aa) This means that the image $T(B_X(0, 1))$ is open.

(a) Hence, T must be an open map. □

One corollary of the open mapping theorem is the inverse mapping theorem, which ensures that the inverse of an invertible bounded linear operator is again bounded.

Corollary 4.1.9 (Inverse Mapping Theorem). *Let X and Y be Banach spaces. Let $T : X \to Y$ be a continuous and bijective linear operator. Then, T^{-1} must be continuous as well.*

Proof. Assume that X and Y are Banach spaces. Assume that $T : X \to Y$ is a continuous and bijective linear operator. Since T is continuous, it is bounded. By the open mapping theorem (4.1.8), T must be open because T is also surjective. So, T is continuous and open. Therefore, T is a homeomorphism between X and Y, when viewed as topological spaces. So, T^{-1} must be continuous as a result. □

4.2 Closed Graph Theorem

The next theorem draws parallels with analogous results in the study of topological spaces. As it is instructive to study and peruse, we will compile these analogous results below:

Theorem 4.2.1 (Graph is closed). *Let $f : X \to Y$ be a function between two topological spaces X and Y. Then, if Y is Hausdorff and f is continuous, then the graph of f, denoted by $\Gamma_f \subseteq X \times Y$, is a closed subset of $X \times Y$.*

Proof. Assume that X and Y are topological spaces. Assume that $f : X \to Y$ is a continuous map. Assume that Y is Hausdorff.

To show: (a) Γ_f is a closed subset of $X \times Y$.

(a) To show: (aa) $(X \times Y)\backslash\Gamma_f$ is an open subset of $X \times Y$.

(aa) The set $(X \times Y)\backslash\Gamma_f$ is defined as follows:

$$(X \times Y)\backslash\Gamma_f = \{(x, y) \in X \times Y \mid y \neq f(x)\}.$$

Assume that $(x, y) \in (X \times Y)\backslash\Gamma_f$. Then, $y \neq f(x)$ by definition. Since Y is Hausdorff, there exist open subsets of Y, U and V, such that $y \in U$ and $f(x) \in V$. Also, $U \cap V = \emptyset$. This means that $(x, y) \in X \times U$, which is an open subset of $X \times Y$, disjoint from $X \times V \subseteq \Gamma_f$. Therefore, $(X \times Y)\backslash\Gamma_f$ is an open set.

(a) So, Γ_f is a closed subset of $X \times Y$. □

Theorem 4.2.2 (Continuous functions and graphs)**.** *Let $f : X \to Y$ be a map between two topological spaces. Then, if Y is compact and Γ_f is closed, then f is a continuous function.*

Proof. Assume that Y is compact and Γ_f is closed.

To show: (a) f is continuous.

(a) First, we consider the projection map $\pi_X : X \times Y \to X$.

To show: (aa) If V is a closed subset of $X \times Y$, then $\pi_X(V)$ is a closed subset of X.

(aa) Assume that V is a closed subset of $X \times Y$. Then, $(X \times Y)\backslash V$ is open in $X \times Y$. Assume that $x_0 \in X\backslash\pi_X(V)$. Then, $x_0 \times Y \not\subseteq V$. So, $(x_0 \times Y) \cap V = \emptyset$. But, $x_0 \times Y$ is homeomorphic to Y. So, $x_0 \times Y$ is compact because Y is compact. Now note for all $y \in Y$, there exists an open set $U_y \times V_y \in X \times Y$ such that $x_0 \in U_y$, $y \in V_y$ and $(U_y \times V_y) \cap V = \emptyset$ because $(x_0 \times Y) \cap V = \emptyset$. In this way, we have an open cover of $x_0 \times Y$. So, there exists a finite subcover $\{U_{i_j} \times V_{i_j}\}_{j=1}^n$ of $x_0 \times Y$, which is disjoint from V. Finally, consider the set $U = U_{i_1} \cap \cdots \cap U_{i_n}$. Note that U is an open subset of $X\backslash\pi_X(V)$, since it is the finite intersection of open sets. Furthermore, $x_0 \in U$. Hence, $X\backslash\pi_X(V)$ is open and as a result, $\pi_X(V)$ is a closed subset of X.

(a) Assume that $x_0 \in X$ and that $f(x_0) \in W$, which is an open subset of Y. Then, $Y\backslash W$ is a closed subset of Y. Due to this, $X \times (Y\backslash W)$ is a closed subset of $X \times Y$. Hence, the set $C = (X \times (Y\backslash W)) \cap \Gamma_f$ is closed. Therefore, $\pi_X(C)$ is a closed subset of X and $X\backslash\pi_X(C)$ is an open subset of X. However,

$$X\backslash\pi_X(C) = \{x \in X \mid (x, f(x)) \notin X \times (Y\backslash W)\}.$$

But, this set is just $f^{-1}(W)$. Hence, $f^{-1}(W)$ is open and so, f must be a continuous function.

□

In particular, 4.2.2 is very similar to the closed graph theorem. First, we will make the required definitions.

Definition 4.2.1. Let X and Y be Banach spaces. Let $\Lambda : X \to Y$ be a linear operator. Λ is said to be **closed** if its graph

$$\Gamma_\Lambda = \{(x, \Lambda x) \mid x \in X\} \subseteq X \times Y$$

is a closed subset of $X \times Y$. Alternatively, Λ is closed if the following holds: If $\{x_n\}$ and $\{y_n\}$ are sequences in X and Y respectively such that $y_n = \Lambda x_n$, $x_n \to x$ and $y_n \to y$, then $\Lambda x = y$.

Every continuous linear operator is closed because linear operators commute with limits. The closed graph theorem states that the converse is also true.

Theorem 4.2.3 (Closed Graph Theorem)**.** *Let X and Y be Banach spaces. Let $\Lambda : X \to Y$ be a closed linear operator. Then, Λ must be continuous.*

Proof. Assume that X and Y are Banach spaces. Assume that $\Lambda : X \to Y$ is a closed linear operator. Let Γ_Λ denote the graph of Λ. Then, Γ_Λ is a closed subset of $X \times Y$.

To show: (a) Λ is continuous.

(a) Consider the projection maps $\pi_1 : \Gamma_\Lambda \to X$ and $\pi_2 : \Gamma_\Lambda \to Y$, defined by $\pi_1(x, \Lambda x) = x$ and $\pi_2(x, \Lambda x) = \Lambda x$. Observe that they are continuous. Furthermore, since Γ_Λ is a closed subspace of a Banach space $X \times Y$, Γ_Λ must also be a Banach space. As a result of this, π_1 is a continuous bijection between two Banach spaces. Hence, π_1^{-1} must be continuous. Finally, we note that $\Lambda = \pi_2 \circ \pi_1^{-1}$. Since Λ is the composite of two continuous maps, Λ must also be a continuous. □

4.3 Adjoint and compact operators

The concept behind an adjoint operator is very similar to that of an adjoint linear transformation in linear algebra. Let X and Y be Banach spaces over a field $\mathbb{K}$ and let X^* and Y^* be their associated dual spaces. Let $\Lambda : X \to Y$ denote a bounded linear operator and $y^* \in Y^*$. Then, there exists a continuous linear functional $x^* \in X^*$ such that for all $x \in X$,

$$x^*(x) = y^*(\Lambda x)$$

The functional x^* is the composite $y^* \circ \Lambda$. Define $\Lambda^* : Y^* \to X^*$ which sends y^* to $y^* \circ \Lambda$. Then, Λ^* is a bounded continuous linear operator as well and is referred to as the **dual** of Λ. Λ^* is also called an **adjoint operator**. The defining characteristic of Λ^* is that for all $x \in X$, $\Lambda^*(y^*)(x) = y^*(\Lambda x)$. This can also be written as the natural pairing

$$\langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle.$$

where $\langle -, -\rangle : X^* \times X \to \mathbb{C}$ sends (f, v) to $f(v)$. The natural pairing written above emphasises the connection of the adjoint operator to adjoint linear transformations on finite dimensional vector spaces. We will now prove a few properties adjoint operators satisfy.

Theorem 4.3.1. *Let X and Y be Banach spaces. Let $\Lambda : X \to Y$ be a bounded linear operator and $\Lambda^* : Y^* \to X^*$ be its associated dual. Then, $\|\Lambda\| = \|\Lambda^*\|$.*

Proof. Assume that X and Y are Banach spaces. Assume that $\Lambda : X \to Y$ is a bounded linear operator and Λ^* is its associated dual. We argue as follows:

$$\begin{aligned}
\|\Lambda^*\| &= \sup_{\|y^*\|=1} \|\Lambda^* y^*\| \\
&= \sup_{\|y^*\|=1} \sup_{\|x\|=1} \|\langle \Lambda^* y^*, x\rangle\| \\
&= \sup_{\|y^*\|=1} \sup_{\|x\|=1} \|\langle y^*, \Lambda x\rangle\| \\
&= \sup_{\|x\|=1} \|\Lambda x\| \\
&= \|\Lambda\|.
\end{aligned}$$

Hence, $\|\Lambda\| = \|\Lambda^*\|$. In particular, this shows that Λ^* is a bounded operator. A direct calculation establishes linearity. □

Theorem 4.3.2. *Let X and Y be Banach spaces. Let $\Lambda : X \to Y$ be a bounded linear operator and $\Lambda^* : Y^* \to X^*$ be its associated dual. Then, $\ker(\Lambda) = [im(\Lambda^*)]^\perp$ and $\ker(\Lambda^*) = [im(\Lambda)]^\perp$.*

Proof. Assume that X and Y are Banach spaces. Assume that $\Lambda : X \to Y$ is a bounded linear operator and Λ^* is its associated dual. For the first equality, we use the definition of Λ^* to obtain for all $y^* \in Y^*$,

$$\begin{aligned}
\ker(\Lambda) &= \{x \in X \mid \Lambda x = 0\} \\
&= \{x \in X \mid \langle y^*, \Lambda x\rangle = 0\} \\
&= \{x \in X \mid \langle \Lambda^* y^*, x\rangle = 0\} \\
&= [\mathrm{im}(\Lambda^*)]^\perp.
\end{aligned}$$

In a similar vein, for all $x \in X$, we have

$$\begin{aligned}\ker(\Lambda^*) &= \{y^* \in Y^* \mid \Lambda^* y^* = 0\} \\ &= \{y^* \in Y^* \mid \langle \Lambda^* y^*, x\rangle = 0\} \\ &= \{y^* \in Y^* \mid \langle y^*, \Lambda x\rangle = 0\} \\ &= [\mathrm{im}(\Lambda)]^{\perp}.\end{aligned}$$

□

We have already encountered the definition of a compact operator briefly. To summarise, a bounded linear operator $\Lambda : X \to Y$ is compact if for all bounded sequences $\{x_n\}$, there exists a subsequence $\{x_{n_k}\}$ such that $\{\Lambda x_{n_k}\}$ converges in Y. We will now prove some characteristic properties of compact operators.

Theorem 4.3.3. *Let X and Y be Banach spaces. Let $\Lambda : X \to Y$ be a bounded linear operator. Then, Λ is compact if and only if for any bounded set $U \subset X$, $\overline{\Lambda(U)}$ is a compact subset of Y.*

Proof. Assume that X and Y are Banach spaces. Assume that $\Lambda : X \to Y$ is a bounded linear operator.

To show: (a) If Λ is compact, then for any bounded set $U \subset X$, $\overline{\Lambda(U)}$ is a compact subset of Y.

(b) If $\overline{\Lambda(U)}$ is a compact subset of Y for any bounded set $U \subset X$, then Λ is a compact operator.

(a) Assume that Λ is a compact operator. Assume that U is a bounded subset of X. Assume that $\{v_n\}$ is a sequence in $\Lambda(U)$. Then, there exists a sequence $\{u_n\}$ such that $\Lambda u_n = v_n$. Since $u_n \in U$, $\{u_n\}$ must be bounded. So, there exists a subsequence $\{u_{n_k}\}$ such that $\{v_{n_k}\}$ converges in $\Lambda(U)$. Since this is a subsequence of $\{v_n\}$, we deduce that $\overline{\Lambda(U)}$ is compact.

(b) Assume that $\overline{\Lambda(U)}$ is a compact subset of Y for any bounded set $U \subset X$. Let $\{u_n\}$ be a sequence in U. Since U is bounded, $\{u_n\}$ must also be bounded. Then, $\{\Lambda u_n\}$ is a sequence in $\overline{\Lambda(U)}$. Since $\overline{\Lambda(U)}$ is compact, there exists a convergent subsequence $\{\Lambda u_{n_k}\}$ in $\Lambda(U)$. So, Λ is a compact operator. □

Theorem 4.3.4. *Let X and Y be Banach spaces. Let $\Lambda : X \to Y$ be a bounded linear operator. If the range/image of Λ is finite dimensional, then Λ must be compact.*

Proof. Assume that X and Y are Banach spaces. Assume that Λ is a bounded linear operator. Assume that Λ has finite dimensional range. Observe that $\overline{\Lambda(U)} \subseteq \text{Im}(\Lambda)$ for all bounded subsets $U \subseteq X$. Since $\overline{\Lambda(U)}$ is a closed and bounded subset of a finite dimensional space, it must be compact. So, Λ must be a compact operator. □

Theorem 4.3.5. *Let X and Y be Banach spaces. Let $\Lambda_n : X \to Y$ be compact operators for all $n \in \mathbb{Z}_{>0}$. If $\lim_{n\to\infty} \|\Lambda_n - \Lambda\| = 0$, then Λ must be compact as well.*

Proof. To prove this, we will use the following fact: since Y is complete, $\overline{\Lambda(B_X(0,1))}$ is compact if and only if $\Lambda(B_X(0,1))$ is **precompact**, which means that for all $\epsilon \in \mathbb{R}_{>0}$, the set $\Lambda(B_X(0,1))$ can be covered by finitely many balls of radius ϵ. Motivated by this, assume $\epsilon \in \mathbb{R}_{>0}$. Assume that $\lim_{n\to\infty} \|\Lambda_n - \Lambda\| = 0$. Then, we can choose $k \in \mathbb{Z}_{>0}$ such that $\|\Lambda - \Lambda_k\| < \epsilon/2$. Since $\overline{\Lambda_k(B_X(0,1))}$ is a compact subset of Y, there exists points $y_i \in Y$ for all $i \in \{1, \ldots, m\}$ such that

$$\overline{\Lambda(B_X(0,1))} \subseteq \bigcup_{i=1}^{m} B_Y(y_i, \frac{\epsilon}{2}).$$

Now, assume $x \in X$ such that $\|x\| \leq 1$. Then, from the definition of the operator norm, $\|\Lambda_k x - \Lambda x\| < \epsilon/2$. Furthermore, since $x \in B_X(0,1)$, there exists $y_j \in Y$ for $j \in \{1, \ldots, m\}$ such that $\|\Lambda_k x - y_j\| < \epsilon/2$. By the triangle inequality,

$$\begin{aligned} \|\Lambda x - y_j\| &\leq \|\Lambda x - \Lambda_k x\| + \|\Lambda_k x - y_j\| \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &< \epsilon. \end{aligned}$$

This means that $\{B_Y(y_i, \epsilon)\}_{i=1}^{m}$ is a finite open cover of $\Lambda(B_X(0,1))$. Hence, from the result we stated at the beginning, $\overline{\Lambda(B_X(0,1))}$ must be compact, which in turn finally reveals that Λ is a compact operator.

□

Another useful property of compact operators is that Λ is compact if and only if its adjoint Λ^* is compact. The proof requires the **Arzela-Ascoli**

theorem, which is important and requires a proof of its own. First, we will make the appropriate definitions.

Definition 4.3.1. Let X be a topological space and (Y, d) be a metric space. Let $\mathcal{F}$ be a family of functions from X to Y. $\mathcal{F}$ is said to be **equicontinuous** at $x \in X$ if for all $\epsilon \in \mathbb{R}_{>0}$ there exists a neighbourhood U_ϵ of x such that

$$d(f(z), f(x)) < \epsilon$$

for all $z \in U_\epsilon$ and all $f \in \mathcal{F}$. $\mathcal{F}$ is called equicontinuous if it is equicontinuous at all points $x \in X$.

Theorem 4.3.6 (Arzela-Ascoli)**.** *Let X be a compact topological space and (M, d) be a complete metric space. These conditions ensure that $Cts(X, M)$ is a complete metric space itself, equipped with the usual d_∞ norm. Let $\mathcal{F} \subseteq Cts(X, M)$. Then, $\overline{\mathcal{F}}$ is compact in $Cts(X, M)$ if and only if the two conditions below hold:*

1. *$\mathcal{F}$ is equicontinuous.*
2. *For all $x \in X$, the set $\mathcal{F}(x) = \{f(x) \mid f \in \mathcal{F}\}$ has compact closure in M.*

Proof. Assume that X is compact and (M, d) is a complete metric space. Assume that $\mathcal{F} \subseteq Cts(X, M)$. Since $Cts(X, M)$ is a complete metric space, $\mathcal{F}$ has compact closure if and only if it is precompact (alternatively, totally bounded). This means that the set $\mathcal{F}$ can be covered by finitely many balls of radius ϵ, where $\epsilon \in \mathbb{R}_{>0}$.

To show: (a) If the two conditions are satisfied, then $\mathcal{F}$ must be precompact.

(a) Assume that the two conditions are satisfied. Assume that $\epsilon \in \mathbb{R}_{>0}$. Then, since $\mathcal{F}$ is equicontinuous, for each $x \in X$, there exists a neighbourhood $V(x)$ of x such that if $y \in V(x)$, then $d(f(x), f(y)) < \epsilon$ for all $f \in \mathcal{F}$. This renders $\{V(x)\}_{x \in X}$ an open cover of X. Since X is compact, there exists a finite subcover $V(x_1), \ldots, V(x_n)$ of X, where $x_i \in X$ for all $i \in \{1, \ldots, n\}$.

By the second condition, the sets $\mathcal{F}(x_j) \subseteq M$ are precompact in M. Hence, the union $\mathcal{F}(x_1) \cup \cdots \cup \mathcal{F}(x_n)$ is also precompact in M. Hence, we can cover this subset with balls of radius ϵ centred at the points $a_1, a_2, \ldots, a_m$.

This means that the set $\{a_1, \dots, a_m\}$ is an ϵ-net for $\mathcal{F}(x_1) \cup \dots \cup \mathcal{F}(x_n)$. For every map φ between the sets $\{1, \dots, n\}$ and $\{1, \dots, m\}$ and for all $j \in \{1, \dots, n\}$, define

$$B_\varphi = \{f \in \mathcal{F} \mid d(f(x_j), a_{\varphi(j)}) < \epsilon\}.$$

Note that there are finitely many sets B_φ and that every $f \in \mathcal{F}$ belongs to one of these sets, due to the fact that $\mathcal{F}$ is equicontinuous. If $f, g \in \mathcal{F}$, then for all $y \in V(x_k)$,

$$d(f(y), g(y)) \leq d(f(y), f(x_k)) + d(f(x_k), a_{\varphi(k)}) + d(a_{\varphi(k)}, g(x_k)) + d(g(x_k), g(y)) < 4\epsilon.$$

Since the $V(x_k)$ forms an open cover of X for all $k \in \{1, \dots, n\}$, this means that $d_\infty(f, g) < 4\epsilon$. So, the diameter of the sets B_φ is at most 4ϵ and they cover $\mathcal{F}$ by the equicontinuous assumption. Hence, $\mathcal{F}$ is precompact and thus, $\overline{\mathcal{F}}$ is compact.

To show: (b) If $\mathcal{F}$ has compact closure, then the two conditions in the statement of the theorem are satisfied.

(b) Assume that $\mathcal{F}$ has compact closure and thus, is precompact. Then, $\mathcal{F}$ has a finite ϵ-net. Note that for all $x \in X$,

$$d(f(x), g(x)) \leq \sup_{x \in X} d(f(x), g(x)) = d_\infty(f, g).$$

Hence, for all $x \in X$, $\mathcal{F}(x)$ is precompact (the same ϵ-net works as before). This proves condition 2. To prove condition 1 holds, since $\mathcal{F}$ is precompact, it has a finite ϵ-net, say $\{f_1, \dots, f_n\}$. So, for all $x \in X$, there exists an open neighbourhood $V(x)$ such that for all $y \in V(x)$, $d(f_j(x), f_j(y)) < \epsilon$ for all $j \in \{1, \dots, n\}$. Now assume that $f \in \mathcal{F}$. Then, we can choose f_k for some $k \in \{1, \dots, n\}$ such that $d_\infty(f_k, f) < \epsilon$. Hence, for all $y \in V(x)$,

$$d(f(x), f(y)) \leq d(f(x), f_k(x)) + d(f_k(x), f_k(y)) + d(f_k(y), f(y)) < 3\epsilon.$$

Therefore, $\mathcal{F}$ is equicontinuous at $x \in X$. Since x was arbitrary, we deduce that $\mathcal{F}$ must be equicontinuous. □

The corollary that we will use from 4.3.6 is stated below:

Corollary 4.3.7. *Let X be a compact topological space and Y a complete metric space. Let $\mathcal{F} \subseteq Cts(X, Y)$ be an equicontinuous family. Then, every sequence of functions in $\mathcal{F}$ has a uniformly convergent subsequence.*

The corollary stems in particular from $\mathcal{F}$ being precompact, given the assumptions of the corollary and Theorem 4.3.6. Now, we will prove the result alluded to regarding compact operators.

Theorem 4.3.8. *Let X and Y be Banach spaces. Let $\Lambda : X \to Y$ be a bounded linear operator. Then, $\Lambda^* : Y^* \to X^*$ is compact if and only if Λ is a compact operator.*

Proof. Assume that X and Y are Banach spaces. Assume that $\Lambda : X \to Y$ is a bounded linear operator.

To show: (a) If Λ is a compact operator, then $\Lambda^* : Y^* \to X^*$ is a compact operator.

(b) If Λ^* is a compact operator, then Λ is a compact operator.

(a) Assume that Λ is a compact operator. If $\{x_n\}_{n=1}^{\infty}$ is a sequence in the closed unit ball

$$B_C(0,1) = \{x \in X \mid \|x\| \leq 1\},$$

then $\{x_n\}$ is bounded and $\{\Lambda x_n\}$ has a convergent subsequence in Y. Since Y is a Banach space, it must be complete. As a result, the closure $K = \overline{\Lambda(B_C(0,1))}$ must be compact.

Now let $\{\phi_n\}_{n=1}^{\infty}$ be a bounded sequence in Y^*. Without loss of generality, we can assume that $\|\phi_n\| \leq 1$ for $n \in \mathbb{Z}_{>0}$.

To show: (aa) $\{\phi_n\}_{n=1}^{\infty}$ is equicontinuous.

(aa) Assume that $\epsilon \in \mathbb{R}_{>0}$ and $n \in \mathbb{Z}_{>0}$. Set $\delta = \epsilon$. If $|y - y'| < \delta$ in Y then

$$|\phi_n(y) - \phi_n(y')| \leq \|\phi_n\|\|y - y'\| \leq \|y - y'\| < \epsilon.$$

Since $n \in \mathbb{Z}_{>0}$, we deduce that the family of functionals $\{\phi_n\}_{n=1}^{\infty}$ is equicontinuous.

(a) The key point here is that due to part (aa), the sequence $\{\phi_n|_K\}_{n=1}^{\infty}$ is equicontinuous, where $\phi_n|_K$ is the restriction of ϕ_n to $K = \overline{\Lambda(B_C(0,1))}$, which is a compact subset of Y.

By the Arzela-Ascoli theorem, there exists a subsequence $\{\phi_{n_k}|_K\}$ of $\{\phi_n|_K\}$, which converges uniformly on K. We now claim that $\{\Lambda^*\phi_{n_k}\}$

converges in X^*.

It suffices to prove that $\{\Lambda^* \phi_{n_k}\}$ is Cauchy. Since $\{\phi_{n_k}|_K\}$ is a convergent sequence in Y^*, $\{\phi_{n_k}|_K\}$ must be Cauchy. Assume that $\epsilon \in \mathbb{R}_{>0}$. Select $N \in \mathbb{Z}_{>0}$ such that if $i, j > N$ then

$$\sup_{k \in K} |\phi_{n_i}(k) - \phi_{n_j}(k)| < \epsilon.$$

We can do this because $\{\phi_{n_k}|_K\}$ converges uniformly on K. Subsequently, if $i, j > N$ then

$$\begin{aligned}
\|\Lambda^* \phi_{n_i} - \Lambda^* \phi_{n_j}\| &= \sup_{\|x\|=1} |\Lambda^* \phi_{n_i}(x) - \Lambda^* \phi_{n_j}(x)| \\
&= \sup_{\|x\|=1} |\phi_{n_i}(\Lambda x) - \phi_{n_j}(\Lambda x)| \\
&\leq \sup_{k \in K} |\phi_{n_i}(k) - \phi_{n_j}(k)| < \epsilon.
\end{aligned}$$

The second last inequality follows from the fact that $\Lambda x \in \Lambda(B_C(0,1)) \subseteq K$. Therefore, the sequence $\{\Lambda^* \phi_{n_k}\}$ is Cauchy in X^*. Since X^* is complete, the sequence $\{\Lambda^* \phi_{n_k}\}$ must converge and thus, qualifies as a convergent subsequence of $\{\Lambda^* \phi_n\}$. So, Λ^* is a compact operator.

(b) Assume that $\Lambda^* : Y \to X$ is a compact operator. By applying part (a) to Λ^*, we deduce that $(\Lambda^*)^* : (X^*)^* \to (Y^*)^*$ is also a compact operator.

Let $\iota_X : X \to (X^*)^*$ and $\iota_Y : Y \to (Y^*)^*$ be the canonical injective isometries. Then,

$$(\Lambda^*)^*|_X = (\Lambda^*)^* \circ \iota_X = \iota_Y \circ \Lambda.$$

Since $(\Lambda^*)^*$ is compact, the composite $(\Lambda^*)^* \circ \iota_X$ is also compact. Hence, $\iota_Y \circ \Lambda$ is a compact operator from X to $(Y^*)^*$.

We claim that because ι_Y is an isometry, Λ must be compact. Let $\{x_n\}$ be a bounded sequence in X. Then, $\{(\iota_Y \circ \Lambda)x_n\}$ must have a convergent subsequence $\{(\iota_Y \circ \Lambda)x_{n_k}\}$ in $(Y^*)^*$. So, this sequence must be Cauchy.

Now assume that $\epsilon \in \mathbb{R}_{>0}$. Select $P \in \mathbb{Z}_{>0}$ such that if $i, j > P$

$$\|(\iota_Y \circ \Lambda)x_{n_i} - (\iota_Y \circ \Lambda)x_{n_j}\| < \epsilon.$$

Using the fact that ι_Y is an isometry, we find that if $i, j > P$ then

$$\begin{aligned}\|\Lambda x_{n_i} - \Lambda x_{n_j}\| &= \|\Lambda(x_{n_i} - x_{n_j})\| \\ &= \|(\iota_Y \circ \Lambda)(x_{n_i} - x_{n_j})\| < \epsilon.\end{aligned}$$

So, $\{\Lambda x_{n_k}\}$ is a Cauchy sequence in Y. Since Y is complete, it must be convergent. So, $\{\Lambda x_{n_k}\}$ is a convergent subsequence of $\{\Lambda x_n\}$ and therefore, Λ is compact as required. □

4.4 Weak Convergence in Hilbert Spaces

We know that every Hilbert space can be converted into a Banach space. So, the concept of weak convergence, which was introduced in the context of normed vector spaces, can be adapted for Hilbert spaces. The crucial difference here is the presence of the Riesz Representation Theorem (3.3.1) for Hilbert spaces. It allows us to express weak convergence with the inner product.

Definition 4.4.1. Let H be a Hilbert space. Let $\{x_n\}$ be a sequence of points in H. We say that $\{x_n\}$ **converges weakly** to a point $x \in H$ if for all $y \in H$,

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

As before, we write $x_n \rightharpoonup x$ to denote weak convergence.

One property of weakly convergent sequences is boundedness.

Theorem 4.4.1. *Let H be a Hilbert space. Let $\{x_n\}$ be a sequence of points in H such that $x_n \rightharpoonup x$. Then, $\{x_n\}$ is a bounded sequence, which means that there exists a constant $C \in \mathbb{K}$ such that $\|x_n\| \leq C$ for all $n \in \mathbb{Z}_{>0}$.*

Proof. Assume that H is a Hilbert space. Assume that $\{x_n\}$ is a sequence of points in H such that $x_n \rightharpoonup x$. This means that for all $y \in H$, $\lim_{n\to\infty}\langle y, x_n\rangle = \langle y, x\rangle$.

To show: (a) $\{x_n\}$ is bounded.

(a) First, we observe that for all $y \in H$, the set $\{\langle y, x_n\rangle \mid n \in \mathbb{Z}_{>0}\}$ is bounded. By the Riesz representation theorem (3.3.1), every $x_n \in H$ corresponds to a linear functional $\phi^{x_n} \in H^*$, which maps $y \in H$ to $\langle y, x_n\rangle$. So, consider the family of linear functionals

$$\Phi = \{\phi^{x_n} \mid n \in \mathbb{Z}_{>0}\}.$$

Each member of Φ is bounded because

$$\|\phi^{x_n}\| = \sup_{\|y\| \leq 1} |\langle y, x_n \rangle| < \infty.$$

Therefore, we can apply the uniform boundedness principle (4.1.3) in order to deduce that Φ is uniformly bounded. The other case cannot happen because $\phi^{x_n}(y)$ is bounded for all $y \in H$. Consequently, $\{x_n\}$ must be a bounded sequence.

□

The next result requires much more work to prove.

Theorem 4.4.2. *Let H be a Hilbert space and $\{x_n\}$ be a bounded sequence of points in H. Then, $\{x_n\}$ has a weakly convergent subsequence $\{x_{n_j}\}$ where $x_{n_j} \rightharpoonup x$ for some $x \in H$.*

Proof. Assume that H is a Hilbert space and $\{x_n\}$ is a bounded sequence of points in H. First, we will construct a candidate subsequence of $\{x_n\}$. Let V be the vector space $\overline{\text{span}\{x_n\}}$ and $V^\perp$ be its orthogonal complement. V is separable because it has a countable, dense subset $\{x_n \mid n \in \mathbb{Z}_{>0}\}$. Hence, by Gram-Schmidt orthogonalisation, we construct an orthonormal basis $\{e_1, e_2, \dots\}$ for V.

Now we will construct the sequence. Consider the sequence $\{\langle e_1, x_n \rangle\}$, where $n \in \mathbb{Z}_{>0}$. Note that this sequence is bounded due to our assumption. So, there exists $I_1 \subset \mathbb{Z}_{>0}$ such that the subsequence $\{\langle e_1, x_n \rangle\}_{n \in I_1}$ converges (remember that this is a sequence in either $\mathbb{R}$ or $\mathbb{C}$. So, the Bolzano-Weierstrass theorem applies). Next, the sequence $\{\langle e_2, x_n \rangle\}_{n \in I_1}$ is also bounded. In a similar vein, there exists $I_2 \subset I_1$ such that the subsequence $\{\langle e_2, x_n \rangle\}_{n \in I_2}$ converges. Continuing in this manner, we find that for all $m \in \mathbb{Z}_{>0}$, there exists a countable set of indices $I_m \subset I_{m-1}$ such that the sequence $\{\langle e_m, x_n \rangle\}_{n \in I_m}$ converges. Now choose a subsequence $n_1 < n_2 < \dots$ with $n_k \in I_k$ for all $k \in \mathbb{Z}_{>0}$. This provides us with the convergence

$$\lim_{k \to \infty} \langle e_m, x_{n_k} \rangle = \overline{\alpha_m}$$

for some $\alpha_m \in \mathbb{R}$ or $\mathbb{C}$ and for all $m \in \mathbb{Z}_{>0}$.

Next, consider the point

$$z = \sum_{m=1}^{\infty} \alpha_m e_m.$$

We have to show that the series actually converges.

To show: (a) z is well defined.

(a) Since the sequence $\{x_n\}$ is bounded, $\|x_n\| \leq C$ for some constant $C \in \mathbb{R}_{>0}$. By the Cauchy-Schwarz inequality, the real sequence $\{|\langle e_m, x_{n_k}\rangle|^2\}_{k \in \mathbb{R}_{>0}}$ satisfies

$$|\langle e_m, x_{n_k}\rangle|^2 \leq \|x_{n_k}\|^2 \leq C^2.$$

So, the dominated convergence theorem justifies this interchange between the limit and sum

$$\sum_{m=1}^{n} \lim_{k\to\infty} |\langle e_m, x_{n_k}\rangle|^2 = \lim_{k\to\infty} \sum_{m=1}^{n} |\langle e_m, x_{n_k}\rangle|^2$$

We think of the above equation as applying the dominated convergence theorem to integrals over the counting measure on $\mathbb{R}$. Hence, for all $n \in \mathbb{Z}_{>0}$,

$$\begin{aligned}
\sum_{m=1}^{n} |\alpha_m|^2 &= \sum_{m=1}^{n} \lim_{k\to\infty} |\langle e_m, x_{n_k}\rangle|^2 \\
&= \lim_{k\to\infty} \sum_{m=1}^{n} |\langle e_m, x_{n_k}\rangle|^2 \\
&\leq \lim_{k\to\infty} \sum_{m=1}^{\infty} |\langle e_m, x_{n_k}\rangle|^2 \\
&\leq \lim_{k\to\infty} \|x_{n_k}\|^2 \quad \text{(Theorem 3.5.2)} \\
&\leq C^2.
\end{aligned}$$

This reveals that the series $\sum_{m=1}^{\infty} |\alpha_m|^2$ is convergent. Since $\{e_1, e_2, \dots\}$ is an orthonormal basis for V, we find that for all $m > n$,

$$\|\sum_{k=n+1}^{m} \alpha_k e_k\|^2 = \sum_{k=n+1}^{m} |\alpha_k|^2 \to 0$$

as $m, n \to \infty$ since $\sum_{m=1}^{\infty} |\alpha_m|^2$ is convergent. So, the sequence of partial sums $\{\sum_{k=1}^{m} \alpha_k e_k\}_{m \in \mathbb{Z}_{>0}}$ is Cauchy and consequently, the series representation of z does indeed have a well-defined sum. Moreover,

$$\|z\|^2 = \sum_{m=1}^{\infty} |\alpha_m|^2 \le C^2.$$

This shows that z is well defined.

The crux of this argument is to show that the subsequence $\{x_{n_k}\}$ converges weakly to z.

To show: (b) $\{x_{n_k}\} \rightharpoonup z$.

(b) Assume that $y \in H$. We have to show that $\lim_{k\to\infty} \langle y, x_{n_k} \rangle = \langle y, z \rangle$. Using the orthogonal complement $V^{\perp}$, we decompose y as the sum $y_1 + y_2$, where $y_2 \in V^{\perp}$ and

$$y_1 = \sum_{m=1}^{\infty} b_m e_m \in V.$$

Assume that $\epsilon \in \mathbb{R}_{>0}$. Choose $N \in \mathbb{Z}_{>0}$ large enough so that

$$\sum_{m>N} |b_m|^2 < \epsilon^2.$$

Then,

$$\begin{aligned}
\langle y, x_{n_k} - z \rangle &= \langle y_1 + y_2, x_{n_k} - z \rangle \\
&= \langle y_1, x_{n_k} - z \rangle \\
&= \langle \sum_{m \le N} b_m e_m, x_{n_k} - z \rangle + \langle \sum_{m>N} b_m e_m, x_{n_k} - z \rangle \\
&= A_k + B_k.
\end{aligned}$$

We control A_k and B_k separately. Firstly, $\lim_{k\to\infty} A_k = 0$ due to our construction of $\{x_{n_k}\}$. We argue that

$$
\begin{aligned}
\lim_{k\to\infty} A_k &= \lim_{k\to\infty} \langle \sum_{m\leq N} b_m e_m, x_{n_k} - z\rangle \\
&= \lim_{k\to\infty} (\sum_{m\leq N} b_m \langle e_m, x_{n_k}\rangle - \sum_{m\leq N} b_m \langle e_m, z\rangle) \\
&= \lim_{k\to\infty} (\sum_{m\leq N} b_m \langle e_m, x_{n_k}\rangle - \sum_{m\leq N} b_m \langle e_m, \sum_{n=1}^{\infty} \alpha_n e_n\rangle) \\
&= \lim_{k\to\infty} (\sum_{m\leq N} b_m \langle e_m, x_{n_k}\rangle - \sum_{m\leq N} b_m \overline{\alpha_m}) = 0.
\end{aligned}
$$

Now, using all the inequalities we have found so far, we control $|B_k|$ as follows:

$$
\begin{aligned}
|B_k| &= |\langle \sum_{m>N} b_m e_m, x_{n_k} - z\rangle| \\
&\leq \| \sum_{m>N} b_m e_m \| \| x_{n_k} - z\| \qquad \text{(Cauchy-Schwarz)} \\
&= (\sum_{m>N} |b_m|^2)^{\frac{1}{2}} \| x_{n_k} - z\| \\
&< \epsilon(\|x_{n_k}\| + \|z\|) \\
&\leq 2C\epsilon.
\end{aligned}
$$

Hence, we have

$$
\lim_{k\to\infty} \sup |\langle y, x_{n_k} - z\rangle| = \lim_{k\to\infty} \sup |A_k + B_k| \leq 2C\epsilon.
$$

Since $\epsilon > 0$ was arbitrary, we conclude that $\limsup_{k\to\infty} |\langle y, x_{n_k} - z\rangle| = 0$, unveiling that the subsequence $\{x_{n_k}\} \rightharpoonup z$ as required. □

A peculiar property of a compact operator is that it maps weakly convergent sequences to *strongly* convergent sequences. The following theorem elucidates this point.

Theorem 4.4.3. *Let H be a Hilbert space. Let $\{x_n\}$ be a weakly convergent sequence such that $x_n \rightharpoonup x$ for some $x \in H$. Let $\Lambda : H \to H$ be a compact operator. Then,*

$$
\lim_{n\to\infty} \|\Lambda x_n - \Lambda x\| = 0,
$$

asserting that the sequence $\{\Lambda x_n\}$ converges strongly to the point $\Lambda x \in H$.

Proof. Assume that H is a Hilbert space and that $\{x_n\}$ is a weakly convergent sequence. Then, $\{x_n\}$ must be a bounded sequence. Since Λ is a compact operator, there exists a subsequence $\{x_{n_j}\}$ such that the sequence $\{\Lambda x_{n_j}\}$ converges strongly to (say) $y \in H$.

To show: (a) $y = \Lambda x$.

(a) We will exploit the uniqueness of the weak limit. Note that for all $v \in H$,

$$\langle \Lambda x_n - \Lambda x, v \rangle = \langle x_n - x, \Lambda^* v \rangle \to 0$$

as $n \to \infty$ because $x_n \rightharpoonup x$. Therefore, $\Lambda x_n \rightharpoonup \Lambda x$. But, we have already shown that $\{\Lambda x_{n_j}\}$ converges strongly to $y \in H$. Since $\{\Lambda x_{n_j}\}$ is strongly convergent, it must also be weakly convergent to the same point in H. Therefore, by the uniqueness of the weak limit, $\Lambda x = y$. □

The next theorem provides another characterisation of strong convergence from weak convergence.

Theorem 4.4.4. *Let $\{x_n\}_{n \in \mathbb{Z}_{>0}}$ be a sequence in a Hilbert space H. Then, $\{x_n\}$ strongly converges to $x \in H$ if and only if $\lim_{n\to\infty} \|x_n\| = \|x\|$ and $x_n \rightharpoonup x$.*

Proof. Assume that H is a Hilbert space and $\{x_n\}_{n \in \mathbb{Z}_{>0}}$ is a sequence in H.

Suppose that $\lim_{n\to\infty} \|x_n\| = \|x\|$ and $x_n \rightharpoonup x$. Then, $\lim_{n\to\infty} \langle x_n, x_n \rangle = \langle x, x \rangle$ and

$$\begin{aligned}
\lim_{n\to\infty} \|x_n - x\|^2 &= \lim_{n\to\infty} \langle x_n - x, x_n - x \rangle \\
&= \lim_{n\to\infty} (\langle x_n, x_n \rangle - \langle x_n, x \rangle - \langle x, x_n \rangle + \langle x, x \rangle) \\
&= \langle x, x \rangle - \langle x, x \rangle - \langle x, x \rangle + \langle x, x \rangle \\
&= 0.
\end{aligned}$$

Thus, $\|x_n - x\| \to 0$ as $n \to \infty$. This demonstrates that $\{x_n\}$ strongly converges to $x \in H$.

For the converse, suppose that $\{x_n\} \to x$. Assume that $y \in H$. Then,

$$\begin{aligned}
\lim_{n\to\infty} \langle y, x_n \rangle &= \lim_{n\to\infty} (\langle y, x_n - x \rangle) + \langle y, x \rangle \\
&= \langle y, x \rangle
\end{aligned}$$

since $\{x_n\}$ strongly converges to x. Hence, $\{x_n\} \rightharpoonup x$. To see that $\lim_{n\to\infty}\|x_n\| = \|x\|$, we note that

$$\begin{aligned}
\lim_{n\to\infty}\|x_n\|^2 &= \lim_{n\to\infty}\langle x_n, x_n\rangle \\
&= \lim_{n\to\infty}(\langle x_n - x, x_n - x\rangle + \langle x_n, x\rangle + \langle x, x_n\rangle) - \langle x, x\rangle \\
&= \langle x, x\rangle = \|x\|^2
\end{aligned}$$

where the last line follows from the fact that $\{x_n\}$ weakly converges to x. Thus, $\lim_{n\to\infty}\|x_n\| = \|x\|$, which completes the proof. □

Chapter 5

Some Spectral Theory

5.1 Fredholm Theorem

The Fredholm Theorem is incredibly important. It asserts that many properties of finite dimensional operators can be adapted to a Hilbert space. Within a Hilbert space H, these properties are only satisfied by operators of the form $I - K$, where $I : H \to H$ is the identity operator and $K : H \to H$ is a compact operator. The expression $I - K$ is reminiscent of the matrices one computes in the process of diagonalisation. In this section, we will go through the proof of this important theorem. We will break the theorem down into multiple components.

Theorem 5.1.1. *Let H be a Hilbert space over the field $\mathbb{R}$. Let $K : H \to H$ be a compact linear operator. Then,* $\ker(I - K)$ *is finite dimensional.*

Proof. Assume that H is a Hilbert space over the real numbers $\mathbb{R}$. Assume that $K : H \to H$ is a compact linear operator. Suppose for the sake of contradiction that $\ker(I - K)$ is infinite dimensional, where $I : H \to H$ is the identity operator. Then, by Gram-Schmidt, one can find an orthonormal basis $\{e_1, e_2, \dots\}$ contained in $\ker(I - K)$. Note that

$$(I - K)(e_n) = I(e_n) - K(e_n) = e_n - K(e_n) = 0.$$

Therefore, $Ke_n = e_n$ for all $n \in \mathbb{Z}_{>0}$. Since H is a Hilbert space, we can use Pythagoras' theorem in tandem with our orthonormal basis $\{e_1, e_2, \dots\}$ to deduce that for $m \neq n$,

$$\|e_m - e_n\|^2 = \|e_m\|^2 + \|e_n\|^2 = 2.$$

Hence, $\|Ke_m - Ke_n\| = \|e_m - e_n\| = \sqrt{2}$ for all $m, n \in \mathbb{Z}_{>0}$ where $m \neq n$. However, this means that the sequence $\{Ke_{n_k}\}$ can never converge, contradicting the assumption that K is a compact operator. Therefore, $\ker(I-K)$ is finite dimensional. □

Theorem 5.1.2. *Let H be a Hilbert space over the field $\mathbb{R}$. Let $K : H \to H$ be a compact linear operator. Then, $Im(I-K)$ is a closed subset of H.*

Proof. Assume that H is a Hilbert space over the real numbers $\mathbb{R}$. Assume that $K : H \to H$ is a compact linear operator. We will begin by proving a preliminary result.

To show: (a) There exists $\beta \in \mathbb{R}_{>0}$ such that $\|u - Ku\| \geq \beta\|u\|$ for all $u \in \ker(I-K)^{\perp}$.

(a) Suppose that the statement is not true. Then, we can find a sequence of points $u_n \in \ker(I-K)^{\perp}$ such that for all $n \in \mathbb{Z}_{>0}$, $\|u_n\| = 1$ and $\|u_n - Ku_n\| < \frac{1}{n}$. Observe that by construction, the sequence $\{u_n\}$ is bounded. So, it must contain a subsequence, which we will denote by $\{u_m\}$, such that $u_m \rightharpoonup u$ ($\{u_m\}$ is weakly convergent). When we apply K to the sequence $\{u_m\}$, we find that the sequence $\{Ku_m\}$ strongly converges to Ku because K is a compact operator on a Hilbert space. Now, we have the following inequality:

$$\|u_m - Ku\| \leq \|u_m - Ku_m\| + \|Ku_m - Ku\| < \frac{1}{m} + \|Ku_m - Ku\| \to 0$$

as $m \to \infty$. So, $\|u\| = \lim_{m\to\infty}\|u_m\| = 1$ and

$$\|u - Ku\| = \lim_{m\to\infty}\|u_m - Ku\| = 0.$$

The latter equality in particular reveals that $u = Ku$ and so, $u \in \ker(I-K)$. Since u is also an element of $\ker(I-K)^{\perp}$, $u = 0$. However, this contradicts the assumption that $\|u\| = 1$. Hence, the statement purported in part a must be true.

To show: (b) $\text{Im}(I-K)$ is a closed subset of H.

(b) It suffices to show that $\overline{\text{Im}(I-K)} \subseteq \text{Im}(I-K)$. Assume that $v \in \overline{\text{Im}(I-K)}$. Then, there exists a sequence of points $\{v_n\}$ in $\text{Im}(I-K)$ such that $v_n \to v$. Our objective is to demonstrate that $v \in \text{Im}(I-K)$.

To show: (ba) There exists $u \in H$ such that $u - Ku = v$.

(ba) Since $v_n \in \text{Im}(I - K)$, there exists $u_n \in H$ such that for all $n \in \mathbb{Z}_{>0}$, $v_n = u_n - Ku_n$. We want to use the result proved in part (a). In order to do this, let $\tilde{u}_n \in \ker(I - K)$ be the perpendicular projection of u_n on $\ker(I - K)$. Then, we define

$$z_n = u_n - \tilde{u}_n \in \ker(I - K)^\perp.$$

So, we have

$$v_n = u_n - Ku_n = z_n + \tilde{u}_n - Kz_n - K\tilde{u}_n = z_n - Kz_n.$$

This is because $K\tilde{u}_n = \tilde{u}_n$. Now, we can apply part (a) in order to deduce that there exists $\beta \in \mathbb{R}_{>0}$ such that

$$\|v_m - v_n\| \geq \beta \|z_m - z_n\|$$

for all $m.n \in \mathbb{Z}_{>0}$ such that $m \neq n$. Since the sequence $\{v_n\}$ converges to v, it must be Cauchy. This reveals that the sequence $\{z_n\}$ is also Cauchy and thus, convergent because H is complete. So, there exists $u \in H$ such that $z_n \to u$ and subsequently, we deduce that

$$u - Ku = \lim_{n\to\infty} (z_n - Kz_n) = \lim_{n\to\infty} v_n = v.$$

□

Theorem 5.1.3. *Let H be a Hilbert space over the field $\mathbb{R}$. Let $K : H \to H$ be a compact linear operator. Then, Im$(I - K) = \ker(I - K^*)^\perp$.*

Proof. Assume that H is a Hilbert space over $\mathbb{R}$. Assume that $K : H \to H$ is a compact linear operator and that $I : H \to H$ is the identity operator. Since $\ker(I - K^*)$ is finite dimensional, it suffices to prove the following:

To show: (a) $\text{Im}(I - K)^\perp = \ker(I - K^*)$.

(a) Assume that $h \in \ker(I - K^*)$. Then, this holds if and only if $(I - K^*)h = 0$. Using the inner product on H, this is true if and only if for all $y \in H$,

$$\langle y, (I - K^*)h \rangle = 0.$$

Using the definition of the adjoint, this subsequently holds if and only if

$$\langle (I - K)y, h \rangle = 0.$$

Since $(I-K)y \in \text{Im}(I-K)$ for all $y \in H$, $h \in \text{Im}(I-K)^\perp$. Hence, $\text{Im}(I-K)^\perp = \ker(I-K^*)$. By taking the orthogonal complement of both sides, we obtain the desired result.

□

Theorem 5.1.4. *Let H be a Hilbert space over the field $\mathbb{R}$. Let $K : H \to H$ be a compact linear operator. Then, $\text{Im}(I-K) = H$ if and only if $\ker(I-K) = \{0\}$.*

Proof. Assume that H is a Hilbert space over $\mathbb{R}$. Assume that $K : H \to H$ is a compact linear operator and that $I : H \to H$ is the identity operator.

To show: (a) If $\ker(I-K) = \{0\}$, then $\text{Im}(I-K) = H$.

(b) If $\text{Im}(I-K) = H$, then $\ker(I-K) = \{0\}$.

(a) Suppose for the sake of contradiction that $\ker(I-K) = \{0\}$, but there exists $h \in H$ such that $h \notin \text{Im}(I-K)$. The idea of this argument is to use induction and Pythagoras' theorem. Let $H_1 = (I-K)(H)$. Note that H_1 is a closed subspace of H. Since $I-K$ is one-to-one, find that $H_2 = (I-K)(H_1) \subset H_1$ is also a closed subspace of H. Hence, by induction, we can construct a sequence of closed subspaces $H_n \subset H_{n-1} \subset \cdots \subset H_1$ such that $H_n = (I-K)^n(H)$.

Now, for each $n \in \mathbb{Z}_{>0}$, we can pick an element $e_m \in H_n \cap H_{n+1}^\perp$ such that $\|e_m\| = 1$. Note that when $m < n$,

$$Ke_m - Ke_n = -(e_m - Ke_m) + (e_n - Ke_n) + (e_m - e_n) = [-(e_m - Ke_m) + (e_n - Ke_n) - e_n] + e_m.$$

Specifically, the element $[-(e_m - Ke_m) + (e_n - Ke_n) - e_n] \in H_{m+1}^\perp$ by definition. This suggests that we can use Pythagoras' theorem to deduce that

$$\begin{aligned} \|Ke_m - Ke_n\|^2 &= \|[-(e_m - Ke_m) + (e_n - Ke_n) - e_n]\|^2 + \|e_m\|^2 \\ &\geq 1. \end{aligned}$$

So, $\|Ke_m - Ke_n\| \geq 1$. As a result, the sequence $\{Ke_n\}$ does not have a convergent subsequence. This contradicts the fact that K is a compact operator. Therefore, $\text{Im}(I-K) = H$.

(b) Assume that $\text{Im}(I - K) = H$. Then, we observe that $\ker(I - K^*) = \text{im}(I - K)^\perp = H^\perp = \{0\}$. Since K is a compact operator, K^* must also be a compact operator. So, $\text{Im}(I - K^*) = H$ by part (a). Applying duality again, we find that $\text{Im}(I - K^*) = \ker(I - K)^\perp = H$. Therefore, $\ker(I - K) = \{0\}$ as required. □

Theorem 5.1.5. *Let H be a Hilbert space over the field $\mathbb{R}$. Let $K : H \to H$ be a compact linear operator. Then,* $\dim \ker(I - K) = \dim \ker(I - K^*)$.

Proof. Assume that H is a Hilbert space over $\mathbb{R}$. Assume that $K : H \to H$ is a compact linear operator.

To show: (a) $\dim \ker(I - K) \geq \dim \text{Im}(I - K)^\perp$.

(a) Suppose for the sake of contradiction that $\dim \ker(I - K) < \dim \text{Im}(I - K)^\perp$. Then, there exists a linear map $A : \ker(I - K) \to \text{Im}(I - K)^\perp$, which is injective, but not surjective. We can extend A to another linear map $B : H \to \text{Im}(I - K)^\perp$, where $Bu = 0$ whenever $u \in \ker(I - K)^\perp$ and $Bu = Au$ otherwise. Note that the range of B is finite dimensional. So, B must be a compact operator and by extension, $K + B$ as well.

We will now demonstrate that $\ker(I - (K + B)) = \{0\}$. Assume that $u \in H$. Then, we can write $u = u_1 + u_2$, where $u_1 \in \ker(I - K)$ and $u_2 \in \ker(I - K)^\perp$. As a result,

$$(I - K - B)(u_1 + u_2) = (I - K)u_2 - Bu_1 \in \text{Im}(I - K) \oplus \text{Im}(I - K)^\perp.$$

Importantly, $(I - K)u_2$ is orthogonal to Bu_1, which means that $(I - K)u_2 - Bu_1 = 0$ if and only if $(I - K)u_2 = 0$ and $Bu_1 = 0$. If $(I - K)u_2 = 0$, then $u_2 \in \ker(I - K)$. But, $u_2 \in \ker(I - K)^\perp$. Therefore, $u_1 = 0$. Furthermore, due to the definition of B and the fact that $u_1 \in \ker(I - K)$, $u_1 = 0$. Hence, $\ker(I - (K + B)) = \{0\}$.

Since $K + B$ is a compact operator, $\ker(I - (K + B)) = \{0\}$ if and only if $\text{Im}(I - (K + B)) = H$. We can construct an element $v \in \text{Im}(I - K)^\perp$ such that $v \notin \text{Im}(B)$. Hence, the equation

$$u - Ku - Au = v$$

cannot be solved, contravening the finding that $\text{Im}(I - (K + B)) = H$. Hence, $\dim \ker(I - K) \geq \dim \text{Im}(I - K)^\perp$.

To show: (b) $\dim \ker(I - K) = \dim \ker(I - K^*)$.

(b) We will use part (a) to deal with this. We already know that for the compact operator K^*,

$$\ker(I - K) = \operatorname{Im}(I - K^*)^{\perp}.$$

Hence, from part (a) we obtain

$$\dim \ker(I - K^*) \geq \dim \operatorname{Im}(I - K^*)^{\perp} = \ker(I - K).$$

Analogously,

$$\ker(I - K^*) = \operatorname{Im}(I - K)^{\perp}.$$

and

$$\dim \ker(I - K) \geq \dim \operatorname{Im}(I - K)^{\perp} = \ker(I - K^*).$$

These two inequalities together show that
$\dim \ker(I - K) = \dim \ker(I - K^*)$.

□

Here is a short summary of Fredholm's theorem in its entirety:

1. $\ker(I - K)$ is finite dimensional.
2. $\operatorname{Im}(I - K)$ is a closed subset of H.
3. $\operatorname{Im}(I - K) = \ker(I - K^*)^{\perp}$.
4. $\ker(I - K) = \{0\}$ if and only if $\operatorname{Im}(I - K) = H$.
5. $\dim \ker(I - K) = \dim \ker(I - K^*)$.

What exactly does this mean for operators of the form $I - K$, where K is a compact operator on a Hilbert space H? It provides us with information about the existence and uniqueness of solutions to the linear equation $u - Ku = f$. There are two possible cases:

Case 1: $\ker(I - K) = \{0\}$

If $\ker(I - K) = \{0\}$, then $\operatorname{Im}(I - K) = H$. This reveals that the operator $I - K$ is bijective. Hence, the equation $(I - K)u = f$ has exactly one solution.

Case 2: $\ker(I - K) \neq \{0\}$

In this case, the equation $(I - K)u = 0$ has a non-trivial solution, since the kernel is non-trivial. In this scenario, the equation $(I - K)u = f$ has solutions if and only if $f \in \ker(I - K^*)^\perp$ as a consequence of 5.1.3.

5.2 Spectrum

We will now delve into the concept of a spectrum, which is pertinent to linear algebra. We are interested in the spectrum of a bounded linear operator $\Lambda : H \to H$, where H is a Hilbert space over $\mathbb{R}$.

Definition 5.2.1. Let H be a Hilbert space over $\mathbb{R}$ and $\Lambda : H \to H$ be a bounded linear operator. Then, the **resolvent set** of Λ is the set of numbers $\eta \in \mathbb{R}$ such that the operator $\eta I - \Lambda$ is both injective and surjective. The resolvent set of Λ is denoted by $\rho(\Lambda)$.

Here is one preliminary thought about the resolvent set. If $\eta \in \rho(\Lambda)$, then the operator $\eta I - \Lambda$ is a bounded, surjective operator on H. By the open mapping theorem (4.1.8), $\eta I - \Lambda$ must be an open map. But, from the definition of the resolvent set, $\eta I - \Lambda$ is also bijective. Since $\eta I - \Lambda$ is bijective and open, it must be a homeomorphism. Consequently, the inverse operator $(\eta I - \Lambda)^{-1}$ is continuous itself (and thus, a bounded linear operator on H).

Definition 5.2.2. Let H be a Hilbert space over $\mathbb{R}$ and $\Lambda : H \to H$ be a bounded linear operator. Then, the **spectrum** of Λ, denoted by $\sigma(\Lambda)$, is the complement of the resolvent set of Λ.

$$\sigma(\Lambda) = \mathbb{R} \backslash \rho(\Lambda).$$

Definition 5.2.3. Let H be a Hilbert space over $\mathbb{R}$ and $\Lambda : H \to H$ be a bounded linear operator. Then, the **point spectrum** of Λ, denoted by $\sigma_p(\Lambda)$, is the set of real numbers $\eta \in \mathbb{R}$ such that $(\eta I - \Lambda)$ is not injective. Alternatively, $\eta \in \sigma_p(\Lambda)$ if there exists a non-zero vector $w \in H$ such that $\Lambda w = \eta w$.

In line with linear algebra, η is called an eigenvalue of Λ and w is the associated eigenvector.

Definition 5.2.4. Let H be a Hilbert space over $\mathbb{R}$ and $\Lambda : H \to H$ be a bounded linear operator. Then, the **essential spectrum** of Λ, denoted by $\sigma_e(\Lambda)$, is defined as the complement

$$\sigma_e(\Lambda) = \sigma(\Lambda) \backslash \sigma_p(\Lambda).$$

Alternatively, it is the set of all real numbers $\delta \in \mathbb{R}$ such that $(\delta I - \Lambda)$ is injective, but not surjective.

Tying in with the previous work on compact operators, what does the spectrum of a compact operator look like? The following theorem reveals the answer.

Theorem 5.2.1. *Let H be an infinite dimensional Hilbert space. Let $K : H \to H$ be a compact linear operator. Then, $0 \in \sigma(K)$ and $\sigma(K) = \sigma_p(K) \cup \{0\}$. Moreover, either $\sigma_p(K)$ is finite or is equal to the set $\{\lambda_k \mid k \in \mathbb{Z}_{>0}\}$ where $\lim_{k\to\infty} \lambda_k = 0$.*

Proof. Assume that H is an infinite dimensional Hilbert space. Assume that $K : H \to H$ is a compact linear operator.

To show: (a) $0 \in \sigma(K)$.

(b) $\sigma(K) = \sigma_p(K) \cup \{0\}$.

(c) Either $\sigma_p(K)$ is finite or is equal to the set $\{\lambda_k \mid k \in \mathbb{Z}_{>0}\}$ where $\lim_{k\to\infty} \lambda_k = 0$.

(a) Suppose for the sake of contradiction that $0 \notin \sigma(K)$. Then, from the definition of $\sigma(K)$, 0 must be an element of the resolvent set $\rho(K)$. By definition of the resolvent set, the operator $-K$ must be bijective. Hence, K is bijective, with a continuous inverse. As a result of composition, $I = K \circ K^{-1}$. Since K is compact and K^{-1} is continuous, I must be a compact operator. This contradicts the fact that H is an infinite dimensional space, in tandem with the fact that the closed unit ball in H is not compact. So, $0 \in \sigma(K)$.

(b) Suppose for the sake of contradiction that $\sigma(K) \neq \sigma_p(K) \cup \{0\}$. Then, there exists $\lambda \in \sigma(K)$ such that $\lambda \neq 0$ and $\lambda \notin \sigma_p(K)$. Since $\lambda \notin \sigma_p(K)$, $\lambda I - K$ must be injective and $\ker(\lambda I - K) = \{0\}$. By Fredholm's theorem, this holds if and only if $\mathrm{Im}(\lambda I - K) = H$. Since $\lambda \in \rho(K)$, the inverse $(\lambda I - K)^{-1}$ is continuous and thus, bounded. This contradicts the fact that

$\lambda \in \sigma(K)$. Hence, $\sigma(K) = \sigma_p(K) \cup \{0\}$.

(c) Assume that $\{\lambda_n\}$ is a sequence of eigenvalues of K, with $\lambda_n \to \lambda$. We will show that $\lambda = 0$. Suppose for the sake of contradiction that $\lambda \neq 0$. Since $\lambda_n \in \sigma_p(K)$, there exists an eigenvector $w_n \in H$ such that $Kw_n = \lambda_n w_n$. Define

$$H_n = span\{w_1, \dots, w_n\}.$$

Note that the set of eigenvectors $\{w_1, \dots, w_n\}$ is linearly independent. So, $H_n \subset H_{n+1}$ for all $n \in \mathbb{Z}_{>0}$. We also observe that $(K - \lambda_n I)H_n \subseteq H_{n-1}$ due to the definition of H_n. Hence, for all $n \in \mathbb{Z}_{>0}$, we choose $e_n \in H_n \cap H_{n-1}^{\perp}$ with $\|e_n\|$. If we fix $m < n$, then we note the following:

1. $Ke_n - \lambda_n e_n \in H_{n-1}$ because $e_n \in H_n$ and $(K - \lambda_n I)H_n \subseteq H_{n-1}$.
2. $Ke_m - \lambda_m e_m \in H_{m-1} \subseteq H_{n-1}$ because $e_m \in H_m$ and $(K - \lambda_m I)H_m \subseteq H_{m-1}$.
3. $e_m \in H_m \subseteq H_{n-1}$.
4. $e_n \in H_{n-1}^{\perp}$.

Therefore,

$$\begin{aligned}
\|Ke_m - Ke_n\|^2 &= \|[(Ke_m - \lambda e_m) - (Ke_n - \lambda e_n) + \lambda_m e_m - \lambda_n e_n]\|^2 \\
&= \|(Ke_m - \lambda e_m) - (Ke_n - \lambda e_n) + \lambda_m e_m\|^2 + \|\lambda_n e_n\|^2. \\
&= \|(Ke_m - \lambda e_m) - (Ke_n - \lambda e_n) + \lambda_m e_m\|^2 + |\lambda_n|^2 \\
&\geq |\lambda_n|^2.
\end{aligned}$$

So, $\|Ke_m - Ke_n\| \geq |\lambda|$. This means that the sequence $\{Ke_n\}$, where $n \in \mathbb{Z}_{>0}$, cannot have any convergent subsequence. However, this contradicts the fact that K is a compact operator. Therefore, $\lambda = 0$.

□

5.3 Hilbert-Schmidt

In this section, we will study symmetric operators on a Hilbert space H over $\mathbb{R}$. We know from linear algebra that a symmetric matrix is diagonalisable. It turns out that a similar result holds for compact symmetric operators on a Hilbert space. This result is the **Hilbert-Schmidt theorem**, which is central to this section.

As usual, we require a few preliminary definitions and results.

Definition 5.3.1. Let H be a Hilbert space over the field $\mathbb{R}$. Let $\Lambda : H \to H$ be a linear operator. Then, Λ is **symmetric** if $\Lambda = \Lambda^*$. If H is a HIlbert space over $\mathbb{C}$, then Λ is called **self-adjoint**.

The following theorem was taken from Kreyszig [EK78]. This theorem is very important for the result we will eventually prove pertaining to the spectrum of a bounded linear symmetric operator. Somehow, this was either assumed knowledge or ignored by Bressan.

Theorem 5.3.1. *Let H be a Hilbert space over $\mathbb{C}$ and $\Lambda : H \to H$ be a bounded self-adjoint linear operator. Then, $\lambda \in \rho(\Lambda)$ if and only if there exists $c \in \mathbb{R}_{>0}$ such that for all $x \in H$,*

$$\|\Lambda x - \lambda x\| \geq c\|x\|.$$

Proof. Assume that H is a Hilbert space over $\mathbb{C}$. Assume that $\Lambda : H \to H$ be a bounded, linear, self-adjoint operator. Assume that $\lambda \in \mathbb{C}$.

To show: (a) If $\lambda \in \rho(\Lambda)$, then there exists $c \in \mathbb{R}_{>0}$ such that for all $x \in H$, $\|\Lambda x - \lambda x\| \geq c\|x\|$.

(b) If there exists $c \in \mathbb{R}_{>0}$ such that for all $x \in H$, $\|\Lambda x - \lambda x\| \geq c\|x\|$, then $\lambda \in \rho(\Lambda)$.

(a) Assume that $\lambda \in \rho(\Lambda)$. By the open mapping theorem, the operator $\Lambda - \lambda I$ must have a bounded inverse. So, there exists a $k \in \mathbb{R}_{>0}$ such that

$$\|(\Lambda - \lambda I)^{-1}\| \leq k.$$

Hence, for all $x \in H$, we have

$$\begin{aligned}\|x\| &= \|(\Lambda - \lambda I)^{-1}(\Lambda - \lambda I)x\| \\ &\leq \|(\Lambda - \lambda I)^{-1}\|\|(\Lambda - \lambda I)x\| \\ &\leq k\|(\Lambda - \lambda I)x\|.\end{aligned}$$

Subsequently, we have

$$\frac{1}{k}\|x\| \leq \|(\Lambda - \lambda I)x\|.$$

Taking the scalar $1/k$ gives the desired result.

(b) Assume that there exists $c \in \mathbb{R}_{>0}$ such that for all $x \in H$,

$$\|\Lambda x - \lambda x\| \geq c\|x\|.$$

To show: (ba) $\Lambda - \lambda I : H \to (\Lambda - \lambda I)(H)$ is bijective.

(bb) The image $(\Lambda - \lambda I)(H)$ is dense in H.

(bc) The image $(\Lambda - \lambda I)(H)$ is a closed subset of H.

(ba) It suffices to show that $\Lambda - \lambda I$ is injective. Assume that $\Lambda x_1 - \lambda x_1 = \Lambda x_2 - \lambda x_2$ for some $x_1, x_2 \in H$. Then,

$$\begin{aligned} 0 &= \|(\Lambda - \lambda I)x_1 - (\Lambda - \lambda I)x_2\| \\ &= \|(\Lambda - \lambda I)(x_1 - x_2)\| \\ &\geq c\|x_1 - x_2\|. \end{aligned}$$

for some $c \in \mathbb{R}_{>0}$ as enforced by our assumption. Therefore, $\|x_1 - x_2\| = 0$ and so, $x_1 = x_2$ as a result. Hence, $\Lambda - \lambda I$ is injective.

(bb) Consider the set $\overline{(\Lambda - \lambda I)(H)}^{\perp}$. Since $(\Lambda - \lambda I)(H)$ is a subset of H, $\overline{(\Lambda - \lambda I)(H)}^{\perp}$ is a closed subset of H. Hence, from orthogonal projection,

$$H = \overline{(\Lambda - \lambda I)(H)} \oplus \overline{(\Lambda - \lambda I)(H)}^{\perp}.$$

To show: (bba) $\overline{(\Lambda - \lambda I)(H)}^{\perp} = \{0\}$.

(bba) Assume that $x_0 \in \overline{(\Lambda - \lambda I)(H)}^{\perp}$. Then, for all $x \in H$,

$$\langle x_0, \Lambda x - \lambda x\rangle = \langle x_0, \Lambda x\rangle - \langle x_0, \lambda x\rangle = 0.$$

We now use the fact that Λ is self-adjoint in order to deduce that

$$\langle x, \Lambda x_0\rangle = \langle \Lambda x, x_0\rangle = \langle x, \overline{\lambda} x_0\rangle.$$

Therefore, $\Lambda x_0 = \overline{\lambda} x_0$. We will now show that $x_0 = 0$. Suppose for the sake of contradiction that $x_0 \neq 0$. Then. the previous equation reveals that $\overline{\lambda}$ is an eigenvalue of Λ. Since Λ is self-adjoint, its eigenvalues must be real. So, $\overline{\lambda} = \lambda$. So,

$$(\Lambda - \lambda I)x_0 = \Lambda x_0 - \lambda x_0 = 0.$$

However, this would contradict our original assumption because

$$0 = \|(\Lambda - \lambda I)x_0\| \geq c\|x_0\|$$

and $c > 0$. Hence, $x_0 = 0$ and consequently, $\overline{(\Lambda - \lambda I)(H)}^{\perp} = \{0\}$.

(bb) Since $\overline{(\Lambda - \lambda I)(H)}^{\perp} = \{0\}$ and

$$H = \overline{(\Lambda - \lambda I)(H)} \oplus \overline{(\Lambda - \lambda I)(H)}^{\perp},$$

$\overline{(\Lambda - \lambda I)(H)} = H$. This shows that the image $(\Lambda - \lambda I)(H)$ is dense in H.

(bc) Assume that $y \in \overline{(\Lambda - \lambda I)(H)}$. We will show that $y \in (\Lambda - \lambda I)(H)$. Since $y \in \overline{(\Lambda - \lambda I)(H)}$, there exists a sequence $\{y_n\}$ such that $y_n \in (\Lambda - \lambda I)(H)$ and $y_n \to y$. Because, $y_n \in (\Lambda - \lambda I)(H)$, there exists $x_n \in H$ such that $\Lambda x_n - \lambda x_n = y_n$. We know from part (a) that

$$\|x_n - x_m\| \leq \frac{1}{c}\|(\Lambda - \lambda I)(x_n - x_m) = \frac{1}{c}\|y_n - y_m\|.$$

This proves that the sequence $\{x_n\}$ is Cauchy because $\{y_n\}$ converges. Since H is complete, $\{x_n\}$ must converge to say $x \in H$. Now observe that $(\Lambda - \lambda I)x \in (\Lambda - \lambda I)(H)$ and due to the uniqueness of limits, we must have $(\Lambda - \lambda I)x = y$. Hence, $y \in (\Lambda - \lambda I)(H)$. So, the image $(\Lambda - \lambda I)(H)$ is closed.

(b) Combining parts (bb) and (bc), we deduce that $(\Lambda - \lambda I)(H) = H$. Therefore, $(\Lambda - \lambda I)(H)$ is bijective. From the definition of a resolvent set, we finally find that $\lambda \in \rho(\Lambda)$.

□

For a continuous, linear, symmetric operator, there are well-defined bounds on its spectrum. We will see this in the following result.

Theorem 5.3.2. *Let H be a Hilbert space over $\mathbb{R}$ and $\Lambda : H \to H$ be a bounded, linear and symmetric operator. Define*

$$m = \inf_{u \in H, \|u\|=1} \langle \Lambda u, u \rangle \text{ and } M = \sup_{u \in H, \|u\|=1} \langle \Lambda u, u \rangle.$$

Then, the spectrum $\sigma(\Lambda) \subseteq [m, M]$*,* $m, M \in \sigma(\Lambda)$ *and* $\|\Lambda\| = \max(-m, M)$*.*

Proof. Assume that H is a HIlbert space over $\mathbb{R}$. Assume that $\Lambda : H \to H$ is a bounded, linear, symmetric operator over H. Assume that $m, M \in \mathbb{R}$ are defined as above.

To show: (a) $\sigma(\Lambda) \subseteq [m, M]$.

(b) $m, M \in \sigma(\Lambda)$.

(c) $\|\Lambda\| = \max(-m, M)$.

(a) It is more tractable for us to work with the resolvent set $\rho(\Lambda)$.

To show: (aa) $(-\infty, m) \cup (M, \infty) \subseteq \rho(\Lambda)$.

(aa) Assume that $\eta \in (M, \infty)$. Define the bilinear functional $B : H \times H \to H$ as

$$B[x, y] = \langle \eta x - \Lambda x, y \rangle.$$

First note that $B[x, y]$ is continuous. We would like to apply the Lax-Milgram theorem to B. This requires B to be positive definite. For this, we observe that for all $u \in H$,

$$B[u, u] = \langle \eta u - \Lambda u, u \rangle = \langle \eta u, u \rangle - \langle \Lambda u, u \rangle \geq (\eta - M)\|u\|^2$$

due to the definition of M. Therefore, B is positive definite. So, we can apply the Lax-Milgram theorem to deduce that for all $h \in H$, there exists a unique $x \in H$ such that for all $y \in H$,

$$B[x, y] = \langle \eta x - \Lambda x, y \rangle = \langle h, y \rangle.$$

This reduces to the statement that for all $h \in H$, there exists a unique $x \in H$ such that

$$(\eta I - \Lambda)x = h.$$

This means that the operator $(\eta I - \Lambda)$ is surjective. Furthermore, due to the uniqueness of $x \in H$, $(\eta I - \Lambda)$ is also injective. Therefore, $\eta \in \rho(\Lambda)$ by definition of the resolvent set. The case where $\eta \in (-\infty, m)$ is similar and uses $-\Lambda$ rather than Λ. So, $(-\infty, m) \cup (M, \infty) \subseteq \rho(\Lambda)$.

(a) Taking complements yields $\sigma(\Lambda) \subseteq [m, M]$.

We will prove part (c) before part (b).

(c) We will assume that $|m| \leq M$. The case where $M < -m$ can be handled by very similar arguments - change Λ to $-\Lambda$ and recycle the arguments which follow.

For all $u, v \in H$, we have

$$\begin{aligned}
4\langle \Lambda u, v\rangle &= \langle \Lambda(u+v), u+v\rangle - \langle \Lambda(u-v), u-v\rangle \\
&\leq M(\|u+v\|^2 + \|u-v\|^2) \\
&= 2M(\|u\|^2 + \|v\|^2).
\end{aligned}$$

The first equality works since $\langle \Lambda u, v\rangle = \langle u, \Lambda v\rangle = \langle \Lambda v, u\rangle$. Now assume that $\Lambda u \neq 0$. Then, if we set

$$v = \frac{\|u\|}{\|\Lambda u\|}\Lambda u,$$

we find that

$$2\|u\|\|\Lambda u\| \leq M(\|u\|^2 + \|u\|^2) = 2M\|u\|^2.$$

and

$$\|\Lambda u\| \leq M\|u\|$$

for all $u \in H$. Note that this inequality also holds when $\Lambda u = 0$. By taking the supremum of both sides and letting $\|u\| = 1$, we deduce that $\|\Lambda\| \leq M$. Next, we will derive another inequality as follows:

$$\begin{aligned}
\|\Lambda\| &= \sup_{\|u\|=1} \|\Lambda u\| \\
&= \sup_{\|u\|=1} \|\Lambda u\|\|u\| \\
&\geq \sup_{\|u\|=1} \langle \Lambda u, u\rangle \quad \text{(Cauchy-Schwarz)} \\
&= M.
\end{aligned}$$

Hence, $\|\Lambda\| \geq M$ and as a result, $\|\Lambda\| = M$. As stated before, the other case where $M < -m$ can be dealt with by similar arguments, which lead to the conclusion that $\|\Lambda\| = m$. Therefore, $\|\Lambda\| = \max(-m, M)$.

(b) Once again, we will only establish that $M \in \sigma(\Lambda)$ since the argument for m is very similar. Since $\sup_{\|u\|=1}\langle \Lambda u, u\rangle = M$ by definition, there exists a sequence $\{u_n\}$ such that $\langle \Lambda u_n, u_n\rangle \to M$ as $n \to \infty$. Furthermore, $\|u_n\| = 1$ for all $n \in \mathbb{Z}_{>0}$. Then, we note that

$$\begin{aligned}
\|\Lambda u_n - Mu_n\|^2 &= \|\Lambda u_n\|^2 - 2M\langle \Lambda u_n, u_n\rangle + M^2\|u_n\|^2 \\
&= \|\Lambda u_n\|^2 - 2M\langle \Lambda u_n, u_n\rangle + M^2 \\
&\leq M^2 - 2M\langle \Lambda u_n, u_n\rangle + M^2 \quad \text{from part (c)} \\
&= 2M^2 - 2M\langle \Lambda u_n, u_n\rangle \\
&\to 0
\end{aligned}$$

as $n \to \infty$. Thus, there is no positive real number c such that

$$\|(\Lambda - MI)u_n\| = \|\Lambda u_n - Mu_n\| \geq c\|x_n\| = c.$$

for all $n \in \mathbb{Z}_{>0}$. By the contrapositive of 5.3.1, $M \notin \rho(\Lambda)$. Hence, $M \in \sigma(\Lambda)$ as required.

□

We have now arrived at the most important result of the section.

Theorem 5.3.3 (Hilbert-Schmidt). *Let H be an infinite dimensional, separable Hilbert space over $\mathbb{R}$. Let $\Lambda : H \to H$ be a compact, symmetric, linear operator. Then, the eigenvectors of Λ form a countable orthonormal basis for H.*

Proof. Assume that H is an infinite dimensional separable Hilbert space over $\mathbb{R}$. Assume that $\Lambda : H \to H$ is a compact symmetric linear operator on H. Let $\eta_0 = 0$ and $\{\eta_1, \eta_2, \dots\}$ be the set of non-zero eigenvalues of Λ. Additionally, we let $H_0 = \ker(\Lambda)$ and $H_i = \ker(\Lambda - \eta_i I)$ for all $i \in \mathbb{Z}_{>0}$. Note that $0 \leq \dim H_0 \leq \infty$ and $0 < \dim H_i < \infty$ for all $i \in \mathbb{Z}_{>0}$.

To show: (a) If $m \neq n$, then H_m and H_n are orthogonal.

(a) Assume that $m \neq n$. Assume that $k_m \in H_m$ and $k_n \in H_n$. Then, from the definition, $\Lambda k_m = \eta_m k_m$ and $\Lambda k_n = \eta_n k_n$. Now, we argue as follows:

$$\begin{aligned}\eta_m\langle k_m, k_n\rangle &= \langle \eta_m k_m, k_n\rangle \\ &= \langle \Lambda k_m, k_n\rangle \\ &= \langle k_m, \Lambda k_n\rangle \\ &= \langle k_m, \eta_n k_n\rangle \\ &= \eta_n\langle k_m, k_n\rangle.\end{aligned}$$

Since $\eta_m \neq \eta_n$, it must be the case that $\langle k_m, k_n\rangle = 0$. Hence, H_m and H_n are orthogonal to each other.

Now we consider the set of linear combinations below:

$$\tilde{H} = \{\sum_{k=1}^{N} \alpha_k u_k \mid N \in \mathbb{Z}_{>0}, u_k \in H_k, \alpha_k \in \mathbb{R}\}.$$

To show: (b) $\tilde{H}^{\perp} \subseteq \ker(\Lambda) = H_0$.

(b) Assume that $u \in \tilde{H}^{\perp}$. Assume that $v \in \tilde{H}$. Then, $\Lambda v \in \Lambda(\tilde{H}) \subseteq \tilde{H}$ and $\langle \Lambda u, v\rangle = \langle u, \Lambda v\rangle = 0$. This reveals that the image $\Lambda(\tilde{H}^{\perp}) \subseteq \tilde{H}^{\perp}$. Define $\tilde{\Lambda}$ to be the restriction of Λ to the subspace $\tilde{H}^{\perp}$. Then, $\tilde{\Lambda}$ is still a compact, symmetric operator. From the previous theorem, we have

$$\|\tilde{\Lambda}\| = \sup_{u\in\tilde{H}^{\perp},\ \|u\|=1} |\langle \tilde{\Lambda} u, u\rangle| = M.$$

Suppose for the sake of contradiction that $M \neq 0$. Then, either $M \in \sigma(\tilde{\Lambda})$ or $-M \in \sigma(\tilde{\Lambda})$. In either case, we note that $\sigma(\tilde{\Lambda}) = \sigma_p(\tilde{\Lambda}) \cup \{0\}$ because $\tilde{\Lambda}$ is a compact operator. Hence, either $M \in \sigma_p(\tilde{\Lambda})$ or $-M \in \sigma_p(\tilde{\Lambda})$. As a result, there exists an eigenvector $w \in \tilde{H}^{\perp}$ such that either

$$\tilde{\Lambda} w = \Lambda w = Mw \text{ or } \tilde{\Lambda} w = \Lambda w = -Mw.$$

This contradicts the fact that all the eigenvectors of K are contained in the union of subspaces H_k. Therefore, $M = 0$ and as a result, $\|\tilde{\Lambda}\| = 0$. Translating back to Λ, we find that $\tilde{H}^{\perp} \subseteq \ker(\Lambda)$.

Since each of the subspaces H_m and H_n are orthogonal to each other whenever $m \neq n$, it follows that $\tilde{H}^{\perp} \subseteq \ker(\Lambda)^{\perp} = H_0^{\perp}$. Combining this result with part (b), we deduce that $\tilde{H}^{\perp} \subseteq H_0^{\perp} \cap H_0 = \{0\}$. So, $\tilde{H} = H$, revealing that $\tilde{H}$ is a dense subset of H (by taking the closure).

For each $k \in \mathbb{Z}_{>0}$, the finite-dimensional subspace H_k has an orthonormal basis $B_k = \{e_{k,1}, e_{k,2}, \ldots, e_{k,N(k)}\}$. Furthermore, since H is separable, the closed subspace $H_0 = \ker(\Lambda)$ has a countable orthonormal basis $B_0 = \{e_{0,1}, e_{0,2}, \ldots\}$. Hence, the union

$$B = \bigcup_{k \in \mathbb{Z}_{>0}} B_k$$

forms an orthonormal basis of H.

□

An important problem is when given a countable set $S = \{u_1, u_2, \ldots\}$ in a Banach space X over $\mathbb{R}$, is it possible to decide whether $span(S)$ is a dense subset of X? There are two theorems that answer this in the affirmative.

1. If X is a separable Hilbert space and there exists a compact, symmetric operator $\Lambda : X \to X$ on X such that $span(S) \subseteq \ker(\Lambda)$ and $span(S)$ contains all the eigenvectors of Λ, then the previous theorem yields $\overline{span(S)} = X$. This can be deduced from the proof of part (b) in 5.3.3.

2. If $X = Cts(E, \mathbb{R})$, where E is a compact metric space, then provided that $span(S)$ is an algebra which separates points and contains the constant functions, $\overline{span(S)} = X$. This is an important result known as the **Stone-Weierstrass Theorem**.

We provide a proof of the Stone-Weierstrass theorem below. Note that the version we give is more general because we only use a locally compact Hausdorff space, rather than a compact space. Also, we do not assume that the subalgebra is unital.

Theorem 5.3.4. *Let X be a locally compact Hausdorff space. Let $Cts_0(X, \mathbb{R}) \subseteq Cts(X, \mathbb{R})$ be the subspace of functions which vanish at infinity. That is,*

$$Cts_0(X, \mathbb{R}) = \{f \in Cts(X, \mathbb{R}) \mid \forall \epsilon \in \mathbb{R}_{\geq 0}, \exists \text{ compact } K \subseteq X \text{ such that } \forall x \notin K, |f(x)| < \epsilon\}.$$

Suppose that $A \subseteq Cts_0(X, \mathbb{R})$ is a non-unital subalgebra of $Cts_0(X, \mathbb{R})$ which separates points and has the property that for all $x \in X$, there exists $f \in A$ such that $f(x) \neq 0$. Then, $\overline{A} = Cts_0(X, \mathbb{R})$.

Proof. In a similar vein to the proof of the usual Stone-Weierstrass theorem, it suffices to prove the statement for the case where the subalgebra A is closed. So, assume that X is locally compact Hausdorff and that A is a closed, non-unital subalgebra of $Cts_0(X, \mathbb{R})$ which separates points. Suppose further that for all $x \in X$, there exists $f \in A$ such that $f(x) \neq 0$. We will break the proof down into multiple steps.

To show: (a) If $x, y \in X$ such that $x \neq y$, then there exists $f \in A$ such that $f(x) \neq f(y)$, $f(x) \neq 0$ and $f(y) \neq 0$.

(a) Assume that $x, y \in X$ such that $x \neq y$. Suppose for the sake of contradiction that there does not exist $f \in A$ such that $f(x) \neq f(y)$, $f(x) \neq 0$ and $f(y) \neq 0$. Then, pick a function $g \in A$ such that $g(x) \neq g(y)$ (we can do this since A separates points). Suppose that $g(x) \neq 0$. Due to our assumption, $g(y) = 0$. Using the second property of A, there exists $h \in A$ such that $h(y) \neq 0$. Suppose for the sake of contradiction that $h(x) \neq h(y)$. Again, we apply our assumption in order to deduce that $h(x) = 0$. Now, we consider the function $f = g + \lambda h$ for some $\lambda \in \mathbb{R}\backslash\{0\}$ which satisfies $g(x) \neq \lambda h(y)$. Since A is a subalgebra, $g + \lambda h \in A$ because $g, h \in A$.

To show: (aa) $f(x) \neq f(y)$.

(ab) $f(x) \neq 0$.

(ac) $f(y) \neq 0$.

(aa) We compute directly from the definition of f that $f(x) = g(x) + \lambda h(x) = g(x)$ and $f(y) = g(y) + \lambda h(y) = \lambda h(y)$. Since $g(x) \neq \lambda h(y)$, we deduce that $f(x) \neq f(y)$.

(ab) From part aa, $f(x) = g(x) \neq 0$ by assumption.

(ac) Again from part aa, $f(x) = \lambda h(y) \neq 0$ since $\lambda, h(y) \neq 0$.

(a) So, we have found a function $f \in A$, which satisfies $f(x) \neq f(y)$, $f(x) \neq 0$ and $f(y) \neq 0$. However, this contradicts our original assumption. As a result of this $h(x) = h(y)$. Next, we consider the function $f_2 = g + h$. Then, $f_2(x) = g(x) + h(x)$ and $f_2(y) = g(y) + h(y) = h(x)$. Since $g(x) \neq 0$, $f_2(x) \neq f_2(y)$. Additionally, since $h(x) = h(y) \neq 0$, $f_2(x) \neq 0$ and $f_2(y) \neq 0$. Once again, this derives a contradiction to our original assumption. For the case where $g(y) \neq 0$, a similar argument to the above once again establishes a contradiction. Hence, there exists $f \in A$ such that $f(x) \neq f(y)$, $f(x) \neq 0$ and $f(y) \neq 0$.

Our next step is to prove the following:

To show: (b) If $x, y \in X$ such that $x \neq y$ and $a, b \in \mathbb{R}$, then there exists $f \in A$ such that $f(x) = a$ and $f(y) = b$.

(b) Assume that $x, y \in X$ such that $x \neq y$. Assume that $a, b \in \mathbb{R}$. Then, from part a, there exists $f \in A$ such that $f(x) \neq f(y)$, $f(x) \neq 0$ and $f(y) \neq 0$. We claim that a function of the form $p = \alpha f + \beta f^2$ does the trick, where $\alpha, \beta \in \mathbb{R}$. Note that since $f \in A$ and A is a subalgebra, $p \in A$ from the definition of p. Since we want $p(x) = a$ and $p(y) = b$, this is equivalent to solving the matrix equation below:

$$\begin{pmatrix} f(x) & [f(x)]^2 \\ f(y) & [f(y)]^2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

The determinant of the 2×2 matrix on the LHS is $f(x)[f(y)]^2 - f(y)[f(x)]^2 = f(x)f(y)(f(y) - f(x))$. Since $f(x) \neq f(y)$, $f(x) \neq 0$ and $f(y) \neq 0$, the determinant is not equal to zero. As a result, we can solve for $\alpha, \beta \in \mathbb{R}$ by multiplying on the left by the inverse. Therefore, we have $p = \alpha f + \beta f^2 \in A$ with $p(x) = a$ and $p(y) = b$ as required.

Now we can finally embark on the proof of Stone-Weierstrass for locally compact Hausdorff spaces.

To show: (c) $A = Cts_0(X, \mathbb{R})$.

(c) Since A is a subalgebra of $Cts_0(X, \mathbb{R})$, $A \subseteq Cts_0(X, \mathbb{R})$. So, it suffices to prove that $Cts_0(X, \mathbb{R}) \subseteq A$. Assume that $f \in Cts_0(X, \mathbb{R})$ and that $\epsilon \in \mathbb{R}_{>0}$. Suppose that $x, y \in X$ such that $x \neq y$. Then, from part b, there exists $f_{x,y} \in A$ such that $f_{x,y}(x) = f(x)$ and $f_{x,y}(y) = f(y)$. Now, note that the function $D_{x,y} : X \to \mathbb{R}, D_{x,y}(z) = |f_{x,y}(z) - f(z)|$ is a continuous function. Hence, the particular preimage below is an open subset of X:

$$D_{x,y}^{-1}((-\infty, \epsilon)) = \{z \in X \mid |f_{x,y}(z) - f(z)| < \epsilon\}.$$

We will denote this open set by $U_{x,y}$. Note that $x \in U_{x,y}$ by definition. Its complement $X \backslash U_{x,y}$ is a closed subset of X. Since X is locally compact, we deduce that $X \backslash U_{x,y}$ is also locally compact. However, we can prove the stronger statement that $X \backslash U_{x,y}$ is compact.

To show: (ca) $X \backslash U_{x,y}$ is compact.

(ca) Since $f \in Cts_0(X, \mathbb{R})$ and $f_{x,y} \in A$, $f - f_{x,y} \in Cts_0(X, \mathbb{R})$. This means that $f - f_{x,y}$ is a function which vanishes at infinity. So, given $\epsilon \in \mathbb{R}_{>0}$, there exists a compact subset $K \subseteq X$ such that for all $z \notin K$, $|f(z) - f_{x,y}(z)| < \epsilon$. As a result, for all $z \in X \backslash K$, $|f(z) - f_{x,y}(z)| < \epsilon$. Due to the definition of $U_{x,y}$, $X \backslash K \subseteq U_{x,y}$. The contrapositive of this statement is that $X \backslash U_{x,y} \subseteq K$. Since $X \backslash U_{x,y}$ is a closed subset of the compact set K, $X \backslash U_{x,y}$ must be a compact subset of $K \subseteq X$ consequently.

(c) Analogously to the proof of the original Stone-Weierstrass theorem, fix a point $x_1 \in X$. Consider the set

$$L = \{U_{x,y} \subseteq X \mid x \in X \backslash U_{x_1,y}\}.$$

Then, L is an open cover of $X \backslash U_{x_1,y}$. Since $X \backslash U_{x_1,y}$ is compact, there exists $x_2, \ldots, x_n \in X$ such that

$$X \backslash U_{x_1,y} = \bigcup_{i=2}^{n} U_{x_i,y}.$$

Now, we let $g_y = \max\{f_{x_1,y}, f_{x_2,y}, \ldots, f_{x_n,y}\}$. Then, $g_y \in Cts_0(X, \mathbb{R})$ and for some arbitrary $s \in X$, there exists $k \in \{1, \ldots, n\}$ such that $s \in U_{x_k,y}$. Due to the definition of the open sets $U_{x_k,y}$, $f(s) - g_y(s) < \epsilon$ for all $s, y \in X$. To establish the other bound, let

$$V_y = U_{x_1,y} \cap U_{x_2,y} \cap \cdots \cap U_{x_n,y}.$$

Then, $y \in V_y$ by definition and since $X \backslash U_{x_i,y}$ is compact for all $i \in \{1, \ldots, n\}$, $X \backslash V_y$ must also be compact because

$$X \backslash V_y = (X \backslash U_{x_1,y}) \cup \cdots \cup (X \backslash U_{x_n,y}).$$

The open sets in $\{V_z \mid z \in X \backslash V_y\}$ form an open cover of $X \backslash V_y$ and so, there exists $V_{y_1}, \ldots, V_{y_m}$ such that

$$X \backslash V_y = V_{y_1} \cup \cdots \cup V_{y_m}.$$

Now set $g = \min\{g_{y_1}, g_{y_2}, \ldots, g_{y_m}\}$. Then, $g \in A$ and for all $s \in X$, $f(s) - g(s) < \epsilon$, revealing that $f(s) - \epsilon < g(s)$. Also, there exists $j \in \{1, \ldots, m\}$ such that $s \in V_{y_j}$ or $s \in V_y$ and as a result,

$$g(s) \leq g_{y_j}(s) < f(s) + \epsilon.$$

Hence, $|f(s) - g(s)| < \epsilon$ for all $s \in X$. This reveals that $f \in A$. Therefore, $Cts_0(X, \mathbb{R}) \subseteq A$ and as a result, $A = Cts_0(X, \mathbb{R})$.

$\square$

As alluded to in many parts of the above proof, the classic version of the Stone-Weierstrass theorem can be proved in a very similar manner.

Chapter 6

Differential Equations and Linear Semigroups

6.1 The Matrix Exponential

The main content of this chapter relies on a working knowledge of the *matrix exponential* (or the exponential of a bounded linear operator). In this section, we will define the exponential of a bounded linear operator and study some of its properties. The main reference for the next two sections is [DM21].

Definition 6.1.1. Let $T : V \to V$ be a bounded linear operator. That is, let $T \in B(V;V)$. Then, the **exponential** of T is defined as the following operator:

$$\exp(T) = \sum_{i=0}^{\infty} \frac{T^i}{i!}$$

The operator T^i denotes the composition of T with itself i times.

With the given definition, we immediately run into a problem. How do we know that $\exp(T) \in B(V;V)$? It is not obvious that the above series converges. Our first main result is to prove that $\exp(T)$ converges absolutely in $B(V;V)$. The following lemma is key to our argument:

Lemma 6.1.1. *Let V be a Banach space and $a_m = \sum_{n=0}^{m} u_n$ be a sequence in V with the property that the corresponding series*

$$\sum_{n=0}^{m} \|u_n\|$$

converges in $\mathbb{R}$ as $m \to \infty$. Then, $\sum_{n=0}^{\infty} u_n$ converges. A series which has the property above is called ***absolutely convergent****.*

Proof. Assume that V is a Banach space and $\{a_m\}$ is the sequence of partial sums in V defined as above for all $m \in \mathbb{Z}_{>0}$. Suppose that the sequence is absolutely convergent. Then, the sequence $\sum_{n=0}^{m} \|u_n\|$ converges as $m \to \infty$.

This means that the sequence $b_m = \sum_{n=0}^{m} \|u_n\|$ is a Cauchy sequence in $\mathbb{R}$. Take $\epsilon \in \mathbb{R}_{>0}$. Then, there exists $N \in \mathbb{Z}_{>0}$ such that for all $m > m' \geq N$,

$$|b_m - b'_m| = |\sum_{n=0}^{m} \|u_n\| - \sum_{n=0}^{m'} \|u_n\|| = |\sum_{n=m'+1}^{m} \|u_n\|| < \epsilon.$$

To see that the sequence $\{a_m\}$ converges in V, we observe that for all $m > m'$,

$$\begin{aligned} \|a_m - a_{m'}\| &= \|\sum_{n=0}^{m} u_n - \sum_{n=0}^{m'} u_n\| \\ &= \|\sum_{n=m'+1}^{m} u_n\| \\ &\leq \sum_{n=m'+1}^{m} \|u_n\| \\ &< \epsilon. \end{aligned}$$

Hence, the sequence of partial sums $\{a_m\}$ converges in V.

□

The above lemma tells us that if we want to prove a certain sequence converges in a Banach space V, we can prove that the norms of each element in the sequence converges in $\mathbb{R}$ instead (which proves that the sequence is absolutely convergent). Unsurprisingly, this is the strategy we will use to prove our first main result of the exponential. The next ingredient for our proof concerns the composition of bounded linear operators.

Lemma 6.1.2. *Let U, V, W be normed vector spaces and $S : V \to W$, $T : U \to V$ be bounded linear operators. Then, $S \circ T$ is also a bounded linear operator satisfying $\|S \circ T\| \leq \|S\|\|T\|$.*

Proof. Assume that U, V, W are normed vector spaces. Assume that $S : V \to W$ and $T : U \to V$ are bounded linear operators.

To show: (a) $\|S \circ T\| \leq \|S\|\|T\|$.

(a) This follows from a quick computation:

$$\begin{aligned}
\|S \circ T\| &= \sup_{\|u\|_U = 1} \|(S \circ T)(u)\|_W \\
&= \sup_{\|u\|_U = 1} \|S(T(u))\|_W \\
&\leq \sup_{\|u\|_U = 1} \|S\|\|T(u)\|_V \\
&= \|S\|\|T\|.
\end{aligned}$$

Hence, $S \circ T$ is a bounded linear operator, with $\|S \circ T\| \leq \|S\|\|T\|$.

□

Now we are ready for the proof of our first main result.

Theorem 6.1.3. *Let V be a Banach space and $T : V \to V$ be a bounded linear operator. Then, the exponential of T*

$$\exp(T) = \sum_{i=0}^{\infty} \frac{T^i}{i!}$$

converges absolutely in $B(V; V)$.

Proof. Assume that V is a Banach space and $T \in B(V; V)$. Since V is a Banach space, $B(V; V)$ is also a Banach space. From 6.1.1, if we want to prove that the sequence

$$a_m = \sum_{i=0}^{m} \frac{T^i}{i!}$$

in $B(V; V)$ is absolutely convergent, it suffices to prove that the sequence

$$b_m = \sum_{i=0}^{m} \|\frac{T^i}{i!}\| = \sum_{i=0}^{m} \frac{\|T^i\|}{i!}$$

converges in $\mathbb{R}$. However, by 6.1.2, the operator $T^i \in B(V; V)$ for all $i \in \mathbb{Z}_{\geq 0}$ with $\|T^i\| \leq \|T\|^i$. Note that $T^0 = id_V$ (the identity operator on V). Therefore, in $\mathbb{R}$,

$$b_m = \sum_{i=0}^{m} \frac{\|T^i\|}{i!} \leq \sum_{i=0}^{m} \frac{\|T\|^i}{i!} \to \exp(\|T\|)$$

as $m \to \infty$. An alternative method of deducing convergence of this sequence would be to use the comparison and ratio tests. Therefore, the sequence $\{b_m\}$ converges in $\mathbb{R}$, revealing that the sequence $\{a_m\}$ converges absolutely in $B(V;V)$.

□

A final application of 6.1.1 finally reveals that $\exp(T)$ is convergent and so, $\exp(T) \in B(V;V)$. Recall that this was a consequence of the uniform boundedness principle (in particular, theorem 4.1.4).

Despite the fact that we are predominantly concerned with infinite dimensional normed vector spaces in these notes, we will revert to discussing finite dimensional normed vector spaces, in order to better understand the specific concept of the matrix exponential - the exponential map when applied to matrices, which denote linear operators/transformations over a finite dimensional vector space.

The next lemma reveals that it is pointless to discuss different norms on a finite dimensional normed vector space.

Lemma 6.1.4. *Let V be a finite dimensional vector space V. Let $\|-\|_a, \|-\|_b$ be two different norms on V. Then, they are Lipschitz equivalent. Alternatively, this means that there exists real numbers $0 < c_1 \leq c_2$ such that*

$$c_1\|x\|_a \leq \|x\|_b \leq c_2\|x\|_a.$$

Proof. Assume that V is a finite dimensional vector space. Assume that $\|-\|_a, \|-\|_b$ are two different norms on V. Suppose that $\{v_1, \dots, v_n\}$ is a basis for V. Since Lipschitz equivalence of norms is an equivalence relation, it suffices to prove that any norm on V, which we will denote by $\|-\|$, is Lipschitz equivalent to the norm $\|-\|_1$, which is defined by

$$\|-\|_1 : V \to \mathbb{K}$$

$$\|\sum_{i=1}^{n} a_i v_i\|_1 = \sum_{i=1}^{n} |a_i|.$$

We will proceed in two distinct steps:

To show: (a) $\|-\|$ is uniformly continuous when V is equipped with the norm $\|-\|_1$.

(a) Assume that $\epsilon \in \mathbb{R}_{>0}$. Set $\delta = \epsilon/C$, where $C = \sup\{\|v_i\| \mid i \in \{1, \ldots, n\}\}$ so that $\|x - x'\|_1 < \delta$. Using the basis of V, set

$$x = \sum_{i=1}^{n} a_i v_i \text{ and } x' = \sum_{i=1}^{n} b_i v_i.$$

Using the reverse triangle inequality on $\mathbb{K}$, we deduce that

$$\begin{aligned} |\|x\| - \|x'\|| &\leq \|x - x'\| \\ &= \|\sum_{i=1}^{n}(a_i - b_i)v_i\| \\ &\leq \sum_{i=1}^{n}|a_i - b_i|\|v_i\| \quad \text{(Linearity of norm)} \\ &\leq C\sum_{i=1}^{n}|a_i - b_i| \\ &= C\|x - x'\|_1 \\ &< C\frac{\epsilon}{C} \\ &= \epsilon. \end{aligned}$$

Hence, the function $\|-\|$ is uniformly continuous from $(V, \|-\|_1)$ to $(\mathbb{K}, |-|)$. From our construction of V with the induced topology from $\|-\|_1$, it must be homeomorphic to $\mathbb{R}^n$. Consequently, the set

$$W = \{v \in V \mid \|v\|_1 = 1\}$$

is compact because the closed unit ball in $\mathbb{R}^n$ is closed and bounded and thus, compact. The extreme value theorem, then tells us that the continuous function $\|-\|$ attains it supremum and infimum on W. So, there exists $v, w \in V$ such that

$$\|v\| = C_1 = \inf\{\|v\| \mid \|v\|_1 = 1\}$$

and

$$\|w\| = C_2 = \sup\{\|v\| \mid \|v\|_1 = 1\}.$$

Since, $\|v\|_1 = \|w\|_1 = 1$ by definition, $v, w \neq 0$. So, $C_1, C_2 \neq 0$ and for all $y \in V$.

$$C_1\|y\|_1 \leq \|y\| \leq C_2\|y\|_1.$$

Note that this follows if $\|y\|_1 = 1$ or $y = 0$. If $\|y\|_1 > 1$, we can multiply by $1/\|y\|_1$ to reduce to the previous case. This completes the proof.

□

Next, we will use the above lemma to prove that bounded operators from a finite dimensional vector space must be bounded.

Lemma 6.1.5. *Let $(V, \|-\|_V)$ and $(W, \|-\|_W)$ be normed vector spaces with V finite dimensional. Then, any linear transformation $T : V \to W$ is bounded.*

Proof. Assume that $(V, \|-\|_V)$ and $(W, \|-\|_W)$ are normed vector spaces with V finite dimensional. By 6.1.4, we can assume that $\|-\|_V = \|-\|_1$ since a linear transformation $T : V \to W$ is bounded with respect to $\|-\|_V$ if and only if it is bounded to any Lipschitz equivalent norm. Let $\{v_1, \ldots, v_n\}$ be a basis for V. Set $x = \sum_{i=1}^n a_i v_i$. Then,

$$\begin{aligned}
\|Tx\|_W &= \|\sum_{i=1}^n a_i T(v_i)\|_W \\
&\leq \sum_{i=1}^n \|a_i T(v_i)\|_W \\
&= \sum_{i=1}^n |a_i| \|T(v_i)\|_W \\
&\leq C\|x\|_1.
\end{aligned}$$

where $C = \sup_{i \in \{1,\ldots,n\}} \|T(v_i)\|_W$. Hence, T is bounded.

□

The final piece of the puzzle in describing the matrix exponential in its most familiar form is the following theorem:

Theorem 6.1.6. *Let $(V, \|-\|_V)$ be a finite dimensional normed vector space. Then, V must be complete.*

Proof. Assume that $(V, \|-\|_V)$ is a finite dimensional normed vector space. By dividing or multiplying by the appropriate constants, we observe that every Cauchy or convergent sequence with respect to the norm $\|-\|_V$ is also Cauchy or convergent with respect to any other Lipschitz equivalent norm on V. Therefore, we can set $\|-\|_V = \|-\|_1$.

Let $\{v_1, \ldots, v_n\}$ be a basis for V and $\{x_m\}$ denote a Cauchy sequence in V. Suppose that $x_m = \sum_{i=1}^n a_{i,m}v_i$.

To show: (a) The sequence $\{x_m\}$ converges with respect to the $\|-\|_1$ norm.

(a) Assume that $\epsilon \in \mathbb{R}_{>0}$. Since $\{x_m\}$ is Cauchy, there exists $N \in \mathbb{Z}_{>0}$ such that for all $m, m' \geq N$,

$$\|x_m - x_{m'}\|_1 < \epsilon.$$

However, we can expand the LHS to get

$$\|\sum_{i=1}^n (a_{i,m} - a_{i,m'})v_i\|_1 = \sum_{i=1}^n |a_{i,m} - a_{i,m'}| < \epsilon$$

So, for all $j \in \{1, \ldots, n\}$,

$$|a_{j,m} - a_{j,m'}| < \epsilon$$

since every summand is non-negative. Thus, the sequence $\{a_{j,m}\}$ converges since $\mathbb{R}$ is complete. Denote the limit by a_j and define

$$x = \sum_{i=1}^n a_i v_i.$$

We will show that the sequence $\{x_m\}$ converges to x. Since $\{a_{j,m}\}$ converges to a_j, there exists $N_j \in \mathbb{Z}_{>0}$ such that for all $m \geq N_j$, $|a_{j,m} - a_j| < \epsilon/n$. Take $P = \max\{N_1, \ldots, N_n\}$. Then, for all $m \geq P$,

$$\begin{aligned}
\|x_m - x\|_1 &= \|\sum_{i=1}^{n}(a_{i,m} - a_i)v_i\|_1 \\
&= \sum_{i=1}^{n}|a_{i,m} - a_i| \\
&< n(\frac{\epsilon}{n}) \\
&= \epsilon.
\end{aligned}$$

Therefore, the sequence $\{x_m\}$ converges to x. So, V must be complete. $\square$

Now we will summarise what the last three lemmas tell us about the matrix exponential. Let V be a finite dimensional vector space with $\dim V = n$ and $A \in B(V\ V)$. Since V is a Banach space from 6.1.6, $B(V\ V)$ is also a Banach space. Since V is finite dimensional, A can be expressed as a $n \times n$ matrix. Furthermore, from 6.1.5, A is bounded, which means that from 6.1.3,

$$\exp(A) = \sum_{i=0}^{\infty} \frac{A^i}{i!}$$

converges absolutely in $B(V; V)$. So, $\exp(A)$ is well-defined for all $A \in M_{n\times n}(\mathbb{K})$ because every matrix $A \in M_{n\times n}(\mathbb{K})$ defines a bounded operator on V.

We will now prove various properties of the exponential. We will relax the assumption that V is finite dimensional. First, we require some technical results. The first one is familiar in the context of real analysis.

Lemma 6.1.7 (Rearrangement)**.** *Let V be a Banach space and $v_n \in V$ such that the sequence of partial sums $\{\sum_{n=0}^{m} v_n\}$ converges absolutely. Let $j : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ be a bijection. Then, the rearranged sequence $\{\sum_{n=0}^{m} v_{j(n)}\}$ converges absolutely and*

$$\sum_{n=0}^{\infty} v_n = \sum_{n=0}^{\infty} v_{j(n)}.$$

Proof. Assume that V is a Banach space and $v_n \in V$ such that the sequence $\{\sum_{n=0}^{m} v_n\}$ converges absolutely. Since the sequence is absolutely convergent, the sequence of partial sums in $\mathbb{R}$

$$\sum_{n=0}^{m} \|v_n\|$$

must converge. Now assume that $j : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ is a bijection. Then, the rearranged sequence in $\mathbb{R}$

$$\sum_{n=0}^{m} \|v_{j(n)}\|$$

converges to the same limit as the previous sequence. By 6.1.1, we know that the sequence of rearranged partial sums $\{\sum_{n=0}^{m} v_{j(n)}\}$ converges in V. It suffices to show that $\sum_{n=0}^{\infty} v_{j(n)} = \sum_{n=0}^{\infty} v_n$.

To show: (a) $\sum_{n=0}^{\infty} v_{j(n)} = \sum_{n=0}^{\infty} v_n$.

(a) Let $L_m = \sum_{n=0}^{m} v_n$ and $R_m = \sum_{n=0}^{m} v_{j(n)}$. We will demonstrate that $\lim_{m\to\infty} \|L_m - R_m\| = 0$. Keeping this in mind, assume $\epsilon \in \mathbb{R}_{>0}$.

Set $S_m = \sum_{n=0}^{m} \|v_n\|$ in $\mathbb{R}$. By absolute convergence, $\{S_m\}$ is a convergent sequence and thus, Cauchy. Hence, we can find $N_1 \in \mathbb{Z}_{>0}$ such that for all $m, m' \geq N_1$,

$$|S_m - S_{m'}| < \frac{\epsilon}{2}.$$

This ensures that the "tail" of the series $\sum_{n=0}^{\infty} \|v_n\|$ is bounded because

$$\sum_{i=m+1}^{m'} \|v_i\| = S_m - S_m' < \frac{\epsilon}{2}$$

and consequently, for all $B \subseteq \mathbb{Z}_{\geq 0} \backslash \{0, \ldots, N_1\}$,

$$\sum_{i \in B} \|v_i\| < \frac{\epsilon}{2}.$$

Now, we have to bound the "front" of this series. Let $N_2 = \max\{j^{-1}(0), \ldots, j^{-1}(N_1)\}$ and in the usual fashion, $N = \max\{N_1, N_2\}$. Then, for all $m > N$, $j_m \notin \{0, \ldots, N_1\}$ because $m > N_2$. Setting $A = \{j^{-1}(0), \ldots, j^{-1}(N_1)\}$, we argue as follows:

$$\begin{aligned}
L_m - R_m &= \sum_{n=0}^{m} v_n - \sum_{n=0}^{m} v_{j(n)} \\
&= \sum_{n=0}^{N_1} v_n + \sum_{n=N_1+1}^{m} v_n - \sum_{n=0}^{m} v_{j(n)} \\
&= \sum_{i\in A} v_{j(i)} + \sum_{n=N_1+1}^{m} v_n - \sum_{n=0}^{m} v_{j(n)} \\
&= \sum_{n=N_1+1}^{m} v_n - \sum_{i\in\{0,\ldots,m\}\setminus A} v_{j(i)}.
\end{aligned}$$

Since the sets $\{N_1+1,\ldots,m\}$ and $j(\{0,\ldots,m\}\setminus A)$ are both finite sets, which are disjoint from $\{0,\ldots,N_1\}$, we deduce that

$$\begin{aligned}
\|L_m - R_m\| &= \sum_{n=N_1+1}^{m} \|v_n\| - \sum_{i\in\{0,\ldots,m\}\setminus A} \|v_{j(i)}\| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}$$

Hence, $\lim_{m\to\infty}\|L_m - R_m\| = 0$ and subsequently, $\sum_{n=0}^{\infty} v_n = \sum_{n=0}^{\infty} v_{j(n)}$. □

We can refine the above lemma slightly to the case where $j : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ is surjective with the preimage $j^{-1}(n)$ is finite for all $n \in \mathbb{Z}_{\geq 0}$. The statement we are concerned with is the following:

Lemma 6.1.8. *Let V be a Banach space and $v_n \in V$ such that the sequence of partial sums $\{\sum_{n=0}^{m} v_n\}$ converges absolutely. Let $j : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ be a surjective map such that the preimage $j^{-1}(n)$ is finite for all $n \in \mathbb{Z}_{\geq 0}$. Then, the sequence of partial sums*

$$\sum_{n=0}^{m}\Big(\sum_{i\in j^{-1}(n)} v_i\Big)$$

converges absolutely and $\sum_{n=0}^{\infty}(\sum_{i\in j^{-1}(n)} v_i) = \sum_{n=0}^{\infty} v_n$.

Proof. Assume that V is a Banach space and $v_n \in V$ such that the sequence of partial sums $\{\sum_{n=0}^{m} v_n\}$ converges absolutely. Assume that $j : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$

be a surjective map such that the preimage $j^{-1}(n)$ is finite for all $n \in \mathbb{Z}_{\geq 0}$.

We will enumerate all the positive integers in $\mathbb{Z}_{\geq 0}$ according to their image under j. Let

$$\alpha(i) = \sum_{a<j(i)} Card(j^{-1}(a)).$$

where $Card$ refers to cardinality. In the case where the elements in the preimage $j^{-1}(j(i))$ are arranged in ascending order contains $\beta(i)$ elements strictly less than i, we define $J(i) = \alpha(i) + \beta(i)$. Since J is an enumeration of $\mathbb{Z}_{\geq 0}$, the previous lemma 6.1.7 tells us that the sequence of partial sums $\{\sum_{n=0}^{m} v_{J(n)}\}$ converges absolutely to $\sum_{n=0}^{\infty} v_n$. The trick here is to realise that the sequence

$$\{\sum_{n=0}^{m} \sum_{i\in j^{-1}(n)} v_i\}$$

is a subsequence of $\{\sum_{n=0}^{m} v_{J(n)}\}$. Therefore, it must converge to the same limit. □

Here is the main property of the exponential map.

Theorem 6.1.9. *Let V be a Banach space and $S, T \in B(V;V)$. Then, if $ST = TS$ (we mean composition), then $\exp(S)\exp(T) = \exp(S+T)$.*

Proof. Assume that V is a Banach space and $S, T \in B(V;V)$. Assume that $ST = TS$. Due to the commutativity of S and T, one can show by induction on n that the usual binomial formula holds:

$$(S+T)^n = \sum_{i=0}^{n} \binom{n}{i} S^{n-i}T^i.$$

Let $\{a_m\}$ be the sequence of partial sums defined by

$$a_m = \sum_{n=0}^{m} \frac{1}{n!}(S+T)^n.$$

Using our expression for $(S+T)^n$, we can write a_m as

$$
\begin{aligned}
a_m &= \sum_{n=0}^{m} \frac{1}{n!}(S+T)^n \\
&= \sum_{n=0}^{m} \frac{1}{n!} \sum_{i=0}^{n} \binom{n}{i} S^{n-i}T^i \\
&= \sum_{n=0}^{m} \sum_{i=0}^{n} (\frac{1}{(n-i)!} S^{n-i})(\frac{1}{i!}T^i) \\
&= \sum_{n=0}^{m} \sum_{a+b=n} (\frac{1}{a!}S^a)(\frac{1}{b!}T^b).
\end{aligned}
$$

We can expand the expression $\exp(S)\exp(T)$ is a similar manner in order to obtain

$$
\begin{aligned}
\exp(S)\exp(T) &= \lim_{k\to\infty} [(\sum_{n=0}^{k} \frac{1}{n!}S^n)(\sum_{n=0}^{k} \frac{1}{n!}T^n)] \\
&= \lim_{k\to\infty} \sum_{a,b=0}^{k} (\frac{1}{a!}S^a)(\frac{1}{b!}T^b).
\end{aligned}
$$

By taking the limit of the sequence a_m as $m \to \infty$ and setting

$$
X_{a,b} = \frac{1}{a!}S^a \frac{1}{b!}T^b,
$$

we deduce the following expressions:

$$
\exp(S+T) = \lim_{m\to\infty} \sum_{a+b\leq m} X_{a,b} \text{ and } \exp(S)\exp(T) = \lim_{m\to\infty} \sum_{a,b\leq m} X_{a,b}.
$$

We want to show that the two expressions are equal. This is equivalent to rearranging the terms in one of the infinite sums to obtain the other. The previous lemmas provide us with a method for dealing with this.

Define $f : \mathbb{Z}_{\geq 0} : \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ to be the enumeration defined by

$$
f(0) = (0,0), f(1) = (0,1), f(2) = (1,0), f(3) = (0,2), f(4) = (1,1), f(5) = (2,0), \ldots
$$

Take $f(i) = (a_i, b_i)$. Then, the sum below

$$\sum_{i=0}^{\infty} \frac{1}{a_i!}\|S\|^{a_i}\frac{1}{b_i!}\|T\|^{b_i}$$

converges because it is the limit of a sequence of partial sums which contains the convergent subsequence

$$\sum_{a+b\leq m} \frac{1}{a!}\|S\|^{a}\frac{1}{b!}\|T\|^{b}$$

An increasing sequence with a convergent subsequence is bounded and therefore, convergent. Hence, the series $\sum_{i=0}^{\infty} X_{f(i)}$ converges absolutely because

$$\sum_{i=0}^{k}\|X_{f(i)}\| \leq \sum_{i=0}^{k} \frac{1}{a_i!}\|S\|^{a_i}\frac{1}{b_i!}\|T\|^{b_i}.$$

In order to apply the previous lemma, we let $\Theta_0 \subseteq \Theta_1 \subseteq \ldots$ be any strictly ascending chain of non-empty finite sets in $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ with $\bigcup_i \Theta_i = \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$. Define $j : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ by

$$j(a) = \inf\{i \mid f(a) \in \Theta_i\}.$$

Note that f is surjective by construction. Hence, from the previous lemma, the series

$$\sum_{n=0}^{\infty}(\sum_{i\in j^{-1}(n)} X_{f(i)})$$

converges absolutely to $\sum_{i=0}^{\infty} X_{f(i)}$. The trick here is that we can take alternatively:

$$\Theta_{2m} = \{(a,b) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \mid a \leq m, b \leq m\}$$

and

$$\Theta_{2m+1} = \{(a,b) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \mid a+b \leq m+1\}.$$

This definition provides the required ascending chain of subsets $\Theta_0 \subseteq \Theta_1 \subseteq \ldots$. So, we can apply our argument above to deduce that

$$\begin{aligned}\exp(S+T) &= \lim_{m\to\infty} \sum_{a+b\leq m} X_{a,b} \\ &= \lim_{m\to\infty} \sum_{a\leq m, b\leq m} X_{a,b} \\ &= \exp(S)\exp(T).\end{aligned}$$

□

We note the following useful consequences of 6.1.9:

Corollary 6.1.10. *Let V be a Banach space. Then,*

1. $\exp(0) = id_V$.
2. *If $\alpha, \beta \in \mathbb{F}$, then for all $S \in B(V;V)$,* $\exp(\alpha S)\exp(\beta S) = \exp((\alpha+\beta)S)$.
3. *The map* $\exp(S)$ *is invertible, with inverse* $\exp(-S)$.

6.2 Logarithms

It is well-known that on the appropriate domain in $\mathbb{R}$, the logarithm function can be defined as an inverse to the exponential map. After the analysis of the matrix exponential conducted in the previous section, a natural question to ask is whether one can define the logarithm of a bounded linear operator on a Banach space. It turns out that just like the ordinary logarithm in $\mathbb{R}$, one can do this in an appropriate radius of convergence. The first result of this section is key to describing the construction of the logarithm.

Lemma 6.2.1. *Let V be a Banach space and $T \in B(V;V)$. If $\|T\| < 1$, then $(I-T)^{-1}$ is a bounded linear operator on V with the expression*

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

Note that I is the identity operator on $B(V;V)$.

Proof. Assume that V is a Banach space and $T \in B(V;V)$. Assume that $\|T\| < 1$. Then, since V is a Banach space, $B(V;V)$ is also a Banach space.

Our first task is to establish the convergence of the sum $\sum_{i=0}^{\infty} T^i$.

To show: (a) The sum $\sum_{i=0}^{\infty} T^i$ converges.

(a) We already know that for all $j \in \mathbb{Z}_{\geq 0}$, $\|T^j\| \leq \|T\|^j$. In light of 6.1.1, we observe that the sequence of partial sums in $\mathbb{R}$, defined by

$$a_m = \sum_{i=0}^{m} \|T^i\| \leq \sum_{i=0}^{m} \|T\|^i$$

converges as $m \to \infty$ because $\|T\| < 1$. The term a_m is bounded above by a geometric series. Hence, from 6.1.1, the sum $\sum_{i=0}^{\infty} T^i$ converges in $B(V; V)$.

Part (a) tells us further that $\sum_{i=0}^{\infty} T^i$ is a bounded linear operator on V. It remains to determine the equation we are after. Let $S = \sum_{i=0}^{\infty} T^i$ and $S_m = \sum_{i=0}^{m} T^i$ for all $i \in \mathbb{Z}_{\geq 0}$.

To show: (b) $S(I - T) = (I - T)S = I$.

(b) We compute directly that

$$\begin{aligned} S_m(I - T) &= (I + T + \cdots + T^m)(I - T) \\ &= I - T^{m+1}. \end{aligned}$$

Notice that $(I - T)S_m$ also evaluates to the same expression, by a similar computation. Taking the limit as $m \to \infty$, we deduce that $S(I - T) = (I - T)S = I$, hence revealing that $S = (I - T)^{-1}$.

□

We want to apply the lemma we just proved to the exponential of a bounded operator $\exp(T)$, where $T \in B(V; V)$ with V a Banach space. However, the lemma 6.2.1 tells us that this can only be done in a certain radius of convergence. In order to define the logarithm, consider the operator $I - \exp(T) \in B(V; V)$. Suppose that $\|T\| < \log 2$. Then,

$$\begin{aligned}
\|I - \exp(T)\| &= \|I - \sum_{i=0}^{\infty} \frac{T^i}{i!}\| \\
&= \|\sum_{i=1}^{\infty} \frac{T^i}{i!}\| \\
&\leq \sum_{i=1}^{\infty} \frac{\|T\|^i}{i!} \\
&= \exp(\|T\|) - 1 \\
&< 1 \quad \text{since } \|T\| < \log 2.
\end{aligned}$$

Hence, 6.2.1 applies, provided that the bounded operator T is in the subset W where

$$W = \{T \in B(V;V) \mid \|T\| < \log 2\}.$$

Thus, $(I - (I - \exp(T)))^{-1} = (\exp(T))^{-1}$ is a bounded linear operator for all $T \in W$. Thus, we can define the logarithm $\log : U \to W$ by

$$\log(T) = (\exp(T))^{-1} = \sum_{i=0}^{\infty} (I - \exp(T))^i.$$

Here, U is the set

$$U = \{T \in B(V;V) \mid \|I - \exp(T)\| < 1\}.$$

Explicitly, the map $I - \exp$ needs to be restricted to a map from W to U, in order to define $\log(T)$ as a consequence of 6.2.1 in the first place.

Returning to the situation in 6.2.1, we can do the following computation to establish a quick bound for the norm of $(I - T)^{-1}$:

$$\begin{aligned}
\|(I - T)^{-1}\| &= \|\sum_{i=0}^{\infty} T^i\| \\
&\leq \sum_{i=0}^{\infty} \|T\|^i \\
&= \frac{1}{1 - \|T\|} \quad \text{since } \|T\| < 1.
\end{aligned}$$

Thus, we have the following corollary:

Corollary 6.2.2. *Let V be a Banach space and $T \in B(V;V)$. If $\|T\| < 1$, then $(I - T)^{-1}$ exists as a bounded linear operator on V, whose norm satisfies the following inequality:*

$$\|(I - T)^{-1}\| \leq (1 - \|T\|)^{-1}.$$

Applying the above corollary to $I - \exp : W \to U$, we deduce that

$$\|\log(T)\| \leq (1 - \|I - \exp(T)\|)^{-1}.$$

6.3 Solutions to ODEs

The problem of determining the existence and uniqueness of solutions to an ordinary differential equation has been very thoroughly studied, alongside iterative methods to construct such solutions. Picard's theorem guarantees the existence and uniqueness of a solution to an ODE over the real numbers within some closed interval (provided that the ODE satisfies a particular condition). The main result of this section reveals that Picard's theorem can be extended to Lipschitz continuous ODEs in Banach spaces. Before we prove this, we require a few definitions and Banach's fixed point theorem, which plays a major role in the proof which follows. Incidentally, Banach's fixed point theorem also plays a large role in the proof of Picard's theorem.

Definition 6.3.1. Let X be a Banach space. Then, the map $g : X \to X$ is a **Lipschitz continuous map** if it satisfies the **Lipschitz condition**: there exists $L \in \mathbb{R}_{>0}$ such that for all $x, y \in X$,

$$\|g(x) - g(y)\| \leq L\|x - y\|.$$

Definition 6.3.2. Let (X, d) be a metric space. Then, the map $g : X \to X$ is called a **contraction mapping** if there exists $\lambda \in (0, 1)$ such that for all $x, y \in X$,

$$d(g(x), g(y)) \leq \lambda d(x, y).$$

λ is called the **contraction factor** of g.

As stated before, the major theorem we require here is Banach's fixed point theorem. We prove this below:

Theorem 6.3.1 (Banach Fixed Point Theorem)**.** *Let (X, d) be a complete metric space and $f : X \to X$ be a contraction mapping with contraction factor $\lambda \in (0, 1)$. Then, f has a unique fixed point. Moreover, if $x_0 \in X$ is arbitrary, then the sequence $\{f^n(x_0)\}$ converges to the fixed point.*

Proof. Assume that (X, d) is a complete metric space and $f : X \to X$ is a contraction mapping. First we will show that a fixed point of f must be unique.

To show: (a) If $x \in X$ such that $f(x) = x$, then x is unique.

(a) Assume that $p, q \in X$ are fixed points of f such that $p \neq q$ and consequently, $d(p, q) > 0$. Then, since f is a contraction mapping,

$$d(f(p), f(q)) \leq \lambda d(p, q) < d(p, q).$$

This contradicts the fact that $d(f(p), f(q)) = d(p, q)$ and that $d(p, q) > 0$. Therefore, the fixed point of f must be unique.

Now we will prove the existence of a fixed point $x \in X$ of f.

To show: (b) There exists $x \in X$ such that $f(x) = x$.

(b) Assume that $x \in X$ and let $a_n = f^n(x)$ (composing f n times). Since f is a contraction mapping, we can prove via induction on n that

$$d(f^n(x), f^n(y)) \leq \lambda^n d(x, y)$$

for all $x, y \in X$.

To show: (ba) The sequence $\{a_n\}_{n \in \mathbb{Z}_{\geq 0}}$ is Cauchy.

(ba) Fix $n \in \mathbb{Z}_{\geq 0}$. Assume that $\epsilon \in \mathbb{R}_{>0}$. Pick $N \in \mathbb{Z}_{>0}$ such that

$$\frac{\lambda^n}{1 - \lambda} d(f(x), x) < \epsilon.$$

Then, for all $m > n \geq N$,

$$
\begin{aligned}
d(a_m, a_n) &\le d(a_m, a_{m-1}) + \cdots + d(a_{n+1}, a_n) \\
&= d(f^{m-1}(f(x)), f^{m-1}(x)) + \cdots + d(f^n(f(x)), f^n(x)) \\
&\le (\lambda^{m-1} + \cdots + \lambda^n) d(f(x), x) \\
&= d(f(x), x) \sum_{i=n}^{m-1} \lambda^i \\
&\le \lambda^n d(f(x), x) \sum_{i=0}^{\infty} \lambda^i \\
&= \frac{\lambda^n}{1-\lambda} d(f(x), x) \\
&< \epsilon.
\end{aligned}
$$

Hence, the sequence $\{a_n\}$ is Cauchy.

(b) Since $\{a_n\}$ is Cauchy and X is a complete metric space, $\{a_n\}$ must converge. Suppose that it converges to $a \in X$. It remains to show that a is a fixed point of f. We argue as follows

$$
\begin{aligned}
f(a) &= f(\lim_{n\to\infty} a_n) \\
&= f(\lim_{n\to\infty} f^n(a)) \\
&= \lim_{n\to\infty} f(f^n(a)) \\
&= \lim_{n\to\infty} f^{n+1}(a) \\
&= a.
\end{aligned}
$$

Therefore, f has a fixed point $a \in X$. □

The ODE we will consider in the next result is termed the **Cauchy problem**. If X is a Banach space, $\overline{x} \in X$ and $g : X \to X$ satisfies the Lipschitz condition, then the Cauchy problem is the ODE

$$
\frac{d}{dt}x(t) = g(x(t)) \text{ with } x(0) = \overline{x}.
$$

The next result states that the Cauchy problem does have a unique solution.

Theorem 6.3.2. *Let X be a Banach space. Let $g : X \to X$ be a Lipschitz continuous map. Then, for all $\overline{x} \in X$, the Cauchy problem has a unique solution $t \mapsto x(t)$, defined for all $t \in \mathbb{R}$.*

Proof. Assume that X is a Banach space and that $g : X \to X$ is a Lipschitz continuous map. First, we fix $T \in \mathbb{R}_{>0}$ and consider the Banach space $Cts([0,T],X)$, not with the usual norm, but equipped with the equivalent norm below:

$$\|w\|_{\dagger} = \sup_{t\in[0,T]} e^{-2Lt}\|w(t)\|.$$

Here, $L \in \mathbb{R}_{>0}$. A function $x : [0,T] \to X$ is a solution to the Cauchy problem if and only if x is a fixed point of the Picard operator

$$\Phi(w)(t) = \overline{x} + \int_0^t g(w(s))\,\mathrm{d}s$$

where $t \in [0,T]$.

To show: (a) The Picard operator Φ is a contraction with respect to the equivalent norm introduced on $Cts([0,T],X)$.

(a) Assume that $u, v \in Cts([0,T],X)$. Then, set $\delta = \|u - v\|_{\dagger}$. So, for all $s \in [0,T]$, we have

$$\begin{aligned}\|u(s) - v(s)\| &\leq \sup_{s\in[0,T]} \|u(s) - v(s)\| \\ &= e^{2Ls} \sup_{s\in[0,T]} e^{-2Ls}\|u(s) - v(s)\| \\ &= \delta e^{2Ls}.\end{aligned}$$

Utilising the properties of integrals and the Lipschitz continuity of g, we deduce that for all $t \in [0,T]$,

$$\begin{aligned}
e^{-2Lt}\|\Phi(u)(t)-\Phi(v)(t)\| &= e^{-2Lt}\|\int_0^t g(u(s))-g(v(s))\,\mathrm{d}s\| \\
&\le e^{-2Lt}\int_0^t \|g(u(s))-g(v(s))\|\,\mathrm{d}s \\
&\le e^{-2Lt}\int_0^t L\|u(s)-v(s)\|\,\mathrm{d}s \\
&\le Le^{-2Lt}\int_0^t \delta e^{2Ls}\,\mathrm{d}s \\
&= Le^{-2Lt}\frac{\delta}{2L}(e^{2Lt}-1) \\
&= \frac{\delta}{2}(1-e^{-2Lt}) \\
&\le \frac{\delta}{2}.
\end{aligned}$$

Therefore, by taking the supremum of both sides over $t \in [0, T]$, we find that

$$\|\Phi(u)-\Phi(v)\|_\dagger \le \frac{\delta}{2} = \frac{1}{2}\|u-v\|_\dagger.$$

Thus, $\Phi : Cts([0, T], X) \to Cts([0, T], X)$ is a contraction mapping with contraction factor $1/2$.

Now, we can apply the Banach fixed point theorem (6.3.1) in order to obtain the existence of a fixed point for Φ. That is, there exists $x \in Cts([0, T], X)$ such that

$$x(t) = \overline{x} + \int_0^t g(x(s))\,\mathrm{d}s$$

for all $t \in [0, T]$. As a result, x becomes the unique solution to the Cauchy problem on the interval $[0, T]$. Observe finally that by similar arguments ("reversing time"), one can construct a unique solution on any time interval of the form $[-T, 0]$. This proves the assertion. □

Presented below are two methods of iteratively constructing an approximate solution to the Cauchy problem. For both of these methods, let $h \in \mathbb{R}_{>0}$ (step size) and define $t_j = jh$ for all $j \in \mathbb{Z}_{\ge 0}$. For the initial condition, we take $x(t_0) = x(0) = \overline{x}$.

1. The more familiar of the two methods is the **forward Euler approximation**, which is more commonly known as the **Euler method**. The values of the approximate solutions at each t_j are inductively defined as follows:

$$x(t_{j+1}) = x(t_j) + hF(x(t_j))$$

2. The second method is called **backward Euler approximation**. The values of the approximate solutions are defined instead as

$$x(t_{j+1}) = x(t_j) + hF(x(t_{j+1})).$$

   In general, a backwards Euler approximation is harder to compute than its forwards counterpart, since we have to solve an implicit equation for $x(t_{j+1})$ when applying the method. However, the upshot here is that backwards Euler approximations often have much better stability and convergence properties.

6.4 Motivating Semigroups

One of the most familiar differential equations is

$$\frac{dx}{dt} = Ax \quad \text{with } x(0) = \overline{x}$$

where $A : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator. The solution to such an equation is well-known:

$$x(t) = e^{tA}\overline{x}$$

where

$$e^{tA} = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!}.$$

Note that the above series is absolutely convergent for every $t \in \mathbb{R}$. This solution is valid for every bounded linear operator A on an arbitrary Banach space X. The important properties of the exponential map in the solution are

1. $e^{0A} = I$, which is just the identity map.

2. For all $\overline{x} \in X$, the map $x(t) = e^{tA}\overline{x}$ is continuous.

3. $e^{sA}e^{tA} = e^{(s+t)A}$.

The last property is particularly important; it is called the **semigroup property**. In a nutshell, the theory of linear semigroups studies the correspondence between a linear operator A and its exponential e^{tA}, where $t \in \mathbb{R}_{\geq 0}$. When A is bounded, its exponential can be computed via the aforementioned convergent series. On the other hand, if we are given a family of exponentials e^{tA}, one can recover A as the limit

$$A = \lim_{t \to 0+} \frac{e^{tA} - I}{t}.$$

Remarkably, there are important cases where the operators e^{tA} are all bounded, whereas A itself is an unbounded operator. These scenarios form the breeding grounds for the most interesting applications of semigroup theory, which are pertinent to the analysis of parabolic and hyperbolic PDEs.

Here are some first examples of the exponential map:

Example 6.4.1. If $A : \mathbb{R}^n \to \mathbb{R}^n$ can be represented by a diagonal matrix $D = diag[\lambda_1, \ldots, \lambda_n]$, then e^{tA} is computed by exponentiating the diagonal. Indeed, we find that e^{tA} can be represented by the matrix

$$e^{tD} = diag[e^{t\lambda_1}, \ldots, e^{t\lambda_n}].$$

If the matrix of A is not diagonal but is diagonalisable (the eigenvalues of A are all distinct), then one has to diagonalise the matrix first so that $A = PDP^{-1}$. In this case, $e^{tA} = Pe^{tD}P^{-1}$. If the matrix of A is not diagonalisable at all, we will first have to reduce the matrix to its Jordan canonical form. So, there exists $U \in GL_n(\mathbb{R})$ such that $A = UJU^{-1}$. Then, we use the fact that J can be decomposed as the sum $D + N$, where D is a diagonal matrix and N is nilpotent. Then,

$$e^{tA} = Ue^{tD}e^{tN}U^{-1}.$$

We use the power series expansion to compute e^{tN}. Since N is nilpotent, there exists $j \in \mathbb{Z}_{>0}$ such that $N^j = 0$. As a consequence, e^{tN} is reduced to a finite sum, thus guaranteeing the existence of the operator e^{tN}. As an example of the most complicated case, let

$$A = \begin{pmatrix} -2 & 1 & 4 \\ -5 & 2 & 5 \\ -1 & 1 & 3 \end{pmatrix}.$$

Then, $A = UJU^{-1}$, where

$$U = \begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 15 \\ 1 & 0 & -2 \end{pmatrix}$$

and

$$J = \begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & -1 \end{pmatrix}$$

is the Jordan canonical form of A. We decompose $J = D + N$, where $D = diag[2, 2, -1]$ and

$$N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is nilpotent, with $N^2 = 0$. Next, we compute

$$e^{tD} = \begin{pmatrix} e^{2t} & & \\ & e^{2t} & \\ & & e^{-t} \end{pmatrix}$$

and

$$\begin{aligned} e^{tN} &= I + Nt \\ &= \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} + \begin{pmatrix} 0 & t & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & t & \\ & 1 & \\ & & 1 \end{pmatrix}. \end{aligned}$$

Putting all of the pieces together, we find that

$$
\begin{aligned}
e^{tA} &= Ue^{tJ}U^{-1} \\
&= Ue^{tD}e^{tN}U^{-1} \\
&= U\begin{pmatrix} e^{2t} & te^{2t} & \\ & e^{2t} & \\ & & e^{-t} \end{pmatrix}U^{-1} \\
&= \begin{pmatrix} (\frac{2}{9}-\frac{5t}{3})e^{2t}+\frac{7}{9}e^{-t} & te^{2t} & (\frac{5t}{3}+\frac{7}{9})e^{2t}-\frac{7}{9}e^{-t} \\ \frac{5}{3}e^{-t}-\frac{5}{3}e^{2t} & e^{2t} & \frac{5}{3}e^{2t}-\frac{5}{3}e^{-t} \\ (\frac{2}{9}-\frac{5t}{3})e^{2t}-\frac{2}{9}e^{-t} & te^{2t} & (\frac{5t}{3}+\frac{7}{9})e^{2t}+\frac{2}{9}e^{-t} \end{pmatrix}.
\end{aligned}
$$

Example 6.4.2. Recall that l^1 is the Banach space of all sequences of complex numbers $x = (x_1, x_2, \dots)$ with norm

$$\|x\|_1 = \sum_{k=1}^{\infty} |x_k|.$$

If we are given a sequence of complex numbers $\{\lambda_k\}$, we can define the linear operator

$$Ax = (\lambda_1 x_1, \lambda_2 x_2, \lambda_3 x_3, \dots).$$

Therefore, the exponential operators are

$$e^{tA} = (e^{t\lambda_1}x_1, e^{t\lambda_2}x_2, \dots).$$

An interesting point here is that $\|A\| = \sup_{k\in\mathbb{Z}_{>0}}|\lambda_k|$ might be infinite, but $\|e^{tA}\| = \sup_{k\in\mathbb{Z}_{>0}}|e^{t\lambda_k}|$ can be bounded for all $t \in \mathbb{R}_{\geq 0}$.

Assume that the real part of all eigenvalues λ_k is uniformly bounded. That is, there exists $\omega \in \mathbb{R}$ such that $Re(\lambda_k) \leq \omega$ for all $k \in \mathbb{Z}_{>0}$. Then,

$$|e^{t\lambda_k}| = |e^{tRe(\lambda_k)}| \leq e^{t\omega}$$

for all $t \in \mathbb{R}_{\geq 0}$. This shows that e^{tA} is a bounded operator, whereas A may not necessarily be a bounded operator itself. For example, if we define the sequence $\{\lambda_k\}$ such that $\lambda_k = 1 + ki$ for all $k \in \mathbb{Z}_{>0}$ and define $A : l^1 \to l^1$ as above, then

$$\|A\| = \sup_{k\in\mathbb{Z}_{>0}} |\lambda_k| = \infty,$$

but for all $t \in \mathbb{R}_{\geq 0}$,

$$\|e^{tA}\| = \sup_{k\in\mathbb{Z}_{>0}} |e^{t\lambda_k}| \leq e^t.$$

In the theory of linear semigroups, one considers two main problems.

1. Given a semigroup of linear operators $\{S_t \mid t \in \mathbb{R}_{\geq 0}\}$, find an operator A such that $S_t = e^{tA}$. A is called a generator for the semigroup.

2. Given a linear operator A, construct its semigroup $\{e^{tA} \mid t \in \mathbb{R}_{\geq 0}\}$ and study its properties.

We will begin our study of semigroup theory from the perspective of the first point.

6.5 Properties of Semigroups

We will now define semigroups.

Definition 6.5.1. Let X be a Banach space A **strongly continuous semigroup** of linear operators on X is a family of linear operators

$$\{S_t : X \to X \mid t \in \mathbb{R}_{\geq 0}\}$$

such that the following properties are satisfied:

1. Each S_t is a bounded linear operator.

2. For all $s, t \in \mathbb{R}_{\geq 0}$, $S_t \circ S_s = S_{t+s}$ where $\circ$ denotes composition of operators. This is the **semigroup property**.

3. $S_0 = I$ (the identity operator).

4. For all $u \in X$, the map $t \mapsto S_t u$ is continuous from $[0, \infty)$ to X.

Definition 6.5.2. Let X be a Banach space and

$$S = \{S_t : X \to X \mid t \in \mathbb{R}_{\geq 0}\}$$

be a strongly continuous semigroup of linear operators. Then, S is a **semigroup of type** ω if the bounded linear operators S_t satisfy the bounds

$$\|S_t\| \leq e^{t\omega}$$

for all $t \in \mathbb{R}_{\geq 0}$. If $\omega = 0$, then S is called a **contractive semigroup**. In this case, $\|S_t\| \leq 1$ and for all $t \in \mathbb{R}_{\geq 0}$ and $u, v \in X$,

$$\|S_t u - S_t v\| \leq \|u - v\|,$$

hence the adjective "contractive".

Definition 6.5.3. Let X be a Banach space and

$$S = \{S_t : X \to X \mid t \in \mathbb{R}_{\geq 0}\}$$

be a strongly continuous semigroup of linear operators. The operator $A : X \to X$ defined by

$$Au = \lim_{t \to 0+} \frac{S_t u - u}{t}$$

is called the **generator of the semigroup** S. It is not too difficult to show that A is a linear operator. The domain of A is

$$Dom(A) = \{u \in X \mid \lim_{t \to 0+} \frac{S_t u - u}{t} \text{ exists}\}.$$

For a given $\overline{u} \in X$, we regard the map $u(t) = S_t \overline{u}$ as the solution to a linear ODE

$$\frac{d}{dt} u(t) = Au(t) \quad \text{with } u(0) = \overline{u}.$$

In the context of ODEs, we are given the solution $u(t)$ and then tasked with the problem of finding the operator A (i.e. finding the equation).

Here are some elementary properties of the semigroup S and its generator A.

Theorem 6.5.1. *Let X be a Banach space and $S = \{S_t \mid t \in \mathbb{R}_{\geq 0}\}$ be a strongly continuous semigroup of linear operators and let A be its generator. Assume that $\overline{u} \in Dom(A)$. Then, for all $t \in \mathbb{R}_{\geq 0}$, $S_t \overline{u} \in Dom(A)$ and $AS_t \overline{u} = S_t A \overline{u}$. Furthermore, the map $u(t) = S_t \overline{u}$ is continuously differentiable and provides a solution to the Cauchy problem.*

Proof. Assume that X is a Banach space and S is the semigroup defined above. Assume that A is the generator of S. Assume that $\overline{u} \in Dom(A)$.

To show: (a) For all $t \in \mathbb{R}_{\geq 0}$, $S_t \overline{u} \in Dom(A)$.

(b) For all $t \in \mathbb{R}_{\geq 0}$, $AS_t\overline{u} = S_tA\overline{u}$.

(a) Assume that $t \in \mathbb{R}_{\geq 0}$. Then, we compute that

$$\begin{aligned}\lim_{s\to 0+} \frac{S_sS_t\overline{u} - S_t\overline{u}}{s} &= \lim_{s\to 0+} \frac{S_{s+t}\overline{u} - S_t\overline{u}}{s} \\ &= \lim_{s\to 0+} \frac{S_tS_s\overline{u} - S_t\overline{u}}{s} \\ &= S_t \lim_{s\to 0+} \frac{S_s\overline{u} - \overline{u}}{s} \\ &= S_tA\overline{u}.\end{aligned}$$

Since the limit we started with exists, we deduce that $S_t\overline{u} \in Dom(A)$ for all $t \in \mathbb{R}_{\geq 0}$.

(b) Note that

$$\lim_{s\to 0+} \frac{S_sS_t\overline{u} - S_t\overline{u}}{s} = AS_t\overline{u}.$$

So, $AS_t\overline{u} = S_tA\overline{u}$ for all $t \in \mathbb{R}_{\geq 0}$.

To show: (c) The map $u(t)$ is continuously differentiable.

(c) We will attempt to compute the left and right derivatives of $S_t\overline{u}$ for all $t \in \mathbb{R}_{\geq 0}$. First, the right derivative is computed as

$$\begin{aligned}\lim_{h\to 0+} \frac{S_{t+h}\overline{u} - S_t\overline{u}}{h} &= \lim_{h\to 0+} \frac{S_tS_h\overline{u} - S_t\overline{u}}{h} \\ &= S_t \lim_{h\to 0+} \frac{S_h\overline{u} - \overline{u}}{h} \\ &= S_tA\overline{u}.\end{aligned}$$

The left derivative is more difficult to compute. We argue as follows:

$$\begin{aligned}\lim_{h\to 0+}[\frac{S_t\overline{u} - S_{t-h}\overline{u}}{h} - S_tA\overline{u}] &= \lim_{h\to 0+}[\frac{S_{t-h}(S_h\overline{u} - \overline{u})}{h} - S_tA\overline{u}] \\ &= \lim_{h\to 0+}[S_{t-h}(\frac{(S_h\overline{u} - \overline{u})}{h} - A\overline{u}) + (S_{t-h}A\overline{u} - S_tA\overline{u})] \\ &= [S_{t-h}(\lim_{h\to 0+}[\frac{(S_h\overline{u} - \overline{u})}{h} - A\overline{u}]) + \lim_{h\to 0+}(S_{t-h}A\overline{u} - S_tA\overline{u})] \\ &= \lim_{h\to 0+}(S_{t-h}A\overline{u} - S_tA\overline{u}) \\ &= 0.\end{aligned}$$

The two calculations together reveals that $\frac{d}{dt}(S_t\overline{u}) = S_tA\overline{u} = AS_t\overline{u}$. Hence, the map $u(t) = S_t\overline{u}$ is differentiable and satisfies the Cauchy problem. Since $A\overline{u} \in X$, the map $t \mapsto S_t(A\overline{u})$ is continuous from $[0, \infty)$ to X. Combining this with the derivative result above, we deduce that the map $u(t) = S_t\overline{u}$ must be continuously differentiable. □

The next result states important properties of generators of semigroups.

Theorem 6.5.2. *Let X be a Banach space and $\{S_t : X \to X \mid t \in \mathbb{R}_{\geq 0}\}$ be a semigroup, with generator $A : X \to X$. Then, $Dom(A)$ is a dense subset of X and A is closed. That is, the graph of A*

$$\Gamma(A) = \{(x, Ax) \mid x \in X\} \subseteq X \times X$$

is a closed subset of $X \times X$.

Proof. Assume that X is a Banach space and $\{S_t \mid t \in \mathbb{R}_{\geq 0}\}$ is a semigroup with generator $A : X \to X$. To prove the first statement, note that $\overline{Dom(A)} \subseteq X$. It suffices to prove the reverse inclusion.

To show: (a) $X \subseteq \overline{Dom(A)}$.

(a) Assume that $u \in X$. For all $\epsilon \in \mathbb{R}_{>0}$, consider

$$U_\epsilon = \frac{1}{\epsilon}\int_0^\epsilon S_s u \, ds.$$

Observe that the limit

$$\lim_{\epsilon \to 0^+} \frac{1}{\epsilon}\int_0^\epsilon S_s u \, ds = u.$$

This is due to the fundamental theorem of calculus. Since $Dom(A)$ is a vector space, it remains to show that

$$\epsilon U_\epsilon = \int_0^\epsilon S_s u \, ds \in Dom(A)$$

for all $\epsilon \in \mathbb{R}_{>0}$. But, for all $0 < h < \epsilon$,

$$\begin{aligned}
\lim_{h\to 0^+}\frac{S_h\epsilon U_\epsilon - \epsilon U_\epsilon}{h} &= \lim_{h\to 0^+}\frac{1}{h}(\int_0^\epsilon S_{h+s}u\ ds - \int_0^\epsilon S_s u\ ds) \\
&= \lim_{h\to 0^+}\frac{1}{h}\int_0^\epsilon S_{h+s}u - S_s u\ ds \\
&= \lim_{h\to 0^+}(\frac{1}{h}\int_h^{\epsilon+h} S_s u\ ds - \frac{1}{h}\int_0^\epsilon S_s u\ ds) \\
&= \lim_{h\to 0^+}(\frac{1}{h}\int_\epsilon^{\epsilon+h} S_s u\ ds + \frac{1}{h}\int_h^\epsilon S_s u\ ds - \frac{1}{h}\int_h^\epsilon S_s u\ ds - \frac{1}{h}\int_0^h S_s u\ ds) \\
&= \lim_{h\to 0^+}(\frac{1}{h}\int_\epsilon^{\epsilon+h} S_s u\ ds - \frac{1}{h}\int_0^h S_s u\ ds) \\
&= S_\epsilon u - u.
\end{aligned}$$

Therefore, $\epsilon U_\epsilon \in Dom(A)$ and so, U_ϵ defines a sequence in $Dom(A)$, which converges to u. Thus, $X \subseteq \overline{Dom(A)}$, revealing that $Dom(A)$ is a dense subset of X.

To see that the graph $\Gamma(A)$ is closed, let $\{(u_k, v_k)\}_{k\in\mathbb{Z}_{>0}}$ be a sequence of points in $\Gamma(A)$ such that $v_k = Au_k$ for all $k \in \mathbb{Z}_{>0}$. Assume that $u_k \to u$ and $v_k \to v$ for some $u, v \in X$.

To show: (b) $u \in Dom(A)$ and $Au = v$.

(b) Since $u_k \in Dom(A)$ for all $k \in \mathbb{Z}_{>0}$,

$$\lim_{h\to 0^+}\frac{S_h u_k - u_k}{h} = Au_k = v_k.$$

Integration yields

$$S_h u_k - u_k = \int_0^h \frac{d}{dt}S_t u_k\ dt = \int_0^h S_t A u_k\ dt.$$

Taking the limit as $k \to \infty$, we obtain

$$S_h u - u = \int_0^h \frac{d}{dt}S_t A u\ dt = \int_0^h S_t v\ dt.$$

Now, we divide both sides by h and take the limit as $h \to 0^+$ to obtain

$$\lim_{h\to 0^+}\frac{S_h u - u}{h} = \lim_{h\to 0^+}\frac{1}{h}\int_0^h S_t v\ dt = S_0 v = v.$$

Therefore, $u \in Dom(A)$ and $v = Au$ by definition.

So, $(u, v) \in \Gamma(A)$ and consequently, $\Gamma(A)$ is a closed subset of $X \times X$. Ergo, A is a closed operator. $\square$

6.6 Resolvent Operators

The next two sections of this chapter are dedicated to exploring whether we there is always an associated semigroup to an operator $A : Dom(A) \to X$. An important link between semigroups and their generators is the *resolvent*. We will first informally discuss why this is so.

Consider the typical "flow" equation a semigroup must satisfy

$$\frac{d}{dt}u(t) = Au(t), \quad u(0) = \overline{u}.$$

Recall that the Cauchy problem can be solved iteratively and approximately by the backward Euler approximation. To enact this, we fix a time step $h \in \mathbb{R}_{>0}$ and solve the following equation for u:

$$u(t + h) = u(t) + Au(t + h).$$

Rearranging this, we obtain for all $u(t) \in X$,

$$u(t + h) = (I - hA)^{-1}u(t).$$

The expression $(I - hA)^{-1}$ is reminiscent of the operators studied in Chapter 5. It is called the **backwards Euler operator** and is denoted by E_h^-. Now, there are two ways to proceed from here, in order to obtain the solution $u(t)$.

First, one could take a limit of backwards Euler approximations. By setting the time step $h = \tau/n$ for some fixed $\tau > 0$ and $n \in \mathbb{Z}_{>0}$ and iterating the approximation procedure n times, we obtain

$$u(\tau) \approx E_{\tau/n}^- \circ \cdots \circ E_{\tau/n}^- = (I - \frac{\tau}{n}A)^{-n}\overline{u}.$$

Now, we fix τ and let $n \to \infty$. Then,

$$u(\tau) = \lim_{n\to\infty} (I - \frac{\tau}{n}A)^{-n}\overline{u}.$$

The second method is to again fix a time step $h \in \mathbb{R}_{>0}$ and set $\lambda = 1/h$. Define the operator

$$A_\lambda : X \to X$$

$$u \mapsto A(I - hA)^{-1}u.$$

We interpret $A_\lambda u$ as the value of A computed at the point $E_h^- u$, which is very close to u. Luckily, for $h \in \mathbb{R}_{>0}$ small enough, $A_\lambda = A_{1/h}$ is a well-defined bounded linear operator. Hence, we can apply the exponential map to tA_λ:

$$\exp(tA_\lambda) = \sum_{k=0}^{\infty} \frac{(tA_\lambda)^k}{k!}.$$

Then, we finally define $u(t) = \lim_{\lambda \to \infty} \exp(tA_\lambda)\overline{u}$. See Bressan [AB10] for an example of this method in action.

Now, we will define the *resolvent operator*.

Definition 6.6.1. Let A be a linear operator on a Banach space X. If $\lambda \in \rho(A)$ (the resolvent set of A), then the **resolvent operator** $R_\lambda : X \to X$ is defined by $R_\lambda u = (\lambda I - A)^{-1}u$.

Notice that the resolvent operator is related to the backward Euler operator by $\lambda R_\lambda = E_{1/\lambda}^-$. Recall that if $\lambda \in \rho(A)$, then R_λ must be a bounded, bijective operator on X. Moreover, its inverse is also a bounded linear operator (check the start of Section 5.2 for the argument).

Since we are interested in the existence of semigroups, it is reasonable to assume that A is a closed operator in light of 6.5.2. By the closed graph theorem (4.2.3), A must be continuous and hence, $R_\lambda = (\lambda I - A)^{-1}$ must also be continuous, for all $\lambda \in \rho(A)$. Furthermore, if $u \in Dom(A)$, then we have $AR_\lambda u = R_\lambda Au$. This is because if $\{S_t\}_{t \in \mathbb{R}_{\geq 0}}$ is the semigroup associated with A, then

$$
\begin{aligned}
R_\lambda A u &= R_\lambda(\lim_{t\to 0^+} \frac{S_t u - u}{t}) \\
&= \lim_{t\to 0^+} R_\lambda(\frac{S_t u - u}{t}) \quad \text{(Continuity)} \\
&= \lim_{t\to 0^+} \frac{R_\lambda S_t u - R_\lambda u}{t} \\
&= \lim_{t\to 0^+} \frac{S_t R_\lambda u - R_\lambda u}{t} \quad (R_\lambda \text{ and } S_t \text{ commute}) \\
&= A R_\lambda u.
\end{aligned}
$$

To see that R_λ and S_t commute for all $t \in \mathbb{R}_{\geq 0}$ and $\lambda \in \rho(A)$, it suffices to note that the identity operator $I : X \to X$ and A both commute with S_t. In the next lemma, we will prove a few more identities about the resolvent operator.

Lemma 6.6.1. *Let X be a Banach space and $A : X \to X$ be a closed linear operator. If $\lambda, \mu \in \rho(A)$, then $R_\lambda - R_\mu = (\mu - \lambda)R_\lambda R_\mu$ and $R_\lambda R_\mu = R_\mu R_\lambda$.*

Proof. Assume that X is a Banach space and $A : X \to X$ is a closed linear operator. Assume that $\lambda, \mu \in \rho(A)$. Then, for all $u \in X$,
$v = R_\lambda u - R_\mu u \in Dom(A)$ (we are implicitly using the fact that $Dom(A)$ is dense in X here). Now, observe that

$$
\begin{aligned}
(\lambda I - A)v &= (\lambda I - A)[(\lambda I - A)^{-1} - (\mu I - A)^{-1}]u \\
&= u - (\lambda I - A)(\mu I - A)^{-1}u \\
&= u - (\lambda I - \mu I + \mu I - A)(\mu I - A)^{-1}u \\
&= (\mu - \lambda)(\mu I - A)^{-1}u.
\end{aligned}
$$

Applying R_λ^{-1} to both sides then gives the desired result. So,
$R_\lambda u - R_\mu u = (\mu - \lambda)R_\lambda R_\mu u$. Now, we can rearrange this identity to show that

$$
R_\lambda R_\mu u = \frac{R_\lambda u - R_\mu u}{\mu - \lambda} = \frac{R_\mu u - R_\lambda u}{\lambda - \mu} = R_\mu R_\lambda u.
$$

□

The most important theorem of this section gives an integral formula for the resolvent operator.

Theorem 6.6.2. *Let $\{S_t\}_{t\in\mathbb{R}_{\geq 0}}$ be a semigroup of type ω and $A : X \to X$ be its generator (X is a Banach space). Then, for all $\lambda > \omega$, $\lambda \in \rho(A)$,*

$$R_\lambda u = \int_0^\infty e^{-t\lambda} S_t u \, dt$$

and

$$\|R_\lambda\| \leq \frac{1}{\lambda - \omega}.$$

Proof. Assume that X is a Banach space and $\{S_t\}_{t\in\mathbb{R}_{\geq 0}}$ is a semigroup of type ω on X. Assume that $A : X \to X$ is its generator. Then, for all $t \in \mathbb{R}_{\geq 0}$, $\|S_t\| \leq e^{t\omega}$. Thus, the integral above is absolutely convergent. We compute for all $\lambda > \omega$

$$\begin{aligned}
\|\int_0^\infty e^{-t\lambda} S_t u \, dt\| &\leq \int_0^\infty e^{-t\lambda} \|S_t\| \|u\| \, dt \\
&= \int_0^\infty e^{t(\omega-\lambda)} \|u\| \, dt \\
&= \frac{1}{\omega - \lambda} \|u\|.
\end{aligned}$$

We can safely define

$$\tilde{R}_\lambda u = \int_0^\infty e^{-t\lambda} S_t u \, dt.$$

The above calculation reveals that $\tilde{R}_\lambda$ is a bounded linear operator, with norm $\|\tilde{R}_\lambda\| \leq 1/(\lambda - \omega)$. We will now show that $\tilde{R}_\lambda = R_\lambda$.

To show: (a) For all $u \in X$, $(\lambda I - A)\tilde{R}_\lambda u = u$.

(a) Assume that $u \in X$. Then, a direct computation reveals that for all $h \in \mathbb{R}_{>0}$,

$$\begin{aligned}
\frac{S_h \tilde{R}_\lambda u - \tilde{R}_\lambda u}{h} &= \frac{1}{h} \int_0^\infty e^{-\lambda t} (S_{t+h} u - S_t u) \, dt \\
&= \frac{1}{h} \int_0^\infty (e^{-\lambda(t-h)} - e^{-\lambda t}) S_t u \, dt - \frac{1}{h} \int_0^h e^{-\lambda(t-h)} S_t u \, dt. \\
&= \frac{e^{\lambda h} - 1}{h} \int_0^\infty e^{-\lambda t} S_t u \, dt - \frac{e^{\lambda h}}{h} \int_0^h e^{-\lambda t} S_t u \, dt.
\end{aligned}$$

Taking the limit as $h \to \infty$, we find that

$$\lim_{h\to\infty} \frac{S_h\tilde{R}_\lambda u - \tilde{R}_\lambda u}{h} = \lambda\tilde{R}_\lambda u - u.$$

Therefore, $\tilde{R}_\lambda u \in Dom(A)$ and so, $A\tilde{R}_\lambda u = \lambda\tilde{R}_\lambda u - u$ for all $u \in X$. Rearranging yields $(\lambda I - A)\tilde{R}_\lambda u = u$.

We have demonstrated that $\lambda I - A$ is surjective for all $\lambda > \omega$ from part (a). Now, we will show that it is injective.

To show: (b) The operator $\lambda I - A$ is injective.

(b) First, we need a preliminary result. If $u \in Dom(A)$, then

$$\begin{aligned} A\tilde{R}_\lambda u &= A\int_0^\infty e^{-\lambda t} S_t u \, dt \\ &= \int_0^\infty e^{-\lambda t} A S_t u \, dt \\ &= \int_0^\infty e^{-\lambda t} S_t A u \, dt \\ &= \tilde{R}_\lambda A u. \end{aligned}$$

Now, suppose that $(\lambda I - A)u = (\lambda I - A)v$ for some $u, v \in X$. Then, $\tilde{R}_\lambda u, \tilde{R}_\lambda v \in Dom(A)$ and so, by using part (a) and the above identity,

$$\begin{aligned} u &= (\lambda I - A)\tilde{R}_\lambda u \quad \text{(Part (a))} \\ &= \lambda I\tilde{R}_\lambda u - A\tilde{R}_\lambda u \\ &= \lambda I\tilde{R}_\lambda u - \tilde{R}_\lambda A u \\ &= \tilde{R}_\lambda(\lambda I - A)u \\ &= \tilde{R}_\lambda(\lambda I - A)v \quad \text{(By assumption)} \\ &= (\lambda I - A)\tilde{R}_\lambda v \\ &= v \quad \text{(Part (a))}. \end{aligned}$$

Therefore, $\lambda I - A$ is an injective operator.

Hence, $\lambda I - A$ is a bijective operator for all $\lambda > \omega$. Therefore, $\lambda \in \rho(A)$ and as a result, $\tilde{R}_\lambda = (\lambda I - A)^{-1} = R_\lambda$. $\square$

How do we interpret this formula? One way is to view the resolvent operators R_λ as the Laplace transform of the semigroup. With $h \in \mathbb{R}_{>0}$, write $h = 1/\lambda$. We can derive the backward Euler approximation as

$$E_h^- u = E_{1/\lambda}^- u = \lambda R_\lambda = \lambda \int_0^\infty e^{-t\lambda} S_t u \, dt = \int_0^\infty \frac{e^{-t/h}}{h} S_t u \, dt.$$

In order to ensure that the integral is convergent, we require $h < \omega^{-1}$, where S is a semigroup of type ω.

6.7 Existence and uniqueness of semigroups

Now, we are ready to tackle one of the main theorems pertaining to semigroups. The question we want to answer is given an operator $A : Dom(A) \to X$, when does there exist a semigroup generated by A?

Theorem 6.7.1 (Existence and uniqueness of a semigroup)**.** *Let X be a Banach space and $A : X \to X$ be a linear operator. Let $\omega \in \mathbb{R}_{\geq 0}$. Then, the following statements are equivalent:*

1. *A is the generator of a unique semigroup of linear operators $\{S_t\}_{t \in \mathbb{R}_{\geq 0}}$ of type ω.*
2. *A is a closed, densely defined operator. For all $\lambda > \omega$, $\lambda \in \rho(A)$ and*

$$\|(\lambda I - A)^{-1}\| \leq \frac{1}{\lambda - \omega}.$$

Proof. Assume that X is a Banach space and $A : X \to X$ is a linear operator. Assume that $\omega \in \mathbb{R}_{\geq 0}$. From the previous results, we have already shown that the first statement implies the second, up to uniqueness of the semigroup. Hence, we will show that A must generate a unique semigroup.

To show: (a) If $\{S_t\}_{t \in \mathbb{R}_{\geq 0}}$ and $\{S'_t\}_{t \in \mathbb{R}_{\geq 0}}$ are two semigroups of bounded linear operators generated by A, then for all $t \in \mathbb{R}_{\geq 0}$, $S_t = S'_t$.

(a) Assume that $u \in Dom(A)$. Then, $S_{t-s} \tilde{S}_s u \in Dom(A)$ for all $s, t \in \mathbb{R}_{\geq 0}$. First, we evaluate the following expression

$$
\begin{aligned}
\frac{d}{ds}[S_{t-s}\tilde{S}_s u] &= \lim_{h\to 0^+} \frac{S_{t-s-h}\tilde{S}_{s+h}u - S_{t-s}\tilde{S}_s u}{h} \\
&= \lim_{h\to 0^+} \frac{S_{t-s-h}\tilde{S}_{s+h}u - S_{t-s-h}\tilde{S}_s u + S_{t-s-h}\tilde{S}_s u - S_{t-s}\tilde{S}_s u}{h} \\
&= \lim_{h\to 0^+} \frac{S_{t-s-h}\tilde{S}_{s+h}u - S_{t-s-h}\tilde{S}_s u}{h} + \lim_{h\to 0^+} \frac{S_{t-s-h}\tilde{S}_s u - S_{t-s}\tilde{S}_s u}{h} \\
&= S_{t-s}(A\tilde{S}_s u) - AS_{t-s}\tilde{S}_s u.
\end{aligned}
$$

The reason why we care about this expression is because

$$
\begin{aligned}
S_t u - \tilde{S}_t u &= \int_0^t \frac{d}{ds}[S_{t-s}\tilde{S}_s u]\, ds \\
&= \int_0^t [S_{t-s}(A\tilde{S}_s u) - AS_{t-s}\tilde{S}_s u]\, ds \\
&= 0.
\end{aligned}
$$

This is because for all $u \in Dom(A)$, $S_{t-s}(A\tilde{S}_s u) = AS_{t-s}\tilde{S}_s u$. Now, since $Dom(A)$ is dense in X, we conclude that $S_t = \tilde{S}_t$ for all $t \in \mathbb{R}_{\geq 0}$ on X. Hence, the semigroup generated by A must be unique.

Now, we will show that statement 2 implies statement 1. This is the more difficult direction. Assume that A is closed, $\overline{Dom(A)} = X$, $\lambda \in \rho(A)$ for all $\lambda > \omega$ (with $\omega \in \mathbb{R}_{\geq 0}$) and

$$
\|(\lambda I - A)^{-1}\| \leq \frac{1}{\lambda - \omega}.
$$

To show: (b) A is the generator of a semigroup of type ω.

(b) From our assumption, the resolvent operator $R_\lambda = (\lambda I - A)^{-1}$ is well-defined for all $\lambda > \omega$. Hence, we can define the following bounded linear operator:

$$
A_\lambda u = -\lambda u + \lambda^2 R_\lambda u = -\lambda(\lambda I - A)R_\lambda u + \lambda^2 R_\lambda u = \lambda A R_\lambda u.
$$

It is bounded because A is bounded via 4.2.3 and R_λ is a bounded linear operator by assumption. Note that by writing $h = 1/\lambda$, we can write

$$
A_\lambda u = A(I - hA)^{-1} = A(E_h^- u).
$$

Since A_λ is bounded for all $\lambda > \omega$, we can consider its exponential map:

$$e^{tA_\lambda} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda^2 t)^k R_\lambda^k}{k!}.$$

Let us establish an upper bound for the norm of e^{tA_λ}:

$$\begin{aligned}
\|e^{tA_\lambda}\| &= \|e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda^2 t)^k R_\lambda^k}{k!}\| \\
&\leq e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda^2 t)^k \|R_\lambda\|^k}{k!} \\
&= \exp(-\lambda t + \lambda^2 t \|R_k\|) \\
&\leq \exp(-\lambda t + \lambda^2 t/(\lambda - \omega)) \\
&= \exp(\frac{\lambda \omega t}{\lambda - \omega}).
\end{aligned}$$

In particular, when $\lambda \geq 2\omega$, $\|e^{tA_\lambda}\| \leq e^{2\omega t}$. The point of defining the bounded operators A_λ is encapsulated in the following parts of the proof:

To show: (ba) For all $u \in Dom(A)$, $\lim_{\lambda \to \infty} A_\lambda u = Au$.

(bb) For all $u \in X$, $\lim_{\lambda \to \infty} A_\lambda u = Au$.

(ba) Assume that $u \in Dom(A)$. Then,

$$\lambda R_\lambda u - u = \lambda R_\lambda u - (\lambda I - A) R_\lambda u = A R_\lambda u = R_\lambda A u.$$

So, the standard argument shows that

$$\begin{aligned}
\|\lambda R_\lambda u - u\| &= \|R_\lambda A u\| \\
&\leq \|R_\lambda\| \|Au\| \\
&\leq \frac{1}{\lambda - \omega} \|Au\| \\
&\to 0 \text{ as } \lambda \to \infty.
\end{aligned}$$

This reveals that $\lim_{\lambda \to \infty} \lambda R_\lambda u = u$. Applying A to both sides and noting that A is a continuous linear operator (by 4.2.3), we find that

$$\lim_{\lambda \to \infty} A_\lambda u = \lim_{\lambda \to \infty} \lambda A R_\lambda u = Au.$$

(bb) Now assume that $v \in X$. Utilising the fact that A is densely defined ($\overline{Dom(A)} = X$), there exists a sequence $\{u_n\}$ in $Dom(A)$ such that $u_n \to v$ as $n \to \infty$. Now, there exists $N \in \mathbb{Z}_{>0}$ such that for all $n > N$, $\|u_n - v\| < \epsilon/2$. Now, we argue as follows:

$$\begin{aligned}
\|\lambda R_\lambda v - v\| &\le \|\lambda R_\lambda v - \lambda R_\lambda u_n\| + \|\lambda R_\lambda u_n - u_n\| + \|u_n - v\| \\
&= \|\lambda R_\lambda (v - u_n)\| + \|\lambda R_\lambda u_n - u_n\| + \frac{\epsilon}{2} \\
&\le \|\lambda R_\lambda\| \|v - u_n\| + \|\lambda R_\lambda u_n - u_n\| + \frac{\epsilon}{2} \\
&< \lambda \|R_\lambda\| \frac{\epsilon}{2} + \|\lambda R_\lambda u_n - u_n\| + \frac{\epsilon}{2}.
\end{aligned}$$

Taking the limit as $\lambda \to \infty$, we find that

$$\lim_{\lambda \to \infty} \|\lambda R_\lambda v - v\| < \lim_{\lambda \to \infty} (\frac{\lambda}{\lambda - \omega} \frac{\epsilon}{2} + \frac{\epsilon}{2}) = \epsilon.$$

Hence, $\lim_{\lambda \to \infty} \lambda R_\lambda v = v$. Applying the operator A, we find that $\lim_{\lambda \to \infty} A_\lambda v = Av$.

Our next claim is that the family of uniformly bounded operators e^{tA_λ} converges to a linear operator, called S_t. In order to see this, we will take $\lambda, \mu > 2\omega$ and estimate the difference $e^{tA_\lambda} - e^{tA_\mu}$.

For all $u \in X$, we have

$$\begin{aligned}
e^{tA_\lambda} u - e^{tA_\mu} u &= \int_0^t \frac{d}{ds} [e^{(t-s)A_\mu} e^{sA_\lambda} u] \, ds \\
&= \int_0^t (A_\lambda - A_\mu) e^{(t-s)A_\mu} e^{sA_\lambda} u \, ds.
\end{aligned}$$

Now, we take the norm of both sides to get

$$\begin{aligned}
\|e^{tA_\lambda} u - e^{tA_\mu} u\| &\le \int_0^t \|A_\lambda u - A_\mu u\| \|e^{(t-s)A_\mu}\| \|e^{sA_\lambda}\| \, ds \\
&= \int_0^t \|A_\lambda u - A_\mu u\| e^{2\omega(t-s)} e^{2\omega s} \, ds \\
&= te^{2\omega t} \|A_\lambda u - A_\mu u\|.
\end{aligned}$$

Using the result proved in part (bb), we can take the limit as $\lambda, \mu \to \infty$ to obtain for all $u \in X$,

$$\limsup_{\lambda,\mu\to\infty}\|e^{tA_\lambda}u - e^{tA_\mu}u\| \leq te^{2\omega t}\limsup_{\lambda,\mu\to\infty}\|A_\lambda u - A_\mu u\| = 0.$$

Therefore, for all $t \in \mathbb{R}_{\geq 0}$ and $u \in X$, the below limit is well-defined

$$S_t u = \lim_{\lambda\to\infty} e^{tA_\lambda}u.$$

We will show that $\{S_t\}_{t\in\mathbb{R}_{\geq 0}}$ is a strongly continuous semigroup of type ω.

To show: (bc) For all $s, t \in \mathbb{R}_{\geq 0}$ and $u \in X$, $S_t S_s u = S_{t+s}u$.

(bd) For all $t \in \mathbb{R}_{\geq 0}$ and $u \in X$, $\|S_t\| \leq e^{t\omega}$.

(bc) Assume that $s, t \in \mathbb{R}_{\geq 0}$ and $u \in X$. Then, by the definition,

$$\begin{aligned}
S_t S_s u &= \lim_{\lambda\to\infty} \exp(tA_\lambda)\exp(sA_\lambda)u \\
&= \lim_{\lambda\to\infty} \exp((t+s)A_\lambda)u \\
&= S_{t+s}u.
\end{aligned}$$

This proves the semigroup property.

(bd) We will compute an upper bound for $\|S_t u\|$:

$$\begin{aligned}
\|S_t u\| &= \|\lim_{\lambda\to\infty} e^{tA_\lambda}u\| \\
&= \lim_{\lambda\to\infty} \|e^{tA_\lambda}u\| \\
&\leq \limsup_{\lambda\to\infty}\|e^{tA_\lambda}\|\|u\| \\
&\leq \lim_{\lambda\to\infty} e^{t\lambda\omega/(\lambda-\omega)}\|u\| \\
&= e^{t\omega}\|u\|.
\end{aligned}$$

Thus, taking the supremum over all $u \in X$ such that $\|u\| \leq 1$, we find that $\|S_t\| < e^{t\omega}$. So, the semigroup must be of type ω.

Now, it remains to prove that A generates the semigroup $\{S_t\}_{t\in\mathbb{R}_{\geq 0}}$. Let B be the operator which generates $\{S_t\}_{t\in\mathbb{R}_{\geq 0}}$. Then, we must show that $A = B$. We will proceed in three steps:

To show: (be) $Dom(A) \subseteq Dom(B)$.

(bf) For all $u \in Dom(A)$, $Bu = Au$.

(bg) $Dom(B) \subseteq Dom(A)$.

(be) Assume that $u \in Dom(A)$. Then, for all $\lambda > \omega$, we must have

$$e^{tA_\lambda}u - u = \int_0^t e^{sA_\lambda}A_\lambda u \, ds.$$

The triangle inequality tells us that

$$\|e^{sA_\lambda}A_\lambda u - S_s Au\| \leq \|e^{sA_\lambda}\|\|A_\lambda u - Au\| + \|e^{sA_\lambda}Au - S_s Au\|.$$

So, as $\lambda \to \infty$, the RHS approaches 0 because $\lim_{\lambda\to\infty} A_\lambda u = Au$ and $\lim_{\lambda\to\infty} e^{sA_\lambda}u = S_s u$. Thus, when $\lambda \to \infty$,

$$\lim_{\lambda\to\infty}(e^{tA_\lambda}u - u) = S_t u - u = \int_0^t S_s Au \, ds$$

for all $t \in \mathbb{R}_{\geq 0}$ and $u \in Dom(A)$. This is equivalent to showing that the limit

$$\lim_{t\to 0^+} \frac{S_t u - u}{t}$$

exists. Thus, $u \in Dom(B)$ and $Dom(A) \subseteq Dom(B)$.

(bf) Expanding the definition of Bu, we find that

$$\begin{aligned} Bu &= \lim_{h\to 0^+} \frac{S_t u - u}{t} \\ &= \lim_{h\to 0^+} \frac{1}{t}\int_0^t S_s Au \, ds \\ &= Au. \end{aligned}$$

This means that $Au = Bu$ for all $u \in Dom(A)$.

(bg) Now choose $\lambda > \omega$. Then, the operators $\lambda I - A : Dom(A) \to X$ and $\lambda I - B : Dom(B) \to X$ are both injective and surjective. By using the previous two parts of the proof, we note that the restriction $(\lambda I - B)|_{Dom(A)} = \lambda I - A$. The restricted map must be surjective. The point here is that the operator $\lambda I - B$ cannot be extended to any domain strictly larger than $Dom(A)$, due to the surjectivity. So, $Dom(A) = Dom(B)$ as required. This finally completes the proof. $\square$

Chapter 7

Sobolev Spaces

7.1 Introducing weak derivatives

Sobolev spaces are an important tool in the study of partial differential equations. The leitmotif of this section is to not look for *strong solutions* of partial differential equations, but rather so-called *weak solutions*, which will be defined later. It turns out that weak solutions to PDEs are contained in appropriate Sobolev spaces. Thus, they are more suitable to the study of PDEs than C^n functions for all $n \in \mathbb{Z}_{>0}$ or Hölder continuous functions.

There are several important ideas which need to be fleshed out. The first of these is the concept of *test functions*. Here is the first definition of the section:

Definition 7.1.1. Let $\Omega \subseteq \mathbb{R}^n$ be an open subset of $\mathbb{R}^n$. Then, define C_c^∞ to be the space of smooth functions $f : \Omega \to \mathbb{R}$ which are compactly supported. That is, for all $f \in C_c^\infty(\Omega)$, the support

$$supp(f) = \overline{\{x \in \Omega \mid f(x) \neq 0\}}$$

is a compact subset of Ω. Functions in C_c^∞ are called **test functions**.

Test functions provide a segue into the definition of a weak derivative. The reason for this is the integration by parts formula. Let $x_1, \dots, x_n$ be coordinates for $\mathbb{R}^n$. If $\Omega \subseteq \mathbb{R}^n$ is open, $u \in C^1(\Omega)$ and $\phi \in C_c^\infty(\Omega)$, then

$$\int_\Omega u \frac{\partial \phi}{\partial x_i} \, dx = -\int_\Omega \frac{\partial u}{\partial x_i} \phi \, dx$$

for all $i \in \{1, \dots, n\}$. In short, since ϕ has compact support in Ω, it must vanish at the boundary $\partial\Omega$.

Note that if $u \in C^k(\Omega)$ for some $k \in \mathbb{Z}_{>0}$, we have the following more general identity for all $i \in \{1, \ldots, n\}$:

$$\int_\Omega u \frac{\partial^k \phi}{\partial x_i^k} \, dx = (-1)^k \int_\Omega \frac{\partial^k u}{\partial x_i^k} \phi \, dx$$

The next definition we will make pertains to locally integrable functions.

Definition 7.1.2. Let $\Omega \subseteq \mathbb{R}^n$ be an open set. We define $L^1_{loc}(\Omega)$ to be the space of **locally integrable functions** on Ω. That is, if $f \in L^1_{loc}(\Omega)$, then f is a Lebesgue measurable function from Ω to $\mathbb{R}$ which is (Lebesgue) integrable when restricted to every compact subset $K \subseteq \Omega$.

For example, the functions $e^x, \log|x|$ and $|x|^{-1/2}$ are all in $L^1_{loc}(\mathbb{R})$, whereas $x^{-1} \notin L^1_{loc}(\mathbb{R})$. On the other hand, if $\Omega = (0, \infty)$, then $x^k \in L^1_{loc}(\Omega)$ for all $k \in \mathbb{Z}$.

The most important point here is that a locally integrable function $f \in L^1_{loc}(\Omega)$ determines the linear functional

$$\Lambda_f : C_c^\infty(\Omega) \to \mathbb{R}$$

$$\varphi \mapsto \int_\Omega f(x)\varphi(x) \, dm(x).$$

It is linear due to the properties of Lebesgue integration. Now, we need some extra notation. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an n-tuple of non-negative integer numbers. The tuple α is commonly known as a **multi-index**. Its length is given by $|\alpha| = \sum_{i=1}^n \alpha_i$. The point is that each multi-index corresponds to a partial differential operator of order $|\alpha|$:

$$D^\alpha f = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} f$$

Now we are ready to define distributions.

Definition 7.1.3. Let $\Omega \subseteq \mathbb{R}^n$ be an open set. A **distribution** is a linear functional $\Lambda : C_c^\infty(\Omega) \to \mathbb{R}$ such that for all compact subsets $K \subseteq \Omega$, there exists $N_K \in \mathbb{Z}_{\geq 0}$ and $C_K \in \mathbb{R}_{>0}$ such that for all $\phi \in C_c^\infty$ such that $supp(\phi) \subseteq K$,

$$|\Lambda(\phi)| \leq C_K \|\phi\|_{C^{N_K}}.$$

Here, we have for all $N \in \mathbb{Z}_{\geq 0}$,

$$\|\phi\|_{C^N} = \sup_{x \in K} \{D^\alpha \phi(x) \mid |\alpha| \leq N\}.$$

It should be stressed here that both N_K and C_K depend on the compact subset K.

Definition 7.1.4. If there exists an integer $N \in \mathbb{Z}_{\geq 0}$ independent o K such that the above definition holds (with the constant C_K still depending on K), then the distribution has **finite order**. The smallest such integer N is the **order** of the distribution.

We already know one example of a distribution.

Example 7.1.1. Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Let $f \in L^1_{loc}(\Omega)$. Then, as discussed before, f induces the linear functional

$$\Lambda_f : C_c^\infty(\Omega) \to \mathbb{R}$$

$$\varphi \mapsto \int_\Omega f(x)\varphi(x)\ dm(x).$$

We claim that this is a distribution. Suppose that $\phi \in C_c^\infty(\Omega)$ such that $supp(\phi) \subseteq K \subseteq \Omega$, where K is a compact subset of Ω. Then,

$$\begin{aligned}
|\Lambda_f(\phi)| &= |\int_\Omega f\phi\ dx| \\
&= |\int_K f\phi\ dx| \\
&\leq \sup_{x\in K}|\phi(x)||\int_K f\ dx| \\
&\leq \sup_{x\in K}|\phi(x)|\int_K |f|\ dx \\
&\leq C\|\phi\|_{C^0}.
\end{aligned}$$

where $C = \int_K |f|\ dx$ and $N = 0$, regardless of the choice of compact subset K. Therefore, Λ_f is a distribution of order zero.

The set of all distributions of Ω is a vector space, as the next lemma demonstrates.

Lemma 7.1.1. *Let $\Omega \subseteq \mathbb{R}^n$. Then, the set of all distributions on Ω forms a $\mathbb{R}$-vector space.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$. Let $\mathcal{D}(\Omega)$ denote the set of all distributions on Ω. We will show that $\mathcal{D}(\Omega)$ is closed under addition and scalar multiplication.

To show: (a) If $\Lambda, \Phi \in \mathcal{D}(\Omega)$, then $\Lambda + \Phi \in \mathcal{D}(\Omega)$.

(b) If $\Lambda \in \mathcal{D}(\Omega)$ and $\alpha \in \mathbb{R}$, then $\alpha\Lambda \in \mathcal{D}(\Omega)$.

(a) Assume that $\Lambda, \Phi \in \mathcal{D}(\Omega)$. Then, for all compact subsets $K \subseteq \Omega$, there exists $C_K, D_K \in \mathbb{R}_{>0}$ and $M_K, N_K \in \mathbb{Z}_{\geq 0}$ such that for all $\varphi \in C_c^\infty(\Omega)$ with $supp(\varphi) \subseteq K$,

$$|\Lambda(\varphi)| \leq C_K \|\varphi\|_{C^{M_K}} \text{ and } |\Phi(\varphi)| \leq D_K \|\varphi\|_{C^{N_K}}.$$

Now observe that

$$\begin{aligned}
|(\Lambda + \Phi)(\varphi)| &= |\Lambda(\varphi) + \Phi(\varphi)| \\
&\leq |\Lambda(\varphi)| + |\Phi(\varphi)| \\
&\leq C_K \|\varphi\|_{C^{M_K}} + D_K \|\varphi\|_{C^{N_K}} \\
&\leq (C_K + D_K) \|\varphi\|_{C^{P_K}}.
\end{aligned}$$

Here, $P_K = \max(C_K, D_K)$. This shows that $\Lambda + \Phi \in \mathcal{D}(\Omega)$.

(b) Now assume that $\alpha \in \mathbb{R}$. Then,

$$\begin{aligned}
|(\alpha\Lambda)(\varphi)| &= |\alpha||\Lambda(\varphi)| \\
&\leq |\alpha| C_K \|\varphi\|_{C^{M_K}}.
\end{aligned}$$

Thus, $\alpha\Lambda \in \mathcal{D}(\Omega)$.

Therefore, $\mathcal{D}(\Omega)$ is a $\mathbb{R}$-vector space.

□

Interestingly, there is more to the vector space $\mathcal{D}(\Omega)$ which meets the eye. It is tied to the idea that while an arbitrary function f may not admit a classical derivative, for a distribution Λ, a "derivative" can *always* be defined. It turns out that $\mathcal{D}(\Omega)$ is also closed under the partial differential operator D^α. We will prove this statement below.

Theorem 7.1.2. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Let $\Lambda : C_c^\infty(\Omega) \to \mathbb{R}$ be a distribution and α a multi-index. Define the map*

$$D^\alpha \Lambda(\phi) = (-1)^{|\alpha|}\Lambda(D^\alpha \phi).$$

Then, $D^\alpha \Lambda$ is itself a distribution.

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is an open set. Assume that $\Lambda \in \mathcal{D}(\Omega)$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index. Assume that $D^\alpha \Lambda$ is defined as above. From its definition, $D^\alpha \Lambda$ must be a linear functional. Assume that $K \subseteq \Omega$ is a compact set and $\varphi \in C_c^\infty(\Omega)$, with $supp(\varphi) \subseteq K$. Now, we argue as follows:

$$\begin{aligned} |D^\alpha \Lambda(\varphi)| &= |(-1)^{|\alpha|}\Lambda(D^\alpha \varphi)| \\ &= |\Lambda(D^\alpha \varphi)| \\ &\leq C_K \|D^\alpha \varphi\|_{C^{N_K}} \\ &\leq C_K \|\varphi\|_{C^{N_K + |\alpha|}}. \end{aligned}$$

So, $D^\alpha \Lambda$ must be a distribution.

□

The reason for the above definition of D^α is that if $\Lambda_f : C_c^\infty(\Omega) \to \mathbb{R}$ is the linear functional associated with $f \in L^1_{loc}(\Omega)$, then for all $\phi \in C_c^\infty(\Omega)$ with $supp(\phi)$ contained in some compact subset of Ω

$$\begin{aligned} D^\alpha \Lambda_f(\phi) &= (-1)^{|\alpha|}\Lambda_f(D^\alpha \phi) \\ &= (-1)^{|\alpha|} \int_\Omega f D^\alpha \phi \, dx \\ &= \int_\Omega D^\alpha f \phi \, dx \\ &= \Lambda_{D^\alpha f}(\phi). \end{aligned}$$

With all of this in mind, we have now reached the definition of a weak derivative.

Definition 7.1.5. Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $f \in L^1_{loc}(\Omega)$. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a multi-index and Λ_f be the linear functional associated with f. If there exists a function $g \in L^1_{loc}(\Omega)$ such that $D^\alpha \Lambda_f = \Lambda_g$ or equivalently

$$\int_\Omega f D^\alpha \phi \, dx = (-1)^{|\alpha|} \int_\Omega g\phi \, dx$$

for all $\phi \in C_c^\infty(\Omega)$, then g is called the **weak α-th derivative** of f.

We will now prove some basic properties of weak derivatives and then discuss a few selected examples. The first result demonstrates that weak derivatives are unique, up to a set of Lebesgue measure zero.

Lemma 7.1.3. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $f \in L^1_{loc}(\Omega)$. Then, if the weak derivative of f exists, it is unique almost everywhere.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is an open set and $f \in L^1_{loc}(\Omega)$. Suppose that $g, \tilde{g} \in L^1_{loc}(\Omega)$ such that for all $\phi \in C_c^\infty(\Omega)$,

$$\int_\Omega f D^\alpha \phi \, dx = (-1)^{|\alpha|} \int_\Omega g\phi \, dx = (-1)^{|\alpha|} \int_\Omega \tilde{g}\phi \, dx$$

for some multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$. Rearranging the above equation, we find that for all $\phi \in C_c^\infty(\Omega)$,

$$\int_\Omega (g - \tilde{g})\phi \, dx = 0.$$

This is enough to demonstrate that $g(x) = \tilde{g}(x)$ for almost all $x \in \Omega$. □

Weak derivative also satisfy a nice convergence property, as depicted below.

Lemma 7.1.4. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $\{u_n\}$ be a sequence in $L^1_{loc}(\Omega)$ such that each u_n has an α-th weak derivative $D^\alpha u_n$ (here, $\alpha = (\alpha_1, \dots, \alpha_n)$ is a multi-index). Then, if $u_n \to u$ and $D^\alpha u_n \to v_\alpha$ in $L^1_{loc}(\Omega)$, then $v_\alpha = D^\alpha u$.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is open and $\{u_n\}$ is a sequence in $L^1_{loc}(\Omega)$. Assume that $u_n \to u$ and $D^\alpha u_n \to v_\alpha$ in $L^1_{loc}(\Omega)$. Then, we compute for every test function $\phi \in C_c^\infty(\Omega)$,

$$\begin{aligned}
\int_\Omega v_\alpha \phi \, dx &= \int_\Omega \lim_{n\to\infty} D^\alpha u_n \phi \, dx \\
&= \lim_{n\to\infty} \int_\Omega D^\alpha u_n \phi \, dx \\
&= \lim_{n\to\infty} (-1)^{|\alpha|} \int_\Omega u_n D^\alpha \phi \, dx \\
&= (-1)^{|\alpha|} \int_\Omega u D^\alpha \phi \, dx.
\end{aligned}$$

Therefore, $v_\alpha = D^\alpha u$.

□

Now, we will cover some examples of weak derivatives.

Example 7.1.2. Set $\Omega = \mathbb{R}$ and consider the function

$$f(x) = \begin{cases} 0, & \text{if } x \le 0, \\ x, & \text{if } x > 0. \end{cases} \quad \in L^1_{loc}(\mathbb{R}).$$

What is the first weak derivative of f? Suppose that $\varphi \in C_c^\infty(\mathbb{R})$. Then, we use integration by parts to compute that

$$\begin{aligned} D^{(1)}\Lambda_f(\varphi) &= -\Lambda_f(D^{(1)}\varphi) \\ &= -\int_{\mathbb{R}} f(x) D^{(1)}\varphi(x)\, dx \\ &= -\int_0^\infty x\varphi'(x)\, dx \\ &= \int_0^\infty \varphi(x)\, dx \quad \text{(Integration by parts)} \\ &= \int_{\mathbb{R}} H(x)\varphi(x)\, dx. \end{aligned}$$

So, we can conclude that the first weak derivative of the function $f(x)$ is the Heaviside step function $H(x)$.

Now, observe that $H(x) \in L^1_{loc}(\mathbb{R})$. What is the first weak derivative of $H(x)$? Again, we can use the definition to calculate

$$\begin{aligned} D^{(1)}\Lambda_H(\varphi) &= -\Lambda_H(D^{(1)}\varphi) \\ &= -\int_{\mathbb{R}} H(x) D^{(1)}\varphi(x)\, dx \\ &= -\int_0^\infty \varphi'(x)\, dx \\ &= \varphi(0) - \varphi(\infty) \\ &= \varphi(0) \quad (\varphi \text{ vanishes at the boundary of } \mathbb{R}). \\ &= \int_{\mathbb{R}} \delta(x)\varphi(x)\, dx. \end{aligned}$$

Note that $\delta(x)$ is the Dirac delta function. In fact, this does **not** constitute a weak derivative of $H(x)$. Intuitively, this is because the Dirac delta function is not actually a function. To see why rigorously, suppose that for some function $g \in L^1_{loc}(\mathbb{R})$ that for all $\varphi \in C_c^\infty(\mathbb{R})$,

$$\int g(x)\varphi(x)\, dx = \varphi(0).$$

Then, the dominated convergence theorem tells us that

$$\lim_{h\to 0} \int_{-h}^{h} |g(x)|\, dx = 0.$$

Choose $\epsilon \in \mathbb{R}_{>0}$ so that $\int_{-\delta}^{\delta} |g(x)|\, dx \leq 1/2$. Let $\eta : \mathbb{R} \to [0,1]$ be a smooth function with $\eta(0) = 1$ and $supp(\eta) \subseteq [-\epsilon, \epsilon]$. Then,

$$\begin{aligned} 1 &= \varphi(0) \\ &= \int_{\mathbb{R}} g(x)\varphi(x)\, dx \\ &= \int_{-\delta}^{\delta} g(x)\varphi(x)\, dx \\ &\leq \sup_{x\in\mathbb{R}} |\varphi(x)| \int_{-\delta}^{\delta} g(x)\, dx \\ &\leq \frac{1}{2}. \end{aligned}$$

This is a contradiction.

The next example highlights the difference between a weak derivative and the classical notion of the derivative.

Example 7.1.3. Let f denote the function

$$f(x) = \begin{cases} 0, & \text{if } x \in \mathbb{Q}, \\ 2 + \sin x, & \text{otherwise.} \end{cases}$$

By definition f is discontinuous at every point $x \in \mathbb{R}$. So, it is not differentiable at any point. However, f does have a first weak derivative. In particular, since $m(\mathbb{Q}) = 0$, we can disregard the behaviour of f on $\mathbb{Q}$ and thus, we have

$$D^{(1)}\Lambda_f(\varphi) = -\int f(x)\varphi'(x)\ dx = -\int (2+\sin x)\varphi'(x)\ dx = \int \cos x\ \varphi(x)\ dx.$$

So, $\cos x$ constitutes the first weak derivative of $f(x)$.

7.2 Mollifications

The primary use of mollifications is to approximate general functions with smooth functions. They can be likened to the role of bump functions and partitions of unity in differential geometry.

Definition 7.2.1. Denote by $J : \mathbb{R}^n \to \mathbb{R}$ the **standard mollifier** on $\mathbb{R}^n$:

$$J(x) = \begin{cases} C_n \exp(-\frac{1}{|x|^2-1}) \text{ if } |x| < 1, \\ 0 \text{ otherwise.} \end{cases}$$

Here the constant C_n is a normalisation constant, ensuring that $\int_{\mathbb{R}^n} J(x)\ dx = 1$.

Definition 7.2.2. For all $\epsilon \in \mathbb{R}_{>0}$ we also define

$$J_\epsilon(x) = \frac{1}{\epsilon^n} J(\frac{x}{\epsilon}).$$

Some properties of the standard mollifier and its rescaled variants include: for all $\epsilon \in \mathbb{R}_{>0}$, $J_\epsilon \in C_c^\infty(\mathbb{R}^n)$ with $supp(J_\epsilon) = \{x \in \mathbb{R}^n \mid |x| \leq \epsilon\}$ and

$$\int_{\mathbb{R}^n} J_\epsilon(x)\ dx = 1.$$

Below, we define the mollification of a locally integrable function f:

Definition 7.2.3. Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $f \in L^1_{loc}(\Omega)$. For all $\epsilon \in \mathbb{R}_{>0}$, the **mollification** of f is defined to be the function

$$f_\epsilon(x) = J_\epsilon(x) * f(x) = \int_{B(x,\epsilon)} J_\epsilon(x-y) f(y)\ dy.$$

Note that due to the properties of J_ϵ and f, the above convolution is well-defined at all points in the subset

$$\Omega_\epsilon = \{x \in \Omega \mid \overline{B(x,\epsilon)} \subseteq \Omega\}.$$

As with any newly defined mathematical object, we will prove some properties mollifications satisfy. In the process, the properties of mollifications will highlight why they are suitable for approximating general functions.

Theorem 7.2.1 (Properties of Mollifiers). *Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $f \in L^1_{loc}(\Omega)$. Let $\epsilon \in \mathbb{R}_{>0}$. Then,*

1. $f_\epsilon \in C^\infty(\Omega_\epsilon)$.
2. $\lim_{\epsilon \to 0} f_\epsilon = f$ *for almost all* $x \in \Omega$.
3. *If f is continuous, then $f_\epsilon \to f$ uniformly on compact subsets of Ω.*
4. *If $p \in [1, \infty)$ and $f \in L^p_{loc}(\Omega)$, then $f_\epsilon \to f$ in $L^p_{loc}(\Omega)$.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is an open set and $f \in L^1_{loc}(\Omega)$. Assume that $\epsilon \in \mathbb{R}_{>0}$.

To show: (a) $f_\epsilon \in C^\infty(\Omega_\epsilon)$.

(b) $\lim_{\epsilon \to 0} f_\epsilon = f$ for almost all $x \in \Omega$.

(c) If f is continuous, then $f_\epsilon \to f$ uniformly on compact subsets of Ω.

(d) If $p \in [1, \infty)$ and $f \in L^p_{loc}(\Omega)$, then $f_\epsilon \to f$ in $L^p_{loc}(\Omega)$.

(a) Assume that $\{e_1, \dots, e_n\}$ is the standard basis for $\mathbb{R}^n$. Assume that $x \in \Omega_\epsilon$ and $i \in \{1, \dots, n\}$. Fix $h \in \mathbb{R}_{>0}$ small enough so that $x + he_i \in \Omega_\epsilon$. Using the definition of f_ϵ, we compute the following quotient:

$$\begin{aligned}\frac{f_\epsilon(x + he_i) - f_\epsilon(x)}{h} &= \frac{1}{h} \int_{B(x,\epsilon)} (J_\epsilon(x + he_i - y) - J_\epsilon(x - y)) f(y)\, dy \\ &= \frac{1}{\epsilon^n} \int_{B(x,\epsilon)} \frac{1}{h} (J(\frac{x + he_i - y}{\epsilon}) - J(\frac{x - y}{\epsilon})) f(y)\, dy.\end{aligned}$$

Since the closed ball $\overline{B(x, \epsilon)}$ is contained in Ω by definition, $f \in L^1(B(x, \epsilon))$. This is because each Ω_ϵ is open in $\mathbb{R}^n$. Now, since J is smooth, we have the following limit as $h \to 0$:

$$\lim_{h \to 0} \frac{1}{h} [J(\frac{x + he_i - y}{\epsilon}) - J(\frac{x - y}{\epsilon})] = \frac{1}{\epsilon} \frac{\partial J}{\partial x_i} (\frac{x - y}{\epsilon})$$

From the previous equation, we can take the limit as $h \to 0$ of both sides to obtain

$$\begin{aligned}
\frac{\partial f_\epsilon}{\partial x_i}(x) &= \frac{1}{\epsilon^n}\int_{B(x,\epsilon)} \frac{1}{\epsilon}\frac{\partial J}{\partial x_i}(\frac{x-y}{\epsilon})f(y)\, dy \\
&= \int_{B(x,\epsilon)} \frac{1}{\epsilon^{n+1}}\frac{\partial J}{\partial x_i}(\frac{x-y}{\epsilon})f(y)\, dy \\
&= \int_{B(x,\epsilon)} \frac{\partial J_\epsilon}{\partial x_i}(x-y)f(y)\, dy.
\end{aligned}$$

By iterating the above argument, we can demonstrate that for all multi-indices α, the partial derivative $D^\alpha f_\epsilon$ exists and is continuous from its integral definition. This is enough to show that $f_\epsilon \in \Omega_\epsilon$.

(b) We will invoke the Lebesgue differentiation theorem, which demonstrates that for almost all $x \in \Omega$, we have

$$\begin{aligned}
0 &= |f(x) - f(x)| \\
&= \lim_{r\to 0} A_r|f(y) - f(x)| \\
&= \lim_{r\to 0} \frac{1}{m(B(x,r))}\int_{B(x,r)} |f(y) - f(x)|\, dy.
\end{aligned}$$

Now, let $x \in \Omega$ be a point such that the above identity holds. Then, we argue as follows:

$$\begin{aligned}
|f_\epsilon(x) - f(x)| &= |\int_{B(x,\epsilon)} J_\epsilon(x-y)f(y)\, dy - f(x)| \\
&= |\int_{B(x,\epsilon)} J_\epsilon(x-y)f(y)\, dy - \int_{B(x,\epsilon)} J_\epsilon(x-y)f(x)\, dy| \\
&\leq |\int_{B(x,\epsilon)} J_\epsilon(x-y)[f(y) - f(x)]\, dy| \\
&\leq \frac{1}{\epsilon^n}\int_{B(x,\epsilon)} J(\frac{x-y}{\epsilon})|f(y) - f(x)|\, dy \\
&\leq \frac{C}{m(B(x,\epsilon))}\int_{B(x,\epsilon)} |f(y) - f(x)|\, dy \quad (C \in \mathbb{R}_{>0}) \\
&\to 0 \quad \text{(Lebesgue differentiation theorem)}
\end{aligned}$$

as $\epsilon \to 0$. Note that the second equality in the above working is due to the normalisation of J_ϵ. This demonstrates that $f_\epsilon \to f$ for almost all $x \in \Omega$.

(c) Assume that f is continuous and $K \subseteq \Omega$ be a compact subset. Choose $\delta \in \mathbb{R}_{>0}$ small enough so that the compact neighbourhood

$$K_\delta = \{x \in \mathbb{R}^n \mid d(x, K) \leq \delta\}$$

which contains K is still contained in Ω. Since f is continuous, it must be uniformly continuous on the compact set K_δ. This means that for all $\epsilon \in \mathbb{R}_{>0}$ and $x \in K$, there exists $\delta \in \mathbb{R}_{>0}$ such that if $|y - x| < \delta$, then

$$y \in K_\delta \text{ and } |f(x) - f(y)| < \epsilon.$$

By applying a similar computation to part (b), we have

$$\begin{aligned} |f_\epsilon(x) - f(x)| &= |\int_\Omega J_\epsilon(x-y)f(y)\,dy - f(x)| \\ &= |\int_\Omega J_\epsilon(x-y)f(y)\,dy - \int_\Omega J_\epsilon(x-y)f(x)\,dy| \\ &\leq \int_\Omega |J_\epsilon(x-y)[f(y) - f(x)]|\,dy < \epsilon. \end{aligned}$$

We carefully note that in the last line, we used uniform continuity on the compact set K_δ. Since $x \in K$ was arbitrary, we obtain uniform convergence $f_\epsilon(x) \to f(x)$ for all $x \in K$.

(d) Assume that $p \in [1, \infty)$ and $f \in L^p_{loc}(\Omega)$. Construct K_δ in the same way as before. The claim is that for all $\epsilon \in (0, p)$,

$$\|f_\epsilon\|_{L^p(K)} \leq \|f\|_{L^p(K_\delta)}$$

Define $q = p/(p-1)$. Then, for all $x \in K$, Hölder's inequality tells us that

$$
\begin{aligned}
|f_\epsilon(x)| &= |\int_{B(x,\epsilon)} J_\epsilon(x-y)f(y)\ dy| \\
&\leq \int_{B(x,\epsilon)} J_\epsilon(x-y)|f(y)|\ dy \\
&= \int_{B(x,\epsilon)} (J_\epsilon(x-y))^{\frac{p-1}{p}}(J_\epsilon(x-y))^{\frac{1}{p}}|f(y)|\ dy \\
&\leq (\int_{B(x,\epsilon)} J_\epsilon(x-y)\ dy)^{\frac{p-1}{p}}(\int_{B(x,\epsilon)} J_\epsilon(x-y)|f(y)|^p\ dy)^{\frac{1}{p}} \quad \text{(Hölder's inequality)} \\
&= (\int_{B(x,\epsilon)} J_\epsilon(x-y)|f(y)|^p\ dy)^{\frac{1}{p}} \quad \text{(Normalisation of } J_\epsilon).
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\|f_\epsilon\|_{L^p(K)} &= \int_K |f_\epsilon(x)|^p\ dx \\
&\leq \int_K (\int_{B(x,\epsilon)} J_\epsilon(x-y)|f(y)|^p\ dy)\ dx \\
&= \int_K |f(y)|^p (\int_{B(x,\epsilon)} J_\epsilon(x-y)\ dx)\ dy \\
&\leq \int_{K_\delta} |f(y)|^p (\int_{B(x,\epsilon)} J_\epsilon(x-y)\ dx)\ dy \\
&= \int_{K_\delta} |f(y)|^p\ dy \\
&= \|f\|_{L^p(K_\delta)}.
\end{aligned}
$$

Now, we use the fact that continuous functions are dense in L^p in order to choose a continuous function $g \in Cts(K_\delta, \mathbb{R})$ such that for all $\delta \in \mathbb{R}_{>0}$, $\|f-g\|_{L^p(K_\delta)} < \delta$. Combining this with the above result, we find that

$$
\begin{aligned}
\|f_\epsilon - f\|_{L^p(K)} &\leq \|f_\epsilon - g_\epsilon\|_{L^p(K)} + \|g_\epsilon - g\|_{L^p(K)} + \|g - f\|_{L^p(K)} \\
&= \|f - g\|_{L^p(K_\delta)} + \|g_\epsilon - g\|_{L^p(K)} + \|g - f\|_{L^p(K)} \\
&< \delta + \|g_\epsilon - g\|_{L^p(K)} + \delta.
\end{aligned}
$$

Since g is continuous, we can use the results in parts (b) and (c) to deduce that $|g_\epsilon - g| \to 0$ uniformly on the compact set K_δ. This means that if we take the limit $\epsilon \to 0$,

$$\limsup_{\epsilon \to 0} \|f_\epsilon - f\|_{L^p(K)} \leq 2\delta.$$

Since $\delta \in \mathbb{R}_{>0}$ was arbitrary, we deduce that $f_\epsilon \to f$ in $L^p(K)$. Since this holds for all compact subsets $K \subseteq \Omega$, we deduce that $f_\epsilon \to f$ in $L^p_{loc}(\Omega)$.

□

Here is a corollary of the above theorem

Corollary 7.2.2. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $f \in L^1_{loc}(\Omega)$ such that for all $\phi \in C_c^\infty(\Omega)$,*

$$\int_\Omega f\phi \, dx = 0.$$

Then, $f(x) = 0$ for almost all $x \in \Omega$.

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is an open set and $f \in L^1_{loc}(\Omega)$ such that for all $\phi \in C_c^\infty(\Omega)$,

$$\int_\Omega f\phi \, dx = 0.$$

Take $\phi(y) = J_\epsilon(x - y)$ for some $\epsilon \in \mathbb{R}_{>0}$. Then, using 7.2.1, we deduce that for almost all $x \in \Omega$,

$$\begin{aligned} f(x) &= \lim_{\epsilon \to 0} f_\epsilon(x) \\ &= \lim_{\epsilon \to 0} \int_{B(x,\epsilon)} J_\epsilon(x - y) f(y) \, dy \\ &= \lim_{\epsilon \to 0} \int_\Omega J_\epsilon(x - y) f(y) \, dy \\ &= 0. \end{aligned}$$

The third equality in the above working is due to the fact that $supp(J_\epsilon) \subseteq \Omega$ for an appropriate choice of $\epsilon \in \mathbb{R}_{>0}$. So, $f(x) = 0$ for almost all $x \in \Omega$.

□

Now we return to thinking about weak derivatives. Another very nice property of mollifications is that since they are test functions, they end up commuting with weak derivatives.

Theorem 7.2.3. *Let $\epsilon \in \mathbb{R}_{>0}$ and $\Omega_\epsilon \subseteq \Omega$ such that*

$$\Omega_\epsilon = \{x \in \Omega \mid \overline{B(x,\epsilon)} \subseteq \Omega\}.$$

Let $f \in L^1_{loc}(\Omega)$ such that its weak derivative $D^\alpha f$ is defined for some multi-index α. Then, for all $x \in \Omega_\epsilon$,

$$D^\alpha(J_\epsilon * f) = J_\epsilon * D^\alpha f.$$

Proof. Assume that $\epsilon \in \mathbb{R}_{>0}$. Then, for all $x \in \Omega_\epsilon$, the function $\phi(y) = J_\epsilon(x-y) \in C_c^\infty(\Omega)$. We write D_x^α and D_y^α to distinguish between differentiation with respect to the variables x and y respectively. Therefore, we compute directly that

$$\begin{aligned} D^\alpha(f_\epsilon)(x) &= D_x^\alpha\Big(\int_\Omega J_\epsilon(x-y)f(y)\,dy\Big) \\ &= \int_\Omega D_x^\alpha(J_\epsilon(x-y))f(y)\,dy \\ &= (-1)^{|\alpha|}\int_\Omega D_y^\alpha(J_\epsilon(x-y))f(y)\,dy \\ &= (-1)^{2|\alpha|}\int_\Omega J_\epsilon(x-y)D_y^\alpha(f(y))\,dy \\ &= \int_\Omega J_\epsilon(x-y)D_y^\alpha(f(y))\,dy \\ &= (D_x^\alpha f)_\epsilon(x). \end{aligned}$$

Hence, for all $x \in \Omega_\epsilon$,

$$D^\alpha(J_\epsilon * f) = J_\epsilon * D^\alpha f.$$

□

The fact that weak derivatives commute with mollifications gives us a method to relate weak derivatives with the classical notion of the derivative. The below example provides a glimpse into why this is true.

Lemma 7.2.4. *Let $\Omega \subseteq \mathbb{R}^n$ be an open connected set and $u \in L^1_{loc}(\Omega)$. Suppose that for almost all $x \in \Omega$ that the first order weak derivatives $D_{x_i}u(x) = 0$ for all $i \in \{1,\dots,n\}$, then u is equal to a constant function almost everywhere.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is an open connected set and $u \in L^1_{loc}(\Omega)$. The mollified function $u_\epsilon = J_\epsilon * u \in C_c^\infty(\Omega_\epsilon)$ whose weak derivatives $D_{x_i}u_\epsilon$ vanish on Ω_ϵ, as a consequence of 7.2.3. Hence, u_ϵ must be a constant function on each connected component of Ω_ϵ.

Interestingly, even if Ω is connected, Ω_ϵ is not necessarily connected. However, this is not a problem. Consider any two points $x, y \in \Omega$. Since Ω is connected, there exists a polygonal path Γ which joins x and y and lies entirely in Ω. Define

$$\delta = \min_{z \in \Gamma} d(z, \partial\Omega).$$

Then, for all $\epsilon \in (0, \delta)$, Γ is contained in the set Ω_ϵ (because it can never extend out of Ω). So, x and y are in the same connected component of Ω_ϵ. So, $u_\epsilon(x) = u_\epsilon(y)$ for all $x, y \in \Omega$.

Finally, set $\tilde{u} = \lim_{\epsilon \to 0} u_\epsilon$. Since u_ϵ is constant for an appropriately small $\epsilon \in \mathbb{R}_{>0}$ as outlined above, $\tilde{u}$ must also be a constant function on Ω. However, we also know from 7.2.1 that $\tilde{u}(x) = u(x)$ for almost all $x \in \Omega$. This completes the proof. □

A direct application of the above lemma is for the following theorem:

Theorem 7.2.5. *Let $\Omega \subseteq \mathbb{R}$ be an open interval. Assume that $u : \Omega \to \mathbb{R}$ has a weak derivative $v \in L^1(\Omega)$. Then, there exists an absolutely continuous function $\tilde{u}$ such that for almost all $x \in \Omega$, $\tilde{u} = u$ and*

$$v(x) = \lim_{h \to 0} \frac{\tilde{u}(x+h) - \tilde{u}(x)}{h}.$$

Proof. Assume that Ω is an open interval in $\mathbb{R}$. Assume that $u : \Omega \to \mathbb{R}$ and v are defined as above. Set

$$\tilde{u}(x) = u(x_0) + \int_{x_0}^{x} v(y)\, dy.$$

Here, x_0 is a Lebesgue point of u. By definition $\tilde{u}$ is absolutely continuous and

$$
\begin{aligned}
\lim_{h\to 0}\frac{\tilde{u}(x+h)-\tilde{u}(x)}{h} &= \lim_{h\to 0}\frac{u(x_0)+\int_{x_0}^{x+h} v(y)\ dy - u(x_0) - \int_{x_0}^{x} v(y)\ dy}{h} \\
&= \lim_{h\to 0}\frac{1}{h}\int_{x}^{x+h} v(y)\ dy \\
&= v.
\end{aligned}
$$

Now, assume that $\epsilon \in \mathbb{R}_{>0}$. Consider the mollifications u_ϵ and v_ϵ. Then, these function are smooth and for all $x \in \Omega$, we have

$$
u_\epsilon(x) = u_\epsilon(x_0) + \int_{x_0}^{x} v_\epsilon(y)\ dy.
$$

because u_ϵ has weak derivative v_ϵ by 7.2.3. Now, as $\epsilon \to 0$, the RHS must converge to $\tilde{u}(x)$ for all $x \in \Omega$ by definition. However, the LHS converges to $u(x)$ for every Lebesgue point of u. This is a consequence of 7.2.1. Hence, $\tilde{u}(x) = u(x)$ for almost all $x \in \Omega$. □

7.3 Properties of Sobolev spaces

Before we arrive at the definition of a Sobolev space, there are a few definitions from measure theory that we will highlight briefly.

Definition 7.3.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a measurable function. The function f is called **summable** if

$$
\int_{\mathbb{R}^n} |f|\ dx < \infty.
$$

It is called **locally summable** if for all open balls $B(x,r) \subseteq \mathbb{R}^n$,

$$
\int_{\mathbb{R}^n} |f \mathbb{1}_{B(x,r)}|\ dx = \int_{B(x,r)} |f|\ dx < \infty.
$$

Here, $\mathbb{1}_{B(x,r)}$ is the characteristic function associated with the open ball $B(x,r)$. The integrals above are with respect to Lebesgue measure on $\mathbb{R}^n$.

Definition 7.3.2. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a measurable function. The **essential supremum** of f is defined by

$$
\text{ess sup } f = \inf\{\alpha \in \mathbb{R} \mid f(x) < \alpha \text{ for almost all } x \in \mathbb{R}^n\}.
$$

There are a multitude of alternative ways of defining the essential supremum. The one above originates from Bressan [AB10]. For instance,

1. ess sup $f = \inf\{\alpha \in \mathbb{R} \mid (f-\alpha)^{-1}((-\infty, 0)) \text{ is conull.}\}$
2. ess sup $f = \inf\{\alpha \in \mathbb{R} \mid (f-\alpha)^{-1}([0, \infty)) \text{ is null.}\}$
3. ess sup $f = \inf\{\alpha \in \mathbb{R} \mid m(\{x \in \mathbb{R}^n \mid f(x) \geq \alpha\}) = 0\}$

In particular, the third alternative definition can be found in Evans [LE98]. It is not too hard to show that all these definitions are equivalent.

Now, we can properly state the definition of a Sobolev space.

Definition 7.3.3. Let $\Omega \subseteq \mathbb{R}^n$. Let $p \in [1, \infty]$ and $k \in \mathbb{Z}_{\geq 0}$. The **Sobolev space** $W^{k,p}(\Omega)$ is the space of all locally summable functions $f : \Omega \to \mathbb{R}$ such that for all multi-indices α with $|\alpha| \leq k$, the weak derivative $D^\alpha f$ exists and belongs to $L^p(\Omega)$.

Recalling the norm on the space $L^p(\Omega)$ for all $p \in [1, \infty]$, we can define the following norms on various Sobolev spaces as follows:

$$\|u\|_{W^{k,p}} = (\sum_{|\alpha| \leq k} \int_\Omega |D^\alpha u|^p \, dx)^{1/p} \quad (p \in [1, \infty))$$

$$\|u\|_{W^{k,\infty}} = \sum_{|\alpha| \leq k} \text{ess } \sup_{x \in \Omega} |D^\alpha u|$$

By using the norms on $L^p(\Omega)$, it is not too difficult to show that the above norms are actually norms. We will show later that with the norms above, the Sobolev spaces $W^{k,p}$ are actually Banach spaces for all $p \in [1, \infty]$ and $k \in \mathbb{Z}_{\geq 0}$.

There are a few more definitions which need to be ironed out before we proceed.

Definition 7.3.4. Let $\Omega \subseteq \mathbb{R}^n$, $p \in [1, \infty]$ and $k \in \mathbb{Z}_{>0}$. The subspace $W_0^{k,p}(\Omega) \subseteq W^{k,p}(\Omega)$ is the closure of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$. Alternatively, $u \in W_0^{k,p}(\Omega)$ if and only if there exists a sequence of functions $\{u_n\}$ in $C_c^\infty(\Omega)$ such that

$$\lim_{n \to \infty} \|u - u_n\|_{W^{k,p}} = 0.$$

One can interpret $W_0^{k,p}(\Omega)$ as the closed subspace of all functions $f \in W^{k,p}(\Omega)$ such that for all multi-indices α such that $|\alpha| \leq k-1$, $D^\alpha u = 0$ on the boundary $\partial\Omega$.

Definition 7.3.5. Let $\Omega \subseteq \mathbb{R}^n$. An open set U is **compactly contained** in Ω if its closure $\overline{U}$ is a compact subset of Ω.

Definition 7.3.6. Let $\Omega \subseteq \mathbb{R}^n$, $p \in [1, \infty]$ and $k \in \mathbb{Z}_{>0}$. Denote by $W^{k,p}_{loc}(\Omega)$ the space of functions which are **locally in** $W^{k,p}$. That is, the functions $u : \Omega \to \mathbb{R} \in W^{k,p}_{loc}(\Omega)$ must satisfy the following property: If U is an open set compactly contained in Ω, then the restriction $u|_U \in W^{k,p}(U)$.

Interestingly, just like L^p spaces, the case where $p = 2$ is particularly special.

Definition 7.3.7. Let $\Omega \subseteq \mathbb{R}^n$ and $p \in [1, \infty]$. The **Hilbert-Sobolev** space is defined by $H^k(\Omega) = W^{k,2}(\Omega)$. It is equipped with the inner product

$$\langle u, v\rangle_{H^k} = \sum_{|\alpha| \leq k} \int_\Omega D^\alpha u D^\alpha v \, dx$$

Similarly, we define $H_0^k(\Omega) = W_0^{k,2}(\Omega)$.

Now we will prove that Sobolev spaces are Banach spaces and Hilbert spaces in the case where $p = 2$.

Theorem 7.3.1. *Let $\Omega \subseteq \mathbb{R}^n$, $p \in [1, \infty]$ and $k \in \mathbb{Z}_{>0}$. Then, $W^{k,p}(\Omega)$ is a Banach space, $W_0^{k,p}(\Omega)$ is a closed subspace of $W^{k,p}(\Omega)$ and hence, a Banach space and $H^k(\Omega)$ and $H_0^k(\Omega)$ are Hilbert spaces.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$, $p \in [1, \infty]$ and $k \in \mathbb{Z}_{>0}$.

To show: (a) $W^{k,p}(\Omega)$ is a vector space.

(b) The function $\|-\|_{W^{k,p}}$ defines a norm on $W^{k,p}(\Omega)$.

(c) $W^{k,p}(\Omega)$ is a Banach space.

(d) $W_0^{k,p}(\Omega)$ is a Banach space.

(a) Assume that $f, g \in W^{k,p}(\Omega)$ and $\lambda, \mu \in \mathbb{R}$. Then, the linear combination $\lambda f + \mu g$ is a locally summable function because f and g are both locally summable. Using the linearity of D^α, the weak derivatives must satisfy

$$D^\alpha(\lambda f + \mu g) = \lambda D^\alpha u + \mu D^\alpha v.$$

Since $L^p(\Omega)$ is a vector space, $D^\alpha(\lambda f + \mu g) \in L^p(\Omega)$ for all $|\alpha| \leq k$. Hence, $W^{k,p}(\Omega)$ is a vector space.

(b) Using the linearity of D^α and the definition of $\|-\|_{W^{k,p}}$, we find that for all $p \in [1, \infty)$ and $u \in W^{k,p}(\Omega)$,

$$\begin{aligned}
\|\lambda u\|_{W^{k,p}} &= (\sum_{|\alpha|\leq k} \int_\Omega |D^\alpha(\lambda u)|^p \, dx)^{1/p} \\
&= (\sum_{|\alpha|\leq k} \int_\Omega |\lambda D^\alpha u|^p \, dx)^{1/p} \\
&= (|\lambda|^p \sum_{|\alpha|\leq k} \int_\Omega |D^\alpha u|^p \, dx)^{1/p} \\
&= |\lambda| (\sum_{|\alpha|\leq k} \int_\Omega |D^\alpha u|^p \, dx)^{1/p} \\
&= |\lambda| \|u\|_{W^{k,p}}.
\end{aligned}$$

A similar computation works for $p = \infty$. Secondly, note again by the definition of $\|-\|_{W^{k,p}}$,

$$\|u\|_{W^{k,p}} \geq \|u\|_{L^p} \geq 0$$

with equality holding if and only if $u = 0$. Finally, to see that the triangle inequality holds in the case where $p \in [1, \infty)$,

$$\begin{aligned}
\|u + v\|_{W^{k,p}} &= (\sum_{|\alpha|\leq k} \|D^\alpha u + D^\alpha v\|_{L^p}^p)^{1/p} \\
&\leq (\sum_{|\alpha|\leq k} (\|D^\alpha u\|_{L^p} + \|D^\alpha v\|_{L^p})^p)^{1/p} \quad \text{(Minkowski)} \\
&\leq (\sum_{|\alpha|\leq k} \|D^\alpha u\|_{L^p}^p)^{1/p} + (\sum_{|\alpha|\leq k} \|D^\alpha v\|_{L^p}^p)^{1/p} \\
&= \|u\|_{W^{k,p}} + \|v\|_{W^{k,p}}.
\end{aligned}$$

For the case where $p = \infty$, the computation is similar:

$$
\begin{aligned}
\|u+v\|_{W^{k,\infty}} &= \sum_{|\alpha|\le k} \text{ess}\ \sup_{x\in\Omega}|D^\alpha(u+v)| \\
&= \sum_{|\alpha|\le k} \|D^\alpha u + D^\alpha v\|_{L^\infty} \\
&\le \sum_{|\alpha|\le k} (\|D^\alpha u\|_{L^\infty} + \|D^\alpha v\|_{L^\infty}) \quad \text{(Minkowski)} \\
&= \|u\|_{W^{k,\infty}} + \|v\|_{W^{k,\infty}}.
\end{aligned}
$$

So, $W^{k,p}(\Omega)$ is a normed vector space for all $k \in \mathbb{Z}_{>0}$ and $p \in [1,\infty]$.

(c) Assume that $\{u_n\}_{n\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in $W^{k,p}(\Omega)$. Then, for all multi-indices α such that $|\alpha| \le k$, the sequence of weak derivatives $\{D^\alpha u_n\}$ is Cauchy in $L^p(\Omega)$, due to 7.1.4. Since $L^p(\Omega)$ is a Banach space, the sequence $\{D^\alpha u_n\}$ must converge to some function $u_\alpha \in L^p(\Omega)$. Similarly, there exists $u \in L^p(\Omega)$ such that

$$
\|u_n - u\|_{W^{k,p}} \to 0.
$$

By 7.1.4, $D^\alpha u = u_\alpha$. Thus, $W^{k,p}(\Omega)$ is a Banach space.

(d) By definition, $W_0^{k,p}(\Omega)$ is a closed subspace of $W^{k,p}(\Omega)$ (it is the closure of $C_c^\infty(\Omega)$). Since $W^{k,p}(\Omega)$ is complete, $W_0^{k,p}(\Omega)$ is also complete and thus, a Banach space.

Finally, in the special case where $p=2$, parts (c) and (d) show that $H^k(\Omega)$ and $H_0^k(\Omega)$ are Hilbert spaces with the inner product

$$
\langle u, v\rangle_{H^k} = \sum_{|\alpha|\le k} \int_\Omega D^\alpha u D^\alpha v \, dx.
$$

□

Let us look at some examples of Sobolev spaces. First, we need a result from measure theory.

Lemma 7.3.2. *Let $(X, \mathcal{A}, \mu)$ be a finite measure space and $p, q \in \mathbb{Z}_{>0}$ such that $1 \le p < q < \infty$. Then, $L^q(X, \mathcal{A}, \mu) \subseteq L^p(X, \mathcal{A}, \mu)$.*

Proof. Assume that $(X, \mathcal{A}, \mu)$ is a finite measure space and $p, q \in \mathbb{Z}_{>0}$ such that $1 \leq p < q < \infty$. Assume that $f \in L^q(X, \mathcal{A}, \mu)$. Then, by Hölder's inequality, we have

$$\begin{aligned}
\int |f|^p \, d\mu &= \int |f|^p \cdot 1 \, d\mu \\
&\leq (\int |f|^{pq/p} \, d\mu)^{p/q} (\int \, d\mu)^{1-p/q} \quad \text{(Hölder)} \\
&= (\int |f|^q \, d\mu)^{p/q} \mu(X)^{1-p/q} \\
&< \infty.
\end{aligned}$$

Therefore, $f \in L^p(X, \mathcal{A}, \mu)$. Note that Hölder's inequality can also prove the converse in a similar vein.

□

Example 7.3.1. Let $\Omega = (a, b) \subseteq \mathbb{R}$ be an open interval, $p \in [1, \infty]$ and $k = 1$. By definition, the Sobolev space $W^{1,p}((a, b))$ consists of all locally summable functions $f : (a, b) \to \mathbb{R}$ whose first order weak derivatives all exist (for all multi-indices α such that $|\alpha| \leq 1$) and are in $L^p((a, b))$. From the above result, $L^p((a, b)) \subseteq L^1((a, b))$. Hence, the functions in $W^{1,p}((a, b))$ are almost everywhere equal to absolutely continuous functions, whose derivatives are in $L^p((a, b))$.

Example 7.3.2. Let $\Omega = B(0, 1) \subseteq \mathbb{R}^n$ be the open ball centred at the origin, with radius 1. Consider the following function for all $\gamma \in \mathbb{R}_{>0}$ and $0 < |x| < 1$

$$u(x) = |x|^{-\gamma} = (\sum_{i=1}^{n} x_i^2)^{-\gamma/2}.$$

Observe that $u \in C^1(\Omega \backslash \{0\})$. with derivative

$$\frac{\partial u}{\partial x_i} = -\frac{\gamma}{2} \cdot 2x_i \cdot (\sum_{i=1}^{n} x_i^2)^{-\gamma/2-1} = \frac{-\gamma x_i}{|x|^{\gamma+2}}.$$

for all $i \in \{1, \ldots, n\}$. Note that the gradient ∇u has norm given by

$$|\nabla u(x)| = (\sum_{i=1}^{n} |\frac{\partial u}{\partial x_i}|^2)^{1/2} = \frac{\gamma}{|x|^{\gamma+1}}.$$

The point of the above computation is that on the open set $\Omega\backslash\{0\}$, u does have weak derivatives of all orders - they are exactly the classical derivatives. The question is: when does the above formula for $\partial_{x_i}u$ (the partial derivatives of u) define the weak derivatives of u over *the entire domain* Ω?

Assume that $\phi \in C_c^\infty(\Omega)$ and $\epsilon \in \mathbb{R}_{>0}$. We begin by using an integration by parts to deduce that

$$\begin{aligned}\int_{\Omega - B(0,\epsilon)} u\partial_{x_i}\phi \, dx &= \int_{\partial(\Omega - B(0,\epsilon))} u\phi\nu^i \, dS - \int_{\Omega - B(0,\epsilon)} \partial_{x_i}u\phi \, dx \\ &= \int_{|x|=\epsilon} u\phi\nu^i \, dS - \int_{\epsilon<|x|<1} \partial_{x_i}u\phi \, dx.\end{aligned}$$

Here, $\nu = (\nu^1, \ldots, \nu^n)$ is the inwards pointing normal on the boundary $|x| = \epsilon$. Note that ϕ vanishes when $|x| = 1$ by definition. When $\gamma + 1 < n$, $|\nabla u(x)| \in L^1(\Omega)$. In this particular case, we can bound the first term as follows

$$\begin{aligned}|\int_{|x|=\epsilon} u\phi\nu^i \, dS| &\leq \int_{|x|=\epsilon} |u\phi\nu^i| \, dS \\ &\leq \|\phi\|_{L^\infty} \int_{|x|=\epsilon} |u\nu^i| \, dS \\ &\leq \|\phi\|_{L^\infty} \int_{|x|=\epsilon} \epsilon^{-\gamma} \, dS \\ &= \|\phi\|_{L^\infty} \epsilon^{n-1-\gamma} \\ &\to 0\end{aligned}$$

as $\epsilon \to 0$. Therefore, in the limit $\epsilon \to 0$, we have

$$\int_\Omega u\partial_{x_i}\phi \, dx = -\int_\Omega \partial_{x_i}u\phi \, dx$$

provided that $\gamma < n - 1$. Furthermore, $|\nabla u(x)| \in L^p(\Omega)$ whenever $(\gamma + 1)p < n$. Therefore, we can conclude that $u \in W^{1,p}(\Omega)$ if $\gamma < (n - p)/p$.

We will now proceed to derive further basic properties of Sobolev spaces. The next result is motivated by the following question: Can we estimate the norm of a function in a Sobolev space by the norm of its weak derivatives?

It turns out that according to [AB10], that this is a deep question. The *Poincaré inequality* provides a small step towards answering the question.

Lemma 7.3.3 (Poincaré Inequality)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be an open set which satisfies the inclusion $\Omega \subseteq (a,b) \times \mathbb{R}^{n-1}$ for some $a, b \in \mathbb{R}$. Then, for all $u \in H_0^1(\Omega)$,*

$$\|u\|_{L^2} \leq 2(b-a)\|D_{x_1}u\|_{L^2}.$$

Proof. From the definition of $H_0^1(\Omega) = W^{1,2}(\Omega)$, it suffices to prove the inequality in the case where u is a test function. Assume first that $u \in C_c^\infty(\Omega)$. Extend u to the entire space $\mathbb{R}^n$ by setting $u(x) = 0$ whenever $x \notin \Omega$. The notation here is that $x = (x_1, x')$, with $x' = (x_2, \ldots, x_n)$. Since $u(x)$ is a test function, it must vanish when $x_1 = a$. So,

$$u^2(x_1, x') = \int_a^{x_1} 2u\partial_{x_1}u(t, x')\ dt.$$

Arguing with integration by parts, we have

$$\begin{aligned}
\|u\|_{L^2}^2 &= \int_{\mathbb{R}^n} u^2(x)\ dx \\
&= \int_{\mathbb{R}^{n-1}} \int_a^b u^2(x_1, x')\ dx_1\ dx' \\
&= \int_{\mathbb{R}^{n-1}} \int_a^b 1 \cdot (\int_a^{x_1} 2u\partial_{x_1}u(t, x')\ dt)\ dx_1\ dx' \\
&= \int_{\mathbb{R}^{n-1}} ([x_1 \int_a^{x_1} 2u\partial_{x_1}u(t, x')\ dt]_a^b - \int_a^b x_1 2u\partial_{x_1}u(t, x')\ dx_1)\ dx' \\
&= \int_{\mathbb{R}^{n-1}} \int_a^b (b - x_1) 2u\partial_{x_1}u(t, x')\ dx_1\ dx' \\
&\leq 2(b-a) \int_{\mathbb{R}^n} |u||\partial_{x_1}u|\ dx \\
&\leq 2(b-a)\|u\|_{L^2}\|\partial_{x_1}u\|_{L^2} \quad \text{(Cauchy-Schwarz)}
\end{aligned}$$

Dividing both sides by $\|u\|_{L^2}$ gives the desired result in the case where $u \in C_c^\infty(\Omega)$.

Now suppose that $f \in H_0^1(\Omega)$. Then, there exists a sequence of functions $f_n \in C_c^\infty(\Omega)$ such that $\|f_n - f\|_{H^1} \to 0$. So,

$$
\begin{aligned}
\|f\|_{L^2} &= \lim_{n\to\infty} \|f_n\|_{L^2} \\
&\leq \lim_{n\to\infty} 2(b-a)\|\partial_{x_1} f_n\|_{L^2} \\
&= 2(b-a)\|\partial_{x_1} f\|_{L^2}.
\end{aligned}
$$

Therefore, $\|f\|_{L^2} \leq 2(b-a)\|D_{x_1} f\|_{L^2}$. This completes the proof.

□

With Sobolev spaces, we can proceed to derive more properties of weak derivatives. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\beta = (\beta_1, \ldots, \beta_n)$ be multi-indices. Then, we can define an order on multi-indices - we say that $\beta \leq \alpha$ if $\beta_i \leq \alpha_i$ for all $i \in \{1, \ldots, n\}$. We also define for multi-indices

$$
\binom{\alpha}{\beta} = \prod_{i=1}^{n} \frac{\alpha_i!}{\beta_i!(\alpha_i - \beta_i)!}
$$

whenever $\beta \leq \alpha$.

Lemma 7.3.4. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set, $p \in [1, \infty]$ and $|\alpha| \leq k$. Let $u \in W^{k,p}(\Omega)$. Then,*

1. *The restriction of u to any open subset $\Omega' \subseteq \Omega$ is contained in the space $W^{k,p}(\Omega')$.*
2. *For all multi-indices α, β such that $|\alpha| \leq k$, $D^\alpha u \in W^{k-|\alpha|,p}(\Omega)$. Moreover, if $|\alpha| + |\beta| \leq k$, then $D^\alpha(D^\beta u) = D^\beta(D^\alpha u) = D^{\alpha+\beta} u$.*
3. *If $\eta \in C^k(\Omega)$, then $\eta u \in W^{k,p}(\Omega)$. Furthermore, the weak derivative is given by the Leibnitz formula*

$$
D^\alpha(\eta u) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta \eta D^{\alpha-\beta} u.
$$

4. *There exists a constant $C \in \mathbb{R}_{>0}$ depending on Ω and on $\|\eta\|_{C^k}$ (but not on u) such that $\|\eta u\|_{W^{k,p}(\Omega)} \leq C\|u\|_{W^{k,p}(\Omega)}$.*
5. *Let $\Omega' \subseteq \mathbb{R}^n$ be an open set and $\varphi : \Omega' \to \Omega$ be a C^k diffeomorphism with a uniformly bounded inverse. Then the composite $u \circ \varphi \in W^{k,p}(\Omega')$.*

6. *There exists a constant $D \in \mathbb{R}_{>0}$ depending on Ω' and on $\|\varphi\|_{C^k}$ (but not on u) such that $\|u \circ \varphi\|_{W^{k,p}(\Omega')} \leq D\|u\|_{W^{k,p}(\Omega)}$.*

Proof. Assume that Ω is an open subset of $\mathbb{R}^n$, $p \in [1, \infty]$ and $|\alpha| \leq k$. Assume that $u \in W^{k,p}(\Omega)$. Then, the first statement follows from the definition of a Sobolev space.

To prove the second statement, we will use the definition of the weak derivative. Since $u \in W^{k,p}(\Omega)$, we have for all test functions $\phi \in C_c^\infty(\Omega)$,

$$\int_\Omega u D^\alpha \phi \, dx = (-1)^{|\alpha|} \int_\Omega D^\alpha u \phi \, dx.$$

where α is a multi-index satisfying $|\alpha| \leq k$. Assume that β is another multi-index such that $|\alpha| + |\beta| \leq k$. Then, $D^\beta \phi \in C_c^\infty(\Omega)$ and

$$\begin{aligned}\int_\Omega D^\alpha u D^\beta \phi \, dx &= (-1)^{|\alpha|} \int_\Omega u D^{\alpha+\beta} \phi \, dx \\ &= (-1)^{|\alpha|+|\alpha|+|\beta|} \int_\Omega (D^{\alpha+\beta} u) \phi \, dx \\ &= (-1)^{|\beta|} \int_\Omega (D^{\alpha+\beta} u) \phi \, dx.\end{aligned}$$

This demonstrates that $D^\alpha \in W^{k-|\alpha|}(\Omega)$ because $|\beta| \leq k - |\alpha|$. Furthermore, the above equation reveals that $D^{\alpha+\beta} u$ is the β^{th} order weak derivative of $D^\alpha u$. So, $D^\alpha(D^\beta u) = D^\beta(D^\alpha u) = D^{\alpha+\beta} u$.

For the third statement, assume that $\eta \in C^k(\Omega)$. Then, from 7.2.1, there exists a sequence $\{u_\epsilon\}$ of smooth functions such that $u_\epsilon \to u$ as $\epsilon \to 0$. The key point here is that smooth functions satisfy the Leibnitz formula. So, for all $\epsilon \in \mathbb{R}_{>0}$,

$$D^\alpha(\eta u_\epsilon) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta \eta D^{\alpha-\beta} u_\epsilon.$$

Therefore, for every test function $\phi \in C_c^\infty(\Omega)$, we must have the following equation:

$$\begin{aligned}\int_\Omega (\eta u_\epsilon) D^\alpha \phi \, dx &= (-1)^{|\alpha|} \int_\Omega D^\alpha(\eta u_\epsilon) \phi \, dx \\ &= \sum_{\beta \leq \alpha} (-1)^{|\alpha|} \binom{\alpha}{\beta} \int_\Omega (D^\beta \eta D^{\alpha-\beta} u_\epsilon) \phi \, dx.\end{aligned}$$

Of course, $|\alpha| \leq k$. Taking the limit as $\epsilon \to 0$, we deduce that

$$\int_\Omega (\eta u) D^\alpha \phi \, dx = \sum_{\beta \leq \alpha} (-1)^{|\alpha|} \binom{\alpha}{\beta} \int_\Omega (D^\beta \eta D^{\alpha-\beta} u) \phi \, dx.$$

This subsequently demonstrates the Leibnitz rule for functions in $W^{k,p}(\Omega)$:

$$D^\alpha(\eta u) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta \eta D^{\alpha-\beta} u.$$

For the next statement, observe that the derivatives of η are bounded so that for all $|\beta| \leq k$,

$$\|D^\beta \eta\|_{L^\infty} \leq \|\eta\|_{C^k}.$$

So, we bound as follows:

$$\begin{aligned}
\|\eta u\|^p_{W^{k,p}(\Omega)} &= \sum_{|\alpha| \leq k} \|D^\alpha(\eta u)\|^p_{L^p} \\
&= \sum_{|\alpha| \leq k} \|\sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta \eta D^{\alpha-\beta} u\|^p_{L^p} \\
&\leq \sum_{|\alpha| \leq k} (\sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \|D^\beta \eta D^{\alpha-\beta} u\|_{L^p})^p \quad \text{(Minkowski)} \\
&\leq \sum_{|\alpha| \leq k} (\sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \|\eta\|_{C^k} \|D^{\alpha-\beta} u\|_{L^p})^p \\
&= \sum_{|\alpha| \leq k} (\sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \|D^{\alpha-\beta} u\|_{L^p})^p \|\eta\|^p_{C^k} \\
&\leq C \|u\|^p_{W^{k,p}(\Omega)}
\end{aligned}$$

where C is a constant which depends on $\|\eta\|^p_{C^k}$ and Ω. Taking the p^{th} root of both sides gives the desired bound.

For the last two statements, we will give a proof by strong induction. Assume that $\Omega' \subseteq \mathbb{R}^n$ is an open set and $\varphi : \Omega' \to \Omega$ be a C^k diffeomorphism, with a uniformly bounded inverse. For the base case, assume that $k = 0$. Then, there are no weak derivatives to worry about and as a result, the composite $u \circ \varphi : \Omega' \to \mathbb{R}$ must be in $W^{0,p}(\Omega')$. Also,

$\|u \circ \varphi\|_{W^{0,p}(\Omega')}\| \leq \|\varphi\|_{C^0(\Omega')}\|u\|_{W^{0,p}(\Omega')}$.

Now, assume that the statements are true for all $k \in \{0, \ldots, m-1\}$. Take the mollification $u_\epsilon = J_\epsilon * u$. Then, by the chain rule,

$$D_{x_i}(u_\epsilon \circ \varphi)(x) = \sum_{j=1}^{n} (D_{x_j} u_\epsilon)(\varphi(x)) D_{x_i} \varphi_j(x)$$

where $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n)$ and $i \in \{1, \ldots, n\}$. So, if $u \in W^{m,p}(\Omega)$, then as $\epsilon \to 0$,

$$\begin{aligned} \|D_{x_i}(u \circ \varphi)\|_{W^{m-1,p}(\Omega')} &\leq C' \|\nabla u\|_{W^{m-1,p}(\Omega)} \|\varphi\|_{C^m(\Omega')} \quad \text{(Chain Rule)} \\ &\leq D\|u\|_{W^{m,p}(\Omega)}. \end{aligned}$$

where D is a constant which depends on Ω' and $\|\varphi\|_{C^m}$. This completes the induction.

□

7.4 Approximations and Extensions

The next theorem is very powerful. Just like how smooth functions are dense in the space of continuous functions, they are also dense in Sobolev spaces! The following proof makes use of a *smooth partition of unity*, a fundamental tool in differential geometry.

Theorem 7.4.1. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $u \in W^{k,p}(\Omega)$ with $p \in [1, \infty)$ and $k \in \mathbb{Z}_{\geq 0}$. Then, $\overline{C^\infty(\Omega)} = W^{k,p}(\Omega)$.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ be an open set. Assume that $u \in W^{k,p}(\Omega)$ with $p \in [1, \infty)$ and $k \in \mathbb{Z}_{\geq 0}$. Consider the locally finite open covering $\{V_j\}_{j \in \mathbb{Z}_{>0}}$, defined by

$$V_1 = \{x \in \Omega \mid d(x, \partial\Omega) > \frac{1}{2}\}$$

$$V_j = \{x \in \Omega \mid \frac{1}{j+1} < d(x, \partial\Omega) < \frac{1}{j-1}\}$$

where $j \in \mathbb{Z}_{\geq 2}$. Define $\{\eta_n\}_{n \in \mathbb{Z}_{>0}}$ to be a partition of unity subordinate to the above cover. Observe that for all $j \in \mathbb{Z}_{>0}$ $\eta_j u \in W^{k,p}(\Omega)$, from the previous theorem. Due to the construction of the partition of unity,

$supp(\eta_j u) \subseteq V_j$.

Now consider the mollifications $J_\epsilon * (\eta_j u)$ for all $\epsilon \in \mathbb{R}_{>0}$ and $j \in \mathbb{Z}_{>0}$. Since mollifications commute with weak derivatives (see 7.2.3), we have for all multi-indices α such that $|\alpha| \leq k$,

$$\lim_{\epsilon \to 0} D^\alpha (J_\epsilon * (\eta_j u)) = \lim_{\epsilon \to 0} J_\epsilon * D^\alpha (\eta_j u) = D^\alpha (\eta_j u)$$

By construction, each η_j has compact support. So, the convergence above occurs in $L^p(\Omega)$. Thus, for all $j \in \mathbb{Z}_{>0}$, we can find $\epsilon_j \in \mathbb{R}_{>0}$ such that

$$\|\eta_j u - J_{\epsilon_j} * (\eta_j u)\|_{W^{k,p}(\Omega)} \leq \epsilon 2^{-k}.$$

With this, we can now define the smooth function $u^\sharp \in C^\infty(\Omega)$

$$u^\sharp = \sum_{j=1}^{\infty} J_{\epsilon_j} * (\eta_j u).$$

It is smooth because each summand is smooth (see 7.2.1). Luckily, we do not have to worry about the convergence of $u^\sharp$. It is pointwise convergent because every compact set $K \subseteq \Omega$ intersects finitely many V_j. When we restrict $u^\sharp$ to K, the above sum therefore has only finitely many non-zero terms, establishing pointwise convergence.

Now let

$$\Omega_{1/n} = \{x \in \Omega \mid d(x, \partial\Omega) > \frac{1}{n}\}$$

Since $\sum_j \eta_j(x) = 1$ (partition of unity), we find that for all $n \geq 1$

$$\begin{aligned} \|u^\sharp - u\|_{W^{k,p}(\Omega_{1/n})} &= \|\sum_{j=1}^{\infty} J_{\epsilon_j} * (\eta_j u) - u\|_{W^{k,p}(\Omega_{1/n})} \\ &= \|\sum_{j=1}^{n+2} (J_{\epsilon_j} * (\eta_j u) - \eta_j u)\|_{W^{k,p}(\Omega)} \\ &\leq \sum_{j=1}^{n+2} \|J_{\epsilon_j} * (\eta_j u) - \eta_j u\|_{W^{k,p}(\Omega)} \\ &\leq \sum_{j=1}^{\infty} \epsilon 2^{-j} \\ &= \epsilon. \end{aligned}$$

Therefore,

$$\|u^\sharp - u\|_{W^{k,p}(\Omega)} = \sup_{n\geq 1}\|u^\sharp - u\|_{W^{k,p}(\Omega_{1/n})} \leq \epsilon.$$

Therefore, the set of smooth functions $C^\infty(\Omega)$ is dense in $W^{k,p}(\Omega)$.

□

We will use the above result to prove a regularity theorem which describes the connection between weak and classical derivatives of a function in an appropriate Sobolev space.

Theorem 7.4.2. *Let $\Omega \subseteq \mathbb{R}^n$ denote the following open set*

$$\Omega = \{x = (x_1, x') \in \mathbb{R}^n \mid x' = (x_2, \ldots, x_n) \in \Omega' \text{ and } \alpha(x') < x_1 < \beta(x')\}$$

where $\Omega' \subseteq \mathbb{R}^{n-1}$ is an open set. Let $u \in W^{1,1}(\Omega)$. Then, there exists a function $\tilde{u}$ such that for almost all $x \in \Omega$, $\tilde{u}(x) = u(x)$. Moreover for almost all $x' \in \Omega' \subseteq \mathbb{R}^{n-1}$, the map $x_1 \mapsto \tilde{u}(x_1, x')$ is absolutely continuous with derivative equal to $D_{x_1}u$ almost everywhere.

Note that we say "almost everywhere", we mean with respect to the usual Lebesgue measure.

Proof. Assume that $\Omega = (\alpha(x'), \beta(x')) \times \Omega'$ as explained above. Assume that $u \in W^{1,1}(\Omega)$. Using 7.4.1, there exists a sequence $\{u_n\}$ in $C^\infty(\Omega)$ such that $\|u_n - u\|_{W^{1,1}(\Omega)} < 2^{-n}$.

Define the functions $f, g : \Omega \to \mathbb{R}$ by

$$f(x) = |u_1(x)| + \sum_{n=1}^{\infty}|u_{n+1}(x) - u_n(x)|$$

and

$$g(x) = |D_{x_1}u_1(x)| + \sum_{n=1}^{\infty}|D_{x_1}u_{n+1}(x) - D_{x_1}u_n(x)|.$$

By the triangle inequality, we have

$$\|u_n - u_{n+1}\|_{W^{1,1}(\Omega)} \leq \|u_n - u\|_{W^{1,1}(\Omega)} + \|u_{n+1} - u\|_{W^{1,1}(\Omega)} < 2^{-n+1}.$$

From the definition of the $W^{1,1}$ norm, we therefore find that $f, g \in L^1(\Omega)$ and that f and g are absolutely convergent for almost all $x \in \Omega$. Thus, they must converge pointwise almost everywhere. Furthermore, we have the following bounds for all $n \in \mathbb{Z}_{>0}$ and $x \in \Omega$:

$$|u_n(x)| \leq f(x) \text{ and } |D_{x_1} u_n(x)| \leq g(x).$$

Now since f and g are integrable over Ω, Fubini's theorem tells us that there exists a null set $N \subseteq \Omega'$ such that for all $x' \in \Omega' \backslash N$,

$$\int_{\alpha(x')}^{\beta(x')} f(x_1, x')\, dx_1 < \infty \text{ and } \int_{\alpha(x')}^{\beta(x')} g(x_1, x')\, dx_1 < \infty.$$

Take a point $x' \in \Omega' \backslash N$. Choose a point $y_1 \in (\alpha(x'), \beta(x'))$ where the pointwise convergence $u_n(y_1, x') \to u(y_1, x')$ holds. Since u_n is smooth, we have for all $\alpha(x') < x_1 < \beta(x')$

$$u_n(x_1, x') = u_n(y_1, x') + \int_{y_1}^{x_1} D_{x_1} u_n(s, x')\, ds.$$

Now consider the limit as $n \to \infty$. Observe that since $|D_{x_1} u_n(x)| \leq g(x)$ and g is integrable, we can use the dominated convergence theorem to show that

$$\begin{aligned}
\lim_{n\to\infty} u_n(y_1, x') + \int_{y_1}^{x_1} D_{x_1} u_n(s, x')\, ds &= u(y_1, x') + \lim_{n\to\infty} \int_{y_1}^{x_1} D_{x_1} u_n(s, x')\, ds \\
&= u(y_1, x') + \int_{y_1}^{x_1} \lim_{n\to\infty} D_{x_1} u_n(s, x')\, ds \\
&= u(y_1, x') + \int_{y_1}^{x_1} D_{x_1} u(s, x')\, ds.
\end{aligned}$$

Now, we can define

$$\tilde{u}(x_1, x') = u(y_1, x') + \int_{y_1}^{x_1} D_{x_1} u(s, x')\, ds.$$

The RHS is an absolutely continuous function of the variable x_1. But, the LHS satisfies for almost all $x_1 \in [\alpha(x'), \beta(x')]$,

$$\tilde{u}(x_1, x') = \lim_{n\to\infty} u_n(x_1, x') = u(x_1, x').$$

□

In the next section of the chapter, we will explore various embeddings of Sobolev spaces. An important technique towards this goal is that under certain circumstances, a function $u \in W^{1,p}(\Omega)$, where $\Omega \subseteq \mathbb{R}^n$ can be extended to a function $Eu \in W^{1,p}(\mathbb{R}^n)$. Again, the proof of this statement relies on a partition of unity.

Theorem 7.4.3. *Let $\Omega \subseteq \tilde{\Omega} \subseteq \mathbb{R}^n$ be bounded open sets, with $\tilde{\Omega}$ containing $\overline{\Omega}$. Assume that $\partial\Omega \in C^1$. Then, there exists a bounded linear operator $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ and a constant $C \in \mathbb{R}_{>0}$ such that $Eu(x) = u(x)$ for almost all $x \in \Omega$, $Eu(x) = 0$ for all $x \notin \tilde{\Omega}$ and*

$$\|Eu\|_{W^{1,p}(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(\Omega)}$$

Proof. Assume that Ω and $\tilde{\Omega}$ are the open sets defined as above. Assume that $\partial\Omega \in C^1$ (a C^1 boundary). We will first prove the statement for the case where

$$\Omega = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1 > 0\}.$$

Int this case we can define $E^\sharp : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ by

$$E^\sharp(u)(x_1, \ldots, x_n) = u(|x_1|, x_2, \ldots, x_n).$$

Intuitively, we can extend the definition of u by reflection to the other half of the plane. To see that this map is well-defined, observe that $D_{x_i}E^\sharp u = D_{x_i}u$ for all $i \in \{2, 3, \ldots, n\}$. For the case where $i = 1$, observe first that on the half-plane Ω,

$$D_{x_1}E^\sharp u(x_1, \ldots, x_n) = D_{x_1}u(x_1, \ldots, x_n)$$

On the other half-plane, we have for all $x_1 > 0$ and $x_2, \ldots, x_n \in \mathbb{R}$,

$$D_{x_1}E^\sharp u(-x_1, \ldots, x_n) = -D_{x_1}u(x_1, \ldots, x_n).$$

This shows that the weak derivatives of $E^\sharp u$ all exist and are in $L^p(\mathbb{R}^n)$. Hence, the map $E^\sharp : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ is well-defined. It can also be verified that it is linear and satisfies

$$
\begin{aligned}
\|E^\sharp u\|_{W^{1,p}(\mathbb{R}^n)} &= (\sum_{|\alpha|\le 1} \|D^\alpha E^\sharp u\|^p_{L^p(\mathbb{R}^n)})^{1/p} \\
&= (\sum_{|\alpha|\le 1} \int_{\mathbb{R}^n} |D^\alpha E^\sharp u(x)|^p \; dx)^{1/p} \\
&= (\sum_{|\alpha|\le 1} \int_{\mathbb{R}^n} |D^\alpha E^\sharp u(x)|^p \; dx)^{1/p} \\
&\le (\sum_{|\alpha|\le 1} \int_{\Omega} |2D^\alpha E^\sharp u(x)|^p \; dx)^{1/p} \\
&= 2(\sum_{|\alpha|\le 1} \int_{\Omega} |D^\alpha u(x)|^p \; dx)^{1/p} \\
&= 2\|u\|_{W^{1,p}(\Omega)}.
\end{aligned}
$$

This establishes the result for the half-plane. To handle the general case, we proceed with the following construction. For all $x \in \overline{\Omega}$, we choose an open ball $B(x, r_x)$ such that

1. If $x \in \Omega$, then $B(x, r_x) \subseteq \Omega$.
2. If $x \in \partial\Omega$, then $B(x, r_x) \subseteq \tilde{\Omega}$.
3. There exists a C^1 bijection, with a C^1 inverse

$$\varphi_x : B(0,1) \to B(x, r_x)$$

which maps the "upper half ball"

$$B^+(0,1) = \{y = (y_1, \ldots, y_n) \mid \sum_{i=1}^n y_i^2 < 1, y_1 > 0\}$$

to the set $B(x, r_x) \cap \Omega$. For a sufficiently small $r_x > 0$, the existence of φ_x follows from the assumption that Ω has a C^1 boundary.

Now since Ω is bounded, its closure $\overline{\Omega}$ is compact. Hence, there exists a finite cover $\{B(x_i, r_i)\}_{i=1}^N$ of $\Omega \subseteq \overline{\Omega}$. Let $\varphi_i : B(0,1) \to B(x_i, r_i)$ be the corresponding bijections. We also define $\{\eta_i\}_{i=1}^N$ to be a partition of unity subordinate to this cover. Then, for all $x \in \Omega$,

$$u(x) = \sum_{i=1}^{N} \eta_i(x)u(x)$$

Now, we partition the indices $\{1, 2, \ldots, N\} = \mathcal{I} \cup \mathcal{J}$, where for all $i \in \mathcal{I}$, $x_i \in \Omega$ and for all $i \in \mathcal{J}$, $x_i \in \partial\Omega$. Observe that for all $i \in \mathcal{J}$, we have

$$\eta_i u \in W^{1,p}(B(x_i, r_i) \cap \Omega) \text{ and } (\eta_i u) \circ \varphi_i \in W^{1,p}(B^+(0, 1)).$$

Applying the operator $E^\sharp$, we find that

$$E^\sharp[(\eta_i u) \circ \varphi_i] \in W^{1,p}(B(0, 1)) \text{ and } E^\sharp[(\eta_i u) \circ \varphi_i] \circ \varphi_i^{-1} \in W^{1,p}(B(x_i, r_i) \cap \Omega).$$

Now, we can sum together all the extensions in order to define the operator $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ by

$$Eu = \sum_{i \in \mathcal{I}} \eta_i u + \sum_{i \in \mathcal{J}} E^\sharp[(\eta_i u) \circ \varphi_i] \circ \varphi_i^{-1}$$

This operator is bounded and linear because E is a finite sum of bounded, linear operators. Moreover, the support $supp(Eu) \subseteq \bigcup_{i=1}^{N} B(x_i, r_i) \subseteq \tilde{\Omega}$.

□

7.5 The embedding theorems of Sobolev spaces

It is emphasised in Bressan [AB10] that in applications to partial differential equations or the calculus of variations, it is important to understand the degree of regularity exhibited by functions $u \in W^{k,p}(\mathbb{R}^n)$. The first embedding theorem we will prove is attributed to Morrey. First, we need to understand the concept of *Hölder continuity*.

Definition 7.5.1. Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $\gamma \in (0, 1]$. A function $f : \Omega \to \mathbb{R}$ is **Hölder continuous** with exponent γ if there exists $C \in \mathbb{R}_{>0}$ such that for all $x, y \in \Omega$,

$$|f(x) - f(y)| \leq C|x - y|^\gamma$$

The notation $C^{0,\gamma}(\Omega)$ is the space of all bounded Hölder continuous functions on Ω. It is a normed vector space, equipped with the norm

$$\|f\|_{C^{0,\gamma}(\Omega)} = \sup_{x \in \Omega}|f(x)| + \sup_{x,y \in \Omega, x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\gamma}.$$

Like Sobolev spaces, Hölder spaces can be extended to functions with Hölder continuous partial derivatives up to order k. This is encapsulated in the following more general definition below.

Definition 7.5.2. For all $k \in \mathbb{Z}_{\geq 0}$, define $C^{k,\gamma}(\Omega)$ to be the space of all continuous functions with Hölder continuous partial derivatives up to order k. This is a normed vector space when equipped with the following norm:

$$\|f\|_{C^{k,\gamma}(\Omega)} = \sum_{|\alpha| \leq k} (\sup_{x \in \Omega} |D^\alpha f(x)|) + \sum_{|\alpha|=k} (\sup_{x,y \in \Omega, x \neq y} \frac{|D^\alpha f(x) - D^\alpha f(y)|}{|x-y|^\gamma}).$$

Note that in the above definition D^α refers to a partial derivative, not a weak derivative. Since D^α is a continuous operator for all multi-indices α, Hölder spaces turn out to be Banach spaces, as exemplified by the following theorem:

Theorem 7.5.1. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Let $k \in \mathbb{Z}_{\geq 0}$ and $\gamma \in (0,1]$. Then, $C^{k,\gamma}(\Omega)$ is a Banach space.*

Proof. Assume that Ω is an open subset of $\mathbb{R}^n$. Assume that $\gamma \in (0,1]$ and $k \in \mathbb{Z}_{\geq 0}$. Suppose that the sequence $\{f_m\}$ in $C^{k,\gamma}(\Omega)$ is Cauchy. Then, for all $x \in \Omega$, the sequence $\{f_m(x)\}$ is a Cauchy sequence in $\mathbb{R}$ and thus, uniformly converges to $f(x)$. Since derivative operators are continuous, the sequence $\{D^\alpha f_m\}$ is also Cauchy and hence, by the same argument, the sequence $\{D^\alpha f_m(x)\}$ in $\mathbb{R}$ converges to $D^\alpha f(x)$ uniformly on Ω. This is enough to show that the expression (which appears in the norm for $C^{k,\gamma}(\Omega)$)

$$\sum_{|\alpha| \leq k} (\sup_{x \in \Omega} |D^\alpha f_m(x) - D^\alpha f(x)|)$$

tends to 0 as $m \to \infty$. Now, it suffices to show that the second summand in the Hölder space norm also converges. This amounts to showing that

$$\lim_{m \to \infty} \sum_{|\alpha|=k} (\sup_{x,y \in \Omega, x \neq y} \frac{|D^\alpha (f_m - f)(x) - D^\alpha (f_m - f)(y)|}{|x-y|^\gamma}) = 0.$$

Since the sequence $\{f_m\}$ is Cauchy, we have

$$\lim_{m,n \to \infty} \sum_{|\alpha|=k} (\sup_{x,y \in \Omega, x \neq y} \frac{|D^\alpha (f_m - f_n)(x) - D^\alpha (f_m - f_n)(y)|}{|x-y|^\gamma}) = 0.$$

Thus, for all $\epsilon \in \mathbb{R}_{>0}$, there exists an index $N \in \mathbb{Z}_{>0}$ such that for all $m, n > N$,

$$\sup_{x,y\in\Omega,x\neq y} \frac{|D^\alpha(f_m - f_n)(x) - D^\alpha(f_m - f_n)(y)|}{|x-y|^\gamma} \leq \epsilon$$

Fix m and let $n \to \infty$. This shows that for all $m \geq N$,

$$\sup_{x,y\in\Omega,x\neq y} \frac{|D^\alpha(f_m - f)(x) - D^\alpha(f_m - f)(y)|}{|x-y|^\gamma} \leq \epsilon$$

Taking the limit as $m \to 0$ the gives us the desired result.

□

Before we dive into the proof of various embedding theorems pertaining to Sobolev spaces, we will highlight the general approach towards such results. The goal is to show that the Sobolev space $W^{k,p}(\Omega)$ lies in some other Banach space X.

1. We first specialise to the case where $u \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$. In this step, we prove that $u \in X$ for some Banach space X and that there exists a constant $C \in \mathbb{R}_{>0}$ which depends on k, p and Ω, but not on u such that for all $u \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$,

$$\|u\|_X \leq C\|u\|_{W^{k,p}(\Omega)}.$$

2. Since $C^\infty(\Omega)$ is dense in $W^{k,p}(\Omega)$, for all $u \in W^{k,p}(\Omega)$, we can find a sequence $\{u_n\}$ of smooth functions such that $\|u - u_n\|_{W^{k,p}(\Omega)} \to 0$ as $n \to \infty$. Using the bound in step 1, this reveals that

$$\limsup_{m,n\to\infty} \|u_m - u_n\|_X \leq \limsup_{m,n\to\infty} C\|u_m - u_n\|_{W^{k,p}(\Omega)} = 0.$$

   Thus, $\{u_n\}$ is also a Cauchy sequence in the space X. Since X was assumed to be a Banach space, $u_n \to \tilde{u}$ for some $\tilde{u} \in X$. Expanding the definition of the norm $W^{k,p}(\Omega)$, we conclude that $\tilde{u}(x) = u(x)$ for almost all $x \in \Omega$. The conclusion here is that, up to a modification on a null set $N \subseteq \Omega$, each function $u \in W^{k,p}(\Omega)$ is also contained in X.

To see how this above proof template works, we will now delve into our first embedding theorem. Once again, the result is attributed to Morrey. The first step is to prove Morrey's inequality.

Lemma 7.5.2. *Let $n < p < \infty$ and set $\gamma = 1 - \frac{n}{p} > 0$. Then, there exists a constant $C \in \mathbb{R}_{>0}$, which depends only on p and n such that for all $u \in C^1(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$,*

$$\|u\|_{C^{0,\gamma}(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(\mathbb{R}^n)}.$$

The proof that we will give for Morrey's inequality originates from [LE98], rather than [AB10].

Proof. Assume that $n < p < \infty$ and $\gamma = 1 - \frac{n}{p} > 0$. Assume that $u \in C^1(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$. Take a ball $B(x,r) \subseteq \mathbb{R}^n$.

To show: (a) There exists a constant $C \in \mathbb{R}_{>0}$, which depends on n, such that

$$\frac{1}{m(B(x,r))}\int_{B(x,r)} |u(y) - u(x)|\ dy \leq C\int_{B(x,r)} \frac{|\nabla u(y)|}{|y-x|^{n-1}}\ dy.$$

(a) Take a point $w \in \partial B(0,1)$. Then observe that for all $0 < r < s$,

$$\begin{aligned} |u(x+sw) - u(x)| &= |\int_0^s \frac{d}{dt}u(x+tw)\ dt| \\ &= |\int_0^s \nabla u(x+tw)w\ dt| \quad \text{(Chain Rule)} \\ &\leq \int_0^s |\nabla u(x+tw)||w|\ dt \\ &= \int_0^s |\nabla u(x+tw)|\ dt. \end{aligned}$$

By integrating over the boundary $\partial B(0,1)$, we have

$$\begin{aligned} \int_{\partial B(0,1)} |u(x+sw) - u(x)|\ dS &\leq \int_0^s \int_{\partial B(0,1)} |\nabla u(x+tw)|\ dSdt \\ &= \int_0^s \int_{\partial B(0,1)} |\nabla u(x+tw)| \frac{t^{n-1}}{t^{n-1}}\ dSdt. \end{aligned}$$

The above integral has a radial component, which constitutes the integral over t, and an angular component, which is the integral over $\partial B(0,1)$. This suggests that we use general polar coordinates to simplify the RHS. If we let $y = x + tw$ so that $t = |x - y|$, we obtain (keep in mind that the corresponding Jacobian is proportional to t^{n-1})

$$\int_{\partial B(0,1)} |u(x+sw)-u(x)|\, dS \le \int_{B(x,s)} \frac{|\nabla u(y)|}{|x-y|^{n-1}}\, dy$$
$$\le \int_{B(x,r)} \frac{|\nabla u(y)|}{|x-y|^{n-1}}\, dy.$$

Now if we multiply by s^{n-1} and integrate both sides from 0 to r with respect to s. We finally obtain

$$\int_{B(x,r)} |u(y)-u(x)|\, dy \le \frac{r^n}{n} \int_{B(x,r)} \frac{|\nabla u(y)|}{|x-y|^{n-1}}\, dy.$$

This gives the intermediate result we are after.

Now, we tackle the first summand of the Hölder norm. Using part (a), we compute that for $x \in \mathbb{R}^n$

$$\begin{aligned}
|u(x)| &\le \frac{1}{m(B(x,1))} \int_{B(x,1)} |u(x)-u(y)|\, dy + \frac{1}{m(B(x,1))} \int_{B(x,1)} |u(y)|\, dy \\
&\le C \int_{B(x,r)} \frac{|\nabla u(y)|}{|x-y|^{n-1}}\, dy + C\|u\|_{L^p(B(x,1))} \quad \text{(Part (a))} \\
&\le C\left(\int_{\mathbb{R}^n} |\nabla u(y)|^p\, dy\right)^{1/p} \left(\int_{\mathbb{R}^n} \frac{1}{|x-y|^{p(n-1)/(p-1)}}\, dy\right)^{(p-1)/p} + C\|u\|_{L^p(\mathbb{R}^n)} \quad \text{(Hölder)} \\
&\le C\|u\|_{W^{1,p}(\mathbb{R}^n)}
\end{aligned}$$

since $\frac{p(n-1)}{p-1} < n$ and as a result,

$$\left(\int_{\mathbb{R}^n} \frac{1}{|x-y|^{p(n-1)/(p-1)}}\, dy\right)^{(p-1)/p} < \infty.$$

Now, if we take the supremum over both sides, we deduce the inequality $\sup_{x\in\mathbb{R}^n}|u(x)| \le C\|u\|_{W^{1,p}(\mathbb{R}^n)}$ for some constant $C \in \mathbb{R}_{>0}$.

Finally, we will bound the second summand in the Hölder norm. Assume that $x, y \in \mathbb{R}^n$. Set $r = |x-y|$ and $W = B(x,r) \cap B(y,r)$. Note that

$$|u(x)-u(z)| \le \frac{1}{m(W)} \int_W |u(x)-u(z)|\, dz + \frac{1}{m(W)} \int_W |u(y)-u(z)|\, dz.$$

Again, we can use part (a) to deduce that

$$\begin{aligned}
\frac{1}{m(W)}\int_W |u(x)-u(z)|\,dz &\le \frac{C}{m(B(x,r))}\int_{B(x,r)} |u(x)-u(z)|\,dz \\
&\le C(\int_{B(x,r)} |\nabla u(y)|^p\,dy)^{1/p}(\int_{B(x,r)} \frac{1}{|x-z|^{p(n-1)/(p-1)}}\,dz)^{(p-1)/p} \\
&\le C(\int_{B(x,r)} |\nabla u(y)|^p\,dy)^{1/p}(\int_{B(x,r)} \frac{1}{r^{p(n-1)/(p-1)}}\,dz)^{(p-1)/p} \\
&= Cr^{1-\frac{n}{p}}\|\nabla u\|_{L^p(\mathbb{R}^n)}.
\end{aligned}$$

The second inequality combines Hölder's inequality and the bound in part (a). The third inequality follows from the fact that the integral is done over the ball $B(x,r)$. Note that the same result can be found by simply replacing x with y. Thus, an upper bound for the quantity $|u(x)-u(y)|$ is

$$|u(x)-u(y)| \le Cr^{1-\frac{n}{p}}\|\nabla u\|_{L^p(\mathbb{R}^n)} = C|x-y|^{1-\frac{n}{p}}\|\nabla u\|_{L^p(\mathbb{R}^n)}$$

Now, we have

$$\begin{aligned}
\|u\|_{C^{0,\gamma}(\mathbb{R}^n)} &= \sup_{x\neq y}\frac{|u(x)-u(y)|}{|x-y|^{1-\frac{n}{p}}} \\
&\le C\|\nabla u\|_{L^p(\mathbb{R}^n)}.
\end{aligned}$$

for some constant $C \in \mathbb{R}_{>0}$, which depends on p and n.

□

Executing the second step in the "proof template", we arrive at the required embedding theorem.

Theorem 7.5.3. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1 boundary. Let $n < p < \infty$ and $\gamma = 1 - \frac{n}{p} > 0$. Then, every function $f \in W^{1,p}(\Omega)$ is almost eveywhere equal to another function $\tilde{f} \in C^{0,\gamma}(\Omega)$. Moreover, there exists a constant $C \in \mathbb{R}_{>0}$ such that for all $f \in W^{1,p}(\Omega)$,*

$$\|\tilde{f}\|_{C^{0,\gamma}} \le C\|f\|_{W^{1,p}}.$$

Proof. Assume that Ω is the open set defined as above. Assume that n, p, γ are defined as above. Define

$$\tilde{\Omega} = \{x \in \mathbb{R}^n \mid d(x,\Omega) < 1\}.$$

Using 7.4.3, we deduce the existence of a bounded extension operator $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ which sends $f \in W^{1,p}(\Omega)$ to $Ef \in W^{1,p}(\mathbb{R}^n)$ whose support is contained in $\tilde{\Omega}$.

Now, we use the fact that $C^1(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$ in order to find a sequence of functions $\{g_n\}$ in $C^1(\mathbb{R}^n)$ such that $g_n \to Ef$ as $n \to \infty$. Morrey's inequality (see 7.5.2) tells us that

$$\limsup_{m,n\to\infty} \|g_m - g_n\|_{C^{0,\gamma}} \leq C \limsup_{m,n\to\infty} \|g_m - g_n\|_{W^{1,p}} = 0$$

because every convergent sequence is Cauchy. Interestingly, this shows that $\{g_n\}$ is a Cauchy sequence in the Hölder space $C^{0,\gamma}(\mathbb{R}^n)$. Since $C^{0,\gamma}(\mathbb{R}^n)$ is a Banach space, the sequence $\{g_n\}$ must converge uniformly to a function $g \in C^{0,\gamma}(\mathbb{R}^n)$ for $x \in \mathbb{R}^n$.

Since $g_n \to Ef$ in $W^{1,p}(\mathbb{R}^n)$, $g(x) = (Ef)(x)$ for almost all $x \in \mathbb{R}^n$. From the definition of E, we must have $g(x) = f(x)$ for almost all $x \in \Omega$. The bound follows from the fact that E is a bounded linear operator and Morrey's inequality (7.5.2). □

Before we proceed, we will highlight a particularly important consequence of Morrey's inequality. If $\Omega \subseteq \mathbb{R}^n$ is a bounded open set with C^1 boundary and $p > n$, then for all $w \in W^{1,p}(\Omega) \subseteq C^{0,\gamma}(\Omega)$, we must have

$$|w(x) - w(y)| \leq C|x-y|^{1-\frac{n}{p}} \left(\int_{B(x,|y-x|)} |\nabla w(z)|^p \, dz \right)^{1/p}.$$

for all $x, y \in \Omega$. This was found during the proof of Morrey's inequality. We will use this bound to demonstrate a differentiability property of the Sobolev space $W^{1,p}(\Omega)$.

Theorem 7.5.4. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1 boundary and $u \in W^{1,p}_{loc}(\Omega)$ for some $p > n$. Then, u is differentiable for almost all $x \in \Omega$ and its classical gradient is equal to its weak gradient*

$$\nabla u = (D_{x_1} u, \ldots, D_{x_n} u).$$

Proof. Assume that Ω is defined as above. Assume that $u \in W^{1,p}_{loc}(\Omega)$ for some $p > n$. Then, its weak derivatives $D_{x_i} u \in L^p_{loc}(\Omega)$ for all $i \in \{1, \ldots, n\}$. Hence, the weak gradient ∇u is also in $L^p_{loc}(\Omega)$.

We begin by applying the Lebesgue differentiation theorem to deduce that for almost all $x \in \Omega$,

$$\lim_{r\to 0}\frac{1}{m(B(x,r))}\int_{B(x,r)}|\nabla u(x)-\nabla u(z)|^p\,dz=0.$$

Take a point $x\in\Omega$ such that the above limit holds. Then, define

$$w(y)=u(y)-u(x)-\nabla u(x)(y-x)\in W^{1,p}_{loc}(\Omega).$$

Then, we bound the quantity $|w(y)-w(x)|$ by using the estimate found in the proof of Morrey's inequality:

$$\begin{aligned}|w(y)-w(x)|&=|w(y)|\\&=|u(y)-u(x)-\nabla u(x)(y-x)|\\&=|\nabla u(x)x-u(x)-(\nabla u(x)y-u(y))|\\&\leq C|x-y|^{1-\frac{n}{p}}(\int_{B(x,|y-x|)}|\nabla u(x)-\nabla u(z)|^p\,dz)^{1/p}.\end{aligned}$$

Dividing both sides by $|y-x|$ and then taking the limit as $|y-x|\to 0$, we find that

$$\lim_{|y-x|\to 0}\frac{|w(y)-w(x)|}{|y-x|}=0$$

due to our choice of $x\in\Omega$. Thus, from the definition of w, we conclude that u is differentiable at x in the classical sense and its classical gradient coincides with the weak gradient $\nabla u(x)$ for almost all $x\in\Omega$.

□

The next result we will prove is the *Gagliardo-Nirenberg inequality*, which is valid in the regime $1\leq p<n$. First, we will make some useful remarks.

We define the *Sobolev conjugate* of p to be

$$p^*=\frac{np}{n-p}\text{ so that }\frac{1}{p^*}=\frac{1}{p}-\frac{1}{n}.$$

Secondly, we need to keep in mind the following application of the generalised Hölder inequality. Once again let $\Omega\subseteq\mathbb{R}^n$ be open. Let $g_1,\ldots,g_{n-1}\in L^1(\Omega)$. Then, for all $i\in\{1,\ldots,n-1\}$, $g_i^{\frac{1}{n-1}}\in L^{n-1}(\Omega)$. By the generalised Hölder inequaltiy, we have

$$\int_\Omega\prod_{i=1}^{n-1}g_i^{\frac{1}{n-1}}\,ds\leq\prod_{i=1}^{n-1}\|g_i^{\frac{1}{n-1}}\|_{L^{n-1}}=\prod_{i=1}^{n-1}\|g_i\|_{L^1}^{\frac{1}{n-1}}$$

Now, we will proceed to prove our second embedding theorem:

Lemma 7.5.5. *Let $p, n \in \mathbb{R}_{>0}$ such that $1 \leq p < n$. Then, there exists a constant $C \in \mathbb{R}_{>0}$, which depends only on p and n such that for all $f \in W^{1,p}(\mathbb{R}^n)$,*

$$\|f\|_{L^{p^*}(\mathbb{R}^n)} \leq C\|\nabla f\|_{L^p(\mathbb{R}^n)}.$$

Proof. Assume that $p, n \in \mathbb{R}_{>0}$ such that $1 \leq p < n$. Utilising the density of smooth functions in $W^{1,p}(\mathbb{R}^n)$, it suffices to prove the above inequality for the special case where $f \in C_c^\infty(\mathbb{R}^n)$. Using the fact that f is a test function, we have for all $i \in \{1, \ldots, n\}$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$

$$f(x) = \int_{-\infty}^{x_i} D_{x_i} f(x_1, \ldots, s_i, \ldots, x_n) \, ds_i.$$

Thus, we have a preliminary bound on $|f(x)|$:

$$|f(x)| = |\int_{-\infty}^{x_i} D_{x_i} f(x_1, \ldots, s_i, \ldots, x_n) \, ds_i| \leq \int_{-\infty}^{\infty} |D_{x_i} f(x_1, \ldots, s_i, \ldots, x_n)| \, ds_i.$$

The point here is that this holds for all $i \in \{1, \ldots, n\}$. So, taking both sides to the power of n, we obtain

$$|f(x)|^n \leq \prod_{i=1}^n \int_{-\infty}^{\infty} |D_{x_i} f(x_1, \ldots, s_i, \ldots, x_n)| \, ds_i$$

So,

$$|f(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^n (\int_{-\infty}^{\infty} |D_{x_i} f(x_1, \ldots, s_i, \ldots, x_n)| \, ds_i)^{\frac{1}{n-1}}.$$

Now, we integrate over $\mathbb{R}$ with respect to the variable x_1. In this case, the first integral in the product does not depend on x_1. So, our inequality becomes

$$\int_{-\infty}^{\infty} |f(x)|^{\frac{n}{n-1}} \, dx_1 \leq (\int_{-\infty}^{\infty} |D_{x_1} f| \, ds_1)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} \prod_{i=2}^n (\int_{-\infty}^{\infty} |D_{x_i} f(x_1, \ldots, s_i, \ldots, x_n)| \, ds_i)^{\frac{1}{n-1}} \, dx_1.$$

We can iterate the above method by integrating with respect to the remaining $n-1$ variables. We obtain

$$\int_{\mathbb{R}^n} |f|^{\frac{n}{n-1}}\ dx \le \prod_{i=1}^{n}(\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|D_{x_i}f|\ dx_1\ldots dx_n)^{\frac{1}{n-1}}$$
$$\le (\int_{\mathbb{R}^n}|\nabla f|\ dx)^{\frac{n}{n-1}}.$$

Therefore,

$$(\int_{\mathbb{R}^n}|f|^{\frac{n}{n-1}}\ dx)^{\frac{n-1}{n}} \le \int_{\mathbb{R}^n}|\nabla f|\ dx$$

which proves the result for the case where $p = 1$ and $p^* = n/(n-1)$. Of course, this holds for $f \in W^{1,p}(\mathbb{R}^n)$ since $C_c^\infty(\mathbb{R}^n)$ is a dense subset.

Now assume that $1 < p < n$ and that $f \in W^{1,p}(\mathbb{R}^n)$ once again. Let us apply the special case above to the function $g = |f|^\beta$ for some $\beta \in \mathbb{R}_{>0}$. Then, we must have

$$\begin{aligned}(\int_{\mathbb{R}^n}|f|^{\frac{\beta n}{n-1}}\ dx)^{\frac{n-1}{n}} &\le \int_{\mathbb{R}^n}|\nabla f^\beta|\ dx \quad \text{(Special Case)}\\ &= \int_{\mathbb{R}^n}\beta|f|^{\beta-1}|\nabla f|\ dx\\ &\le \beta(\int_{\mathbb{R}^n}|f|^{\frac{p(\beta-1)}{p-1}}\ dx)^{\frac{p-1}{p}}(\int_{\mathbb{R}^n}|\nabla f|^p\ dx)^{\frac{1}{p}} \quad \text{(Hölder)}\end{aligned}$$

If we select $\beta = p(n-1)/(n-p)$, then we have

$$\frac{\beta n}{n-1} = \frac{p(\beta-1)}{p-1} = \frac{np}{n-p} = p^*.$$

Making this substitution, our inequality reduces to

$$(\int_{\mathbb{R}^n}|f|^{p^*}\ dx)^{\frac{n-1}{n}} \le \beta(\int_{\mathbb{R}^n}|f|^{p^*}\ dx)^{\frac{p-1}{p}}(\int_{\mathbb{R}^n}|\nabla f|^p\ dx)^{\frac{1}{p}}$$

Now it suffices for us to note that

$$\frac{n-1}{n} - \frac{p-1}{p} = \frac{n-p}{np} = \frac{1}{p} - \frac{1}{n} = \frac{1}{p^*}.$$

Hence, by dividing both sides by $(\int_{\mathbb{R}^n}|f|^{p^*}\ dx)^{\frac{p-1}{p}}$, we deduce that

$$(\int_{\mathbb{R}^n} |f|^{p^*} \, dx)^{\frac{1}{p^*}} \le \beta(\int_{\mathbb{R}^n} |\nabla f|^p \, dx)^{\frac{1}{p}}$$

as required.

□

Recall that if $\Omega \subseteq \mathbb{R}^n$ is bounded, then $L^q(\Omega) \subseteq L^{p^*}(\Omega)$ for all $q \in [1, p^*]$. Now, we can establish our next embedding theorem.

Theorem 7.5.6. *Let $\Omega \subseteq \mathbb{R}^n$ denote a bounded open set with C^1 boundary and let $1 \le p < n$. Then, for all $q \in [1, p^*]$ (recall that $p^* = np/(n-p)$), there exists a constant $C \in \mathbb{R}_{>0}$ such that for all $f \in W^{1,p}(\Omega)$,*

$$\|f\|_{L^q(\Omega)} \le C\|f\|_{W^{1,p}(\Omega)}.$$

Proof. The proof is very similar to the first embedding theorem. Again, let

$$\tilde{\Omega} = \{x \in \mathbb{R}^n \mid d(x, \Omega) < 1\}.$$

Using 7.4.3, we deduce the existence of a bounded extension operator $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ which sends $f \in W^{1,p}(\Omega)$ to $Ef \in W^{1,p}(\mathbb{R}^n)$ whose support is contained in $\tilde{\Omega}$.

Now apply the Gagliardo-Nirenberg inequality to Ef in order to deduce the existence of constants $C_1, C_2, C_3 \in \mathbb{R}_{>0}$ such that

$$\begin{aligned} \|f\|_{L^q(\Omega)} &\le C_1\|f\|_{L^{p^*}(\Omega)} \quad (L^q(\Omega) \subseteq L^{p^*}(\Omega)) \\ &\le C_2\|Ef\|_{L^{p^*}(\mathbb{R}^n)} \\ &\le C_3\|f\|_{W^{1,p}(\Omega)} \quad \text{(Gagliardo-Nirenberg)}. \end{aligned}$$

□

So far, we have established embeddings for the space $W^{1,p}(\Omega)$, where $p < \infty$. We would like to push this agenda forwards and describe embeddings for the Sobolev space $W^{k,p}(\Omega)$ for all $k \in \mathbb{Z}_{>0}$. We will be making extensive use of 7.5.2 and 7.5.5. First, we will make the necessary definitions.

Definition 7.5.3. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1 boundary. Let $u \in W^{k,p}(\Omega)$. Then, the **net smoothness** of u is defined by $k - \frac{n}{p}$.

Define $m \in \mathbb{Z}_{>0}$ and $\gamma \in [0, 1)$ to be the **integer part** and **fractional part** of $k - \frac{n}{p}$ respectively. That is, in the usual notation,

$$m = \lfloor k - \frac{n}{p} \rfloor \text{ and } \gamma = \{k - \frac{n}{p}\}.$$

Definition 7.5.4. Let X and Y be Banach spaces, with norms $\|-\|_X$ and $\|-\|_Y$ respectively. We say that X is **continuously embedded** in Y if $X \subseteq Y$ and there exists a constant C such that for all $u \in X$, $\|u\|_Y \leq C\|u\|_X$.

The next theorem reveals a whole class of general Sobolev embeddings.

Theorem 7.5.7 (General Sobolev Embeddings)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1 boundary. Consider the space $W^{k,p}(\Omega)$ and set*

$$m = \lfloor k - \frac{n}{p} \rfloor \textit{ and } \gamma = \{k - \frac{n}{p}\}.$$

Then,

1. *If $k - \frac{n}{p} < 0$, then $W^{k,p}(\Omega) \subseteq L^q(\Omega)$ where*

$$\frac{1}{q} = \frac{1}{p} - \frac{k}{n} = \frac{1}{n}(\frac{n}{p} - k).$$

2. *If $k - \frac{n}{p} = 0$, then $W^{k,p}(\Omega) \subseteq L^q(\Omega)$ for all $1 \leq q < \infty$.*

3. *If $m \geq 0$ and $\gamma > 0$, then $W^{k,p}(\Omega) \subseteq C^{m,\gamma}(\Omega)$.*

4. *If $m \geq 1$ and $\gamma = 0$, then for all $\gamma' \in [0, 1)$, we have the inclusion $W^{k,p}(\Omega) \subseteq C^{m-1,\gamma'}(\Omega)$.*

Before we state the proof below, it should be emphasised that the statement $W^{k,p}(\Omega) \subseteq C^{m,\gamma}(\Omega)$ means that for all $u \in W^{k,p}(\Omega)$, there exists a function $\tilde{u} \in C^{m,\gamma}(\Omega)$ such that for almost all $x \in \Omega$, $u(x) = \tilde{u}(x)$. Furthermore, there exists a constant $C \in \mathbb{R}_{>0}$, which depends on k, p, m, γ, but not on u, such that

$$\|u\|_{C^{m,\gamma}(\Omega)} \leq C\|\tilde{u}\|_{W^{k,p}(\Omega)}.$$

In other words, the above equation implies a continuous embedding. Similar statement apply for the other inclusions.

Proof. We will first prove statement 1. So, assume that the net smoothness $k - \frac{n}{p} < 0$. Assume that $u \in W^{k,p}(\Omega)$. From the definition of a Sobolev space, $D^\alpha u \in L^p(\Omega)$ for all multi-indices α such that $|\alpha| \leq k$. Now, we can apply the Gagliardo-Nirenberg inequality to find that for all multi-indices β with $|\beta| \leq k - 1$,

$$\|D^\beta u\|_{L^{p^*}(\Omega)} \leq C\|\nabla D^\beta u\|_{L^p(\Omega)} \leq C\|u\|_{W^{k,p}(\Omega)}.$$

Therefore, $D^\beta u \in L^{p^*}(\Omega)$ for all β satisfying $|\beta| \leq k - 1$. So, $u \in W^{k-1,p^*}(\Omega)$. Recall that

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}.$$

The main point in the proof is that we can iterate the above argument. Set $p_1 = p^*$ and $p_j = p_{j-1}^*$ for all $j \in \mathbb{Z}_{>0}$ so that

$$\frac{1}{p_1} = \frac{1}{p} - \frac{1}{n} \text{ and } \frac{1}{p_j} = \frac{1}{p_{j-1}} - \frac{1}{n} = \frac{1}{p} - \frac{j}{n}.$$

In order to iterate the above argument, we will apply the Gagliardo-Nirenberg inequality k times in order to obtain the following chain of inclusions:

$$W^{k,p}(\Omega) \subseteq W^{k-1,p_1}(\Omega) \subseteq \cdots \subseteq W^{k-j,p_j}(\Omega) \subseteq \cdots \subseteq W^{0,p_k}(\Omega).$$

Therefore, $u \in W^{0,p_k}(\Omega) = L^{p_k}(\Omega)$. This completes the proof of statement 1.

In order to prove statement 2, assume that $kp = n$. Using the same argument as above, we find that if $u \in W^{k,p}(\Omega)$, then

$$u \in W^{1,p_{k-1}}(\Omega) = W^{1,n}(\Omega) \subseteq W^{1,n-\epsilon}(\Omega)$$

for all $\epsilon \in (0, n]$. Note that $p_{k-1} = n$ because

$$\frac{1}{p_{k-1}} = \frac{1}{p} - \frac{k-1}{n} = \frac{1}{p} - \frac{k-1}{kp} = \frac{1}{kp} = n.$$

Utilising the Gagliardo-Nirenberg inequality one more time, we find that $u \in W^{1,n-\epsilon}(\Omega) \subseteq L^q(\Omega)$, where

$$q = \frac{n(n-\epsilon)}{n-(n-\epsilon)} = \frac{n^2 - \epsilon n}{\epsilon}.$$

Since $\epsilon \in (0, n]$ was arbitrary, we find that statement 2 is true for all $q \in [1, \infty)$.

To prove statement 3, assume that $m \geq 0$ and $\gamma > 0$. Assume that $u \in W^{k,p}(\Omega)$. Choose $j \in \mathbb{Z}_{>0}$ to be the smallest integer such that $p_j > n$. Then, $u \in W^{k,p}(\Omega) \subseteq W^{k-j,p_j}(\Omega)$, which means that for all multi-indices α with $|\alpha| \leq k - j - 1$, we can apply 7.5.2 to deduce that

$$D^\alpha u \in W^{1,p_j}(\Omega) \subseteq C^{0,\gamma}(\Omega)$$

where

$$\gamma = 1 - \frac{n}{p_j} = 1 - \frac{n}{p} + j.$$

The important point here which motivates our choice of $j \in \mathbb{Z}_{>0}$ is that $\gamma \in (0, 1]$ and

$$k - \frac{n}{p} = (k - j - 1) + (1 - \frac{n}{p} + j) = (k - j - 1) + \gamma.$$

So, γ is the fractional part of $k - \frac{n}{p}$ and $k - j - 1$ is the integer part of $k - \frac{n}{p}$. Now since $D^\alpha u \in C^{0,\gamma}(\Omega)$ where $|\alpha| \leq k - j - 1$, we deduce that $u \in C^{k-j-1,\gamma}(\Omega)$, which completes the proof of the third statement (with $m = k - j - 1$).

To see that the fourth and final statement holds, assume that $j = n/p \in \mathbb{Z}_{>0}$. Assume that $u \in W^{k,p}(\Omega)$. Then, the Gagliardo-Nirenberg inequality, as depicted in statement 2, can be used to show that for all multi-indices α with $|\alpha| \leq j - 1$,

$$D^\alpha u \in W^{k-j,p}(\Omega) \subseteq W^{1,q}(\Omega)$$

for all $q \in [1, \infty)$. This is a result of statement 2. Now, Morrey's inequality gives

$$D^\alpha u \subseteq W^{1,q}(\Omega) \subseteq C^{0,1-\frac{n}{q}}(\Omega)$$

and $u \in C^{j-1,1-\frac{n}{q}}(\Omega)$. Since $q \in [1, \infty)$ was arbitrary, we deduce statement 4.

□

7.6 Rellich-Kondrachov Theorem

As the title suggests, the main result of this section is the Rellich-Kondrachov theorem, which is yet another embedding theorem of Sobolev spaces. However, the statement of this theorem is actually a compact embedding - we will clear this up later.

First, we will make a preliminary observation with the purpose of motivating the theorem. Suppose that $\Omega \subseteq \mathbb{R}^n$ is a bounded open set with C^1 boundary. Observe that if $p > n$, then

$$m = \lfloor 1 - \frac{n}{p} \rfloor = 0 \text{ and } \gamma = \{1 - \frac{n}{p}\} = 1 - \frac{n}{p} > 0.$$

By 7.5.7 (in particular, the third statement), we have the inclusion $W^{1,p}(\Omega) \subseteq C^{0,\gamma}(\Omega)$, revealing that for all $u \in W^{1,p}(\Omega)$, u is Hölder continuous. More specifically, if $\{u_m\}_{m \in \mathbb{Z}_{>0}}$ is a bounded sequence in $W^{1,p}(\Omega)$, then each function u_m must be equicontinuous and uniformly bounded, due to the norm on $C^{0,\gamma}(\Omega)$. From the corollary of 4.3.6, we can extract a subsequence $\{u_{m_j}\}$ which converges uniformly to a continuous function u on Ω. Now, since Ω is bounded, we must have $\|u_{m_j} - u\|_{L^q(\Omega)} \to 0$ for all $q \in [1, \infty]$. Thus, the embedding $W^{1,p}(\Omega) \subseteq L^q(\Omega)$ is compact for all $p > n$ and $q \in [1, \infty]$.

Definition 7.6.1. Let X and Y be Banach spaces. We say that X is **compactly embedded** in Y if $X \subseteq Y$ and the canonical inclusion map $\iota : X \to Y$ is a compact linear operator. We use the notation $X \subset\subset Y$ to denote the statement "X is compactly embedded in Y".

The Rellich-Kondrachov theorem focuses on the case where $p < n$. We will now state this important theorem below:

Theorem 7.6.1 (Rellich-Kondrachov)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with C^1 boundary. Suppose that $1 \leq p < n$. The for all $q \in [1, p^*)$, where $p^* = \frac{np}{n-p}$, we have the compact embedding $W^{1,p}(\Omega) \subset\subset L^q(\Omega)$.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is a bounded open set with C^1 boundary. Assume that $1 \leq p < n$ and p^* is defined as above. Assume that $q \in [1, p^*)$. The theorem 7.4.3 tells us that we can assume that every function $u \in W^{1,p}(\Omega)$ is defined over all $\mathbb{R}^n$ and vanishes outside a fixed bounded open set $\tilde{\Omega} \subseteq \mathbb{R}^n$.

Let $\{u_m\}_{m \in \mathbb{Z}_{>0}}$ be a bounded sequence in $W^{1,p}(\tilde{\Omega})$. Then, since $q < p^*$ and Ω is bounded, we have the following upper bound:

$$
\begin{aligned}
\|u_m\|_{L^q(\mathbb{R}^n)} &= \|u_m\|_{L^q(\tilde{\Omega})} \quad (\text{support contained in } \tilde{\Omega}) \\
&\leq C\|u_m\|_{L^{p^*}(\tilde{\Omega})} \quad (\text{since } L^q(\tilde{\Omega}) \subseteq L^{p^*}(\tilde{\Omega})) \\
&\leq C'\|u_m\|_{W^{1,p}(\tilde{\Omega})} \quad (\text{Gagliardo-Nirenberg})
\end{aligned}
$$

Here, $C, C' \in \mathbb{R}_{>0}$ are constants. The main message from the above reasoning is that the sequence $\{u_m\}$ is uniformly bounded in $L^q(\tilde{\Omega})$ and that $W^{1,p}(\Omega) \subseteq L^q(\Omega)$, due to the above bound.

Now assume that $\epsilon \in \mathbb{R}_{>0}$. Consider the mollified sequence $\{u_m^\epsilon\}$, where $u_m^\epsilon = J_\epsilon * u_m$. These functions are all supported inside $\tilde{\Omega}$.

To show: (a) $\|u_m^\epsilon - u_m\|_{L^q(\tilde{\Omega})} \to 0$ as $\epsilon \to 0$. This convergence is uniform with respect to m.

(a) Using the fact that smooth functions are dense in $W^{1,p}(\tilde{\Omega})$ (see 7.4.1), it suffices to demonstrate the bound in the case where $u_m \in C^\infty(\tilde{\Omega})$. Using the definition of a mollification, we argue as follows:

$$
\begin{aligned}
u_m^\epsilon(x) - u_m(x) &= \int_{B(y',\epsilon)} J_\epsilon(y')[u_m(x-y') - u_m(x)]\, dy' \\
&= \int_{B(y,\epsilon)} J(y)[u_m(x-\epsilon y) - u_m(x)]\, dy \quad (y' = \epsilon y) \\
&= \int_{B(y,\epsilon)} J(y)(\int_0^1 \frac{d}{dt}(u_m(x-\epsilon t y))\, dt)\, dy \\
&= -\epsilon \int_{B(y,\epsilon)} J(y)(\int_0^1 \nabla(u_m(x-\epsilon t y)) \cdot y\, dt)\, dy \quad (\text{Chain Rule})
\end{aligned}
$$

If we integrate both sides with respect to $x \in \tilde{\Omega}$, we obtain

$$
\int_{\tilde{\Omega}} |u_m^\epsilon(x) - u_m(x)|\, dx \leq \epsilon \int_{\tilde{\Omega}} \int_{B(y,\epsilon)} J(y)(\int_0^1 |\nabla(u_m(x-\epsilon t y))|\, dt)\, dy\, dx
$$

Setting $z = x - \epsilon t y$, the inequality simplifies even further to obtain

$$
\int_{\tilde{\Omega}} |u_m^\epsilon(x) - u_m(x)|\, dx \leq \epsilon \int_{\tilde{\Omega}} |\nabla u_m(z)|\, dz.
$$

Using 7.4.1, we conclude that the above estimate also holds for $u_m \in W^{1,p}(\tilde{\Omega})$. Hence,

$$\|u_m^\epsilon - u_m\|_{L^1(\tilde{\Omega})} \leq \epsilon \|\nabla u_m\|_{L^1(\tilde{\Omega})} \leq \epsilon C \|u_m\|_{W^{1,p}(\tilde{\Omega})}$$

for some constant $C \in \mathbb{R}_{>0}$. This proves the statement in part (a).

For the next argument, we will utilise an interpolation inequality (see [LE98, p. 623]). Select $\theta \in (0,1)$ such that

$$\frac{1}{q} = (\theta \cdot 1) + (1-\theta) \cdot \frac{1}{p^*}.$$

Then, we must have

$$\begin{aligned}
\|u_m^\epsilon - u_m\|_{L^q(\tilde{\Omega})} &= \left(\int_{\tilde{\Omega}} |u_m^\epsilon(x) - u_m(x)|^q \, dx\right)^{1/q} \\
&\leq \|u_m^\epsilon - u_m\|_{L^1(\tilde{\Omega})}^\theta \|u_m^\epsilon - u_m\|_{L^{p^*}(\tilde{\Omega})}^{(1-\theta)} \\
&\leq C_0 \epsilon^\theta.
\end{aligned}$$

The constant $C_0 \in \mathbb{R}_{>0}$ is independent of m. The L^{p^*} norm is bounded above by a constant, due to the Gagliardo-Nirenberg inequality once again (7.5.5). Now fix $\delta \in \mathbb{R}_{>0}$ and choose $\epsilon \in \mathbb{R}_{>0}$ small enough so that

$$\|u_m^\epsilon - u_m\|_{L^q(\tilde{\Omega})} \leq C_0 \epsilon^\theta \leq \frac{\delta}{2}.$$

To see that the sequence $\{u_m^\epsilon\}_{m \in \mathbb{Z}_{>0}}$ is equicontinuous, we observe that

$$|u_m^\epsilon| \leq \|J_\epsilon\|_{L^\infty} \|u_m\|_{L^1} \leq C_1$$

and

$$|\nabla u_m^\epsilon| \leq \|\nabla J_\epsilon\|_{L^\infty} \|u_m\|_{L^1} \leq C_2.$$

Here, the constants C_1 and C_2 depend on ϵ, but are independent of m. The above bounds originate once again from Hölder's inequality, with $p = 1$ and $q = \infty$.

Now, we can exploit this finding and use the Arzela-Ascoli theorem (4.3.6) to obtain a subsequence $\{u_{m_j}^\epsilon\}$ which converges uniformly in $\tilde{\Omega}$ to a continuous function u. Hence, by the triangle inequality

$$\begin{aligned}\limsup_{j,k\to\infty}\|u_{m_j}-u_{m_k}\|_{L^q(\tilde{\Omega})} &\le \limsup_{j,k\to\infty}(\|u_{m_j}-u^{\epsilon}_{m_j}\|_{L^q(\tilde{\Omega})}+\|u^{\epsilon}_{m_j}-u\|_{L^q(\tilde{\Omega})}\\ &\quad+\|u-u^{\epsilon}_{m_k}\|_{L^q(\tilde{\Omega})}+\|u^{\epsilon}_{m_k}-u_{m_k}\|_{L^q(\tilde{\Omega})})\\ &\le \delta.\end{aligned}$$

Now, we finish off the proof with a diagonalisation argument. Using the above bound, we can find an infinite set of indices I_1 such that the subsequence $\{u_m\}_{m\in I_1}$ satisfies for all $\ell, m \in I_1$,

$$\limsup_{\ell,m\to\infty}\|u_\ell - u_m\|_{L^q(\Omega)} \le 2^{-1}.$$

By induction, we can repeat this argument and construct an infinite set of indices $I_j \subseteq I_{j-1}$ for all $j \in \mathbb{Z}_{>0}$ such that for all $\ell, m \in I_j$

$$\limsup_{\ell,m\to\infty}\|u_\ell - u_m\|_{L^q(\Omega)} \le 2^{-j}.$$

Then, we choose a strictly increasing sequence of integers $\{m_j\}$ such that $m_j \in I_j$ for all $j \in \mathbb{Z}_{>0}$. The subsequence $\{u_{m_j}\}$ must satisfy

$$\limsup_{j,k\to\infty}\|u_{m_j} - u_{m_k}\|_{L^q(\Omega)} = 0$$

due to the fact that $I_j \subseteq I_{j-1}$ for all $j \in \mathbb{Z}_{>0}$. So, $\{u_{m_j}\}$ is a Cauchy sequence, which then converges to some limit $u \in L^q(\Omega)$. This is enough to show that $W^{1,p}(\Omega) \subset\subset L^q(\Omega)$.

□

To top off the entire chapter, we will look at an application of 7.6.1, which is also referred to as Poincaré's inequality.

Lemma 7.6.2. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded, connected and open set with C^1 boundary. Let $p \in [1, \infty]$. Then, there exists a constant $C \in \mathbb{R}_{>0}$ such that for all $u \in L^p(\Omega)$,*

$$\|u - \frac{1}{m(\Omega)}\int_\Omega u\, dx\|_{L^p(\Omega)} \le C\|\nabla u\|_{L^p(\Omega)}.$$

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is a bounded, connected and open set with C^1 boundary. Assume that $p \in [1, \infty]$. Suppose for the sake of contradiction that such a constant $C \in \mathbb{R}_{>0}$ does not exist. Then, there exists a sequence of functions $u_k \in W^{1,p}(\Omega)$ such that

$$\|u_k - \frac{1}{m(\Omega)} \int_\Omega u_k \, dx\|_{L^p(\Omega)} > k\|\nabla u_k\|_{L^p(\Omega)}$$

for all $k \in \mathbb{Z}_{>0}$. Define the sequence

$$v_k = \frac{u_k - \frac{1}{m(\Omega)} \int_\Omega u_k \, dx}{\|u_k - \frac{1}{m(\Omega)} \int_\Omega u_k \, dx\|_{L^p(\Omega)}}.$$

Note that the sequence $\{v_k\}$ in the Sobolev space $W^{1,p}(\Omega)$ must satisfy

$$\frac{1}{m(\Omega)} \int_\Omega v_k \, dx = 0, \|v_k\|_{L^p(\Omega)} = 1$$

and for all $k \in \mathbb{Z}_{>0}$ and $i \in \{1, \ldots, n\}$,

$$\begin{aligned}
\|D_{x_i} v_k\|_{L^p(\Omega)} &= (\int_\Omega |D_{x_i} v_k(x)|^p \, dx)^{1/p} \\
&\leq (\int_\Omega |\nabla v_k(x)|^p \, dx)^{1/p} \\
&= \|\nabla v_k\|_{L^p(\Omega)} \\
&\leq \frac{\|\nabla u_k\|_{L^p(\Omega)}}{\|u_k - \frac{1}{m(\Omega)} \int_\Omega u_k \, dx\|_{L^p(\Omega)}} \\
&< \frac{1}{k}.
\end{aligned}$$

The Rellich-Kondrachov theorem (7.6.1) tells us that $W^{1,p}(\Omega) \subset\subset L^p(\Omega)$. Since $\{v_k\}$ is a bounded sequence in $W^{1,p}(\Omega)$, there exists a subsequence $\{v_{k_j}\}$ which converges uniformly in $L^p(\Omega)$ to some continuous function $v \in L^p(\Omega)$. We also have $\nabla v_k \to 0$ in $L^p(\Omega)$. By 7.1.4, the zero function is the weak gradient of the limit function v and as a result,

$$\frac{1}{m(\Omega)} \int_\Omega v \, dx = \lim_{k \to \infty} \frac{1}{m(\Omega)} \int_\Omega v_k \, dx = 0.$$

Because $\nabla v = 0 \in L^p(\Omega)$, it must be constant on the connected set Ω and subsequently, $v(x) = 0$ for almost all $x \in \Omega$, since its average value is 0. But, this contradicts the fact that

$$\|v\|_{L^p(\Omega)} = \lim_{k \to \infty} \|v_k\|_{L^p(\Omega)} = 1.$$

□

Chapter 8

Applications to PDEs

8.1 Second order elliptic equations

As a fitting end to all of the theory developed in the previous chapters, we will apply the theory in order to understand partial differential equations. In particular, we will be looking at second order linear partial differential equations. As we will see later, we will be looking at three different types of PDEs.

Below is a summary of the main results we will be using to analyse PDEs.

1. Rellich-Kondrachov Theorem (7.6.1)
2. Lax-Milgram Theorem (3.6.3)
3. Characterisation of semigroups (6.7.1)
4. Hilbert-Schmidt Theorem (5.3.3)
5. The Fredholm alternative

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set and $a^{ij}, b^i, c : \Omega \to \mathbb{R}$ be measurable functions (with respect to Lebesgue measure on $\mathbb{R}^n$). Then, we define the operator L by

$$Lu = -\sum_{i,j=1}^{n} (a^{ij}(x)u_{x_i})_{x_j} + \sum_{i=1}^{n} b^i(x)u_{x_i} + c(x)u.$$

The boundary value problem that we are most concerned with is

$$Lu = f \text{ for all } x \in \Omega \text{ and } u = 0 \text{ for all } x \in \partial\Omega. \tag{8.1}$$

Here, $f \in L^2(\Omega)$. In order to study the existence and uniqueness of solutions to this boundary value problem, we need to make some tighter assumptions. First, we will assume that the coefficients of the differential operator L satisfy

$$a^{ij} = a^{ji}, b^i, c \in L^\infty(\Omega).$$

The next definition is essentially the namesake of this section's title.

Definition 8.1.1. Let L be the partial differential operator defined as above. The operator L is called **uniformly elliptic** in Ω if there exists a constant $\theta \in \mathbb{R}_{>0}$ such that for all $x \in \Omega$ and $\xi \in \mathbb{R}^n$,

$$\sum_{i,j=1}^{n} a^{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2.$$

Here is another way of reinterpreting the above definition. Define the matrix $A(x) = (a^{ij}(x)) \in M_{n\times n}(\mathbb{R})$. Since $a^{ij} = a^{ji}$ for all $i, j \in \{1, \ldots, n\}$, $A(x)$ must be a symmetric matrix, with real eigenvalues. The above equation shows that $A(x)$ is a strictly positive definite matrix, whose smallest eigenvalue is greater than or equal to θ.

Example 8.1.1. Assume that $\Omega \subseteq \mathbb{R}^n$ is a bounded open set. Consider the following boundary value problem:

$$\begin{cases} -\Delta u + u = f & x \in \Omega, \\ u = 0 & x \in \partial\Omega. \end{cases}$$

Here, $\Delta u = \sum_{i=1}^{n} u_{x_i x_i}$ is the Laplacian. The operator $-\Delta u$ is uniformly elliptic because firstly, the matrix $A(x) = I_n$ (the $n \times n$ identity matrix) and secondly, if we select $\theta \in (0, 1]$, then for all $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$,

$$\sum_{i,j=1}^{n} a^{ij}(x)\xi_i\xi_j = \sum_{i,j=1}^{n} \delta_{ij}\xi_i\xi_j = |\xi|^2 \geq \theta|\xi|^2.$$

At the moment, it is not clear why we require this definition. This will be elucidated in the later results. Moreover, we do not expect 8.1 to have a classical solution a priori - a function $u \in C^2(\Omega)$ which satisfies 8.1 and the boundary conditions at every point in Ω. So, how can we study solutions to 8.1? Following the theme of the previous chapter, we will study weak solutions instead.

Definition 8.1.2. Consider the boundary value problem 8.1. A **weak solution** of 8.1 is a function $u \in H_0^1(\Omega) = W_0^{1,2}(\Omega)$ such that for all $v \in H_0^1(\Omega)$,

$$\int_\Omega (\sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b^i u_{x_i} v + cuv)\, dx = \int_\Omega fv\, dx.$$

Before we proceed, let us analyse the above definition in more detail. First of all, why do we want $u \in H_0^1(\Omega)$? Recall that $H_0^1(\Omega)$ is defined as the closure of $C_c^\infty(\Omega)$ in $H^1(\Omega)$ - a closed subspace of $H^1(\Omega)$ such that for all multi-indices α such that $|\alpha| \leq 0$, $D^\alpha u = 0$ on the boundary $\partial\Omega$. This is equivalent to saying that $u = 0$ on $\partial\Omega$. Hence, setting $u \in H_0^1(\Omega)$ incorporates the boundary condition of 8.1.

The second mystery here is the first summand on the LHS. To see where this originates from, consider the equation

$$\int_\Omega (Lu)v\, dx = \int_\Omega fv\, dx.$$

Now expand the RHS to get

$$\begin{aligned}
\int_\Omega (Lu)v\, dx &= \int_\Omega -\sum_{i,j=1}^n (a^{ij}(x)u_{x_i})_{x_j} v + \sum_{i=1}^n b^i(x) u_{x_i} v + c(x)uv\, dx \\
&= \int_\Omega -\sum_{i,j=1}^n (a^{ij} u_{x_i})_{x_j} v\, dx + \int_\Omega \sum_{i=1}^n b^i u_{x_i} v + cuv\, dx \\
&= -\sum_{i,j=1}^n ([a^{ij} u_{x_i} v]_{\partial\Omega} - \int_\Omega a^{ij} u_{x_i} v_{x_j}\, dx) + \int_\Omega \sum_{i=1}^n b^i u_{x_i} v + cuv\, dx \\
&= \int_\Omega (\sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b^i u_{x_i} v + cuv)\, dx \quad (\text{since } v \in H_0^1(\Omega)).
\end{aligned}$$

The point here is that since $v \in H_0^1(\Omega)$, the boundary term $[a^{ij} u_{x_i} v]_{\partial\Omega}$, which appears as the result of an integration by parts, must vanish.

It is reasonable to ask whether we can deal with non-homogeneous boundary value problems, such as

$$Lu = f \text{ for all } x \in \Omega \text{ and } u = g \text{ for all } x \in \partial\Omega.$$

where $g \in H^1(\Omega)$. Fortunately, we can make the substitution $\tilde{u} = u - g$ in order to reduce the above problem to the case 8.1.

Another alternative method of expressing the concept of a weak solution is to define the following bilinear form on the Hilbert space $H_0^1(\Omega)$:

$$B[u, v] = \int_\Omega (\sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b^i u_{x_i} v + cuv)\, dx$$

A function $u \in H_0^1(\Omega)$ is a weak solution of 8.1 if for all $v \in H_0^1(\Omega)$, $B[u, v] = \langle f, v \rangle_{L^2}$. This is a useful reformulation of the concept of a weak solution, due to the Lax-Milgram theorem (3.6.3).

The first step towards the existence and uniqueness of uniformly elliptic PDEs is to show that the equation 8.1.1 has a unique solution.

Lemma 8.1.1. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. Then, for all $f \in L^2(\Omega)$, the BVP 8.1.1 has a unique weak solution $u \in H_0^1(\Omega)$. Moreover, the map $f \mapsto u$ from $L^2(\Omega)$ to $H_0^1(\Omega)$ is a compact linear operator.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is a bounded open set. We first apply the Rellich-Kondrachov theorem (7.6.1) to deduce that $H_0^1(\Omega) \subset\subset L^2(\Omega)$. This means that the canonical embedding $\iota : H_0^1(\Omega) \to L^2(\Omega)$ is a compact linear operator.

As a result, the adjoint operator $\iota^* : L^2(\Omega) \to H_0^1(\Omega)$ must also be a compact linear operator. Here, we have implicitly used the Riesz representation theorem (3.3.1) to deduce that $[L^2(\Omega)]^* = L^2(\Omega)$ and $[H_0^1(\Omega)]^* = H_0^1(\Omega)$.

Assume that $f \in L^2(\Omega)$. Then, from the definition of the adjoint operator, we have for all $v \in H_0^1(\Omega)$,

$$\langle \iota^* f, v \rangle_{H^1} = \langle f, \iota v \rangle_{L^2} = \langle f, v \rangle_{L^2}.$$

From the definition of a weak solution, we deduce that $\iota^* f \in H_0^1(\Omega)$ is the unique weak solution to 8.1.1. $\square$

Now, we will press on towards a greater generalisation. Consider the following elliptic boundary value problem:

$$\begin{cases} L_0 u = f & x \in \Omega, \\ u = 0 & x \in \partial\Omega. \end{cases} \tag{8.2}$$

Here, the differential operator L_0 contains only second order terms:

$$L_0 u = -\sum_{i,j=1}^{n} (a^{ij}(x) u_{x_i})_{x_j}.$$

In the case of 8.2, a weak solution is a function $u \in H_0^1(\Omega)$ such that for all $v \in H_0^1(\Omega)$,

$$B_0[u,v] = \int_\Omega \sum_{i,j=1}^{n} a^{ij} u_{x_i} v_{x_j} \, dx = \langle f, v \rangle_{L^2} = \langle \iota^* f, v \rangle_{H^1}.$$

Theorem 8.1.2. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set. Assume that L_0 is the operator defined in 8.2. Assume further that L_0 is uniformly elliptic. Then, for all $f \in L^2(\Omega)$, the BVP 8.2 has a unique weak solution $u \in H_0^1(\Omega)$. Furthermore, its corresponding operator $L_0^{-1} : L^2(\Omega) \to H_0^1(\Omega)$, which sends f to u is a compact linear operator.*

Proof. Assume that $\Omega \subseteq \mathbb{R}^n$ is a bounded open set. Assume that L_0 is a uniformly elliptic differential operator, defined as in 8.2. In order to show that 8.2 has a unique solution, we will demonstrate that the bilinear form $B_0 : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ satisfies the conditions of the Lax-Milgram theorem (3.6.3).

To show: (a) B_0 is a continuous functional.

(b) B_0 is a positive definite functional.

(a) To see that B_0 is continuous, we argue as follows:

$$\begin{aligned}
|B_0[u,v]| &= \left| \int_\Omega \sum_{i,j=1}^{n} a^{ij} u_{x_i} v_{x_j} \, dx \right| \\
&\leq \sum_{i,j=1}^{n} \int_\Omega |a^{ij} u_{x_i} v_{x_j}| \, dx \\
&\leq \sum_{i,j=1}^{n} \|a^{ij}\|_{L^\infty} \|u_{x_i}\|_{L^2} \|v_{x_j}\|_{L^2} \quad \text{(Hölder)} \\
&\leq C \|u\|_{H^1} \|v\|_{H^1}
\end{aligned}$$

for some constant $C \in \mathbb{R}_{>0}$. This inequality is enough to show that B_0 is a continuous functional.

(b) Here, we will use the fact that Ω is bounded, alongside the Poincaré inequality, in order to deduce the existence of a constant $\kappa \in \mathbb{R}_{>0}$ such that for all $u \in H_0^1(\Omega)$,

$$\|u\|_{L^2(\Omega)}^2 \leq \kappa \int_\Omega |\nabla u|^2 \, dx.$$

Here is where the uniform elliptic assumption on L_0 comes into play. For all $u \in H_0^1(\Omega)$,

$$B_0[u,u] = \int_\Omega \sum_{i,j=1}^n a^{ij} u_{x_i} u_{x_j} \, dx \geq \int_\Omega \theta \sum_{i=1}^n u_{x_i}^2 \, dx = \theta \int_\Omega |\nabla u|^2 \, dx.$$

Now, we can combine the two inequalities to obtain

$$\begin{aligned}
\|u\|_{H^1(\Omega)}^2 &= \sum_{|\alpha|\leq 1} \int_\Omega |D^\alpha u|^2 \, dx \\
&= \int_\Omega |u|^2 \, dx + \sum_{i=1}^n \int_\Omega |u_{x_i}|^2 \, dx \\
&= \|u\|_{L^2(\Omega)}^2 + \int_\Omega \sum_{i=1}^n |u_{x_i}|^2 \, dx \\
&= \|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2 \\
&\leq (\kappa+1)\|\nabla u\|_{L^2(\Omega)}^2 \quad \text{(First inequality)} \\
&\leq \frac{\kappa+1}{\theta} B_0[u,u] \quad \text{(Second inequality)}
\end{aligned}$$

So, $B_0[u,u] \geq \frac{\theta}{\kappa+1}\|u\|_{H^1(\Omega)}^2$ for all $u \in H_0^1(\Omega)$. This is enough to show that B_0 is positive definite.

Combining parts (a) and (b), we can apply the Lax-Milgram theorem (3.6.3) in order to deduce that for all $\tilde{f} \in H_0^1(\Omega)$, there exists a unique element $u \in H_0^1(\Omega)$ such that for all $v \in H_0^1(\Omega)$, $B_0[u,v] = \langle \tilde{f}, v\rangle_{H^1(\Omega)}$. Moreover, we also have the bound $\|u\|_{H^1(\Omega)} \leq \beta^{-1}\|\tilde{f}\|_{H^1(\Omega)}$ for some $\beta \in \mathbb{R}_{>0}$. This is enough to demonstrate that the following operator is continuous:

$$\Lambda : H_0^1(\Omega) \to H_0^1(\Omega)$$

$$\tilde{f} \mapsto u$$

Now set $\tilde{f} = \iota^* f \in H_0^1(\Omega)$, where $f \in L^2(\Omega)$ and $\iota : H_0^1(\Omega) \to L^2(\Omega)$ is the compact canonical embedding. Then, for all $v \in H_0^1(\Omega)$,

$$B_0[u, v] = \langle \iota^* f, v \rangle_{H^1(\Omega)} = \langle f, v \rangle_{L^2(\Omega)}.$$

Thus, u qualifies as a weak solution to 8.2.

Finally, to see that $L_0^{-1} : f \mapsto u$ is a compact operator, note that it can be written as the following composite:

$$L^2(\Omega) \xrightarrow{\iota^*} H_0^1(\Omega) \xrightarrow{\Lambda} H_0^1(\Omega).$$

The first operator ι^* is compact, whereas Λ is continuous. Hence, $L_0^{-1} = \Lambda \circ \iota^*$ must also be a compact operator.

□

The next result provides us with more properties about the solution operator L_0^{-1}.

Lemma 8.1.3. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded, open set. Then, the linear operator $L_0^{-1} : L^2(\Omega) \to L^2(\Omega)$ is compact, injective and self-adjoint. Thus, $L^2(\Omega)$ has an orthonormal basis $\{\phi_k\}_{k \in \mathbb{Z}_{>0}}$, consisting of eigenfunctions of L_0^{-1}. Suppose that λ_k is the corresponding eigenvalue of ϕ_k. Then, $\lambda_k > 0$ for all $k \in \mathbb{Z}_{>0}$ and $\lim_{k \to \infty} \lambda_k = 0$.*

Proof. Assume that $L_0^{-1} : L^2(\Omega) \to H_0^1(\Omega)$ is the solution operator defined in 8.1.2. To see that $L_0^{-1} : L^2(\Omega) \to L^2(\Omega)$ is compact, note that it is the composite $L_0^{-1} : L^2(\Omega) \to H_0^1(\Omega)$ followed by the inclusion $\iota : H_0^1(\Omega) \to L^2(\Omega)$, which is also compact.

To show: (a) L_0^{-1} is injective.

(b) L_0^{-1} is self-adjoint.

(a) Assume that $f \in \ker L_0^{-1} \subseteq L^2(\Omega)$. Then, $L_0^{-1} f = 0$. Since $L_0^{-1} f$ is a weak solution to 8.2, we must have for all $v \in H_0^1(\Omega)$

$$B_0[L_0^{-1} f, v] = \langle f, v \rangle_{L^2(\Omega)} = \int_\Omega f v \, dx = 0.$$

Notably, the above equation holds for all $v \in C_c^\infty(\Omega)$. Therefore, $f(x) = 0$ for almost all $x \in \Omega$. This demonstrates that $\ker L_0^{-1} = \{0\}$ and that L_0^{-1} is injective.

(b) To see that L_0^{-1} is self-adjoint, assume that $f, g \in L^2(\Omega)$ and that $u = L_0^{-1} f$ and $v = L_0^{-1} g$. Then, $u, v \in H_0^1(\Omega)$ and

$$\begin{aligned}
\langle L_0^{-1} f, g \rangle_{L^2(\Omega)} &= \int_\Omega (L_0^{-1} f) g \, dx \\
&= \int_\Omega ug \, dx \\
&= B_0[u, v] \\
&= \int_\Omega \sum_{i,j=1}^{n} a^{ij} u_{x_i} v_{x_j} \, dx \\
&= \int_\Omega fv \, dx \\
&= \langle f, L_0^{-1} g \rangle.
\end{aligned}$$

Hence, L_0^{-1} must be self-adjoint.

Combining parts (a) and (b), we find that L_0^{-1} is compact and self-adjoint. By the Hilbert-Schmidt theorem (5.3.3), $L^2(\Omega)$ has a countable orthonormal basis of eigenvectors of L_0^{-1}, which we will denote by $\{\phi_k\}_{k \in \mathbb{Z}_{>0}}$. Let $\lambda_k \in \mathbb{C}$ denote the eigenvalue corresponding to ϕ_k. Since L_0^{-1} is compact, we must have $\lim_{k \to \infty} \lambda_k = 0$.

Finally, observe that for all $k \in \mathbb{Z}_{>0}$

$$1 = \langle \phi_k, \phi_k \rangle_{L^2(\Omega)} = B_0[L_0^{-1} \phi_k, \phi_k] = B_0[\lambda_k \phi_k, \phi_k].$$

Since B_0 is bilinear, we therefore have

$$\lambda_k = \frac{1}{B_0[\phi_k, \phi_k]} > 0$$

due to the fact that B_0 is positive definite. □

Now that we have studied the special case 8.2, we want to move onto the more general BVP in 8.1. As the following example demonstrates, we will need to tread carefully because the loosening of assumptions required to go from 8.2 to 8.1 means that 8.1.2 is no longer valid.

Example 8.1.2. Assume that $\Omega = (0, \pi) \subseteq \mathbb{R}$. Consider the operator $Lu = -u_{xx} - 4u$. This operator is uniformly elliptic in a similar vein to 8.1.1. However, the corresponding bilinear form $B : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$, defined by

$$B[u, v] = \int_0^\pi u_x v_x - 4uv \, dx$$

is not positive definite. For example, if we take $u = v = \sin x$, we find that

$$B[u, u] = \int_0^\pi \cos^2 x - 4\sin^2 x \, dx = -\frac{3\pi}{2}.$$

Now observe that if we take $f(x) = \sin(2x) \in L^2(\Omega)$, we find that the boundary value problem

$$\begin{cases} -u_{xx} - 4u = \sin 2x & x \in (0, \pi), \\ u(x) = 0 & x \in \{0, \pi\}. \end{cases}$$

has no weak solutions. If we take $v(x) = \sin 2x$, then for all $u \in H_0^1(\Omega)$,

$$B[u, v] = \int_0^\pi u_x v_x - 4uv \, dx = \int_0^\pi (2u_x \cos 2x - 4u \sin 2x) \, dx = 0 \neq \langle f, v \rangle_{L^2((0,\pi))}.$$

Now, we will proceed to analyse 8.1. We will restate the BVP below for convenience.

$$\begin{cases} Lu = L_0 u + L_1 u = f & x \in \Omega, \\ u = 0 & x \in \partial\Omega. \end{cases}$$

Here, L_0 is the second order differential operator

$$L_0 u = -\sum_{i,j=1}^n (a^{ij}(x) u_{x_i})_{x_j}.$$

and L_1 is the differential operator which contains the rest of the terms:

$$L_1 u = \sum_{i=1}^n b^i(x) u_{x_i} + c(x) u.$$

A function $u \in H_0^1(\Omega)$ is a solution to the BVP if and only if for all $v \in H_0^1(\Omega)$,

$$B_0[u, v] = \langle f, v \rangle_{L^2(\Omega)} - \langle L_1 u, v \rangle_{L^2(\Omega)}.$$

The above equality is from the definition of a weak solution. Hence, the above equality is true if and only if $u = L_0^{-1}(f - L_1 u)$. This can be rewritten in the form $(I + K)u = g$, where $g = L_0^{-1} f$ and $Ku = L_0^{-1} L_1 u$. Now, observe that L_1 is a bounded linear operator and L_0^{-1} is compact from 8.1.2. So, K must be a compact operator and consequently, the Fredholm alternative applies. We now have the following two scenarios:

1. In the first case, $\ker(I + K) = \{0\}$ and the equation $(I + K)u = g$ has exactly one unique solution in $H_0^1(\Omega)$ for all $g \in H_0^1(\Omega)$.
2. In the second case, $\ker(I + K) \neq \{0\}$ and the equation $(I + K)u = 0$ has a non-trivial solution in $H_0^1(\Omega)$.

The second situation is equivalent to saying that $u \in H_0^1(\Omega)$ is a weak solution to the homogeneous boundary value problem

$$\begin{cases} Lu = L_0 u + L_1 u = 0 \quad x \in \Omega, \\ u = 0 \quad x \in \partial\Omega. \end{cases} \tag{8.3}$$

Following through with the first case, we can now confidently state when 8.1 has a unique solution. The proof is a consequence of the above application of the Fredholm alternative.

Theorem 8.1.4. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set and let L denote the differential operator in the BVP 8.1. Suppose that $a^{ij}, b^i, c \in L^\infty(\Omega)$ and L is uniformly elliptic. Then, the BVP 8.1 has a unique solution for all $f \in L^2(\Omega)$ if and only if the homogeneous BVP 8.3 has $u(x) = 0$ as the only solution.*

8.2 Parabolic PDEs

Once again, let $\Omega \subseteq \mathbb{R}^n$ denote a bounded open set and let L be the operator in 8.1. We will also assume that L is uniformly elliptic, $a^{ij} = a^{ji} \in W^{1,\infty}(\Omega)$ and $b^i, c \in L^\infty(\Omega)$.

Below is the **parabolic** BVP that we will analyse in this section:

$$\begin{cases} \frac{\partial u}{\partial t} + Lu = 0 \quad t > 0, x \in \Omega, \\ u(t, x) = 0 \quad t > 0, x \in \partial\Omega, \\ u(0, x) = g(x) \quad x \in \Omega. \end{cases} \tag{8.4}$$

By defining $A = -L$, we can rewrite the above BVP as a Cauchy problem in $L^2(\Omega)$:

$$\frac{d}{dt}u = Au,\ u(0) = g$$

Here, $Dom(A) = \{u \in H_0^1(\Omega) \mid Lu \in L^2(\Omega)\}$. Alternatively, $u \in Dom(A)$ if u is a solution to 8.1 for some $f \in L^2(\Omega)$.

The main goal of this section is to construct solutions to 8.4 using semigroup theory. As in the previous section, we will start with some extra assumptions first.

Theorem 8.2.1. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set and the operator L be defined as in 8.4. Suppose that the bilinear form $B : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ corresponding to L*

$$B[u,v] = \int_\Omega (\sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b^i u_{x_i} v + cuv)\, dx$$

is positive definite. Then, the operator $A = -L$ generates a contractive semigroup $\{S_t\}_{t \in \mathbb{Z}_{\geq 0}}$ of linear operators on $L^2(\Omega)$.

Proof. From 6.7.1, we have to demonstrate that the operator $A : L^2(\Omega) \to L^2(\Omega)$ satisfies the following properties:

To show: (a) The set $Dom(A)$ is dense in $L^2(\Omega)$.

(b) The graph of A is closed.

(c) For all $\lambda > 0$, $\lambda \in \rho(A)$ and

$$\|(\lambda I - A)^{-1}\| \leq \frac{1}{\lambda}.$$

(a) Assume that $\varphi \in C_c^\infty(\Omega)$. Then, by definition $L\varphi \in L^2(\Omega)$ since Ω is bounded. So, $\varphi \in Dom(A)$, suggesting that $C_c^\infty(\Omega) \subseteq Dom(A)$. Since $C_c^\infty(\Omega)$ is dense in $L^2(\Omega)$, $Dom(A)$ must also be dense in $L^2(\Omega)$.

(b) Assume that B is positive definite. Since it is also continuous, as $|B[u,v]| \leq C\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)}$ for some constant $C \in \mathbb{R}_{>0}$, we can apply the Lax-Milgram theorem (3.6.3) to deduce that for all $f \in L^2(\Omega)$, there exists a unique solution $u \in H_0^1(\Omega)$ such that for all $v \in H_0^1(\Omega)$,

$$B[u,v] = \langle f, v\rangle_{L^2(\Omega)}.$$

Moreover, the map L^{-1} which sends f to u is a bounded linear operator from $L^2(\Omega)$ to $L^2(\Omega)$.

Suppose that $\Gamma(A)$ denotes the graph of A. We now observe that since $A = -L$, the pair $(u, f) \in L^2(\Omega) \times L^2(\Omega)$ is in $\Gamma(A)$ if and only if $(-f, u) \in \Gamma(L^{-1})$. Note that since L^{-1} is continuous (as it is bounded), $\Gamma(L^{-1})$ must be closed. Therefore, $\Gamma(A)$ must also be closed.

(c) Assume that $\lambda \in \mathbb{R}_{>0}$. It suffices to show that the operator $\lambda I - A$ has a bounded inverse with operator norm $\|(\lambda I - A)^{-1}\| \leq 1/\lambda$. This is equivalent to the statement that for all $f \in L^2(\Omega)$, the BVP

$$\begin{cases} \lambda u + Lu = f \quad x \in \Omega, \\ u = 0 \quad x \in \partial\Omega, \end{cases}$$

has a weak solution $u \in H_0^1(\Omega)$ which satisfies $\|u\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)}/\lambda$ (note that $u = (\lambda I - A)^{-1} f$).

Since B is positive definite and continuous, the Lax-Milgram theorem tells us that there exists a unique $u \in H_0^1(\Omega)$ such that for all $v \in H_0^1(\Omega)$,

$$\langle \lambda u, v\rangle_{L^2(\Omega)} + B[u,v] = \langle f, v\rangle_{L^2(\Omega)}$$

If we set $u = v$, we deduce that for all $v \in H_0^1(\Omega)$,

$$\lambda \|v\|_{L^2(\Omega)} + B[v,v] = \langle f, v\rangle_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}$$

by the Cauchy-Schwarz inequality. Since B is positive definite, $B[v,v] \geq 0$ for all $v \in H_0^1(\Omega)$. Therefore, $\lambda \|u\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)}$ and consequently,

$$\|u\|_{L^2(\Omega)} = \|(\lambda I - A)^{-1} f\|_{L^2(\Omega)} \leq \frac{1}{\lambda} \|f\|_{L^2(\Omega)}.$$

So, by the definition of the operator norm, we must have

$$\|(\lambda I - A)^{-1}\| \leq \sup_{\|f\|_{L^2(\Omega)}=1} \|(\lambda I - A)^{-1} f\|_{L^2(\Omega)} \leq \frac{1}{\lambda}.$$

This also reveals that $\lambda \in \rho(A)$. By combining all parts of the proof and using 6.7.1, we find that $A = -L$ must generate a contractive semigroup. $\square$

As with the elliptic BVPs, the next logical question we will explore is whether the solutions to 8.4 form an orthonormal basis of $L^2(\Omega)$ (whether the Hilbert-Schmidt theorem applies). We will return to the scenario in 8.1.3, where we have an orthonormal basis $\{\phi_k\}_{k\in\mathbb{Z}_{>0}}$ consisting of eigenfunctions of the compact, self-adjoint operator L_0^{-1}. Since $L_0^{-1}\phi_k = \lambda_k\phi_k$ for all $k \in \mathbb{Z}_{>0}$, we must have $\phi_k \in Dom(L_0)$ because

$$L_0\phi_k = \frac{1}{\lambda_k}\phi_k.$$

Define $\mu_k = \frac{1}{\lambda_k} > 0$. Note that $\mu_k \to \infty$ as $k \to \infty$. Now define

$$u(t) = e^{-\mu_k t}\phi_k.$$

The point of this definition is that $u(t)$ satisfies the following Cauchy problem:

$$\frac{du}{dt} = -L_0 u(t) \quad u(0) = \phi_k.$$

Notice that if we apply the previous theorem to L_0, then in tandem with 8.1.2, the operator $-L_0$ must generate a contractive semigroup $\{S_t\}_{t\in\mathbb{Z}_{>0}}$. Recall from 6.7.1 that the semigroup generated by $-L_0$ must be unique. Therefore, $S_t\phi_k = e^{-\mu_k t}\phi_k$. Extending by linearity, we also have

$$S_t(\sum_{k=1}^{N} c_k\phi_k) = \sum_{k=1}^{N} c_k e^{-\mu_k t}\phi_k.$$

Using the fact that $\{\phi_k\}$ is an orthonormal basis for $L^2(\Omega)$, we must have for all $g \in L^2(\Omega)$ and for all $t \in \mathbb{R}_{\geq 0}$ that

$$S_t g = \sum_{k=1}^{\infty} \langle g, \phi_k\rangle_{L^2(\Omega)} e^{-\mu_k t}\phi_k. \tag{8.5}$$

The above equation is the subject of our next result.

Lemma 8.2.2. *Assume that L_0 is the operator in 8.2. The for all $g \in L^2(\Omega)$, the formula 8.5 defines a map*

$$\varphi_g : (0, \infty) \to L^2(\Omega)$$

$$t \mapsto S_t g.$$

The map φ_g is continuous for all $t \in (0, \infty)$ and continuously differentiable for all $t \in \mathbb{R}_{>0}$. Moreover, $u(t) = e^{-\mu_k t}\phi_k \in Dom(L_0) \subseteq H_0^1(\Omega)$ for all $t \in \mathbb{R}_{>0}$ and

$$\frac{d}{dt}u(t) = L_0 u(t).$$

Proof. Assume that $g \in L^2(\Omega)$. To see that

$$S_t g = \sum_{k=1}^{\infty} \langle g, \phi_k \rangle_{L^2(\Omega)} e^{-\mu_k t}\phi_k.$$

is uniformly convergent for $t \in \mathbb{R}_{\geq 0}$, observe that for all $t \in \mathbb{R}_{\geq 0}$,

$$\begin{aligned}\sum_{k=1}^{\infty} |e^{-\mu_k t}\langle g, \phi_k \rangle_{L^2(\Omega)}|^2 &\leq \sum_{k=1}^{\infty} \langle g, \phi_k \rangle_{L^2(\Omega)}^2 \quad (\mu_k > 0)\\ &= \|g\|_{L^2(\Omega)}^2 \quad \text{(Bessel's Inequality)}\\ &< \infty.\end{aligned}$$

Since the coefficients converge, we deduce that the series representation of $S_t g$ must converge uniformly for $t \in \mathbb{R}_{\geq 0}$. Furthermore, all of the partial sums are continuous functions with respect to t. This confirms that the map $t \mapsto S_t g$ is also continuous.

Our next claim is that for all $t > 0$, $S_t g \in Dom(L_0) \subseteq H_0^1(\Omega)$, even if $g \notin H_0^1(\Omega)$. Similarly to the previous argument, a function $u = \sum_k c_k \phi_k$ is in $Dom(L_0)$ if and only if $\sum_k c_k^2 \mu_k^2 < \infty$. If we set $c_k(t) = e^{-\mu_k t}\langle g, \phi_k \rangle_{L^2(\Omega)}$, we have the bound

$$\sum_{k=1}^{\infty} (c_k \mu_k)^2 \leq \sup_{k \in \mathbb{Z}_{>0}} (\mu_k e^{-\mu_k t})^2 \sum_{k=1}^{\infty} \langle g, \phi_k \rangle_{L^2(\Omega)}^2 = \sup_{k \in \mathbb{Z}_{>0}} (\mu_k e^{-\mu_k t})^2 \sum_{k=1}^{\infty} \|g\|_{L^2(\Omega)}^2.$$

Define the function $f(\zeta) = \zeta e^{-\zeta t}$, where $\zeta \geq 0$. It attains its global maximum when $\zeta = 1/t$ and $f(1/t) = 1/et$. Therefore,

$$\sum_{k=1}^{\infty} (c_k \mu_k)^2 \leq \frac{1}{e^2 t^2} \|g\|_{L^2(\Omega)}^2 < \infty.$$

This demonstrates that the series defining $L_0 u(t)$ is convergent and thus, $u(t) \in Dom(L_0)$ for each $t > 0$.

Finally, we can differentiate the series in 8.5 term by term. One should find that the series of derivatives is also convergent. This completes the proof. □

Now, we will extend 8.2.1 in order to remove the assumption that B is positive definite. In order to illustrate the motivation behind what proceeds, we will return to the finite dimensional case. Let $L \in M_{n\times n}(\mathbb{R})$ and consider the following linear ODE

$$\frac{d}{dt}x(t) = -Lx(t).$$

If L is positive definite (that is, for all $x \in \mathbb{R}^n$, $\langle Lx, x\rangle \geq 0$), then $-L$ must generate a contractive semigroup. Observe also that the Euclidean norm of a solution does not increase with time t because

$$\frac{d}{dt}|x(t)|^2 = 2\langle \frac{d}{dt}x(t), x(t)\rangle = 2\langle -Lx, x\rangle \leq 0.$$

Now, let $L \in M_{n\times n}(\mathbb{R})$ be an arbitrary matrix. Then, there exists $\gamma \in \mathbb{R}_{\geq 0}$ so that $L + \gamma I$ is positive definite and thus, generates a contractive semigroup. In this case, if $x(t) = e^{-tL}x(0)$ is a solution to the linear ODE above, we can write $-L = \gamma I - (L + \gamma I)$ and obtain

$$|x(t)| = |e^{-Lt}x(0)| = |e^{[\gamma I-(L+\gamma I)]t}x(0)| = e^{\gamma t}|e^{-(L+\gamma I)t}x(0)| \leq e^{\gamma t}|x(0)|.$$

Hence, the operator $-L$ must generate a semigroup of type γ.

Now, we will transfer the above reasoning to the case where L is a uniformly elliptic operator in 8.1 and the corresponding bilinear form $B[u, v]$ is not necessarily positive definite. The following lemma is required to carry out the first step.

Lemma 8.2.3. *Let the operator L be uniformly elliptic with $a^{ij}, b{\cdot}c \in L^\infty(\Omega)$. Then, there exists constants $\alpha, \beta, \gamma > 0$ such that for all $u, v \in H_0^1(\Omega)$,*

$$|B[u, v]| \leq \alpha\|u\|_{H^1}\|v\|_{H^1} \text{ and } \beta\|u\|_{H^1}^2 \leq B[u, u] + \gamma\|u\|_{L^2}^2.$$

Proof. Assume that $u, v \in H_0^1(\Omega)$. The first inequality can be shown directly from the definition of B and Hölder's inequality as follows:

$$
\begin{aligned}
|B[u,v]| &= |\int_\Omega (\sum_{i,j=1}^n a^{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b^i u_{x_i} v + cuv)| \\
&\leq \sum_{i,j=1}^n \|a^{ij}\|_{L^\infty} \|u_{x_i}\|_{L^2} \|v_{x_j}\|_{L^2} + \sum_{i=1}^n \|b^i\|_{L^\infty} \|u_{x_i}\|_{L^2} \|v\|_{L^2} \\
&\quad + \|c\|_{L^\infty} \|u\|_{L^2} \|v\|_{L^2} \quad \text{(Hölder)} \\
&\leq \alpha \|u\|_{H^1} \|v\|_{H^1}.
\end{aligned}
$$

For the second inequality, we will use the fact that L is uniformly elliptic and the inequality

$$
ab \leq \frac{\theta}{2} a^2 + \frac{1}{2\theta} b^2.
$$

We then obtain

$$
\begin{aligned}
\theta \sum_{i=1}^n \|u_{x_i}\|_{L^2}^2 &= \theta \int_\Omega \sum_{i=1}^n u_{x_i}^2 \, dx \\
&\leq \int_\Omega \sum_{i,j=1}^n a^{ij} u_{x_i} u_{x_j} \, dx \quad \text{(Uniformly Elliptic)} \\
&= B[u,u] - \int_\Omega (\sum_{i=1}^n b^i u_{x_i} u + cu^2) \, dx \\
&\leq B[u,u] + \sum_{i=1}^n \|b^i\|_{L^\infty} \|u_{x_i}\|_{L^2} \|u\|_{L^2} + \|c\|_{L^\infty} \|u\|_{L^2}^2 \\
&\leq B[u,u] + (\sum_{i=1}^n \frac{\theta}{2} \|u_{x_i}\|_{L^2}^2 + \sum_{i=1}^n \frac{1}{2\theta} \|b^i\|_{L^\infty}^2 \|u\|_{L^2}^2) + \|c\|_{L^\infty} \|u\|_{L^2}^2.
\end{aligned}
$$

So, for all $u \in H_0^1(\Omega)$,

$$
B[u,u] \geq \frac{\theta}{2} \sum_{i=1}^n \|u_{x_i}\|_{L^2}^2 - C\|u\|_{L^2}^2
$$

for some constant $C \in \mathbb{R}_{>0}$. So, we can take $\beta = \theta/2$ and $\gamma = C + \theta/2$ in order to deduce the second inequality.

□

Taking the constant γ from the above lemma, we define

$$L_\gamma u = Lu + \gamma u \text{ and } B_\gamma[u,v] = B[u,v] + \gamma\langle u,v\rangle_{L^2(\Omega)}.$$

So, our parabolic PDE 8.4 can be written as

$$\frac{\partial u}{\partial t} = -L_\gamma u + \gamma u.$$

The lemma reveals that B_γ is positive definite and therefore, the operator $A_\gamma = -L_\gamma$ generates a contractive semigroup, which we will call $\{S_t^\gamma \mid t \in \mathbb{R}_{\geq 0}\}$. Therefore $A = -L = \gamma I - L_\gamma$ must generate a semigroup of type γ by using the same analysis as the finite dimensional case.

The above arguments culminate in the following theorem:

Theorem 8.2.4. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set and the operator L be defined as in 8.4. Then, the operator $A = -L$ generates a semigroup $\{S_t\}_{t\in\mathbb{Z}_{\geq 0}}$ of linear operators on $L^2(\Omega)$.*

After constructing a semigroup $\{S_t\}_{t\in\mathbb{R}_{\geq 0}}$, one lingering thought is when the function $t \mapsto S_t f$ provides a solution to the BVP 8.4. In the case where $L = L_0$ (in 8.2), we can use 8.2.2 to deduce that for all $g \in L^2(\Omega)$, the map $t \mapsto S_t g$ is a C^1 map (continuously differentiable). Furthermore, $u(t) = e^{-\mu_k t}\phi_k \in Dom(L_0)$ and satisfies the required differential equation

$$\frac{d}{dt}u(t) = L_0 u(t)$$

for all $t \in \mathbb{R}_{\geq 0}$. Recall that this works because of 8.5.

A similar result can be established for the more general elliptic operator L, but this requires more work.

8.3 Hyperbolic PDEs

The final type of BVP that we will study is a linear hyperbolic BVP of the following form:

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} + L_0 u = 0 \quad t \in \mathbb{R}, x \in \Omega, \\ u(t,x) = 0 \quad t \in \mathbb{R}, x \in \partial\Omega, \\ u(0,x) = f(x), \frac{\partial u}{\partial t}(0,x) = g(x) \quad x \in \Omega. \end{cases} \tag{8.6}$$

Here, we have the usual assumptions. As a recap, $\Omega \subseteq \mathbb{R}^n$ is a bounded open set, $f, g \in L^2(\Omega)$, L_0 is the second order uniformly elliptic operator

$$L_0 u = -\sum_{i,j=1}^{n} (a^{ij}(x) u_{x_i})_{x_j}$$

with $a^{ij} = a^{ji} \in W^{1,\infty}(\Omega)$ for all $i, j \in \{1, \ldots, n\}$. By 8.1.3 and 8.2.2, we find that $L^2(\Omega)$ admits an orthonormal basis $\{\phi_k\}_{k \in \mathbb{Z}_{>0}}$ such that for all $k \in \mathbb{Z}_{>0}$, $\phi_k \in Dom(L_0)$ and $L_0 \phi_k = \mu_k \phi_k$ where $\lim_{k\to\infty} \mu_k \to \infty$.

Our strategy for dealing with 8.6 is to rewrite it as a first order system. Define $v = \partial u / \partial t$. Then, our hyperbolic BVP can be rewritten in the product space $X = H_0^1(\Omega) \times L^2(\Omega)$ as

$$\frac{d}{dt}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & I \\ -L_0 & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}, \quad \text{where } \begin{pmatrix} u \\ v \end{pmatrix}(0) = \begin{pmatrix} f \\ g \end{pmatrix}.$$

Using the eigenfunctions ϕ_k, we will construct a semigroup of solutions. First consider the special case where $f = a_k \phi_k$ and $g = b_k \phi_k$ for some $k \in \mathbb{Z}_{>0}$ with $a_k, b_k \in \mathbb{R}$. An explicit solution is

$$u(t) = a(t)\phi_k \text{ and } v(t) = a'(t)\phi_k$$

where the coefficient $a(t)$ must satisfy

$$a''(t) + \mu_k a(t) = 0, a(0) = a_k \text{ and } a'(0) = b_k.$$

Fortunately, the above ODE admits a nice solution:

$$a(t) = a_k \cos(\sqrt{\mu_k} t) + \frac{b_k}{\sqrt{\mu_k}} \sin(\sqrt{\mu_k} t).$$

Substitution then yields the identities of $u(t)$ and $v(t)$, in terms of the following matrix equation:

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} \cos(\sqrt{\mu_k} t) & \frac{1}{\sqrt{\mu_k}} \sin(\sqrt{\mu_k} t) \\ -\sqrt{\mu_k} \sin(\sqrt{\mu_k} t) & \cos(\sqrt{\mu_k} t) \end{pmatrix}\begin{pmatrix} a_k \phi_k \\ b_k \phi_k \end{pmatrix}.$$

The map $t \mapsto (u(t), v(t))$ is a continuously differentiable map from $\mathbb{R}$ to X, which satisfies 8.6. Of course, we can extend this further to arbitrary $f, g \in L^2(\Omega)$ via linearity and the fact that $\{\phi_k\}$ is an orthonormal basis of $L^2(\Omega)$.

With the above analysis, we are finally able to describe the solutions of 8.6.

Theorem 8.3.1. *Let $f, g \in L^2(\Omega)$ as in 8.6 and $X = H_0^1(\Omega) \times L^2(\Omega)$. For all $t \in \mathbb{R}$, define the map $S_t : X \to X$ by*

$$S_t \begin{pmatrix} f \\ g \end{pmatrix} = \sum_{k=1}^{\infty} \begin{pmatrix} \cos(\sqrt{\mu_k} t) & \frac{1}{\sqrt{\mu_k}} \sin(\sqrt{\mu_k} t) \\ -\sqrt{\mu_k} \sin(\sqrt{\mu_k} t) & \cos(\sqrt{\mu_k} t) \end{pmatrix} \begin{pmatrix} \langle f, \phi_k \rangle_{L^2(\Omega)} \phi_k \\ \langle g, \phi_k \rangle_{L^2(\Omega)} \phi_k \end{pmatrix}.$$

Then, $\{S_t\}_{t \in \mathbb{R}}$ is a strongly continuous semigroup of bounded linear operators on X. Moreover, each S_t is an isometry with respect to the norm

$$\|(u, v)\|_X = (B_0[u, u] + \|v\|_{L^2(\Omega)}^2)^{\frac{1}{2}}.$$

Proof. Suppose that $u \in H_0^1(\Omega)$, $v \in L^2(\Omega)$ and $\|(u, v)\|_X$ is defined as above. Recall from 8.1.2 that B_0 is a strictly positive definite bilinear form, which means that there exists a constant $\beta \in \mathbb{R}_{>0}$ such that for all $u \in H_0^1(\Omega)$, $B_0[u, u] \geq \beta \|u\|_{H^1(\Omega)}^2$.

Now consider the standard product norm on X, defined by

$$\|(u, v)\|_{H^1 \times L^2} = (\|u\|_{H^1(\Omega)}^2 + \|v\|_{L^2}^2)^{\frac{1}{2}}.$$

Due to the fact that B_0 is positive definite, both norms must be (Lipschitz) equivalent norms on X.

Now let $f = \sum_{k=1}^{\infty} a_k \phi_k$ and $g = \sum_{k=1}^{\infty} b_k \phi_k$ be elements of $L^2(\Omega)$. Then, a direct computation yields

$$B_0[f, f] = \langle L_0 f, f \rangle_{L^2(\Omega)} = \sum_{k=1}^{\infty} \langle \mu_k a_k \phi_k, a_k \phi_k \rangle = \sum_{k=1}^{\infty} \mu_k a_k^2$$

and

$$\|g\|_{L^2(\Omega)}^2 = \langle g, g \rangle_{L^2(\Omega)} = \sum_{k=1}^{\infty} b_k^2.$$

This suggests that $(f, g) \in X$ if and only if the quantities $\sum_{k=1}^{\infty} \mu_k a_k^2$ and $\sum_{k=1}^{\infty} b_k^2$ are both finite.

Now a tedious calculation reveals that for all $t \in \mathbb{R}$,

$$\|S_t \begin{pmatrix} f \\ g \end{pmatrix}\|_X^2 = \sum_{k=1}^{\infty} \mu_k a_k(t)^2 + \sum_{k=1}^{\infty} b_k(t)^2.$$

Recalling the preliminary computation done before the statement of this theorem, we have

$$a_k(t) = \cos(\sqrt{\mu_k}t)a_k + \frac{1}{\sqrt{\mu_k}}\sin(\sqrt{\mu_k}t)$$

and

$$b_k(t) = a_k'(t) = -\sqrt{\mu_k}\sin(\sqrt{\mu_k}t)a_k + \cos(\sqrt{\mu_k}t).$$

Conveniently, if $(f, g) \in X$, then for all $t \in \mathbb{R}$, the series define S_t must be convergent because

$$\|S_t \begin{pmatrix} f \\ g \end{pmatrix}\|_X^2 = \sum_{k=1}^{\infty} \mu_k a_k(t)^2 + \sum_{k=1}^{\infty} b_k(t)^2.$$

is convergent. Substituting the expressions for $a_k(t)$ and $b_k(t)$, we discover that

$$\|S_t \begin{pmatrix} f \\ g \end{pmatrix}\|_X^2 = \|\begin{pmatrix} f \\ g \end{pmatrix}\|_X^2.$$

So, for all $t \in \mathbb{R}$, each $S_t : X \to X$ is an isometry.

It remains to demonstrate that $\{S_t\}_{t\in\mathbb{R}}$ satisfies the properties of a strongly continuous semigroup. When $t = 0$, $a_k(t) = a_k$, $b_k(t) = b_k$ and

$$S_0 \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}.$$

The defining semigroup property

$$S_t S_s \begin{pmatrix} f \\ g \end{pmatrix} = S_{t+s} \begin{pmatrix} f \\ g \end{pmatrix}$$

for all $s, t \in \mathbb{R}$ follows from a long calculation. It remains to show that the map $t \mapsto S_t(f, g)$ is continuous from $\mathbb{R}$ to X, for all $f, g \in L^2(\Omega)$. However, we observe that for all $m \in \mathbb{Z}_{>0}$, the map

$$t \mapsto S_t \begin{pmatrix} \sum_{k=1}^{m} \langle f, \phi_k \rangle_{L^2(\Omega)} \phi_k \\ \sum_{k=1}^{m} \langle g, \phi_k \rangle_{L^2(\Omega)} \phi_k \end{pmatrix}$$

is continuous because each component function is a linear combination of trigonometric functions in the variable t, which is continuous. Taking the limit of the above maps as $m \to \infty$ then gives us the desired conclusion that $t \mapsto S_t(f,g)$ is continuous from $\mathbb{R}$ to X because the convergence is uniform (holds for all $t \in \mathbb{R}$).

$\square$

A consequence of the fact that the map $t \mapsto S_t(f,g)$ is continuous is that $\|u(t) - f\|_{H^1(\Omega)} \to 0$ and $\|v(t) - g\|_{L^2(\Omega)} \to 0$ as $t \to 0$ where

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = S_t \begin{pmatrix} f \\ g \end{pmatrix}.$$

The reason for this is because for all $t \in \mathbb{R}$, S_t is an isometry as determined by the theorem above. Thus, the initial conditions in 8.6 are all satisfied. Moreover, $u(t) \in H_0^1(\Omega)$ for all $t \in \mathbb{R}$, by definition of X, unveiling that the boundary conditions of 8.6 are also satisfied.

Chapter 9

Epilogue

This chapter concludes the notes on functional analysis and the close reading of [AB10] with a few comments. These notes were constructed over a span of 2 years, due to university coursework and research taking priority. Although I have learnt a lot about functional analysis, these notes ultimately reflect one person's view on the subject (Bressan). There are still many aspects of functional analysis I want to learn as a result. In order to obtain a more holistic grasp of the material, it was necessary to consult multiple references on the subject (see the bibliography). I highly recommend using multiple references to supplement a close reading of a particular text.

I wrote these notes with the intent of conveying the theory in a clear and comprehensible manner. Simply copying the notes in [AB10] is not conducive to learning the theory because it is not an active way of thinking about what was written or read. It is important to fill in crucial details in the proofs (see 5.3.1 for a concrete example). In this way, it forced me to think more conscientiously about what the result is trying to tell me.

One weakness of these notes is that it does not contain any completed exercises from [AB10] or any of the other references. Once I take this subject at university, as well as the closely related subjects on partial differential equations, I will tackle a wide variety of exercises in order to further internalise the material. I emphasise that it is important to do exercises when learning any mathematics subject.

I intend to do more close readings of texts that I am interested in. However, there are plenty of higher priority tasks to deal with first! I will update these notes once I find mistakes or better methods of explaining the theory.

Bibliography

[AB10] A. Bressan. *Lecture Notes on Functional Analysis and Linear Partial Differential Equations*, October 8th 2010.

[Bab71] I. Babuška. *Error-Bounds for Finite Element Method*, Numer. Math. **16**, 1971, Pages 322-333, MR0288971.

[DM21] D. Murfet. *Background 1: Matrix Exponentials*, 2021, Available at: http://therisingsea.org/notes/mast90132/background1.pdf

[EK78] E. Kreyszig. *Introductory Functional Analysis with Applications*, 1978, New York, John Wiley and Sons Inc, ISBN 0-471-50731-8, MR0467220.

[LE98] L. Evans. *Partial Differential Equations*, 1998, American Mathematical Society, Graduate Studies in Mathematics, Volume 19, ISBN 0-8218-0772-2, MR1625845.