Functional Analysis

Matthew M. Jones

October 2019

Contents

Chapter 1

Complete metric spaces

1.1 Introduction

This course is an abstract course studying, amongst other things, the geometry of infinite dimensional vector spaces, spaces of functions, and analysis on these spaces.

It is however also concerned with concrete applications and we will study some of these in the workshops. As always it is important, and will be emphasised whenever I can, the links between these areas and others you have studied or are currently studying.

The notion of *dimension* for example comes from Linear Algebra, whereas Metric Spaces are studied in any analysis course. It is expected that you know all of the theory from these two areas. If you have forgotten some of the theory then you should revise it in your own time. Having said that we will swiftly cover the main ideas below before extending these to include the notion of **completeness** which is fundamentally important in Functional Analysis.

1.2 Metric spaces

Definition 1.1 – Metric Space

Suppose X is a set. Then a metric on X is a function $d\colon X \times X \to \mathbb{R}$ that satisfies the following axioms for $x, y, z \in X$.

M1. $d(x, y) \geq 0$

M2. $d(x, y) = 0$ if and only if $x = y$

M3. $d(x, y) = d(y, x)$

M4. $d(x, y) \leq d(x, z) + d(z, y)$

The following important concepts will be used in this module and will be reviewed in the workshops.

* Open sets
* Closed sets
* Compact sets
* Limit points
* The closure of a set
* Continuous functions at a point y
* Uniformly continuous functions

Themes and results from the study of metric spaces that will be used include the following.

* The function $f\colon (X, d_X) \to (Y, d_Y)$ is continuous throughout X if and only if
  1. whenever O is open in Y, $f^{-1}(O)$ is open in X
  2. whenever C is closed in Y, $f^{-1}(C)$ is closed in X
* A continuous function maps compact sets to compact sets
* A continuous function on a compact set is uniformly continuous

1.3 Cauchy sequences

Definition 1.2 – Cauchy sequences

Suppose (x_n) is a sequence in a metric space (X, d). We say that (x_n) is a **Cauchy sequence** if for all $\epsilon > 0$ there is a $N > 0$ such that

$$n, m \geq N \qquad \Rightarrow \qquad d(x_n, x_m) < \epsilon.$$

Example 1.3. *Let us consider the sequence* $(1/n^2)_{n=1}^{\infty}$ *in* $(\mathbb{R}, d)$ *where* $d(x, y) = |x - y|$.

Without loss of generality we assume $m > n$. Then

$$\begin{aligned} d(x_n, x_m) &= \left| \frac{1}{n^2} - \frac{1}{m^2} \right| \\ &= \frac{1}{n^2} - \frac{1}{m^2} \\ &= \frac{m^2 - n^2}{n^2 m^2} \\ &\leq \frac{m^2}{n^2 m^2} && \text{since } m > n \\ &= \frac{1}{n^2} \end{aligned}$$

It follows that

$$n, m > \lfloor \epsilon^{-1/2} \rfloor \qquad \Rightarrow \qquad d(x_n, x_m) < \epsilon$$

This calculation is similar to how you prove that $1/n^2 \to 0$ as $n \to \infty$. Indeed the definition of Cauchy sequences differs only slightly from the definition of a convergent sequence.

In fact if $x_n \to x$ in the metric space (X, d) then

$$0 \leq d(x_n, x_m) \leq d(x_n, x) + d(x_m, x).$$

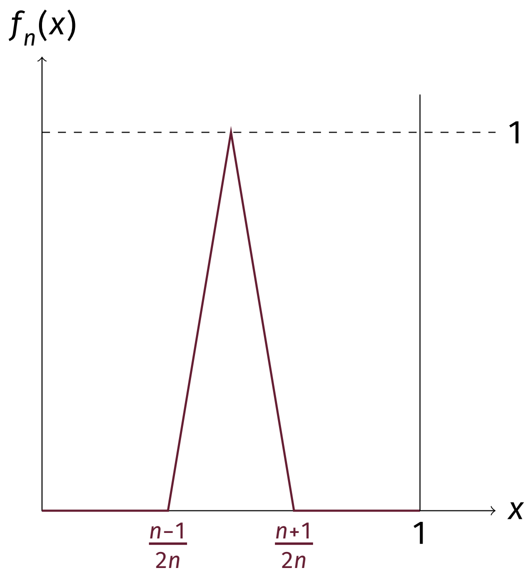
Therefore since the right hand inequality tends to 0 we have that the sequence (x_n) is a Cauchy sequence. In other words **If a sequence converges in a metric space then it is a Cauchy sequence.**

Unsurprisingly the converse is false in general.

Example 1.4. *Let* $C[0, 1]$ *be the metric space of continuous functions with metric*

$$d(f, g) = \int_0^1 |f(t) - g(t)| dt.$$

Let f_n be defined, for $n = 1, 2, \ldots$, as the function having graph



A calculation (Check!) shows that the area under this triangle is $1/(2n)$ and it follows that

$$d(f_n, f_m) = \frac{1}{2}\left(\frac{1}{n} - \frac{1}{m}\right).$$

Hence it is easy to check that $(f_n)_{n=1}^{\infty}$ is a Cauchy sequence. However the sequence **does not converge** to a continuous function.

Definition 1.5 – Complete metric space

A metric space (X, d) is **complete** if every Cauchy sequence converges.

The previous example shows that $C[0, 1]$ is not complete.

Theorem 1.6. *The metric space* $(\mathbb{R}^n, d)$ *is complete, where* $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$ *is the euclidean metric.*

We will prove this for $\mathbb{R}$ only, the proof is the same for $n > 1$. The structure will be as follows

If (x_n) is a Cauchy sequence in $\mathbb{R}$ we will:

1. show that (x_n) is bounded;
2. deduce that (x_n) has a convergent subsequence (Heine-Borel theorem);

3. show that (x_n) converges to this limit.

Proof. Suppose (x_n) is a Cauchy sequence. Then we will show that (x_n) converges.

1. Let $\epsilon > 0$ then there is a $N > 0$ such that

$$n, m > N \qquad \Rightarrow \qquad |x_n - x_m| < \epsilon.$$

In particular letting $m = N + 1$, for all $n > N$,

$$|x_n| - |x_{N+1}| \leq |x_n - x_{N+1}| < \epsilon.$$

Therefore,

$$|x_n| \leq |x_{N+1}| + \epsilon, \qquad \text{whenever } n > N.$$

Let $M = \max\{|x_1|, |x_2|, \ldots, |x_N|\}$. Then

$$|x_n| \leq \max\{M, |x_{N+1}| + \epsilon\}$$

for all n and hence (x_n) is a bounded sequence.

2. It follows by the Heine-Borel theorem (or the Bolzano-Weierstaß theorem) that (x_n) has a convergent subsequence, say

$$x_{n_k} \to x \qquad k \to \infty.$$

3. Let $\epsilon > 0$. Since (x_n) is a Cauchy sequence we may find a $N' > 0$ such that

$$n_k, n > N \qquad \Rightarrow \qquad |x_n - x_{n_k}| < \epsilon/2.$$

Since $x_{n_k} \to x$ as $n \to \infty$ we may pick $n_k > N'$ such that

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

Therefore for any $n > N'$

$$|x_n - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x| < \epsilon.$$

Hence $x_n \to x$ as $n \to \infty$ and we have shown that every Cauchy sequence converges.

$\square$

In any standard analysis class you will have learned that **every** metric on $\mathbb{R}^n$ is equivalent. So we have the following consequence.

Corollary 1.7. *The metric space $(\mathbb{R}^n, d)$ is complete with **any** metric d.*

Note that $\mathbb{R}^n$ is also a **vector space** whose dimension is n. It seems then that completeness in finite dimensions is automatic. The problem is then that many spaces that we come across in mathematics are *infinite dimensional*.

1.4 Density and separable metric spaces

We will use the notion of closures from metric spaces. Recall if A is a subset of a metric space (X, d) then its closure $\overline{A}$ is the smallest closed set containing A. In other words **any** closed set $B \supset A$ satisfies $B \supset \overline{A}$, or, equivalently

$$\overline{A} = \bigcap_{\substack{B \text{ closed} \\ A \subset B}} B.$$

Definition 1.8 – Density

A subset A of a metric space (X, d) is said to be **dense** if $\overline{A} = X$.

Example 1.9. *The metric space $(\mathbb{Q}^n, d)$ is dense in $(\mathbb{R}^n, d)$ for any metric d.*

Definition 1.10 – Separable metric space

A metric space (X, d) is **separable** if it contains a *countable* dense set A.

Note this means there is a sequence (x_n) in (X, d) such that for all $\epsilon > 0$ and $y \in (X, d)$ there is a n so that

$$d(x_n, y) < \epsilon.$$

In the example above since $\mathbb{Q}^n$ is countable, the metric space $(\mathbb{R}^n, d)$ is separable. We will meet non-separable metric spaces later on.

An important example of a separable metric space is $C[a, b]$, the space of continuous functions on the closed interval $[a, b]$ with metric

$$d_\infty(f, g) = \sup_{x \in [a,b]} |f(x) - g(x)|.$$

This follows from the following classic result

Theorem 1.11 – Weierstraß approximation theorem

Let $f \in (C[a,b], d_\infty)$. For all $\epsilon > 0$ there is a polynomial $p(x) = a_0 + a_1 x + \cdots a_n x^n$ such that $d(f, p) < \epsilon$.

Problems

1.1. Prove that for all $x, y, z \in (X, d)$, a metric space,

$$d(x, y) \geq d(x, z) - d(y, z).$$

Deduce that

$$d(x, y) \geq |d(x, z) - d(y, z)|\,.$$

1.2. Let (X, d) be a metric space and $z \in X$ a **fixed** element. Show that the map

$$\begin{aligned} f\colon X &\to X \\ x &\mapsto d(x, z) \end{aligned}$$

is continuous. Deduce that if $A \subset X$ is a compact set and $z \in X$ then there is a $a \in A$ such that

$$\min_{x \in A} d(x, z) = d(a, z).$$

Is a **unique** in this statement?

1.3. Let (x_n) be a sequence in a metric space (X, d). Suppose

$$d(x_{n+1}, x_n) \leq 2^{-n}, \qquad \text{for all } n = 1, 2, \ldots.$$

Show that (x_n) is a Cauchy sequence.

1.4. Prove that if

$$\sum_{n=1}^{\infty} d(x_{n+1}, x_n)$$

converges then (x_n) is a Cauchy sequence.

1.5. Show that $\mathbb{Q}$ is a dense subset of $\mathbb{R}$.

1.6. Show how the Weierstraß approximation theorem means that $(C[a, b], d_\infty)$ is separable (you may use the fact that a countable union of countable sets is countable).

Chapter 2

Normed vector spaces

2.1 Vector spaces

A vector space over $\mathbb{R}$ is a set V of **vectors v** along with two binary operations:[1]

Scalar multiplication	$\mathbb{R} \times V \to V$	$(\lambda, \mathbf{v}) \mapsto \lambda\mathbf{v}$
Vector addition	$V \times V \to V$	$(\mathbf{v}, \mathbf{u}) \mapsto \mathbf{v} + \mathbf{u}$

Careful! Do not confuse vector spaces with sets of vectors of the type you learn about in school (i.e. lines with arrows on them). A vector space is an abstract algebraic structure that can consist of other objects such as functions.

Example 2.1. *The set $C[0, 1]$ of continuous functions is a vector space.*

How do we define addition and scalar multiplication in this example?

Example 2.2. *The space $\mathbb{R}^n$ of n-tuples is a vector space.*

We saw in the last chapter that $\mathbb{R}^n$ is **also** a metric space. In order to be able to treat $\mathbb{R}^n$ as a metric space we needed the magnitude of the vector, $|\mathbf{v}|$.

In order to give a general vector space a metric space structure we also need a notion of magnitude. This is done using **norms**.

[1] It is also possible to define a vector space over any Field K by replacing $\mathbb{R}$ with K in the definition.

2.2 Norms and normed vector spaces

Definition 2.3 – Norm

Suppose V is a vector space. A **norm** on V is a function $\|\cdot\| : V \to \mathbb{R}$ that satisfies the following axioms for $\mathbf{u}, \mathbf{v} \in V$ and $\lambda \in \mathbb{R}$.

N1. $\|\mathbf{v}\| \geq 0$

N2. $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$

N3. $\|\mathbf{v} + \mathbf{u}\| \leq \|\mathbf{v}\| + \|\mathbf{u}\|$

N4. $\|\lambda\mathbf{v}\| = |\lambda|\,\|\mathbf{v}\|$

The first three axioms here should be compared to the axioms of a metric space. Indeed N3 is also known as the **triangle inequality**.

Definition 2.4 – Normed vector spaces

A vector space V together with a norm $\|\cdot\|$ is a **normed vector space**. We write this as the pair $(V, \|\ \|)$.

Theorem 2.5. *A normed vector space* $(V, \|\ \|)$ *is a metric space with metric*

$$d(\mathbf{v}, \mathbf{u}) = \|\mathbf{v} - \mathbf{u}\|.$$

Proof. Exercise: confirm the axioms for the metric d using the axioms for the norm N1, N2, N3 and N4. □

Example 2.6. *In* $\mathbb{R}^n$ *we may define the following norms for* $\mathbf{v} = (v_1, v_2, \ldots, v_n)^T$.

$$\begin{aligned} \|\mathbf{v}\|_1 &= \sum_{i=1}^{n} |v_i| \\ \|\mathbf{v}\|_2 &= \left(\sum_{i=1}^{n} |v_i|^2\right)^{1/2} \\ \|\mathbf{v}\|_\infty &= \max_{i=1,2,\ldots,n} |v_i| \end{aligned} \tag{2.1}$$

Note: It can be shown that

$$\|\mathbf{v}\|_\infty \leq \|\mathbf{v}\|_1 \leq \|\mathbf{v}\|_2 \leq \sqrt{n}\|\mathbf{v}\|_\infty.$$

In the same way that you did with metrics, we say that these norms are **equivalent**.

If V is a finite dimensional vector space then *every* norm is equivalent to every other norm. Consequently every finite dimensional vector space is *isomorphic* to $\mathbb{R}^n$ for some n. As metric spaces the notions of open / closed / compact sets are therefore valid and all the same regardless of the norm.

We can now express concepts of analysis in terms of norms. For example, the sequence $\mathbf{u}_n$, $n = 1, 2, \ldots$, converges to $\mathbf{u} \in V$ if

$$\lim_{n\to\infty} \|\mathbf{u}_n - \mathbf{u}\| = 0.$$

We can also define convergence of a series. Suppose we have a sequence of vectors $(\mathbf{v}_n)_{n=1}^{\infty}$. Then what does

$$\sum_{n=1}^{\infty} \mathbf{v}_n$$

mean? If we generalise straight from real analysis you should guess that this means there is a vector $\mathbf{s}$ such that

$$\mathbf{s} = \lim_{N\to\infty} \sum_{n=1}^{N} \mathbf{v}_n.$$

Or, alternatively,

$$\lim_{N\to\infty} \left\| \mathbf{s} - \sum_{n=1}^{N} \mathbf{v}_n \right\| = 0. \tag{2.2}$$

As in real analysis, $\sum_{n=1}^{N} \mathbf{v}_n$ are the **partial sums** of the series.

2.3 Inner products

The most natural norm on $\mathbb{R}^n$ is $\|\mathbf{v}\|_2$ defined by (2.1). Often this is called the *Euclidean norm* since it has a geometric interpretation as the Euclidean distance from the origin to the point defined by the vector $\mathbf{v}$. There is another reason why this norm is more natural – it can also be defined using an **inner product**.

For two vectors $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)^T$ the inner product is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \|\mathbf{u}\|_2 \|\mathbf{v}\|_2 \cos\theta,$$

where θ is the angle between the two vectors $\mathbf{u}$ and $\mathbf{v}$.

The inner product on $\mathbb{R}^n$ adds more geometric structure to the vector space since it allows us to find the *angle* between two vectors.

The Euclidean norm, $\|\cdot\|_2$, can be defined by this inner product since

$$\|\mathbf{v}\|_2 = \sqrt{\mathbf{v}\cdot\mathbf{v}}.$$

In general, an inner product on a vector space over $\mathbb{R}$ can take any number of forms. The following definition for an inner product gives enough information for it to be useful.

Definition 2.7 – Inner product

Let V be a vector space. An **inner product** is a mapping from $V \times V \to \mathbb{R}$, written as

$$\langle \mathbf{u}, \mathbf{v} \rangle,$$

and satisfying the following axioms for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\lambda \in \mathbb{R}$.

IP1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

IP2. $\langle \lambda\mathbf{u}, \mathbf{v} \rangle = \lambda\langle \mathbf{u}, \mathbf{v} \rangle$

IP3. $\langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle$

IP4. $\langle \mathbf{u}, \mathbf{u} \rangle > 0$ when $\mathbf{u} \neq \mathbf{0}$

Note now that with this definition, any vector space that has an inner product can be made into a normed vector space where the norm is defined by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

Definition 2.8 – Banach and Hilbert spaces

1. A **Banach space** is a complete, normed vector space
2. A **Hilbert space** is a complete, normed vector space where the norm is induced by an inner product

Note: some authors and textbooks also require that the vector spaces are infinite dimensional. We won't but will highlight the differences when they occur.

Theorem 2.9. *If U is a Banach space and $V \subset U$ is a **closed** subspace then V is also a Banach space.*

Proof. Any Cauchy sequence in V is also a Cauchy sequence in U. Therefore the sequence converges (since U is complete).

Since V is closed the limit must also be an element of V. Hence V is complete and therefore a Banach space. □

We will state here the following extremely useful inequality.

Theorem 2.10 (Cauchy-Schwarz inequality). *Let U be a vector space with inner product $\langle \mathbf{u}, \mathbf{v} \rangle$. Then for all $\mathbf{u}, \mathbf{v} \in U$,*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|.$$

The proof can be found in any of the textbooks.

2.4 Sequence spaces: a concrete example

Let $0 < p < \infty$. We define ℓ^p to be the set of sequences $(a_n)_{n=1}^{\infty}$ such that the series

$$\sum_{n=1}^{\infty} |a_n|^p$$

converges. These spaces are called *sequence spaces*.

Example 2.11. *The sequence $(1/n)_{n=1}^{\infty}$ is an element of ℓ^p whenever $p > 1$ since*

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges if and only if $p > 1$.

We define the norm on ℓ^p as

$$\|(a_n)\|_p = \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{1/p}.$$

This is sometimes called the p-norm.

Example 2.12. *The sequence $(1/n)_{n=1}^{\infty}$ has 2-norm*

$$\|(1/n)\|_2 = \frac{\pi}{\sqrt{6}}.$$

We will prove the following theorem in this section and in the process meet a few useful results and techniques.

Theorem 2.13

If $1 \le p < \infty$ then ℓ^p is a Banach space. Furthermore ℓ^2 is a Hilbert space.

These are very important spaces called sequence spaces.

Lemma 2.14 (Hölder's Inequality). *Suppose $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\left|\sum_{k=1}^{n} a_k b_k\right| \le \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |b_k|^q\right)^{1/q} \tag{2.3}$$

Proof. Given non-negative real numbers a and b it is possible to prove the following inequality,

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}. \tag{2.4}$$

Let

$$A^p = \sum_{k=1}^{n} |a_k|^p,$$
$$B^q = \sum_{k=1}^{n} |b_k|^q.$$

We assume without loss of generality that both A and B are positive. Then define $A_k = a_k/A$ and $B_k = b_k/B$. So that

$$\sum_{k=1}^{n} |A_k|^p = 1, \qquad \sum_{k=1}^{n} |B_k|^q = 1.$$

Now

$$\begin{aligned}\left|\sum_{k=1}^{n} A_k B_k\right| &\le \sum_{k=1}^{n} |A_k B_k| \\ &\le \sum_{k=1}^{n}\left(\frac{A_k}{p} + \frac{B_k}{q}\right) \\ &= \frac{1}{p}\sum_{k=1}^{n} |A_k| + \frac{1}{q}\sum_{k=1}^{n} |B_k| \\ &= \frac{1}{p} + \frac{1}{q} = 1 \\ &= \left(\sum_{k=1}^{n} |A_k|^p\right)^{1/p}\left(\sum_{k=1}^{n} |B_k|^q\right)^{1/q}.\end{aligned}$$

Returning to our original sequence note first that

$$\begin{aligned}\left|\sum_{k=1}^{n} a_k b_k\right| &= AB\left|\sum_{k=1}^{n} A_k B_k\right| \\ &\le AB\left(\sum_{k=1}^{n} |A_k|^p\right)^{1/p}\left(\sum_{k=1}^{n} |B_k|^q\right)^{1/q} \\ &= \left(\sum_{k=1}^{n} |AA_k|^p\right)^{1/p}\left(\sum_{k=1}^{n} |BB_k|^q\right)^{1/q} \\ &= \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p}\left(\sum_{k=1}^{n} |b_k|^q\right)^{1/q}.\end{aligned}$$

□

Note: this inequality and the next concern *finite* sequences (i.e. n-tuples) – we will need to take limits to generalise them to the norms.

When $p = q = 2$ this is just the Cauchy-Schwartz inequality, so Hölder's inequality should be thought of as a generalisation of the Cauchy-Schwartz inequality. It will be used to prove the triangle inequality N3 on ℓ^p.

Lemma 2.15 (Minkovski's Inequality). *If (a_k) and (b_k) are real sequence and $p > 1$ then*

$$\left(\sum_{k=1}^{n} |a_k + b_k|^p\right)^{1/p} \le \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |b_k|^p\right)^{1/p} \tag{2.5}$$

Proof. First we write

$$\begin{aligned}\sum_{k=1}^{n}|a_k+b_k|^p &= \sum_{k=1}^{n}|a_k+b_k|^{p-1}|a_k+b_k| \\ &\leq \sum_{k=1}^{n}|a_k+b_k|^{p-1}(|a_k|+|b_k|) \\ &= \sum_{k=1}^{n}|a_k|\,|a_k+b_k|^{p-1}+\sum_{k=1}^{n}|b_k|\,|a_k+b_k|^{p-1}.\end{aligned}$$

Now with q defined as in Hölder's inequality we have that

$$\sum_{k=1}^{n}|a_k|\,|a_k+b_k|^{p-1} \leq \left(\sum_{k=1}^{n}|a_k+b_k|^{q(p-1)}\right)^{1/q}\left(\sum_{k=1}^{n}|a_k|^p\right)^{1/p}$$

and similarly,

$$\sum_{k=1}^{n}|b_k|\,|a_k+b_k|^{p-1} \leq \left(\sum_{k=1}^{n}|a_k+b_k|^{q(p-1)}\right)^{1/q}\left(\sum_{k=1}^{n}|b_k|^p\right)^{1/p}$$

So that

$$\sum_{k=1}^{n}|a_k+b_k|^p \leq \left(\sum_{k=1}^{n}|a_k+b_k|^{p}\right)^{1/q}\left(\left(\sum_{k=1}^{n}|a_k|^p\right)^{1/p}+\left(\sum_{k=1}^{n}|b_k|^p\right)^{1/p}\right).$$

Dividing both sides by $\left(\sum_{k=1}^{n}|a_k+b_k|^p\right)^{1/q}$ (assuming without loss of generality that this is positive) we have

$$\left(\sum_{k=1}^{n}|a_k+b_k|^{p}\right)^{1-1/q} \leq \left(\sum_{k=1}^{n}|a_k|^p\right)^{1/p}+\left(\sum_{k=1}^{n}|b_k|^p\right)^{1/p}.$$

The inequality follows since $1-\dfrac{1}{q}=\dfrac{1}{p}$. □

Minkovski's inequality is the main technical difficulty with proving that ℓ^p is a Banach space ($p \geq 1$). The rest is proving completeness.

The structure of the proof of Theorem 2.13 will be as follows.

1. Verify that ℓ^p is a vector space

2. Verify that the p-norm is a norm on ℓ^p

3. Show that ℓ^p is complete

Although parts 1 and 2 can be interwoven to some extent.

Proof of Theorem 2.13 parts 1 and 2: that ℓ^p is a normed vector space. **The space ℓ^p is a normed vector space**

If $(a_n) \in \ell^p$ then we define scalar multiplication as

$$\lambda(a_n) = (\lambda a_n), \qquad \lambda \in \mathbb{R}.$$

We also define addition as

$$(a_n) + (b_n) = (a_n + b_n)$$

If $(a_n) \in \ell^p$ then

$$\sum |\lambda a_n|^p = |\lambda|^p \sum |a_n|^p.$$

So $\lambda(a_n) \in \ell^p$ and

$$\|\lambda(a_n)\|_p = |\lambda| \, \|(a_n)\|_p.$$

We will simultaneously prove the triangle inequality N3 and the closure under addition of ℓ^p.

Note in order to use Minkovski's inequality we will use the fact that

$$\sum_{n=1}^{\infty} |a_n|^p = \lim_{N\to\infty} \sum_{n=1}^{N} |a_n|^p,$$

and similar relations for (b_n) and $(a_n + b_n)$ etc.

We will also use the fact that since $x \mapsto x^{1/p}$ is continuous

$$\|(a_n)\|_p = \lim_{N\to\infty} \left(\sum_{n=1}^{N} |a_n|^p \right)^{1/p}$$

whenever the limit on the right exists.

N1 and N4 follow from these observations.

To prove N2 suppose $\|(a_n)\| = 0$. Then

$$\sum_{n=1}^{\infty} |a_n|^p = 0.$$

Let $S_N = \sum_{n=1}^{N} |a_n|^p$. Then

$$S_1 = |a_1|^p \geq 0, \qquad S_{N+1} - S_N = |a_{N+1}|^p \geq 0.$$

Therefore (S_N) is an increasing sequence:

$$0 \leq S_1 \leq S_2 \leq \cdots \leq S_N \leq S_{N+1} \leq \cdots$$

By elementary properties of increasing sequences,

$$0 = \lim_{N\to\infty} S_N = \sup_N S_N.$$

Therefore $S_N = 0$ for all N and

$$|a_{N+1}| = \left(S_{N+1} - S_N\right)^{1/p} = 0.$$

Hence $(a_n) = (0, 0, 0, \ldots) = \mathbf{0}$ and we have verified N2.

To prove N3 the triangle inequality we use Minkovski's inequality, (2.5),

$$\left(\sum_{n=1}^{N} |a_n + b_n|^p\right)^{1/p} \leq \left(\sum_{n=1}^{N} |a_n|^p\right)^{1/p} + \left(\sum_{n=1}^{N} |b_n|^p\right)^{1/p}.$$

The left-hand side increases with N and the right-hand side converges as $N \to \infty$ to

$$\|(a_n)\|_p + \|(b_n)\|_p$$

if $(a_n), (b_n) \in \ell^p$. Therefore

$$\begin{aligned}
\|(a_n + b_n)\|_p &= \lim_{N\to\infty} \left(\sum_{n=1}^{N} |a_n + b_n|^p\right)^{1/p} \\
&= \sup_N \left(\sum_{n=1}^{N} |a_n + b_n|^p\right)^{1/p} \\
&\leq \|(a_n)\|_p + \|(b_n)\|_p.
\end{aligned}$$

We have therefore verified N3. This also coroborates that ℓ^p is closed under addition. We have thus shown that ℓ^p is a normed vector space. □

We also need to show that ℓ^p is complete. This means we must show that *any* Cauchy sequence converges in ℓ^p. This, and all proofs of completeness, take the following form.

STEP 1. Find a candidate for a limit (a potential limit), normally by some weaker notion of a limit

STEP 2. Show that the Cauchy sequence does converge to this limit

STEP 3. Show that the candidate limit is in the normed vector space

This is a common strategy throughout analysis (it's how we find solutions to differential equations as well).

Careful: the notation in this proof gets a bit hairy, make sure you're following which index is which in the sequences!

Proof of Theorem 2.13 par 3: that ℓ^p is complete. For each $m = 1, 2, \ldots$ define

$$\mathbf{a}^{(m)} = (a_1^{(m)}, a_2^{(m)}, a_3^{(m)}, \ldots) = (a_n^{(m)})_{n=1}^{\infty}.$$

We will assume that $(\mathbf{a}^{(m)}) = (\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots)$ is a Cauchy sequence in ℓ^p. We must show that it converges in ℓ^p using the strategy outlined above.

STEP 1. Find a candidate for a limit.

First note that for any $N, M \geq 1$ and for **fixed** n

$$\begin{aligned}|a_n^{(N)} - a_n^{(M)}| &= \left(|a_n^{(N)} - a_n^{(M)}|^p\right)^{1/p} \\ &\leq \left(\sum_{n=1}^{\infty} |a_n^{(N)} - a_n^{(M)}|^p\right)^{1/p} = \|\mathbf{a}^{(N)} - \mathbf{a}^{(M)}\|_p.\end{aligned}$$

Therefore since $(\mathbf{a}^{(m)})_{m=1}^{\infty}$ is a Cauchy sequence then so is $(a_n^{(m)})_{m=1}^{\infty}$ for each n. But this latter sequence is a sequence in $\mathbb{R}$ and so since $\mathbb{R}$ is complete (Theorem 1.6) we have that the limit

$$\lim_{m \to \infty} a_n^{(m)} = a_n$$

exists. The sequence $\mathbf{a} = (a_n)_{n=1}^{\infty}$ is our candidate limit.

STEP 2. Show that the Cauchy sequence does converge to this limit.

Let $\epsilon > 0$. Then there is a N_0 such that for $N, M > N_0$

$$\|\mathbf{a}^{(N)} - \mathbf{a}^{(M)}\|_p < \epsilon.$$

But for any M_0

$$\sum_{n=1}^{M_0} |a_n^{(N)} - a_n^{(M)}|^p \le \sum_{n=1}^{\infty} |a_n^{(N)} - a_n^{(M)}|^p = \|\mathbf{a}^{(N)} - \mathbf{a}^{(M)}\|_p^p < \epsilon^p.$$

Letting $M \to \infty$ on the left-hand side (which we can do since it's a finite sum) gives us

$$\sum_{n=1}^{M_0} |a_n^{(N)} - a_n|^p \le \epsilon^p.$$

This sum is increasing in M_0 and bounded by ϵ^p. It therefore converges and

$$\sum_{n=1}^{\infty} |a_n^{(N)} - a_n|^p \le \epsilon^p.$$

Hence we have shown that

$$\|\mathbf{a}^{(N)} - \mathbf{a}\|_p \le \epsilon. \tag{2.6}$$

Therefore

$$\lim_{N\to\infty} \mathbf{a}^{(N)} = \mathbf{a}.$$

STEP 3. Show that the candidate limit is in ℓ^p.

This bit's easy in fact. We know that for each N, $\mathbf{a}^{(N)}$ is in ℓ^p. We also now know, by (2.6), that

$$\mathbf{a}^{(N)} - \mathbf{a} \in \ell^p.$$

Since ℓ^p is a vector space and

$$\mathbf{a} = \mathbf{a}^{(N)} - \left(\mathbf{a}^{(N)} - \mathbf{a}\right)$$

we have that $\mathbf{a} \in \ell^p$ as required.

□

We will finish this section with two other sequence spaces.

Example 2.16. *The space ℓ^∞ is the Banach space of bounded sequences (a_n). The norm is*

$$\|(a_n)\|_\infty = \sup_n |a_n|.$$

Example 2.17. *The space c_0 is the normed vector space of sequences that tend to 0 (with the ∞-norm defined above). This is a subspace of ℓ^∞ and is* ***closed****. Therefore it is also a Banach space by Theorem 2.9.*

Note we have the following strict inclusions

$$\ell^p \subset \ell^{p'} \subset c_0 \subset \ell^\infty$$

whenever $p < p'$. In other words the sequence spaces ℓ^p get bigger as p gets bigger.

Question: Strict means that for each $p < p'$ there is a sequence in $\ell^{p'}$ that is not in ℓ^p. Find one.

2.5 Spaces of functions: another concrete example

Define $C[a, b]$ to be the set of continuous functions defined on the **compact** interval $[a, b]$.

The **uniform norm** on $C[a, b]$ is defined as

$$\|f\|_\infty = \sup_{a \le t \le b} |f(t)|, \qquad f \in C[a, b].$$

Theorem 2.18. *Let $[a, b]$ be a* ***compact*** *interval. Then $(C[a, b], \| \, \|_\infty)$ is a Banach space.*

Being a Banach space implies that $C(I)$ is closed, meaning that if $f_n \to f$ in the uniform norm then $f \in C(I)$. This is a special type of convergence called **uniform convergence**.

Finally we will also be interested in the so-called Lebesgue spaces.

Definition 2.19. A set $E \in \mathbb{R}$ is a **set of measure 0** or a **null set** if for all $\epsilon > 0$ there is an open covering of E by intervals I_k, $k = 1, 2, \ldots$ such that

$$\sum_{i=1}^{\infty} |I_k| < \epsilon,$$

where $|I_k|$ denotes the size of the interval I_k (i.e. if $I = (a, b)$ then $|I| = b - a$).

Examples of sets of measure 0 include all finite sets, and all countable sets of points.

Definition 2.20 – Lebesgue spaces

Let $1 \leq p < \infty$ and $I \subset \mathbb{R}$. The **Lebesuge space** $L^p(I)$ is the normed vector space of functions f for which the norm

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

Here two functions f and g are considered the same if $f(x) = g(x)$ for all points x outside a set of measure 0.

Warning. The integral above has to be carefully defined – the Riemann integral is not good enough for this. Instead the integral will be understood as the **Lebesgue integral**. We won't describe in detail here what this is since it is beyond the scope of these notes. Suffice it to say that this is a generalise, more flexible definition of the integral that means, amongst many other things, that the Lebesgue spaces below are complete.

The Lebesgue spaces have similar properties to the sequence spaces ℓ^p. For example $L^p(I)$ is a Banach space for $p \geq 1$. There is also a version of the Hölder inequality, for $f \in L^p(I)$ and $g \in L^q(I)$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$,

$$\int_I f(t)g(t)dt \leq \left(\int_I |f(t)|^p\right)^{1/p} \left(\int_I |g(t)|^q\right)^{1/q}.$$

Finally, like ℓ^2, $L^2(I)$ is a Hilbert space with inner product

$$\langle f, g \rangle = \int_I f(t)g(t)dt.$$

Proving these spaces are complete requires a deeper understanding of the Lebesgue integral so we won't worry too much about it yet since you've not yet seen it. But these spaces are important and we will come back to them over the rest of the module.

Problems

2.1. Prove Theorem 2.5.

2.2. Let $(U, \|\cdot\|)$ be a normed vector space. Show that for any $\mathbf{u}, \mathbf{v} \in U$, the triangle inequality implies,

$$\|\mathbf{u} - \mathbf{v}\| \geq |\,\|\mathbf{u}\| - \|\mathbf{v}\|\,|\,.$$

2.3. Use your answer to the previous problem to show that the function

$$\begin{aligned} f\colon U &\to U \\ \mathbf{u} &\mapsto \|\mathbf{u}\| \end{aligned}$$

is continuous.

2.4. Suppose $K \subset U$ is a **compact** subset. Show that there is a $\mathbf{u} \in K$ such that

$$\|\mathbf{u}\| = \sup\{\|\mathbf{v}\| : \mathbf{v} \in K\}.$$

2.5. Let H be a vector space with inner product $\langle \mathbf{u}, \mathbf{v}\rangle$. Prove the **parallelogram rule**:

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2.$$

2.6. Let H be a vector space with inner product $\langle \mathbf{u}, \mathbf{v}\rangle$. Show that if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal (i.e. $\langle \mathbf{u}, \mathbf{v}\rangle = 0$) then they satisfy **Pythagoras's theorem**,

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

2.7. Let $a_n = 2^{-n}$, $n = 1, 2, \ldots$. Show that $(a_n) \in \ell^p$ for all p and

$$\|(a_n)\|_p = (2^p - 1)^{-1/p}.$$

2.8. Let $(\rho_k)_{k=0}^{\infty}$ be a real-sequence. Define the sequence $(a_n)_{n=1}^{\infty}$ as

$$a_n = 2^{-k}\rho_k, \qquad \text{when } 2^k \leq n < 2^{k+1}.$$

(a) Find a sequence (ρ_k) so that (a_n) has the properties:

$$\sum_{n=1}^{\infty} a_n = \infty, \qquad \text{whereas } na_n \to 0, \qquad (n \to \infty).$$

(b) Find a sequence (ρ_k) so that (a_n) has the property:

$$(a_n) \in \ell^p \qquad \Longleftrightarrow \qquad p \geq 1.$$

2.9. Show that on $L^2[0,1]$,

$$\langle f, g \rangle = \int_0^1 f(t)g(t)\mathrm{d}t$$

defines an inner product.

2.10. Let H be a vector space with inner product and suppose $\mathbf{v} \in H$ is fixed. Use the Cauchy-Schwartz inequality to show that the mapping

$$\begin{aligned} f\colon H &\to H \\ \mathbf{u} &\mapsto \langle \mathbf{u}, \mathbf{v} \rangle \end{aligned}$$

is continuous.

2.11. Let I be a **compact** subset of $\mathbb{R}$. Show that $L^{p_2}(I) \subset L^{p_1}(I)$ whenever $1 \leq p_1 < p_2$, (hint: use Hölder's inequality).

2.12. Consider the following two functions:

$$f(x) = \begin{cases} \frac{1}{\sqrt{x}} & 0 < x < 1; \\ 0 & \text{otherwise.} \end{cases} \qquad\qquad g(x) = \begin{cases} \frac{1}{x} & x \geq 1; \\ 0 & \text{otherwise.} \end{cases}$$

Decide whether or not each function belongs to $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$. What does this tell you about $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$?

Chapter 3

Bases

3.1 Finite-dimensional bases

If V is a finite dimensional vector space this means it has a **basis** of vectors

$$E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}, \qquad n = \dim V.$$

We want to have a similar idea for normed vector spaces that may be infinite dimensional. Before we do that we need to make clear what we mean by an *infinite dimensional* vector space.

Definition 3.1. A vector space V is said to be infinite dimensional if there is no finite set that spans V.

We will in this module write the span of a (finite) set of vectors as

$$\operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\} = \{\lambda_1\mathbf{e}_1 + \lambda_2\mathbf{e}_2 + \dots + \lambda_n\mathbf{e}_n : \lambda_i \in \mathbb{R}\}.$$

A **basis** E is required to have the following two properties:

1. **(Spanning set)** $\operatorname{span} E = V$
2. **(Linear independence)** If $\lambda_1\mathbf{e}_1 + \lambda_2\mathbf{e}_2 + \dots + \lambda_n\mathbf{e}_n = \mathbf{0}$ then $\lambda_i = 0$ for each i

These can be understood in the following way: the first says that every vector in V can be written as a linear combination of the basis vectors, and the second says that this representation is unique.

3.2 Linear spans and Hamel bases

Let E denote a set of vectors in a normed vector space $(V, \| \ \|)$. Then we define

$$\operatorname{span} E = \left\{ \sum_{i=1}^{n} \lambda_i \mathbf{e}_i : \mathbf{e}_i \in E, n \in \mathbb{N} \right\}.$$

Think of E as an infinite set of vectors. Then the span is the set of all possible *finite* linear combinations of these vectors.

Letting $E_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ we could also write this as

$$\operatorname{span} E = \bigcup_{n>0} \operatorname{span} E_n.$$

The set E is *linearly independent* if *for any n*

$$\sum_{i=1}^{n} \lambda_i \mathbf{e}_i = \mathbf{0} \qquad \Rightarrow \qquad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Definition 3.2 – Hamel Basis

A set E is a **Hamel basis** for a normed vector space $(V, \| \ \|)$ if

- $\operatorname{span} E = V$
- E is linearly independent

For *finite dimensional* vector spaces V a Hamel basis is the same as a basis. Each element of V in this case can be written as

$$\mathbf{v} = \alpha_1 \mathbf{e}_1 + \dots + \alpha_n \mathbf{e}_n.$$

What about infinite dimensional V?

Example 3.3. *In the Hilbert space ℓ^2 define the vector $\mathbf{e}_i$ to be the sequence*

$$\mathbf{e}_i = (0, 0, \dots, 0, 1, 0, \dots)$$

where the 1 *is in the ith position.*

Is $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots\}$ a Hamel basis? It seems like it should be a good candidate for a basis since the equivalent vectors in $\mathbb{R}^n$ are in fact the standard basis.

There is a problem however. For example the sequence $(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{16}, \ldots)$ is in ℓ^2 but is not in span E.

It turns out that Hamel bases, being the most obvious generalisation of a basis in finite dimensions, are not very useful.

Instead we start with a definition that encapsulates what we want out of a basis.

Definition 3.4 – Schauder Basis

A set $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ is a **Schauder basis** for a normed vector space $(V, \| \ \|)$ if every vector $\mathbf{v} \in V$ can be written **uniquely** as an infinite series[a]

$$\mathbf{v} = \sum_{i=1}^{\infty} \alpha_i \mathbf{e}_i, \qquad \alpha_i \in \mathbb{R}.$$

[a] Recall infinite series are defined as in (2.2).

It can be shown that every vector space has a Hamel basis but not necessarily a Schauder basis. The existence of a Schauder basis does, however, have a useful consequence.

Definition 3.5 – Separable Normed Vector Space

Let $(U, \| \cdot \|)$ be a normed vector space. We say that $(U, \| \cdot \|)$ is **separable** if the metric space induced by the norm is separable.

Theorem 3.6

If the Banach space $(U, \| \cdot \|)$ has a Schauder basis then it is separable.

Remark: Whether or not the converse to this statement was true in general was a famous open problem for a number of years, first appearing in Banach's book Banach (1955); it was solved in 1973 Per Enflo constructed a separable Banach space with no Schauder basis, (Enflo 1973, Theorem 1). The proof involves a delicate and brilliant construction of a Banach space by gluing together infinitely many finite dimensional subspaces whose dimension tends to infinity.

Example 3.7. *For the sequence spaces ℓ^p, $1 \le p < \infty$, define the sequence $\mathbf{e}_n$ as having* 1 *as its nth term and all other terms* 0.

$$\mathbf{e}_n = (0, 0, \ldots, 0, 1, 0, \ldots).$$

Then $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ *forms a Schauder basis for* ℓ^p.

We will see in Problem 3.5. that ℓ^∞ is **not** separable.

3.3 Orthonormal bases for a Hilbert space

When we are in the Hilbert space setting there is a special and elegant way of finding a Schauder basis. Furthermore a Hilbert space is separable if and only if it has a Schauder basis.

In this section we let H denote a Hilbert space with inner product $\langle \cdot, \cdot \rangle$.

Two vectors $\mathbf{u}, \mathbf{v} \in H$ are orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Definition 3.8. A set $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ is said to be an **orthogonal sequence** if

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0, \qquad \text{whenever } i \neq j.$$

If, in addition, $\|\mathbf{e}_i\| = 1$ for all i then E is said to be an **orthonormal sequence**.

Our aim will be to show that orthonormal sequences can be Schauder bases for H and the coefficients (that we will call Fourier coefficients) can be calculated explicitly.

We will need some preliminary results that are interesting in their own right and are unique to Hilbert space.

Recall that $A \subset H$ is a subspace if whenever $\mathbf{a}, \mathbf{b} \in A$ so is $\lambda \mathbf{a} + \mu \mathbf{b}$ for any $\lambda, \mu \in \mathbb{R}$.

Also remember the parallelogram formula:

$$\|\mathbf{u} - \mathbf{v}\|^2 + \|\mathbf{u} + \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2.$$

Lemma 3.9 (Closest point property)**.** *Let A be a closed, non-empty subspace of a Hilbert space H.*

For every $\mathbf{v} \in H$ *there is a* ***unique*** $\mathbf{a} \in A$ *such that* $\mathbf{a}$ *is closer to* $\mathbf{v}$ *than any other point.*

Proof. Let $M = \inf_{\mathbf{b} \in A} \|\mathbf{b} - \mathbf{v}\|$. Then $M < \infty$ since A is non-empty, and M is the distance from $\mathbf{v}$ to A.

It follows that there is a sequence of points $\mathbf{y}_n$ in A, $n = 1, 2, \ldots$, such that

$$\|\mathbf{v} - \mathbf{y}_n\|^2 \leq M^2 + \frac{1}{n}.$$

We will show that $(\mathbf{y}_n)$ is a Cauchy sequence.

By the parallelogram formula,

$$\begin{aligned}\|(\mathbf{v}-\mathbf{y}_n)+(\mathbf{v}-\mathbf{y}_m)\|^2+\|(\mathbf{v}-\mathbf{y}_n)-(\mathbf{v}-\mathbf{y}_m)\|^2 &= 2\|\mathbf{v}-\mathbf{y}_n\|^2+2\|\mathbf{v}-\mathbf{y}_m\|^2\\ &\le 4M^2+2\left(\frac{1}{n}+\frac{1}{m}\right).\end{aligned}$$

Now the left-hand side is

$$\|2\mathbf{v}-(\mathbf{y}_n+\mathbf{y}_m)\|^2+\|\mathbf{y}_n-\mathbf{y}_m\|^2 = 4\left\|\mathbf{v}-\frac{1}{2}(\mathbf{y}_n+\mathbf{y}_m)\right\|^2+\|\mathbf{y}_n-\mathbf{y}_m\|^2.$$

But since A is a subspace and $\mathbf{y}_n, \mathbf{y}_m \in A$, $\frac{1}{2}(\mathbf{y}_n+\mathbf{y}_m) \in A$. So

$$\left\|\mathbf{v}-\frac{1}{2}(\mathbf{y}_n+\mathbf{y}_m)\right\| \ge M,$$

and so

$$\|(\mathbf{v}-\mathbf{y}_n)+(\mathbf{v}-\mathbf{y}_m)\|^2+\|(\mathbf{v}-\mathbf{y}_n)-(\mathbf{v}-\mathbf{y}_m)\|^2 \ge 4M^2+\|\mathbf{y}_n-\mathbf{y}_m\|^2.$$

Hence,

$$4M^2+\|\mathbf{y}_n-\mathbf{y}_m\|^2 \le 4M^2+2\left(\frac{1}{n}+\frac{1}{m}\right),$$

so that

$$\|\mathbf{y}_n-\mathbf{y}_m\|^2 \le 2\left(\frac{1}{n}+\frac{1}{m}\right)$$

Therefore $(\mathbf{y}_n)$ is a Cauchy sequence as we wanted to prove.

Since A is a closed subspace of a Hilbert space it is itself a Hilbert space, Theorem 2.9. Therefore $(\mathbf{y}_n)$ converges to a point $\mathbf{a} \in A$.

We have

$$M \le \|\mathbf{a}-\mathbf{v}\| = \lim_{n\to\infty}\|\mathbf{y}_n-\mathbf{v}\| \le \lim_{n\to\infty}\left(M^2+\frac{1}{n}\right)^{1/2} = M,$$

so that $\|\mathbf{a}-\mathbf{v}\| = M = \inf_{\mathbf{b}\in A}\|\mathbf{b}-\mathbf{v}\|$ as required.

To prove uniqueness suppose $\mathbf{v}_1$ and $\mathbf{v}_2$ are both closest points, $\|\mathbf{v}-\mathbf{v}_1\| = \|\mathbf{v}-\mathbf{v}_2\| = M$. Then by the parallelogram rule,

$$\begin{aligned}\|\mathbf{v}_1-\mathbf{v}_2\|^2 &= \|(\mathbf{v}-\mathbf{v}_2)-(\mathbf{v}-\mathbf{v}_1)\|^2\\ &= 2\|(\mathbf{v}-\mathbf{v}_2)\|^2+2\|(\mathbf{v}-\mathbf{v}_2)\|^2-\|(\mathbf{v}-\mathbf{v}_2)+(\mathbf{v}-\mathbf{v}_2)\|^2\\ &= 2M^2+2M^2-4\left\|\mathbf{v}-\frac{1}{2}(\mathbf{v}_1+\mathbf{v}_2)\right\|^2\\ &\le 2M^2+2M^2-4M^2 = 0,\end{aligned}$$

since $\frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2) \in A$. Hence $\mathbf{v}_1 = \mathbf{v}_2$. □

Definition 3.10 – Fourier coefficients

Let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ be an orthonormal sequence in a Hilbert space H. Let $\mathbf{v} \in H$. Then the nth **Fourier coefficient** is

$$\langle \mathbf{v}, \mathbf{e}_n \rangle.$$

We also define the following.

Definition 3.11 – Fourier series

Let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ be an orthonormal sequence in a Hilbert space H. Let $\mathbf{v} \in H$. The **Fourier series** of $\mathbf{v}$ is the sum

$$\sum_{n=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_n \rangle \mathbf{e}_n.$$

There are many questions here that need answering. Does the Fourier series converge? What does the Fourier series converge to?

What we would like is for the Fourier series to be $\mathbf{v}$. It turns out that this is the case, plus much more. And we will prove this in the rest of this chapter.

We will need the following two Lemmas. We won't prove them because they are simply a matter of expanding inner products.

Lemma 3.12. *If $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ is a finite orthogonal sequence in a Hilbert space H then*

$$\left\| \sum_{i=1}^{n} \mathbf{x}_i \right\|^2 = \sum_{i=1}^{n} \|\mathbf{x}_i\|^2.$$

Lemma 3.13. *Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ be an* ***orthonormal*** *sequence in a Hilbert space H and let $\mathbf{v} \in H$.*

Let $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$. Then

$$\left\| \mathbf{v} - \sum_{i=1}^{n} \lambda_i \mathbf{e}_i \right\|^2 = \|\mathbf{v}\|^2 + \sum_{i=1}^{n} |\lambda_i - c_i|^2 - \sum_{i=1}^{n} |c_i|^2, \tag{3.1}$$

where $c_i = \langle \mathbf{v}, \mathbf{e}_i \rangle$ is the ith Fourier coefficient.

As a consequence of the previous Lemma note that as $\lambda_1, \lambda_2, \ldots, \lambda_n$ vary the vector $\sum_{i=1}^n \lambda_i \mathbf{e}_i$ varies over all the elements of $E_n = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.

Therefore by taking the minimum over all possible values of the λs the left-hand side of (3.1) gives the distance from $\mathbf{v}$ to span E_n.

But the minimum of the right-hand side of (3.1) when the λs vary must occur when $|\lambda_i - c_i| = 0$ for all i. This minimum is then

$$\|\mathbf{v}\|^2 - \sum_{i=1}^n |c_i|^2.$$

We have the following consequence then.

Corollary 3.14. *Let $E_n = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be an* ***orthonormal*** *sequence in a Hilbert space H and let $\mathbf{v} \in H$.*

The closest point of span E_n *to* $\mathbf{v}$ *is*

$$\mathbf{u} = \sum_{i=1}^n \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i,$$

and the distance d satisfies

$$d^2 = \|\mathbf{v}\|^2 - \sum_{i=1}^n |\langle \mathbf{v}, \mathbf{e}_i \rangle|^2.$$

We now turn our attention to infinite orthonormal sequences.

Theorem 3.15 – Bessel's Inequality

Let $\mathbf{e}_1, \mathbf{e}_2, \ldots$ be an orthonormal sequence in a Hilbert space H. The for any $\mathbf{v} \in H$,

$$\sum_{i=1}^\infty |\langle \mathbf{v}, \mathbf{e}_i \rangle|^2 \leq \|\mathbf{v}\|^2 \tag{3.2}$$

Note that as a consequence the sum on the left converges.

Proof. By Corollary 3.14, for any m

$$\|\mathbf{v}\|^2 - \sum_{i=1}^m |\langle \mathbf{v}, \mathbf{e}_i \rangle|^2 = d^2 \geq 0.$$

Rearranging we get

$$\sum_{i=1}^{m} |\langle \mathbf{v}, \mathbf{e}_i \rangle|^2 \le \|\mathbf{v}\|^2.$$

Since this is true for all m, we have Bessel's inequality (3.2). □

The following remarkable theorem is unique to Hilbert spaces. It tells you a lot about the rigidity of Hilbert space, by which I mean they are all pretty similar.

Theorem 3.16

Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, \dots$ be an orthonormal sequence in a Hilbert space H and let $\lambda_i \in \mathbb{R}$ for all $n = 1, 2, \dots$.
Then

$$\sum_{i=1}^{\infty} \lambda_i \mathbf{e}_i \text{ converges} \qquad \Longleftrightarrow \qquad \sum_{i=1}^{\infty} \lambda_i^2 \text{ converges.}$$

Proof. Suppose first that $\mathbf{v} = \sum_{i=1}^{\infty} \lambda_i \mathbf{e}_i$ converges.
For $m \ge k$,

$$\left\langle \sum_{i=1}^{m} \lambda_i \mathbf{e}_i, \mathbf{e}_k \right\rangle = \lambda_k.$$

It follows by continuity that

$$\begin{aligned} \langle \mathbf{v}, \mathbf{e}_k \rangle &= \left\langle \lim_{m\to\infty} \sum_{i=1}^{m} \lambda_i \mathbf{e}_i, \mathbf{e}_k \right\rangle \\ &= \lim_{m\to\infty} \left\langle \sum_{i=1}^{m} \lambda_i \mathbf{e}_i, \mathbf{e}_k \right\rangle \\ &= \lambda_k. \end{aligned}$$

Hence

$$\sum_{i=1}^{\infty} \lambda_i^2 = \sum_{i=1}^{\infty} |\langle \mathbf{v}, \mathbf{e}_k \rangle|^2 \le \|\mathbf{v}\|^2 < \infty,$$

by Bessel's inequality, (3.2).
To prove the converse we start by assuming $\sum_{i=1}^{\infty} \lambda_i^2$ converges.
Let $S_m = \sum_{i=1}^{m} \lambda_i \mathbf{e}_i$. We must show this is a Cauchy sequence in m.

For $n, m \in \mathbb{N}$, with $n > m$,

$$\|S_n - S_m\|^2 = \left\| \sum_{i=m+1}^{n} \lambda_i \mathbf{e}_i \right\|^2 = \sum_{i=m+1}^{n} \|\lambda_i \mathbf{e}_i\|^2,$$

by Lemma 3.12,

$$= \sum_{i=m+1}^{n} \lambda_i^2 \|\mathbf{e}_i\|^2 = \sum_{i=m+1}^{n} \lambda_i^2 = \sigma_n - \sigma_m$$

where $\sigma_n = \sum_{i=1}^{n} \lambda_i^2$ are the partial sums of the series $\sum_{i=1}^{\infty} \lambda_i^2$.
Since we assume this sum converges, its partial sums form a Cauchy sequence. Therefore we have that S_n, $n = 1, 2, \ldots$ forms a Cauchy sequence and we are done. □

To summarise, we know that if $\mathbf{v}$ is any vector in H then

$$\sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle^2$$

converges by Bessel's inequality. Therefore we have just shown that

$$\sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i$$

also converges. Is this equal to $\mathbf{v}$? Not necessarily.

Example 3.17. *In ℓ^2 consider the elements*

$$\mathbf{e}_i = (0, 0, \ldots, 0, 1, 0, \ldots), \qquad i \geq 2$$

were the 1 *appears in the ith position.*

These form an orthonormal sequence. But the element of ℓ^2

$$\mathbf{v} = (2, 0, 0, \ldots)$$

cannot be written as a linear combination of the elements $\mathbf{e}_2, \mathbf{e}_3, \ldots$.

The vector $\mathbf{v}$ is orthogonal to **each** of the elements:

$$\langle \mathbf{v}, \mathbf{e}_i \rangle = 0, \qquad i = 2, 3, \ldots$$

This turns out to be the crux of the problem. So to solve it we have the following definition.

Definition 3.18. Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ be an orthonormal sequence in a Hilbert space H. It is **complete** if the only vectors orthogonal to **all** the $\mathbf{e}_i$s is the zero vector.

We can now give the following result that brings this all together.

Theorem 3.19

Let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ be a complete orthonormal sequence in a Hilbert space H. Then E is a Schauder basis for H.
Furthermore, if $\mathbf{v} \in H$ then

$$\mathbf{v} = \sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i, \qquad \text{and} \qquad \|\mathbf{v}\|^2 = \sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle^2.$$

Proof. Let

$$\mathbf{u} = \mathbf{v} - \sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i.$$

Then

$$\begin{aligned}
\langle \mathbf{u}, \mathbf{e}_j \rangle &= \langle \mathbf{v}, \mathbf{e}_j \rangle - \left\langle \sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i, \mathbf{e}_j \right\rangle \\
&= \langle \mathbf{v}, \mathbf{e}_j \rangle - \sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle \langle \mathbf{e}_i, \mathbf{e}_j \rangle \\
&= \langle \mathbf{v}, \mathbf{e}_j \rangle - \langle \mathbf{v}, \mathbf{e}_j \rangle \\
&= 0.
\end{aligned}$$

So $\mathbf{u}$ is orthogonal to all the $\mathbf{e}_j$s. Hence $\mathbf{u} = \mathbf{0}$, since E is complete, and

$$\mathbf{v} = \sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i.$$

We can use this to determine the norm of $\mathbf{v}$.

First by Theorem 3.12,

$$\left\| \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i \right\|^2 = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{e}_i \rangle^2.$$

Since E is **orthonormal**. Therefore,

$$\begin{aligned} \sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle^2 &= \lim_{n\to\infty} \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{e}_i \rangle^2 \\ &= \lim_{n\to\infty} \left\| \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i \right\|^2 \\ &= \left\| \sum_{i=1}^{\infty} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i \right\|^2. \end{aligned}$$

Where we have used the continuity of the norm here. □

The preceding result shows that if H is a Hilbert space with complete orthonormal sequence

$$E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$$

then

$$H = \left\{ \sum_{k=1}^{\infty} \lambda_k \mathbf{e}_k : (\lambda_k) \in \ell^2 \right\}.$$

Furthermore,

$$\left\| \sum_{k=1}^{\infty} \lambda_k \mathbf{e}_k \right\|^2 = \sum_{k=1}^{\infty} |\lambda_k|^2.$$

This means that H contains a Schauder basis. The converse of this statement is also true: if a Hilbert space has a Schauder basis then it has a complete orthonormal sequence. To prove this we construct an orthonormal sequence from the Schauder basis via the Gram-Schmidt orthgonalisation process which is the subject of Problem 3.1.

Problems

3.1. **(Gram-Schmidt orthgonalisation)** Let H be a Hilbert space and suppose $\{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$ is a linearly independent set in H. Define a sequence $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ as follows.

$$\begin{aligned}
\mathbf{e}_1 &= \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 \\
\mathbf{f}_k &= \mathbf{v}_k - \sum_{j=1}^{k-1} \langle \mathbf{v}_k, \mathbf{e}_j \rangle \mathbf{e}_j, \qquad\qquad k = 2, 3, \ldots \\
\mathbf{e}_k &= \frac{1}{\|\mathbf{f}_k\|}\mathbf{f}_k
\end{aligned} \tag{3.3}$$

(a) Use induction to show that E is an orthonormal sequence

(b) Prove that $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, for all n.

(c) Deduce that the closure of $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$ is equal to the subspace

$$\left\{ \sum_{k=1}^{\infty} \lambda_k \mathbf{e}_k : (\lambda_k) \in \ell^2 \right\}.$$

3.2. **(Lagrange polynomials)** In $L^2[-1,1]$ the polynomials $1, x, x^2, \ldots$ form a linearly independent set. Use the Gram-Schmidt orthogonalisation process to show that the first three elements of the corresponding orthonormal sequence are

$$\mathbf{e}_1(x) = \frac{1}{\sqrt{2}}, \quad \mathbf{e}_2(x) = \sqrt{\frac{2}{3}}x, \quad \mathbf{e}_3(x) = x^2 - \frac{1}{3}.$$

3.3. A **separable** Hilbert space is one with a countable, complete orthonormal sequence. Show that between **any** two separable Hilbert spaces H and K there is a bijection.

3.4. Use your answer to the previous question to deduce that if $K \subset H$ where H is a separable Hilbert space then either K is finite dimensional or there is a bijection between K and H.

3.5. **(ℓ^∞ is not separable)**

(a) Let $A \subset \mathbb{N}$. Define the sequence $\mathbf{1}_A = (a_n)$ where

$$a_n = \begin{cases} 1, & n \in A; \\ 0, & n \notin A. \end{cases}$$

Given $\|(a_n) - (b_n)\|_\infty = \sup_n |a_n - b_n|$ show that if $A, B \subset \mathbb{N}$ and $A \neq B$ then

$$\|\mathbf{1}_A - \mathbf{1}_B\|_\infty = 1.$$

(b) Deduce that there are **uncountably** many elements of ℓ^∞ of distance 1 apart.

(c) Deduce that ℓ^∞ is **not** separable.

Chapter 4

Linear Functionals and Duality

4.1 Linear functionals

Definition 4.1 – Linear functional

A linear functional on a normed vector space $(V, \| \ \|)$ is a function $F \colon V \to \mathbb{R}$ such that for all $\mathbf{u}, \mathbf{v} \in V$, and $\lambda \in \mathbb{R}$,

$$\begin{aligned} F(\mathbf{u} + \mathbf{v}) &= F(\mathbf{u}) + F(\mathbf{v}) \\ F(\lambda \mathbf{u}) &= \lambda F(\mathbf{u}) \end{aligned} \tag{4.1}$$

Example 4.2. *On the space* $(C[0, 1], \| \cdot \|_\infty)$*, the following is a linear functional:*

$$F(f) = \int_0^1 f(t)\mathrm{d}t.$$

In fact this example can be generalised.

Example 4.3. *On the space* $(C[0, 1], \| \cdot \|_\infty)$*, if* $0 \leq a < b \leq 1$*, the following is a linear functional:*

$$F(f) = \int_a^b f(t)\mathrm{d}t.$$

We will work with sequence spaces mostly in this chapter. The following examples build some of the main ideas. The first of these is an example of a **projection**.

Example 4.4. *On* ℓ^p *define*

$$F((a_i)_{i=1}^\infty) = a_1.$$

The next two examples illustrate a more general way of defining linear functionals on ℓ^1 that we will generalise.

Example 4.5. *On ℓ^1 define*

$$G((a_i)_{i=1}^{\infty}) = \sum_{i=1}^{\infty} (-1)^i a_i.$$

Example 4.6 (Important example!). *Let $(b_i)_{i=1}^{\infty} \in \ell^{\infty}$ (the bounded sequences). Then*

$$H((a_i)_{i=1}^{\infty}) = \sum_{i=1}^{\infty} a_i b_i$$

defines a linear functional on ℓ^1.

We should check the definition makes sense.
If $(a_i)_{i=1}^{\infty} \in \ell^1$ then $\sum_{i=1}^{\infty} |a_i|$ converges.
Let $S_n = \sum_{i=1}^{n} |a_n b_n|$ and recall

$$\|(b_i)_{i=1}^{\infty}\|_{\infty} = \sup_i |b_i|.$$

Then

$$\begin{aligned} S_n &= \sum_{i=1}^{n} |a_i|\,|b_i| \\ &\leq \sum_{i=1}^{n} |a_i| \max_{i=1,\dots,n} |b_i| \\ &\leq \|(b_i)_{i=1}^{\infty}\|_{\infty} \sum_{i=1}^{n} |a_i| \\ &\leq \|(b_i)_{i=1}^{\infty}\|_{\infty} \|(a_i)_{i=1}^{\infty}\|_1. \end{aligned}$$

However S_n is an increasing sequence in n (since each term is non-negative). We have shown that it is also bounded above, so that the limit exists and hence

$$\sum_{i=1}^{\infty} |a_i b_i|$$

converges. This means this series is absolutely convergent and hence is convergent and

$$H((a_i)_{i=1}^{\infty}) = \sum_{i=1}^{\infty} a_i b_i$$

is well-defined. You can check that this is also a linear functional from the axioms (4.1).

Note: it follows from the definition of a linear functional that $F(\mathbf{0}) = 0$, since $\mathbf{0} = \mathbf{v} - \mathbf{v}$ so that

$$F(\mathbf{0}) = F(\mathbf{v} - \mathbf{v}) = F(\mathbf{v}) - F(\mathbf{v}) = 0.$$

Question: how many linear functionals are there? What form do they take — i.e. what do they look like?

For the finite dimensional vector space we know what the answer to this last question is. What is it?

4.2 Continuous linear functionals

Consider an arbitrary linear functional $F: V \to \mathbb{R}$ where $(V, \| \ \|)$ is a normed vector space.

When is F a continuous function? If $\mathbf{v}$ is some element of V then F is continuous at $\mathbf{v}$ if and only if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|\mathbf{u} - \mathbf{v}\| < \delta \qquad \Rightarrow \qquad |F(\mathbf{u}) - F(\mathbf{v})| < \epsilon.$$

Equivalently we may use the limit definition:

$$\mathbf{v}_n \to \mathbf{v} \qquad \Rightarrow \qquad F(\mathbf{v}_n) \to F(\mathbf{v})$$

as $n \to \infty$.

With the limit definition we may exploit the linearity of F as follows. Define $\mathbf{u}_n = \mathbf{v}_n - \mathbf{v}$ then

$$\begin{aligned} F(\mathbf{u}_n) &= F(\mathbf{v}_n - \mathbf{v}) = F(\mathbf{v}_n) - F(\mathbf{v}) \\ &\to 0 = F(\mathbf{0}), \end{aligned}$$

as $n \to \infty$. It follows that if a linear functional is continuous at some $\mathbf{v}$ then it is continuous at $\mathbf{0}$. The converse is also true (check!). We have thus proved part of the following.

Theorem 4.7 – Continuous linear functionals

Let $(V, \| \ \|)$ be a normed vector space and let F be a linear functional on V. The following statements are equivalent:

1. F is continuous at $\mathbf{v}$ for some $\mathbf{v} \in V$;
2. F is continuous at $\mathbf{v}$ for all $\mathbf{v} \in V$;
3. F is continuous at $\mathbf{0}$;
4. There is a real constant $A > 0$ such that for all $\mathbf{u} \in V$

$$|F(\mathbf{u})| \leq A\|\mathbf{u}\|. \tag{4.2}$$

Proof. We only need to prove 3 $\Leftrightarrow$ 4 since we already showed that 1 $\Leftrightarrow$ 2 $\Leftrightarrow$ 3.

Suppose 3 is true. Since F is continuous at $\mathbf{0}$ (taking $\epsilon = 1$ in the definition), there is a $\delta > 0$ such that

$$\|\mathbf{u}\| \leq \delta \qquad \Rightarrow \qquad |F(\mathbf{u})| \leq 1.$$

Now let $\mathbf{v} \in V$ be an arbitrary element. Then let

$$\mathbf{w} = \frac{\delta}{\|\mathbf{v}\|}\mathbf{v} \in V,$$

and

$$\|\mathbf{w}\| = \left\| \frac{\delta}{\|\mathbf{v}\|}\mathbf{v} \right\| = \delta \frac{\|\mathbf{v}\|}{\|\mathbf{v}\|} = \delta.$$

It follows that

$$|F(\mathbf{w})| \leq 1.$$

Hence

$$\begin{aligned} 1 \geq |F(\mathbf{w})| &= \left| F\left(\frac{\delta}{\|\mathbf{v}\|}\mathbf{v} \right) \right| \\ &= \left| \frac{\delta}{\|\mathbf{v}\|} F(\mathbf{v}) \right| = \frac{\delta}{\|\mathbf{v}\|} |F(\mathbf{v})|. \end{aligned}$$

Therefore,

$$|F(\mathbf{v})| \leq \delta^{-1} \|\mathbf{v}\|$$

and 4 holds with $A = \delta^{-1}$.

Suppose 4 is true. For any $\epsilon > 0$ we choose $\delta = \epsilon/A$. Then

$$\|\mathbf{v} - \mathbf{0}\| < \delta \qquad \Rightarrow \qquad |F(\mathbf{v}) - F(\mathbf{0})| < A\delta = \epsilon.$$

Therefore 3 holds. □

4.3 The dual

A linear functional that satisfies (4.2) is said to be **bounded**. Here boundedness is equivalent to continuity and so we may sometimes interchange these ideas.

Given a bounded linear functional we define its norm to be the smallest possible value for A in (4.2). If F is the linear functional then we will simply denote its norm in the usual way as $\|F\|$.

The norm of a linear functional has a few interpretations that are useful for calculations.

$$\begin{aligned} \|F\| &= \inf\{A : |F(\mathbf{v})| \le A\|\mathbf{v}\|\} \\ &= \sup_{\mathbf{v}\in V} \frac{|F(\mathbf{v})|}{\|\mathbf{v}\|} \\ &= \sup_{\substack{\mathbf{v}\in V \\ \|\mathbf{v}\|=1}} |F(\mathbf{v})| \\ &= \sup_{\substack{\mathbf{v}\in V \\ \|\mathbf{v}\|\le 1}} \frac{|F(\mathbf{v})|}{\|\mathbf{v}\|} \end{aligned} \tag{4.3}$$

Example 4.8. *Remember Example 4.6. Look through the calculation again to see that*

$$\|H\| \le \|(b_i)\|_\infty.$$

What we want to do now is think of the set of linear functionals as a space itself. This last example shows that every sequence in ℓ^∞ gives rise to a linear functional on ℓ^1. But ℓ^∞ is a Banach space itself – so actually the set of linear functionals on ℓ^1 has also some structure there.

Theorem 4.9

Let $(V, \|\ \|)$ be a normed vector space. Then the set of bounded (continuous) linear functionals is also a normed vector space.
It is denoted V^* and called the **dual** of V.

Here we must define what it means to add two linear functionals and to multiply by a scalar (that's what's needed for the vector space part). Let $F, G \in V^*$ and $\lambda \in \mathbb{R}$. Then we define $F + G$ to be the linear functional that satisfies

$$(F + G)(\mathbf{v}) = F(\mathbf{v}) + G(\mathbf{v}).$$

Similarly (λF) is the linear functional that satisfies

$$(\lambda F)(\mathbf{v}) = \lambda F(\mathbf{v}).$$

Proof. We need only prove it's a vector space and that (4.3) defines a norm. Most of it is pretty straightforward so I'll only go through a couple of the ideas.

To prove that V^* is closed under additions suppose that $F, G \in V^*$. Really we should check that $F + G$ is a linear functional but it's easy and not very interesting. We do however need to show that it is bounded. Let $A = \|F\|$ and $B = \|G\|$. Then

$$\begin{aligned} |(F + G)(\mathbf{v})| &= |F(\mathbf{v}) + G(\mathbf{v})| \leq |F(\mathbf{v})| + |G(\mathbf{v})| \\ &\leq A\|\mathbf{v}\| + B\|\mathbf{v}\| = (A + B)\|\mathbf{v}\|. \end{aligned}$$

Therefore $F + G$ is a bounded linear functional as required.

Note also that the calculation above proves that

$$\|F + G\| \leq \|F\| + \|G\|$$

which is the triangle inequality for the norm of linear functionals.

The rest of the axioms are also true and you should check them (no, really you should, even if only to practise). □

What else can be said about the dual – well it turns out quite a lot.

Theorem 4.10

Let $(V, \| \ \|)$ be a normed vector space. Then the dual V^* is a Banach space.

The proof follows the following familiar structure.

Take a Cauchy sequence (F_n) of linear functionals

1. Find a candidate for the limit
2. Show that the sequence converges to this candidate

3. Show that the limit is a bounded linear functional

Proof. Suppose (F_n) is a Cauchy sequence in V^*.

Find a candidate for the limit.

Let $\mathbf{v} \in V$ be arbitrary. Then

$$|F_n(\mathbf{v}) - F_m(\mathbf{v})| \leq \|F_n - F_m\| \cdot \|\mathbf{v}\|.$$

Hence $(F_n(\mathbf{v}))$ is also a Cauchy sequence. But $\mathbb{R}$ is complete so this sequence has a limit and we can define our candidate linear functional as

$$F(\mathbf{v}) = \lim_{n\to\infty} F_n(\mathbf{v}).$$

Show that the sequence converges to this candidate.

Let $\epsilon > 0$ then there is a $N > 0$ so that $n, m > N$ implies $\|F_n - F_m\| < \epsilon$.

Therefore for $n > N$,

$$\begin{aligned} |F_n(\mathbf{v}) - F(\mathbf{v})| &= |F_n(\mathbf{v}) - \lim_{m\to\infty} F_m(\mathbf{v})| \\ &= \lim_{m\to\infty} |F_n(\mathbf{v}) - F_m(\mathbf{v})| \\ &\leq \limsup_{m\to\infty} \|F_n - F_m\| \cdot \|\mathbf{v}\| \leq \epsilon\|\mathbf{v}\|. \end{aligned} \tag{4.4}$$

Thus $\|F_n - F\| \leq \epsilon$ and we have shown $F_n \to F$ as $n \to \infty$.

Show that the limit is a bounded linear functional

To see that F is a linear functional, for $\mathbf{u}, \mathbf{v} \in V$,

$$F(\mathbf{u} + \mathbf{v}) = \lim_{n\to\infty} F_n(\mathbf{u} + \mathbf{v}) = \lim_{n\to\infty} F_n(\mathbf{u}) + F_n(\mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v}).$$

Similarly $F(\lambda\mathbf{v}) = \lambda F(\mathbf{v})$ for $\lambda \in \mathbb{R}$.

To show that F is bounded, note that from (4.4) we have, for $n > N$,

$$\begin{aligned} |F(\mathbf{v})| &\leq |F_n(\mathbf{v})| + |F(\mathbf{v}) - F_n(\mathbf{v})| \\ &\leq \|F_n\| \cdot \|\mathbf{v}\| + \epsilon\|\mathbf{v}\| = (\|F_n\| + \epsilon)\|\mathbf{v}\|. \end{aligned}$$

Therefore F is a bounded linear functional as required. □

4.4 A concrete example: the sequence spaces

We want an example that corroborates the theorem above. We turn to the sequence spaces ℓ^p.

Theorem 4.11. *Suppose* $1 < p < \infty$. *Then* $F \in (\ell^p)^*$ *if and only if*

$$F((a_i)) = \sum_{i=1}^{\infty} a_i b_i, \tag{4.5}$$

for some $(b_i) \in \ell^q$ *where* $\frac{1}{p} + \frac{1}{q} = 1$. *Furthermore in this case we have* $\|F\| = \|(b_i)\|_p$.

Remember that p and q are related as in Hölder's inequality, (2.3).

Proof. First suppose F is defined by (4.5). By Hölder's inequality

$$\sum_{i=1}^{n} |a_i b_i| \leq \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |b_i|^q\right)^{1/q} \leq \|(a_i)\|_p \,\|(b_i)\|_q. \tag{4.6}$$

Since $\sum_{i=1}^{n} |a_i b_i|$ is an increasing sequence and bounded above, it converges. Therefore this is an absolutely convergent series and

$$\sum_{i=1}^{\infty} a_i b_i$$

is well-defined. It follows that F is a bounded linear functional, that is $F \in (\ell^p)^*$.

Conversely suppose $F \in (\ell^p)^*$. Recall that the sequences

$$\mathbf{e}_n = (0, 0, \ldots, 0, 1, 0, \ldots),$$

where the 1 is in the nth position, forms a Schauder basis for ℓ^p. This means that every element of ℓ^p can be written as

$$\sum_{i=1}^{\infty} a_i \mathbf{e}_i.$$

We define $b_i = F(\mathbf{e}_i)$.

First since $(b_i) \in \ell^q$ if and only if $(|b_i|) \in \ell^q$, and $\|(b_i)\|_q = \|(|b_i|)\|_q$, we can assume $b_i \geq 0$ for all i.

Now if $(b_i) \notin \ell^q$ then $\sum_{i=1}^{\infty} b_i^q$ diverges. However since the b_is are non-negative this means that for any $M > 0$ there is a $n > 0$ such that

$$\sum_{i=1}^{n} b_i^q = u > M.$$

Define

$$\mathbf{a} = b_1^{q/p}\mathbf{e}_1 + b_2^{q/p}\mathbf{e}_2 + \cdots + b_n^{q/p}\mathbf{e}_n = \sum_{i=1}^{n} b_i^{q/p}\mathbf{e}_i.$$

Then

$$\|\mathbf{a}\|_p = \left(\sum_{i=1}^{n}(b_i^{q/p})^p\right)^{1/p} = \left(\sum_{i=1}^{n} b_i^q\right)^{1/p} = u^{1/p}.$$

On the other hand,

$$|F(\mathbf{a})| = \sum_{i=1}^{n} b_i^{q/p} b_i = \sum_{i=1}^{n} b_i^{q/p+1} = \sum_{i=1}^{n} b_i^q = u.$$

Therefore

$$\frac{|F(\mathbf{a})|}{\|\mathbf{a}\|_p} = u^{1-1/p} = u^{1/q} > M^{1/q}.$$

But M was arbitrary and so this shows that F is not a bounded linear functional which contradicts our assumption. Hence $(b_i) \in \ell^q$ and we have that

$$F((a_i)) = \sum_{i=1}^{\infty} a_i b_i,$$

as required.

Finally to show that $\|F\| = \|(b_i)\|_p$ note first that by (4.6)

$$\left|\sum_{i=1}^{\infty} a_i b_i\right| \leq \sum_{i=1}^{\infty} |a_i b_i| \leq \|(a_i)\|_p \, \|(b_i)\|_q$$

so that $\|F\| \leq \|(b_i)\|_p$.

Conversely, as before we may, without loss of generality, assume $b_i \geq 0$ for all i. Define $a_i = b_i^{q/p}$. Then

$$\begin{aligned}|F((a_i))| &= \left|\sum_{i=1}^{\infty} a_i b_i\right| = \left|\sum_{i=1}^{\infty} b_i^{q/p+1}\right| \\ &= \left|\sum_{i=1}^{\infty} b_i^q\right| = \|(b_i)\|_q^q.\end{aligned}$$

However,

$$\|(a_i)\|_p = \left(\sum_{i=1}^{\infty} b_i^q\right)^{1/p} = \|(b_i)\|_q^{q/p}.$$

Hence,

$$\frac{|F((a_i))|}{\|(a_i)\|_p} = \frac{\|(b_i)\|_q^q}{\|(b_i)\|_q^{q/p}}$$
$$= \|(b_i)\|_q^{q-q/p} = \|(b_i)\|_q,$$

so that $\|F\| \geq \|(b_i)\|_p$ as required. □

This theorem says that there is a bijective mapping from $(\ell^p)^*$ to ℓ^q. This mapping preserves the linearity of the dual space and so we say these two Banach spaces are **isomorphic** and write this as

$$(\ell^p)^* \cong \ell^q.$$

We will make this more formal later on

Note: It is possible to show that $(\ell^1)^* \cong \ell^\infty$. But it is not true that the dual of ℓ^∞ is isomorphic to ℓ^1. The dual of ℓ^2 is isomorphic to ℓ^2 itself.

4.5 Linear functionals on Hilbert space

A Hilbert space has an inner product, and

$$\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle.$$

Example 4.12. *On ℓ^2 define*

$$\langle (a_i), (b_i) \rangle = \sum_{i=1}^{\infty} a_i b_i.$$

This defines the inner product.

By Theorem 4.11 with $p = 2$ we have that every linear functional on ℓ^2 can be written as

$$F((a_i)) = \langle (a_i), (b_i) \rangle$$

for some $(b_i) \in \ell^2$. This turns out to be true for all Hilbert spaces H and sets up an isomorphism $H^* \cong H$.

Definition 4.13. Let $(V, \|\ \|)$ be a normed vector space and let $F \in V^*$. The **kernel** of F is the closed subspace

$$\ker F = \{\mathbf{v} \in V : F(\mathbf{v}) = 0\}.$$

That $\ker F$ is closed does need some work but it follows from the continuity of F.

Definition 4.14. Let H be a Hilbert space. If E is a closed subspace of H then the **orthogonal complement** of E is the set

$$E^{\perp} = \{\mathbf{v} \in H : \langle \mathbf{v}, \mathbf{u} \rangle = 0, \ \forall \mathbf{u} \in E\}.$$

Definition 4.15. Suppose E and F are subspaces of a normed vector space.

- $E + F = \{\mathbf{u} + \mathbf{v} : \mathbf{u} \in E, \ \mathbf{v} \in F\}$,
- If in addition $E \cap F = \{\mathbf{0}\}$ then we call this the **direct sum** of E and F and write it as
$$E \oplus F.$$

Theorem 4.16

If E is a closed subspace of a Hilbert space H then

$$H = E \oplus E^{\perp}.$$

This theorem fails dramatically when we aren't in the Hilbert space setting. Given a closed subspace of a Banach space, for example, there needn't be any other closed subspace that allows us to write the Banach space as a direct sum.

Proof. I'll leave you to show that $E \cap E^{\perp} = \{\mathbf{0}\}$. We will only prove that for any $\mathbf{v} \in H$ we can write this as

$$\mathbf{v} = \mathbf{u} + \mathbf{u}^{\perp}, \qquad \mathbf{u} \in E, \ \mathbf{u}^{\perp} \in E^{\perp}.$$

Recall the closest point property for Hilbert space, Lemma 3.9.

Define $\mathbf{u}$ to be the closest point of E to $\mathbf{v}$, in particular $\mathbf{u} \in E$. Let $\mathbf{u}^{\perp} = \mathbf{v} - \mathbf{u}$ then we must show that $\mathbf{u}^{\perp} \in E^{\perp}$.

Suppose it is not. Then there is a $\mathbf{w} \in E$ with $\langle \mathbf{u}^{\perp}, \mathbf{w} \rangle \neq 0$. In fact by replacing $\mathbf{w}$ with $-\mathbf{w}$ if necessary we may assume

$$\langle \mathbf{u}^{\perp}, \mathbf{w} \rangle > 0.$$

Since $\mathbf{u}$ is the closest point in E to $\mathbf{v}$ we have that for all $t > 0$

$$\begin{aligned}\|\mathbf{u}^\perp\|^2 &= \|\mathbf{v}-\mathbf{u}\|^2 \\ &\le \|\mathbf{v}-(\mathbf{u}+t\mathbf{w})\|^2 \\ &= \|\mathbf{v}-\mathbf{u}\|^2 + t^2\|\mathbf{w}\|^2 - 2t\langle \mathbf{v}-\mathbf{u},\mathbf{w}\rangle \\ &= \|\mathbf{u}^\perp\|^2 + t^2\|\mathbf{w}\|^2 - 2t\langle \mathbf{u}^\perp,\mathbf{w}\rangle.\end{aligned}$$

Rearranging we get

$$t \ge \frac{\langle \mathbf{u}^\perp,\mathbf{w}\rangle}{\|\mathbf{w}\|^2} > 0.$$

If we choose $t < \dfrac{\langle \mathbf{u}^\perp,\mathbf{w}\rangle}{\|\mathbf{w}\|^2}$ we have a contradiction. Hence there is no such $\mathbf{w}$ and we are done. □

A consequence of this proof is the following.

Corollary 4.17. *Suppose E is a closed subspace of a Hilbert Space H. Let $\mathbf{v} = \mathbf{u} + \mathbf{u}^\perp$ be the decomposition of $\mathbf{v} \in H$ from Theorem 4.16. Then the distance from $\mathbf{v}$ to E is $\|\mathbf{u}^\perp\|$, i.e.*

$$\inf_{\mathbf{w}\in E}\|\mathbf{v}-\mathbf{w}\| = \|\mathbf{v}-\mathbf{u}\| = \|\mathbf{u}^\perp\|.$$

In order to prove that $H^* \cong H$ we must have some sort of candidate for the linear functionals in H^*. We define, for $\mathbf{w} \in H$ the linear functional

$$\theta_{\mathbf{w}}(\mathbf{v}) = \langle \mathbf{v},\mathbf{w}\rangle.$$

For a fixed $\mathbf{w}$, $\theta_{\mathbf{w}}$ is a linear functional. Furthermore by the Cauchy-Schwartz inequality,

$$|\theta_{\mathbf{w}}(\mathbf{v})| = |\langle \mathbf{v},\mathbf{w}\rangle| \le \|\mathbf{v}\|\cdot\|\mathbf{w}\|.$$

Hence $\|\theta_{\mathbf{w}}\| \le \|\mathbf{w}\|$. In fact since $|\theta_{\mathbf{w}}(\mathbf{w})| = \|\mathbf{w}\|^2$ it follows that

$$\|\theta_{\mathbf{w}}\| = \|\mathbf{w}\|.$$

These are our *candidate* linear functionals.

Theorem 4.18 – The Riesz Representation Theorem

Let H be a Hilbert space and $F \in H^*$. Then there is a $\mathbf{w} \in H$ such that

$$F = \theta_{\mathbf{w}}.$$

Proof. If $F = 0$ then $F = \theta_{\mathbf{0}}$.

If $F \neq 0$ then we let $E = \ker F$. Since $F \neq 0$, $E \neq H$. Therefore $E^{\perp} \neq \{\mathbf{0}\}$ by Theorem 4.16 and $\dim E^{\perp} \geq 1$.

We claim that also $\dim E^{\perp} \leq 1$.

Suppose not. Then there are two linearly independent vectors in $E^{\perp}$, call them $\mathbf{a}$ and $\mathbf{b}$. Since they are linearly independent

$$\lambda\mathbf{a} + \mu\mathbf{b} \neq \mathbf{0}$$

unless $\lambda = \mu = 0$.

Setting $\mathbf{v} = F(\mathbf{b})\mathbf{a} - F(\mathbf{a})\mathbf{b}$ it follows, since $F(\mathbf{a}) \neq 0$ and $F(\mathbf{b}) \neq 0$, that $\mathbf{v} \neq \mathbf{0}$.

However,

$$F(\mathbf{v}) = F(\mathbf{b})F(\mathbf{a}) - F(\mathbf{a})F(\mathbf{b}) = 0$$

so $\mathbf{v} \in \ker F = E$.

Hence, by Theorem 4.16, $\mathbf{v} \notin E^{\perp}$. This contradicts the assumption that these two vectors were linear independent elements in $E^{\perp}$. Hence $\dim E^{\perp} \leq 1$ as required.

We have thusfar shown that $E^{\perp}$ is a 1-dimensional subspace of H. We assume that it is spanned by the single element $\mathbf{w}$, and by multiplying by a scalar we will normalise so that

$$F(\mathbf{w}) = \|\mathbf{w}\|^2.$$

Now any $\mathbf{v} \in H$ can be written in the form $\mathbf{v} = \mathbf{u} + \rho\mathbf{w}$ for some $\rho \in \mathbb{R}$ and $\mathbf{u} \in \ker F$.

Then

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u} + \rho\mathbf{w}, \mathbf{w} \rangle = \rho\|\mathbf{w}\|^2.$$

Whereas,

$$F(\mathbf{v}) = F(\mathbf{u} + \rho\mathbf{w}) = \rho F(\mathbf{w}) = \rho\|\mathbf{w}\|^2.$$

It follows that

$$F(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} \rangle = \theta_{\mathbf{w}}(\mathbf{v}).$$

□

Corollary 4.19. *The map $\Theta \colon \mathbf{w} \mapsto \theta_{\mathbf{w}}$ defines an isomorphism between H and H^*. Furthermore $\|\Theta(\mathbf{w})\| = \|\mathbf{w}\|$.*

The last property means we call Θ an **isometric isomorphism**.

Note: Since $(\ell^p)^* \cong \ell^q$, this result is unique to Hilbert space.

4.6 The Hahn-Banach Theorem

Suppose $(V, \| \ \|)$ is a normed vector space, and M is a **subspace** of V. We would like to know more about the linear functionals on M and, importantly, how they relate to the linear functionals on V.

First note that if $F \in V^*$ then we can simply restrict this to vectors in M and obtain a linear functional on M. We write this as

$$F|_M.$$

Notice that

$$\|F|_M\| \leq \|F\|.$$

In this section we will prove the following.

Theorem 4.20 – The Hahn-Banach Theorem

Let $(V, \| \ \|)$ be a normed vector space and let M be a subspace of V. Suppose $F \in M^*$. Then there is a $G \in V^*$ such that

$$G|_M = F, \qquad \text{and} \qquad \|G\| = \|F\|.$$

The linear functional G here should be thought of as an *extension* of F to the whole vector space V.

That the extension exists is not the point here – actually you can extend it quite easily. Clearly we must have $\|G\| \geq \|F\|$, so the big deal here is that you can extend F *without* increasing its norm.

We will prove this in the case that V is separable. When it is not the proof requires *Zorn's Lemma* which is an extension of the normal set theory. The proof can be found in any of the textbooks available if you want to read it.

Our proof will consist of the following stepwise construction.

1. Extend F to $M_1 = M + \operatorname{span}\{\mathbf{v}_1\}$
2. Use induction to extend F to $M_n = M + \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$
3. Finally extend F to the closure of $\bigcup M_n$

Lemma 4.21. *Let M be a subspace of a normed vector space $(V, \| \ \|)$ and suppose $\mathbf{v}_1 \notin M$. Define $M_1 = M + \operatorname{span}\{\mathbf{v}_1\}$.*
If $F \in M^*$ then there is a $G \in M_1^*$ such that

$$G|_M = F, \qquad \text{and} \qquad \|G\| = \|F\|.$$

Proof. First if $F = 0$ then we can simply let $G = 0$ and we are done. We can therefore assume $F \neq 0$.

Since every vector $\mathbf{w} \in M_1$ is of the form $\mathbf{w} = \lambda\mathbf{v}_1 + \mathbf{z}$ where $\mathbf{z} \in M$, we define G in the only way possible:

$$G(\mathbf{w}) = G(\lambda\mathbf{v}_1 + \mathbf{z}) = \lambda G(\mathbf{v}_1) + F(\mathbf{z}).$$

The only part of this that we can define ourselves is $G(\mathbf{v}_1) = \gamma$. So the result will follow if we can choose γ so that $\|G\| = \|F\|$.

From now on we assume $\|F\| = 1$ just to make the following calculation easier to follow.

We need to choose γ so that $|G(\mathbf{w})| \leq \|\mathbf{w}\|$ in the above definition. In other words we need

$$-\|\lambda\mathbf{v}_1 + \mathbf{z}\| \leq |\lambda\gamma + F(\mathbf{z})| \leq \|\lambda\mathbf{v}_1 + \mathbf{z}\|, \qquad \forall\, \mathbf{z} \in M.$$

For $\lambda \neq 0$ these inequalities become

$$-\left\|\mathbf{v}_1 + \frac{1}{\lambda}\mathbf{z}\right\| - F\left(\frac{1}{\lambda}\mathbf{z}\right) \leq \gamma \leq \left\|\mathbf{v}_1 + \frac{1}{\lambda}\mathbf{z}\right\| - F\left(\frac{1}{\lambda}\mathbf{z}\right).$$

Since M is a subspace we can simply replace $\frac{1}{\lambda}\mathbf{z}$ with $\mathbf{z}$ without loss of generality. Then we need to show that we can choose γ so that whenever $\mathbf{z} \in M$

$$-\|\mathbf{v}_1 + \mathbf{z}\| - F(\mathbf{z}) \leq \gamma \leq \|\mathbf{v}_1 + \mathbf{z}\| - F(\mathbf{z}).$$

It turns out more is true. Suppose $\mathbf{z}_1, \mathbf{z}_2 \in M$. Then

$$\begin{aligned} F(\mathbf{z}_1) - F(\mathbf{z}_2) = F(\mathbf{z}_1 - \mathbf{z}_2) &\leq \|\mathbf{z}_1 - \mathbf{z}_2\| \\ &= \|\mathbf{z}_1 + \mathbf{v}_1 - (\mathbf{z}_2 + \mathbf{v}_1)\| \\ &\leq \|\mathbf{v}_1 + \mathbf{z}_1\| + \|\mathbf{v}_1 + \mathbf{z}_2\|. \end{aligned}$$

Rearranging we get, for all $\mathbf{z}_1, \mathbf{z}_2 \in M$,

$$-\|\mathbf{v}_1 + \mathbf{z}_2\| - F(\mathbf{z}_2) \leq \|\mathbf{v}_1 + \mathbf{z}_1\| - F(\mathbf{z}_1).$$

In particular we can take the supremum of the left-hand side and the infinum of the right-hand side and we have

$$C_- = \sup_{\mathbf{z}_2 \in M} -\|\mathbf{v}_1 + \mathbf{z}_2\| - F(\mathbf{z}_2) \leq \inf_{\mathbf{z}_1 \in M} \|\mathbf{v}_1 + \mathbf{z}_1\| - F(\mathbf{z}_1) = C_+.$$

It follows that there is a $\gamma \in [C_-, C_+]$. If we choose this γ for our extension G then the lemma is proved. □

We will also need the following lemma which you should be able to prove yourselves.

Lemma 4.22. *Suppose M is a subspace of the normed vector space $(V, \| \ \|)$ and let $F \in M^*$. Then we can extend F to the closure $\overline{M}$ without increasing its norm.*

Proof of the Hahn-Banach Theorem. We assume that V is separable and so

$$V = \overline{M + \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots\}}.$$

Let $M_n = V + \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Then we may inductively extend F to M_n without increasing its norm by Lemma 4.21.

If there are countably many vectors $\mathbf{e}_k$ then we can extend F to $F_n \in M_n^*$ indefinitely. This means that we have extended F to $\bigcup M_n$. To complete the proof we must use Lemma 4.22 to extend this to

$$\overline{\bigcup M_n} = \overline{M + \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots\}} = V.$$

□

There are a number of consequences to this theorem that are important. The proof of the first shows how we can use this result to construct linear functionals with certain properties.

Corollary 4.23. *Let $\mathbf{v}_0 \in V$. Then there is a linear functional $F \in V^*$ with*

$$\|F\| = 1, \qquad \textit{and} \qquad F(\mathbf{v}_0) = \|\mathbf{v}_0\|.$$

Proof. Let $M = \operatorname{span}\{\mathbf{v}_0\}$. Define F as

$$F(\mathbf{v}_0) = \|\mathbf{v}_0\|.$$

Then

$$|F(\lambda \mathbf{v}_0)| = |\lambda| |F(\mathbf{v}_0)| = |\lambda| \|\mathbf{v}_0\| = \|\lambda \mathbf{v}_0\|.$$

So $\|F\| = 1$. Now simply extend F to the whole of V by the Hahn-Banach theorem. □

Corollary 4.24. *Let* $\mathbf{v} \in V$. *Then*

$$\|\mathbf{v}\| = \sup\{|F(\mathbf{v})| : F \in V^*, \ \|F\| \le 1\}.$$

Proof. Let $\alpha = \sup\{|F(\mathbf{v})| : F \in V^*, \ \|F\| \le 1\}$.

Since $|F(\mathbf{v})| \le \|\mathbf{v}\|$ whenever $\|F\| \le 1$, $\alpha \le 1$.

On the other hand, let F be the linear functional of Corollary 4.23. Then $\alpha \ge |F(\mathbf{v})| = \|\mathbf{v}\|$. □

4.7 Quotient spaces

A vector space is a group under vector addition. Given a *closed* subspace M we define the Quotient Space V/M to be the set of **cosets**

$$\mathbf{v} + M$$

under the usual equivalence relation

$$\mathbf{v} + M \sim \mathbf{u} + M \qquad \Longleftrightarrow \qquad \mathbf{v} - \mathbf{u} \in M.$$

If M is closed we can define a norm on V/M.

$$\|\mathbf{v} + M\| = \inf_{\mathbf{u} \in M} \|\mathbf{v} - \mathbf{u}\|.$$

In this section we will characterise M^* for a **closed** subspace M of $(V, \| \ \|)$.

Definition 4.25. Suppose M is a closed subspace of a normed vector space $(V, \| \ \|)$. The **anihilator** of M is the subspace $M^\perp$ of V^* defined by

$$M^\perp = \{F \in V^* : \forall \mathbf{v} \in M, \ F(\mathbf{v}) = 0\}.$$

Note that $M^\perp$ is a closed subspace of the Banach space V^*. So we can define the quotient space

$$V^*/M^\perp$$

to be the cosets $F + M^\perp$, for $F \in V^*$.

We have two isomorphism theorems concerning linear functionals on M^*. The first is the following consequence of the Hahn-Banach theorem.

Theorem 4.26

Let M be a closed subspace of a normed vector space $(V, \| \ \|)$. Then

$$M^* \cong V^*/M^\perp$$

under the isometric isomorphism

$$\begin{aligned} \rho : V^*/M^\perp &\to V^*/M^\perp \\ F + M^\perp &\mapsto F|_M \end{aligned}.$$

To prove this we simply need to check the following.

1. ρ is linear
2. ρ is injective
3. ρ is surjective
4. $\|\rho(F + M^\perp)\| = \|F + M^\perp\|$

Throughout the proof we will use the fact that if $G \in M^\perp$ then

$$(F + G)|_M = F|_M.$$

Proof. 1. ρ **is linear**

Let $F, G \in V^*$. Then
$$(F + G)|_M = F|_M + G|_M.$$
And this remains true if F and G are replaced by any linear functionals in $F + M^\perp$ or $G + M^\perp$.

2. ρ **is injective**

It is enough to show that $\rho(F + M^\perp) = 0$ if and only if $F = 0$, since ρ is linear.

Suppose $\rho(F + M^\perp) = 0$. Then $F|_M = 0$. In particular $F \sim 0$ as required.

3. ρ **is surjective**

Let $F \in M^*$. Then by the Hahn-Banach theorem there is a $G \in V^*$ such that $G|_M = F$. So that
$$\rho(G + M^\perp) = G|_M = F$$
and ρ is surjective.

4. $\|\rho(F + M^\perp)\| = \|F + M^\perp\|$

   Let $F \in V^*$ and $G \in M^\perp$. Then

$$\|F|_M\| = \|(F - G)|_M\| \le \|F - G\|.$$

   Taking the infinmum over all $G \in M^\perp$ we get

$$\|F|_M\| \le \|F + M^\perp\|.$$

   Conversely, suppose $F \in M^*$. Then by the Hahn-Banach theorem there is a $G \in V^*$ such that

$$G|_M = F, \qquad \text{and} \qquad \|G\| = \|F\|.$$

   Hence $F = \rho(G + M^\perp)$ and

$$\begin{aligned}\|F\| &= \|G\| \\ &= \|G - 0\| \ge \|G + M^\perp\|.\end{aligned}$$

□

The second theorem is as follows. We omit the proof.

Theorem 4.27

Let M be a closed subspace of a normed vector space $(V, \|\ \|)$. Then

$$M^\perp \cong (V/M)^*.$$

4.8 Reflexivity

We have seen a number of different relationships between Banach spaces and their duals. We have seen that the dual of ℓ^p is isomorphic to ℓ^q when $p > 1$ and $1/p + 1/q = 1$. We have also seen that the dual of a Hilbert space is isomorphic to itself.

Since the dual of any normed vector space is a Banach space, we can define the dual of the dual of a normed vector space. We write this as V^{**} and call it the **second dual** of V.

Theorem 4.28 – Canonical embedding of V in V^{**}

Let $\mathbf{v} \in V$. The mapping $\hat{\mathbf{v}} \colon V^* \to \mathbb{R}$ defined by

$$\hat{\mathbf{v}}(F) = F(\mathbf{v})$$

is a bounded linear functional on V^*, i.e. $\hat{\mathbf{v}} \in V^{**}$, and $\|\hat{\mathbf{v}}\| = \|\mathbf{v}\|$.

Proof. First we check that $\hat{\mathbf{v}}$ is linear. Let $F, G \in V^*$. Then

$$\hat{\mathbf{v}}(F+G) = (F+G)(\mathbf{v}) = F(\mathbf{v}) + G(\mathbf{v}) = \hat{\mathbf{v}}(F) + \hat{\mathbf{v}}(G).$$

Similarly $\hat{\mathbf{v}}(\lambda F) = \lambda \hat{\mathbf{v}}(F)$.

To prove $\|\hat{\mathbf{v}}\| = \|\mathbf{v}\|$ note that for all $F \in V^*$,

$$|\hat{\mathbf{v}}(F)| = |F(\mathbf{v})| \leq \|F\|\,\|\mathbf{v}\|.$$

Hence

$$\frac{|\hat{\mathbf{v}}(F)|}{\|F\|} \leq \|\mathbf{v}\| \qquad \Rightarrow \qquad \|\hat{\mathbf{v}}\| \leq \|\mathbf{v}\|.$$

The reverse inequality follows from Corollary 4.23. With the linear functional F of that corollary we have that

$$|\hat{\mathbf{v}}(F)| = |F(\mathbf{v})| = \|F\|\,\|\mathbf{v}\|, \qquad \text{since } \|F\| = 1.$$

Hence $\|\hat{\mathbf{v}}\| \geq \|\mathbf{v}\|$ are required. □

This mapping is called a **canonical mapping**. The following is immediate.

Corollary 4.29. *For any normed vector space V, the canonical mapping $\mathbf{v} \mapsto \hat{\mathbf{v}}$ is an isometric isomorphism of V onto a* ***subspace*** *of V^{**}.*

You should think of this as saying that $V \subset V^{**}$.

The second dual of ℓ^1 is c_0, so the inclusion above may be strict.

Note however that in the case of ℓ^p for $1 < p < \infty$ the second dual of ℓ^p is isomorphic to ℓ^p. Furthermore for any Hilbert space the second dual is isomorphic to itself by the Riesz representation theorem.

It turns out that this is a property that distinguishes different types of Banach space.

Definition 4.30 – Reflexivity

A Banach space $(V, \|\cdot\|)$ is said to be **reflexive** if V^{**} is isomorphic to V.

Problems

4.1. Show that if F and G are linear functionals then so is $F + G$ and λF for $\lambda \in \mathbb{R}$.

4.2. Let $\alpha \in C^1[0,1]$ (i.e. α is a differentiable function on $[0,1]$ with **continuous** derivative α'). Define on $C[0,1]$

$$F_\alpha : f \mapsto \int_0^1 f(t)\alpha'(t)\mathrm{d}t.$$

Show that F_α is a bounded linear functional and find its norm.

4.3. Generalise your answer to the previous question to the Banach space $C([a,b])$ where $b > a$.

4.4. Prove that $E \cap E^\perp = \{\mathbf{0}\}$ for E a closed subspace of a Hilbert space H.

4.5. Show that if E is a closed subspace of the Hilbert space H then each $\mathbf{v} \in H$ can be written **uniquely** in the form $\mathbf{u} + \mathbf{u}^\perp$ where $\mathbf{u} \in E$ and $\mathbf{u}^\perp \in E^\perp$.

4.6. In $\mathbb{R}^4$ find the distance from the vector $(1,1,1,1)^T$ to the two dimensional subspace spanned by $(1,0,0,1)$ and $(0,2,0,0)$ in the 2-norm.

4.7. Use your solution to Problem 2 in Chapter 3 to find

$$\min_{a,b,c\in\mathbb{R}} \int_{-1}^1 |x^3 - a - bx - cx^2|^2 \mathrm{d}x.$$

4.8. Let H be a Hilbert space and suppose $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ is a complete orthonormal sequence in H. Define on H the linear functional

$$F_n\left(\sum_{k=1}^{\infty} \lambda_k \mathbf{e}_k\right) = \lambda_n.$$

Show that $\|F_n\| = 1$ and find the vector $\mathbf{v} \in H$ such that $F_n(\mathbf{u}) = \langle \mathbf{u}, \mathbf{v} \rangle$ for all $\mathbf{u} \in H$.

4.9. Show that the definition of $E^\perp$ for a subspace of a Hilbert space H corresponds to the **anihillator** of E defined in the Banach space setting.

4.10. **(Linear functionals on ℓ^∞ are not of the form (4.5))** Let c be the subspace of ℓ^∞ consisting of sequences that converge.

(a) Defne F as $F((a_n)) = \lim_{n\to\infty} a_n$. Show that F is a bounded linear functional on c

(b) Show that $\mathbf{e}_n \in c$ for all n and $F(\mathbf{e}_n) = 0$.

(c) Let G be the extension of F to ℓ^∞ guaranteed by the Hahn-Banach theorem. Show that if G has the form (4.5) then $G = 0$.

Chapter 5

Linear operators

5.1 Bounded linear operators

Definition 5.1 – Linear operator

Let U and V be normed vector spaces. A function $T\colon U \to V$ is called a **linear operator** if for all $\mathbf{u}, \mathbf{v} \in U$ and $\lambda \in \mathbb{R}$ the following hold.

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$$
$$T(\lambda \mathbf{u}) = \lambda T(\mathbf{u})$$

Suppose we take a function $f \in C[0,1]$. Here $C[0,1]$ is the Banach space of continuous functions on $[0,1]$ with uniform norm,

$$\|f\| = \sup_{t\in[0,1]} |f(t)|.$$

If we integrate this function as follows then we get another continuous function.

$$g(x) = \int_0^x f(t)\mathrm{d}t.$$

This is an **operator** on $C[0,1]$ since it maps continuous functions to continuous functions. Let's write this as

$$I(f)(x) = \int_0^x f(t)\mathrm{d}t.$$

Then we recognise from elementary calculus the following properties:

$$\int_0^x f(t) + g(t)\mathrm{d}t = \int_0^x f(t)\mathrm{d}t + \int_0^x g(t)\mathrm{d}t$$
$$\int_0^x \lambda f(t)\mathrm{d}t = \lambda \int_0^x f(t)\mathrm{d}t.$$

Or,

$$I(f + g) = I(f) + I(g)$$
$$I(\lambda f) = \lambda I(f).$$

So integration is an example of a linear operator on $C[0,1]$.

Example 5.2. *Since $\mathbb{R}$ itself is a normed vector space, every linear functional is an example of a linear operator.*

Example 5.3. *On ℓ^p, $1 \le p \le \infty$, define*

$$B((a_1, a_2, a_3, \dots)) = (a_2, a_3, \dots).$$

This is a linear operator.

This is an important linear operator called the *Backward shift operator*.

There is also a *forward shift operator*. It is also a linear operator.

Example 5.4. *On ℓ^p, $1 \le p \le \infty$, define*

$$S((a_1, a_2, a_3, \dots,)) = (0, a_1, a_2, \dots).$$

As with linear functionals we can define a norm on linear operators.

Definition 5.5 – The operator norm

Let $(U, \|\ \|_U)$ and $(V, \|\ \|_V)$ be normed vector spaces. Suppose T is a linear operator from U to V. Then the **operator norm** is defined as

$$\|T\| = \sup\left\{\frac{\|T(\mathbf{u})\|_V}{\|\mathbf{u}\|_U} : \mathbf{u} \in U\right\}.$$

If a linear operator has a finite norm then we call it a **bounded linear operator** from U to V.

Example 5.6. *The forward shift operator on ℓ^p, $1 \le p \le \infty$, was described above. You should justify to yourself that*

$$\|S((a_n))\|_p = \|(a_n)\|_p$$

for all $(a_n) \in \ell^p$. Therefore

$$\|S\| = 1.$$

The calculation of the norm in the previous example is unusually easy. We normally use the following strategy.

To find the norm of a linear operator there are two steps:

1. Find an upper bound in general so that $\|T(\mathbf{v})\| \le M\|\mathbf{v}\|$ for all $\mathbf{v}$
2. Give an example of an element $\mathbf{u}$ such that $\|T(\mathbf{u})\| \ge M\|\mathbf{u}\|$

Example 5.7. *The operator norm for the Backward shift is also 1 on ℓ^p, $1 \le p < \infty$.*

To prove this statement we start with the observation,

$$\|B((a_1, a_2, a_3, \ldots))\|_p = \left(\sum_{i=2}^{\infty} |a_i|^p\right)^{1/p} = S_0^{1/p}$$

$$\|(a_1, a_2, a_3, \ldots)\|_p = \left(\sum_{i=1}^{\infty} |a_i|^p\right)^{1/p} = (S_0 + |a_1|^p)^{1/p}.$$

Therefore

$$\|B\| = \sup \frac{\|B((a_n))\|_p}{\|(a_n)\|_p} = \sup\left(\frac{S_0}{S_0 + |a_1|^p}\right)^{1/p} \le 1.$$

On the other hand define

$$\mathbf{a} = (0, 1, 1/2, 1/4, 1/8, \ldots).$$

Then

$$\|B(\mathbf{a})\|_p = \|\mathbf{a}\|_p,$$

and it follows that

$$\|B\| = \sup \frac{\|B((a_n))\|_p}{\|(a_n)\|_p} \ge \frac{\|B(\mathbf{a})\|_p}{\|\mathbf{a}\|_p} = 1.$$

Hence $\|B\| = 1$.

Example 5.8. *If we consider $U = \mathbb{R}^n$ and $V = \mathbb{R}^m$. Let A be a $m \times n$ matrix. Then we can define the linear operator*

$$T(\mathbf{u}) = A\mathbf{u}, \qquad \mathbf{u} \in \mathbb{R}^n.$$

This is an important example. Every linear operator from one finite dimensional vector space to another arises as a matrix (with respect to some basis). Furthermore for finite dimensions every linear operator is bounded. So the study of linear operators should be thought of as a generalisation of the study of matrices in Linear Algebra.

Definition 5.9

Let $(U, \| \ \|_U)$ and $(V, \| \ \|_V)$ be normed vector spaces. The set of **bounded linear operators** from U to V is denoted by $B(U, V)$.
If $U = V$ this is shortened to $B(U)$.

Note: The dual U^* can be written as $B(U, \mathbb{R})$ (although shouldn't be).

Some of the results for linear functionals have corresponding results for bounded linear operators.

- If $(V, \| \ \|_V)$ is a Banach space then $B(U, V)$ is a Banach space.
- A linear operator T is continuous if and only if it is bounded.
- A linear operator T is continuous at all $\mathbf{v}$ if and only if it is continuous at $\mathbf{0}$.
- The operator norm has a number of other useful equivalent formulations, for example,

$$\|T\| = \sup_{\|\mathbf{v}\|=1} \|T(\mathbf{v})\|.$$

If T and S are two operators then TS is an operator defined by $(TS)(\mathbf{u}) = T(S(\mathbf{u}))$ for all $\mathbf{u}$. In other words it's the operator you obtain by composing T and S. In addition to the results above we have the following.

Lemma 5.10. *For two bounded linear operators T and S*

$$\|TS\| \leq \|T\| \, \|S\|.$$

5.2 The adjoint operator

Suppose $T: U \to V$ is a bounded linear operator and take any $F \in V^*$ a bounded linear functional. Then

$$\mathbf{u} \mapsto F(T(\mathbf{u}))$$

is a bounded linear functional on U. That is $F \circ T \in U^*$.

The linearity follows from the linearity of F and T, and the boundedness follows since for each $\mathbf{u} \in U$,

$$\begin{aligned} |F \circ T(\mathbf{u})| &= |F(T(\mathbf{u}))| \\ &\leq \|F\| \cdot \|T(\mathbf{u})\| \leq \|F\| \cdot \|T\| \cdot \|\mathbf{u}\|. \end{aligned}$$

Now consider F varying through all the linear functionals in V^*. The corresponding mapping

$$F \mapsto F \circ T$$

maps V^* to U^*.

We call this the adjoint of T. As long as it exists this will be a useful tool.

Theorem 5.11 – The adjoint operator

Suppose $(U, \| \ \|_U)$ and $(V, \| \ \|_V)$ are Banach spaces. Let $T \in B(U, V)$. Then there is a **unique** operator $S \in B(V^*, U^*)$ such that

$$\text{for all } F \in V^*, \qquad S \circ F = F \circ T.$$

Definition 5.12. The operator S above is called the **adjoint** of T and is written T^*.

We will prove this only in the following simplified case:

- H is a Hilbert space
- $T \in B(H)$

Proof of Theorem 5.11. Note first that by the Riesz representation theorem we need to prove that there is an operator $S \in B(H)$ such that for all $\mathbf{u}, \mathbf{v} \in H$,

$$\langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, S(\mathbf{v}) \rangle.$$

Fix $\mathbf{v}$. Let $F_{\mathbf{v}}(\mathbf{u}) = \langle T(\mathbf{u}), \mathbf{v}\rangle$. Then $F_{\mathbf{v}}$ is a linear functional on H. and

$$\|F_{\mathbf{v}}\| \leq \|T(\mathbf{u})\| \cdot \|\mathbf{v}\| \leq \|T\| \cdot \|\mathbf{u}\| \cdot \|\mathbf{v}\|.$$

Hence $F_{\mathbf{v}} \in H^*$. By the Riesz representation theorem this means that there is a $\mathbf{w} \in H$ such that

$$F_{\mathbf{v}} = \theta_{\mathbf{w}}.$$

Or,

$$\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle.$$

We *define* S to be the operator such that $S(\mathbf{v}) = \mathbf{w}$.

It is now simply a matter of checking S is a bounded linear operator. First,

$$\begin{aligned}\langle \mathbf{u}, S(\mathbf{v}+\mathbf{z})\rangle &= \langle T(\mathbf{u}), \mathbf{v}+\mathbf{z}\rangle = \langle T(\mathbf{u}), \mathbf{v}\rangle + \langle T(\mathbf{u}), \mathbf{z}\rangle \\ &= \langle \mathbf{u}, S(\mathbf{v})\rangle + \langle \mathbf{u}, S(\mathbf{z})\rangle.\end{aligned}$$

Since this is true for all $\mathbf{u} \in H$ we have that

$$S(\mathbf{v}+\mathbf{z}) = S(\mathbf{v}) + S(\mathbf{z})$$

as required.

Checking that $S(\lambda\mathbf{v}) = \lambda S(\mathbf{v})$ is similar. To see that S is bounded, suppose $\mathbf{u}$ satisfies $\|\mathbf{u}\| = 1$. Then

$$\begin{aligned}\|S(\mathbf{u})\|^2 &= \langle S(\mathbf{u}), S(\mathbf{u})\rangle \\ &= \langle TS(\mathbf{u}), \mathbf{u}\rangle \\ &\leq \|TS(\mathbf{u})\|\ \|\mathbf{u}\| = \|TS(\mathbf{u})\|\end{aligned}$$

by the Cauchy-Schwarz inequality, and the fact that $\|\mathbf{u}\| = 1$,

$$\leq \|T\|\ \|S(\mathbf{u})\|.$$

Therefore

$$\|S(\mathbf{u})\| \leq \|T\|.$$

Taking the supremum over all vectors $\mathbf{u}$ with $\|\mathbf{u}\| = 1$ we get

$$\|S\| = \sup_{\|\mathbf{u}\|=1} \|S(\mathbf{u})\| \leq \|T\|,$$

and hence S is a bounded linear operator.

The uniqueness of the operator S follows from the uniqueness in the Hahn-Banach theorem. □

We state the following without proof (the proofs can be found in the textbook).

In general finding an exact description of the adjoint of a given linear operator is difficult. However there are a number of examples where we can do just that.

Example 5.13. *On ℓ^2 the adjoint of the forward shift operator is the backward shift operator, $S^* = B$.*

If $E \subset H$ is a closed subset of a Hilbert space H recall that every element $\mathbf{v} \in H$ can be written as

$$\mathbf{v} = \mathbf{u} + \mathbf{u}^\perp, \qquad \mathbf{u} \in E, \mathbf{u}^\perp \in E^\perp.$$

A **orthogonal projection** onto E is the linear operator $P_E : H \to H$ such that, with the notation above,

$$P_E(\mathbf{v}) = \mathbf{u}.$$

Example 5.14. *In any Hilbert space H an orthogonal projection is* ***self-adjoint****: $P_E^* = P_E$.*

5.3 The algebra of bounded linear operators on a Hilbert space

In the Hilbert space setting the space $B(H)$ has some interesting properties worth looking at.

First note if H is a Hilbert space $H^* \cong H$, and

$$T \in B(H) \qquad \Rightarrow \qquad T^* \in B(H).$$

Moreover for all $\mathbf{u}, \mathbf{v} \in H$,

$$\langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, T^*(\mathbf{v}) \rangle.$$

This follows from the Riesz representation theorem.

Theorem 5.15

Suppose H is a Hilbert space. Let $T, S \in B(H)$.

1. $TS \in B(H)$ and $\|TS\| \leq \|T\| \cdot \|S\|$
2. $(TS)^* = S^*T^*$
3. $(T^*)^* = T$
4. $\|T\| = \|T^*\|$
5. $\|T^*T\| = \|T\|^2$

Proof. 1. You should try to prove this yourself.

2. Let $\mathbf{u}, \mathbf{v} \in H$. Then
$$\langle (TS)\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, (TS)^*\mathbf{v}\rangle.$$
Whereas
$$\begin{aligned}\langle TS\mathbf{u}, \mathbf{v}\rangle &= \langle T(S(\mathbf{u})), \mathbf{v}\rangle \\ &= \langle S(\mathbf{u}), T^*(\mathbf{v})\rangle \\ &= \langle \mathbf{u}, S^*T^*(\mathbf{v})\rangle.\end{aligned}$$
Since this is true for all $\mathbf{u}, \mathbf{v} \in H$ we have that $(TS)^* = T^*S^*$.

3. Again if $\mathbf{u}, \mathbf{v} \in H$ then first,
$$\langle T^*(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{v}, T^*(\mathbf{u})\rangle = \langle T(\mathbf{v}), \mathbf{u}\rangle = \langle \mathbf{u}, T(\mathbf{v})\rangle.$$
And second
$$\langle T^*(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, (T^*)^*(\mathbf{v})\rangle.$$
Therefore $(T^*)^* = T$.

4. and 5. Let $\mathbf{u} \in H$ and $\|\mathbf{u}\| \leq 1$. Then
$$\begin{aligned}\|T(\mathbf{u})\|^2 = \langle T(\mathbf{u}), T(\mathbf{u})\rangle &= \langle \mathbf{u}, T^*T(\mathbf{u})\rangle \\ &\leq \|\mathbf{u}\| \cdot \|T^*T(\mathbf{u})\| \\ &\leq \|T^*T\| \leq \|T^*\| \cdot \|T\|.\end{aligned}$$

Hence,

$$\|T\|^2 \leq \|T^*T\| \leq \|T^*\| \cdot \|T\| \tag{5.1}$$

Dividing by $\|T\|$ we get

$$\|T\| \leq \|T^*\|.$$

However if we replace T with T^* in the above inequality we get the opposite:

$$\|T^*\| \leq \|(T^*)^*\| = \|T\|$$

So $\|T\| = \|T^*\|$ and going back to (5.1) we now have equality throughout:

$$\|T\|^2 = \|T^*T\| = \|T^*\| \cdot \|T\|$$

as required.

□

Note: We won't pursue this but the properties in this theorem give $B(H)$ an algebraic structure called a C^*-algebra. The mathematical field of **operator theory** is a study of C^*-algebras.

5.4 Invertible linear operators

Suppose U and V are normed vector spaces. A bounded linear operator $T \in B(U, V)$ is said to be **invertible** if there is a **bounded** linear operator $S \in B(V, U)$ such that

$$TS = I_V, \qquad ST = I_U$$

where I_U and I_V are the identity operators on U and V respectively, i.e. $I_U(\mathbf{u}) = \mathbf{u}$ for all $\mathbf{u} \in U$ and similarly for V. We normally write the inverse of T as T^{-1}.

Definition 5.16

We let $\mathbf{Inv}(U, V)$ be the set of invertible bounded linear operators in $B(U, V)$. If $U = V$ then this is called the **general linear group** of U and is written $GL(U)$.

Note: We won't make use of the notation $GL(U)$ again.

Theorem 5.17

Suppose $T \in \mathbf{Inv}(U, V)$ and $S \in \mathbf{Inv}(V, W)$. Then $ST \in \mathbf{Inv}(U, W)$ and

$$(ST)^{-1} = T^{-1}S^{-1}.$$

The proof is just algebra and Il won't be given here.

What about adjoints?

Theorem 5.18

If $T \in \mathbf{Inv}(U, V)$ then $T^* \in \mathbf{Inv}(V^*, U^*)$ and

$$(T^*)^{-1} = (T^{-1})^*.$$

Again, the proof is just some algebra and so I'll leave it.

5.5 Isomorphisms

Previously we have shown that a Hilbert space H is isomorphic to its dual H^*. This is the content of the Riesz representation theorem. With the concepts discussed in this chapter we can make the notion of an isomorphism formal.

Definition 5.19 – Isomorphisms

Let $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ be normed vector spaces. An **isomorphism** is a linear operator $T \in \mathbf{Inv}(U, V)$. If there is an isomorphism between U and V then they are said to be **isomorphic**.
If, in addition, for all $\mathbf{u} \in U$,

$$\|T(\mathbf{u})\|_V = \|\mathbf{u}\|_U,$$

we say that U and V are **isometrically isomorphic**, and T is an **isometry**.

5.6 The open mapping theorem

Recall that since bounded linear operators are continuous functions they map compact sets to compact sets, but not necesarily open sets to open sets or closed sets to closed sets.

However it is a consequence of the linearity that a bounded linear operator from one Banach space to another that is surjective does preserve open sets. A function that does this (preserves open sets) is called an **open mapping**. The following theorem is one of the fundamental theorems of Functional Analysis. It has a number of important consequences.

Theorem 5.20 – The open mapping theorem

Suppose U and V are Banach spaces and $T \in B(U, V)$. If T is surjective then T is an open mapping. That is for any open set $O \subset U$, $T(O)$ is an open set in V.

Here, of course,

$$T(O) = \{T(\mathbf{u}) : \mathbf{u} \in O\}.$$

The proof of the open mapping theorem relies on a result called the Baire Category Theorem. Although it is not that difficult to prove once you have learned about this theorem, it would mean we would have to spend quite some time discussing minutiae of topology and I would rather focus on Functional Analysis instead. It could be an interesting project if you want though.

We will give two important results as consequences of the open mapping theorem.

Theorem 5.21 – The inverse mapping theorem

Suppose U and V are Banach spaces. If $T \in B(U, V)$ is a bijection, then $T \in \mathbf{Inv}(U, V)$.

The proof relies on the theory of metric spaces. In particular the fundamental result that a function $f : X \to Y$ is continuous if and only if whenever $A \subset Y$ is open, $f^{-1}(A)$ is open.

Proof. Since T is bijective we may define an inverse mapping $T^{-1} : V \to U$. We only need to prove that $T^{-1} \in B(V, U)$.

However if $O \subset U$ is an open set then

$$(T^{-1})^{-1}(O) = T(O)$$

is open in V by the open mapping theorem. Therefore by the result mentioned T^{-1} is continuous. Hence T^{-1} is bounded and the result is proved. □

This result justifies the use of the term **invertible** for linear operators since it coincides with the usual notion of invertible functions.

5.7 The closed graph theorem

The second consequence of the open mapping theorem is the closed graph theorem.

Before stating the theorem we must understand some of the ideas. Suppose $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ are Banach spaces. Then we define the normed vector space

$$U \otimes V = \{(\mathbf{u}, \mathbf{v}) : \mathbf{u} \in U, \mathbf{v} \in V\}, \qquad \|(\mathbf{u}, \mathbf{v})\|_{U \otimes V} = \|\mathbf{u}\|_U + \|\mathbf{v}\|_V.$$

This is also a Banach space. Indeed $(\mathbf{u}_n, \mathbf{v}_n) \to (\mathbf{u}, \mathbf{v})$ as $n \to \infty$ if and only if both $\mathbf{u}_n \to \mathbf{u}$ and $\mathbf{v}_n \to \mathbf{v}$ in their respective Banach spaces.

A closed set in $E \subset U \otimes V$ is a set that includes all its limit points. If $(\mathbf{u}_n, \mathbf{v}_n) \to (\mathbf{u}, \mathbf{v})$ and $(\mathbf{u}_n, \mathbf{v}_n) \in E$ for all n, then $(\mathbf{u}, \mathbf{v}) \in E$.

The graph of a function $f : \mathbb{R} \to \mathbb{R}$ is the set

$$G_f = \{(x, f(x)) : x \in \mathbb{R}\}.$$

This is the proper definition of the line you draw in school when you are asked to plot a function.

Definition 5.22 – The graph of an operator

Suppose U and V are Banach spaces and T is a linear operator from U to V. The **graph** of T is the set

$$\textbf{graph}\, T = \{(\mathbf{u}, T(\mathbf{u})) : \mathbf{u} \in U\}.$$

The graph of T is a subset of $U \otimes V$.

Theorem 5.23 – The closed graph theorem

Suppose U and V are Banach spaces. Let $T : U \to V$ be a linear operator (not necessarily bounded).

If **graph** T is closed then T is a bounded linear operator.

Proof. Let $G = \textbf{graph}\, T$. Since G is a closed subset of the Banach space $U \otimes V$, it is itself a Banach space.

Let $P : G \to U$ be defined by $P((\mathbf{u}, T(\mathbf{u})) = \mathbf{u}$. Then P is a bounded linear operator. To see it is bounded note that

$$\|P((\mathbf{u}, T(\mathbf{u})))\| = \|\mathbf{u}\|$$
$$\|(\mathbf{u}, T(\mathbf{u}))\| = \|\mathbf{u}\| + \|T(\mathbf{u})\|.$$

So that,

$$\frac{\|P((\mathbf{u}, T(\mathbf{u})))\|}{\|(\mathbf{u}, T(\mathbf{u}))\|} = \frac{\|\mathbf{u}\|}{\|\mathbf{u}\| + \|T(\mathbf{u})\|} \leq 1.$$

Moreover, P is bijective. (You should check this).

It follows by the inverse mapping theorem that P has a bounded inverse P^{-1}.

Now define $Q: G \to V$ by $Q((\mathbf{u}, T(\mathbf{u}))) = T(\mathbf{u})$. Then again this is a bounded linear functional. (Check)

However note that $T = QP^{-1}$. Since these two linear operators are bounded, so is T. □

Note: The condition that **graph** T is closed is equivalent to the following statement which is easier to work with.

$$\text{If } \mathbf{u}_n \to \mathbf{u} \text{ and } T(\mathbf{u}_n) \to \mathbf{v} \text{ as } n \to \infty \text{ then } T(\mathbf{u}) = \mathbf{v}.$$

5.8 The principle of uniform boundedness

Suppose we have a sequence of linear operators $T_n \in B(U, V)$, $n = 1, 2, \ldots$, for some normed spaces U and V. We say these are **uniformly bounded** if there is a constant $K > 0$ such that

$$\|T_n\| \leq K, \qquad \text{for all } n.$$

For any uniformly bounded sequence of linear operators we have the fact that for any $\mathbf{u} \in U$,

$$\|T_n(\mathbf{u})\| \leq K\|\mathbf{u}\|.$$

So that $T_n(\mathbf{u})$ is also uniformly bounded in this sense. This can be useful when deciding if $T_n(\mathbf{u})$ converges in any sense at all. The fact that $T_n(\mathbf{u})$ is uniformly bounded is seemingly a weaker statement than T_n being uniformly bounded.

It turns out however that if U is a Banach space then these two statements are equivalent.

Theorem 5.24 – The principle of uniform boundedness

Let $(U, \|\cdot\|)$ be a Banach space and $(V, \|\cdot\|)$ a normed vector space. Suppose $T_n \in B(U, V)$, $n = 1, 2, \ldots,$. If for each $\mathbf{u} \in U$

$$\sup \|T_n(\mathbf{u})\| < \infty,$$

then

$$\sup \|T_n\| < \infty,$$

that is T_n, $n = 1, 2, \ldots$, is uniformly bounded.

The proof we will work through here was proved by Alan Sokal, see Sokal (2011). The history of the proof is somewhat interesting, Hahn proved it using an idea similar the one presented here but over the years shorter proofs that used more sophisticated techniques emerged. This proof uses only the ideas we have gone through and some ingenuity.

We define the ball

$$B(\mathbf{u}, r) = \{\mathbf{v} : \|\mathbf{v} - \mathbf{u}\| < r\}.$$

We will need the following geometric result.

Lemma 5.25. *Let $T \in B(U, V)$ where U and V are normed vector spaces. Then for any $\mathbf{u} \in U$*

$$\sup_{\mathbf{v} \in B(\mathbf{u}, r)} \|T(\mathbf{v})\| \geq \|T\| r.$$

Proof. Let $\mathbf{w} \in U$. Since

$$\frac{1}{2}(\|T(\mathbf{u} + \mathbf{w})\| + \|T(\mathbf{u} - \mathbf{w})\|)$$

is the average of $\{\|T(\mathbf{u} + \mathbf{w})\|, \|T(\mathbf{u} - \mathbf{w})\|\}$ we have that

$$\frac{1}{2}(\|T(\mathbf{u} + \mathbf{w})\| + \|T(\mathbf{u} - \mathbf{w})\|) \leq \max\{\|T(\mathbf{u} + \mathbf{w})\|, \|T(\mathbf{u} - \mathbf{w})\|\}.$$

Now,

$$\begin{aligned} 2\|T(\mathbf{w})\| &= \|2T(\mathbf{w})\| = \|T(\mathbf{u} + \mathbf{w}) - T(\mathbf{u} - \mathbf{w})\| \\ &\leq \|T(\mathbf{u} + \mathbf{w})\| + \|T(\mathbf{u} - \mathbf{w})\|. \end{aligned}$$

Combining these we get the estimate

$$\|T(\mathbf{w})\| \leq \max\{\|T(\mathbf{u} + \mathbf{w})\|, \|T(\mathbf{u} - \mathbf{w})\|\} \tag{5.2}$$

which holds for all $\mathbf{w}, \mathbf{u} \in U$. Note that

$$\|T\| = \sup_{\mathbf{w}' \in B(\mathbf{0},1)} \|T(\mathbf{w}')\|.$$

Therefore if $r > 0$, taking the supremum of the left-hand side of (5.2) over $B(\mathbf{0}, r)$ and writing $\mathbf{w} = r\mathbf{w}'$, $\mathbf{w}' \in B(\mathbf{0}, 1)$,

$$\sup_{\mathbf{w} \in B(\mathbf{0},r)} \|T(\mathbf{w})\| = \sup_{\mathbf{w}' \in B(\mathbf{0},1)} \|T(r\mathbf{w}')\| = r \sup_{\mathbf{w}' \in B(\mathbf{0},1)} \|T(\mathbf{w}')\| = r\|T\|.$$

For the right-hand side we again take the supremum,

$$\begin{aligned} \sup_{\mathbf{w} \in B(\mathbf{0},r)} \max\{\|T(\mathbf{u}+\mathbf{w})\|, \|T(\mathbf{u}-\mathbf{w})\|\} &= \sup_{\mathbf{w} \in B(\mathbf{0},r)} \|T(\mathbf{u}+\mathbf{w})\| \\ &= \sup_{\mathbf{w} \in B(\mathbf{u},r)} \|T(\mathbf{w})\|. \end{aligned}$$

The result follows. □

We will use Lemma 5.25 to construct a sequence. Note that if

$$\sup_{\mathbf{v} \in B(\mathbf{u},r)} \|T(\mathbf{v})\| \geq \|T\| r,$$

this means that we can always find a $\mathbf{v} \in B(\mathbf{u}, r)$ so that

$$\|T(\mathbf{v})\| > (1-\epsilon)\|T\| r$$

for any $0 < \epsilon < 1$. In the proof we simply take $\epsilon = 1/3$.

Proof of the principle of uniform boundedness. We prove the contrapositive of the statement: if $\sup \|T_n\| = \infty$ then there is a $\mathbf{u} \in U$ such that $\sup \|T_n(\mathbf{u})\| = \infty$.

If $\sup \|T_n\| = \infty$ then if needed we can switch to a subsequence and assume

$$\|T_n\| \geq 4^n. \tag{5.3}$$

We will define a sequence $(\mathbf{u}_n)_{n=0}^{\infty}$ as follows. First let $\mathbf{u}_0 = \mathbf{0}$. By Lemma 5.25 for any $r > 0$ we can find $\mathbf{v}$ such that $\mathbf{v} \in B(\mathbf{u}_0, r)$ and

$$\|T_1(\mathbf{v})\| \geq \frac{2}{3}\|T_1\| r.$$

In particular if $r = 1/3$ we let $\mathbf{u}_1 \in U$ be such that $\mathbf{u}_1 \in B(\mathbf{u}_0, 1/3)$ and

$$\|T_1(\mathbf{u}_1)\| \geq \frac{2}{3}\|T_1\| 3^{-1}.$$

Now pick $\mathbf{u}_2 \in B(\mathbf{u}, 1/3^2)$ such that

$$\|T_2(\mathbf{u}_2)\| \geq \frac{2}{3}\|T_2\|3^{-2}.$$

We can continue in this way picking $\mathbf{u}_n$ so that $\mathbf{u}_n \in B(\mathbf{u}_{n-1}, 1/3^n)$ and

$$\|T_n(\mathbf{u}_n)\| \geq \frac{2}{3}\|T_n\|3^{-n}.$$

Claim: $(\mathbf{u}_n)$ is a Cauchy sequence. If $m > n > 0$ then

$$\begin{aligned}
\|\mathbf{u}_m - \mathbf{u}_n\| &= \|\mathbf{u}_m - \mathbf{u}_{m-1} + \mathbf{u}_{m-1} - \mathbf{u}_{m-2} + \cdots + \mathbf{u}_{n+1} - \mathbf{u}_n\| \\
&\leq \|\mathbf{u}_m - \mathbf{u}_{m-1}\| + \|\mathbf{u}_{m-1} - \mathbf{u}_{m-2}\| + \cdots + \|\mathbf{u}_{n+1} - \mathbf{u}_n\| \\
&\leq 3^{-m} + 3^{-(m-1)} + \cdots + 3^{-(n+1)} \\
&\leq 3^{-(n+1)} \sum_{k=0}^{m-(n+1)} 3^{-k} \\
&\leq 3^{-(n+1)} \sum_{k=0}^{\infty} 3^{-k} = 3^{-n}\frac{1}{3}\left(\frac{1}{1-1/3}\right) = \frac{1}{2}3^{-n}.
\end{aligned}$$

Therefore $(\mathbf{u}_n)$ is a Cauchy sequence and we may define

$$\mathbf{u} = \lim_{n\to\infty} \mathbf{u}_n.$$

Note also that if we let $m \to \infty$ in the estimate above we get

$$\|\mathbf{u} - \mathbf{u}_n\| \leq \frac{1}{2}3^{-n}.$$

Now bringing this all together

$$\begin{aligned}
\|T_n(\mathbf{u})\| &= \|T_n(\mathbf{u} - \mathbf{u}_n) + T_n(\mathbf{u}_n)\| \\
&\geq \left|\,\|T_n(\mathbf{u}_n)\| - \|T_n(\mathbf{u} - \mathbf{u}_n)\|\,\right| \\
&\geq \frac{2}{3}\|T_n\|3^{-n} - \|T_n\| \cdot \|\mathbf{u} - \mathbf{u}_n\| \\
&\geq \frac{2}{3}3^{-n}\|T_n\| - \frac{1}{2}3^{-n}\|T_n\| \\
&= \frac{1}{6}3^{-n}\|T_n\| = \frac{1}{6}\left(\frac{4}{3}\right)^n.
\end{aligned}$$

It follows that $\|T_n(\mathbf{u})\| \to \infty$ and the result follows. □

Problems

5.1. Prove Lemma 5.10.

5.2. Let $I\colon C[0,1] \to L^1[0,1]$ be defined by

$$I(f)(x) = \int_0^x f(t)\mathrm{d}t.$$

Show that I is a bounded linear operator and that $\|I\| = 1/2$.

5.3. Show that for any bounded linear operator $T \in B(U,V)$, $T(\mathbf{0}_U) = \mathbf{0}_V$, where $\mathbf{0}_U$ and $\mathbf{0}_V$ are the zero vectors in U and V respectively.

5.4. Show that $T \in B(U,V)$ is injective if and only if $\ker T = \{\mathbf{0}\}$.

5.5. Show that if $T, S \in B(U,V)$ then $(T+S)^* = T^* + S^*$.

5.6. Suppose H and K are Hilbert spaces and $T \in B(H,K)$. Prove that the two relations hold:

$$T(H)^\perp = \ker T^*,$$
$$T^*(K)^\perp = \ker T.$$

5.7. Let $T \in B(\ell^p, \ell^p)$ $(1 < p < \infty)$ be defined as

$$T(a_1, a_2, a_3, \ldots) = (a_1, 0, a_2, 0, a_3, 0, \ldots).$$

Show that $\|T(a_n)\|_p = \|(a_n)\|_p$ and find T^*.

5.8. Let U and V be Banach spaces and suppose $T \in B(U,V)$. Assume that $\ker T = \{\mathbf{0}\}$. Prove that $T(U)$ is closed if and only if there is a $\delta > 0$ such that

$$\forall \mathbf{u} \in U, \qquad \|T(\mathbf{u})\| \geq \delta \|\mathbf{u}\|.$$

5.9. **(Banach-Steinhaus Theorem)** Suppose T_n, $n = 1, 2, \ldots$ is a sequence of bounded linear operators from the Banach space $(U, \|\cdot\|)$ to the normed vector space $(V, \|\cdot\|)$ such that for all $\mathbf{u} \in U$, the limit $\lim_{n\to\infty} T_n(\mathbf{u})$ exists. Show that the linear operator

$$T(\mathbf{u}) = \lim_{n\to\infty} T_n(\mathbf{u})$$

is also bounded and $\|T\| \leq \limsup_{n\to\infty} \|T_n\|$.

Bibliography

Banach, S. (1955), *Théorie des opérations linéaires*, Chelsea Publishing Co., New York.

Enflo, P. (1973), 'A counterexample to the approximation problem in Banach spaces', *Acta Math.* **130**, 309–317.

Sokal, A. D. (2011), 'A really simple elementary proof of the uniform boundedness theorem', *Amer. Math. Monthly* **118**(5), 450–452.