Functional Analysis, Part 1

Paul D. Mitchener

September 16, 2013

Contents

Chapter 1

Normed Spaces

Many problems, for example those dealing with the existence and uniqueness of solutions to systems of differential equations, naturally involve dealing with infinite-dimensional vector spaces. Typically the elements of such spaces are *functions*.

We want to be able to do analysis on the vector speces we consider. Therefore we assume that all vector spaces are over either the field, $\mathbb{R}$, of real numbers, or the field, $\mathbb{C}$, of complex numbers. Throughout these notes we write $\mathbb{F}$ to denote either $\mathbb{R}$ or $\mathbb{C}$.

1.1 Examples

Here are a few examples of infinite-dimensional vector spaces.

Example 1.1 *Let S be a metric space. Let $C(S)$ denote the set of continuous functions $f\colon S \to \mathbb{F}$. Then $C(S)$ is a vector space under the operations of* pointwise *addition and scalar multiplication of functions. By this, we mean that addition and scalar multiplication are defined by writing*

$$(f+g)(x) = f(x) + g(x) \qquad f, g \in C(S),\ x \in S,$$

and

$$(\alpha f)(x) = \alpha f(x) \qquad \alpha \in \mathbb{F},\ f \in C(S),\ x \in S,$$

respectively.

The vector space $C(S)$ is infinite-dimensional unless the topological space S is finite.

Example 1.2 *Let $p \geq 1$. Let l^p denote the set of sequences (a_n) in the field $\mathbb{F}$ such that the series*

$$\sum_{n=0}^{\infty} |\alpha_n|^p$$

converges. Then l^p is an infinite-dimensional vector space. The operations of addition and scalar multiplication are defined by writing

$$(a_n) + (b_n) = (a_n + b_n) \qquad (a_n), (b_n) \in l^p$$

and

$$\alpha(a_n) = (\alpha a_n) \qquad \alpha \in \mathbb{F},\ (a_n) \in l^p$$

respectively.

Example 1.3 *Let c_0 be the set of sequences, (a_n), in the field $\mathbb{F}$ such that $a_n \to 0$ as $n \to \infty$. The operations of addition and scalar multiplication are defined as in the above example.*

Example 1.4 *Let $\Omega \subseteq \mathbb{C}$ be a non-empty open set. Define $H(\Omega)$ to be the set of holomorphic functions $f\colon \Omega \to \mathbb{C}$. Then $H(\Omega)$ is an infinite-dimensional subspace of the vector space $C(\Omega)$ defined in example 1.1*

1.2 Norms

In order to perform any meaningful analysis on infinite-dimensional vector spaces it is helpful to have some additional structure. One of the simplest and most useful structures to have available is that of a *norm*.

Definition 1.5 *Let X be a vector space over the field $\mathbb{F}$.*[1] *Then a* norm *on the space X is a function*

$$\| - \| \colon X \to \mathbb{R}^{\geq 0}$$

such that:

- $\|\alpha x\| = |\alpha| \|x\|$ *for all scalars $\alpha \in \mathbb{F}$ and vectors $x \in X$*
- $\|x + y\| \leq \|x\| + \|y\|$ *for all vectors $x, y \in X$*
- *Given a vector $x \in X$, if $\|x\| = 0$ then $x = 0$*

The second of the above axioms is sometimes called the *triangle inequality*.

A vector space equipped with a norm is referred to as a *normed vector space* or merely as a *normed space*. If X is a normed vector space, and $Y \leq X$ is a linear subspace, then Y is also a normed vector space; the norm on the space Y is defined by restriction of the norm on the space X.

The following handy result is a consequence of the triangle inequality; we leave the proof as an exercise.

Proposition 1.6 *Let X be a normed vector space. Let $x, y \in X$. Then:*

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

□

[1] Recall that the symbol $\mathbb{F}$ denotes either the real numbers, $\mathbb{R}$, or the complex numbers, $\mathbb{C}$.

We present an easy proposition as an example of playing with the axioms.

Proposition 1.7 *Let X be a normed vector space. Then there is a metric on the space d defined by the formula*

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

Proof: Certainly the supposed metric, d, is a function $d: X \times X \to \mathbb{R}^{\geq 0}$. We need to check that the axioms of a metric hold.

- Let $x, y \in X$. Then

$$\|x-y\| = \|(-1)(y-x)\| = |-1|\|y-x\| = \|y-x\|$$

  Hence $d(x,y) = d(y,x)$

- Let $x, y, z \in X$. Then

$$\|x-z\| = \|(x-y)+(y-z)\| \leq \|x-y\| + \|y-z\|$$

  Hence $d(x,z) \leq d(x,y) + d(y,z)$

- Suppose that $x, y \in X$ and $d(x,y) = 0$. Then $\|x-y\| = 0$ and so $x = y$.

□

The above metric is called the metric *induced* by the the norm. Because any normed vector space is also a metric space, we can talk about limits in normed vector spaces, and things such as open sets, and functions being continuous. The following is left as an exercise.

Proposition 1.8 *Let V be a normed vector space. Then we have a continuous map $f: V \to \mathbb{R}$ defined by the formula $f(v) = \|v\|$.* □

Definition 1.9 *Let X be a vector space. Let d be a metric on the set X.*

- *The metric d is said to be* translation-invariant *if $d(x+z, y+z) = d(x,y)$ for all vectors $x, y, z \in X$*
- *The metric d is said to be* dilation-invariant *if $d(\alpha x, \alpha y) = |\alpha| d(x,y)$ for all scalars $\alpha \in \mathbb{F}$ and vectors $x, y \in X$*

Proposition 1.10 *Let X be a normed vector space. Then the induced metric is translation-invariant and dilation-invariant.*

Proof: Let d denote the induced metric.

- Let $x, y, z \in X$. Then

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

- Let $\alpha \in \mathbb{F}$, $x, y \in X$. Then

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

□

Some of the examples contained in section 1.1 are normed vector spaces.

Example 1.11 *Let K be a compact metric space. Then there is a norm on the vector space $C(K)$ defined by the formula*

$$\|f\| = \sup\{|f(x)| \mid x \in K\}.$$

We will check that the axioms required for the stated function in the above example to be a norm are true.

- Let $f \in C(K)$ be a function, and let $\alpha \in \mathbb{F}$ be a scalar. Then

$$\begin{aligned} \|\alpha f\| &= \sup\{|\alpha f(x)| \mid x \in K\} \\ &= \sup\{|\alpha||f(x)| \mid x \in K\} \\ &= |\alpha|\|f\| \end{aligned}$$

- Let $f, g \in C(K)$ be functions. Then for each point $x \in K$ we have the triangle inequality

$$|f(x) + g(x)| \leq |f(x)| + |g(x)|$$

Hence

$$\begin{aligned} \|f + g\| &= \sup\{|f(x) + g(x)| \mid x \in K\} \\ &\leq \sup\{|f(x)| + |g(x)| \mid x \in K\} \\ &\leq \sup\{|f(x)| \mid x \in K\} + \sup\{|g(y)| \mid y \in K\} \\ &= \|f\| + \|g\| \end{aligned}$$

- Suppose that $f \in C(K)$ and $\|f\| = 0$. Then

$$\sup\{|f(x)| \mid x \in K\} = 0$$

and so $f(x) = 0$ for all points $x \in K$. Thus $f = 0$.

Example 1.12 *Let l^1 be the set of sequences, (α_n), in the field $\mathbb{F}$ such that $\sum_{n=1}^{\infty} |\alpha_n|$ converges. Then we can define a norm on the vector space l^1 by the formula*

$$\|(\alpha_n)\| = \sum_{n=1}^{\infty} |\alpha_n|.$$

Again, we check the axioms.

- Let $(\alpha_n) \in l^1$, and $\alpha \in \mathbb{F}$. Then

$$\|\alpha(\alpha_n)\| = \sum_{n=1}^{\infty} |\alpha\alpha_n| = |\alpha| \sum_{n=1}^{\infty} |\alpha_n| = |\alpha|\|(\alpha_n)\|.$$

- Let $(\alpha_n), (\beta_n) \in l^1$. Then for each natural number n we have the triangle inequality

$$|\alpha_n + \beta_n| \leq |\alpha_n| + |\beta_n|$$

  Taking sums, we see

$$\|(\alpha_n) + (\beta_n)\| = \sum_{n=1}^{\infty} |\alpha_n + \beta_n| \leq \sum_{n=1}^{\infty} |\alpha_n| + \sum_{n=1}^{\infty} |\beta_n| = \|(\alpha_n)\| + \|(\beta_n)\|.$$

- Suppose that $(\alpha_n) \in l^1$ and $\|(\alpha_1)\| = 0$. Then

$$\sum_{n=1}^{\infty} |\alpha_n| = 0$$

  and so $\alpha_n = 0$ for all points n, since $|\alpha_n| \geq 0$. Thus $(\alpha_n) = 0$.

Example 1.13 *Let $a < b$. Then there is a norm on the space $C[a,b]$ defined by the formula*

$$\|f\| = \int_a^b |f(x)| \, dx$$

As above, the only potentially unclear part of proving that the function $\|-\|$ is a norm is checking the triangle inequality.

To do this, let $f, g \in C[a,b]$. Observe

$$\begin{array}{rcl} \|f+g\| & = & \int_a^b |f(x) + g(x)| \, dx \\ & \leq & \int_a^b |f(x)| + |g(x)| \, dx \\ & = & \int_a^b |f(x)| \, dx + \int_a^b |g(x)| \, dx \\ & = & \|f\| + \|g\| \end{array}$$

and we are done.

1.3 Banach Spaces

The nicest type of normed vector spaces are those which are *complete* as metric spaces. To see what completeness involves we need to look at the notion of convergence in a normed vector space.

Definition 1.14 *Let V be a normed vector space. Then a sequence, (x_n), in the space V* converges in norm *to a* limit *$x \in V$ if for all real numbers $\varepsilon > 0$ there exists a natural number N such that $\|x_n - x\| < \varepsilon$ whenever $n \geq N$.*

We can rephrase the above definition by saying that the sequence (x_n) converges (in norm) to a limit $x \in X$ if

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

The proof of the following result is left as an exercise.

Proposition 1.15 *Let V be a normed vector space. Let (x_n) and (y_n) be sequences in V converging to points x and y respectively. Let $\alpha, \beta \in \mathbb{F}$. Then the sequence $(\alpha x_n + \beta y_n)$ converges to $\alpha x + \beta y$.* □

Definition 1.16 *Let V be a normed vector space. Then a sequence, (x_n), in the space X is called a* Cauchy sequence *if for all real numbers $\varepsilon > 0$ there exists a natural number N such that $\|x_m - x_n\| < \varepsilon$ whenever $m, n \geq N$.*

As a simple example of analysis in the world of normed spaces we present the following result.

Proposition 1.17 *Let (x_n) be a norm-convergent sequence in a normed vector space X. Then the sequence (x_n) is a Cauchy sequence.*

Proof: Let x be the limit of the sequence (x_n). Let $\varepsilon > 0$. Then we can find a natural number N such that if $n \geq N$ then $\|x_n - x\| < \frac{\varepsilon}{2}$. In particular, if $m, n \geq N$ then

$$\|x_m - x_n\| \leq \|x_m - x\| + \|x_n - x\| < \varepsilon$$

□

These notions can be generalised to the world of metric spaces. A sequence (x_n) in a metric space X is said to *converge* to a *limit* $x \in X$ if for all real numbers $\varepsilon > 0$ there exists a natural number N such that $d(x_n, x) < \varepsilon$ whenever $n \geq N$. A sequence (x_n) in a metric space X is said to be a *Cauchy sequence* if for all real numbers $\varepsilon > 0$ there exists a natural number N such that $d(x_m, x_n) < \varepsilon$ whenever $m, n \geq N$.

Proposition 1.18 *Let X be a metric space. Let (x_n) be a Cauchy sequence in the space X, and suppose that the sequence (x_n) has a convergent subsequence. Then the sequence (x_n) converges.* □

The proof of this result is left as an exercise.

Definition 1.19 *A normed vector space is called* complete *if every Cauchy sequence converges in norm to some limit. A* Banach space *is a complete normed vector space.*

Any closed subspace of a Banach space is also a Banach space.

More generally, a metric space is also called *complete* if every Cauchy sequence converges to some limit. For example, one of the defining properties of the field of real numbers, $\mathbb{R}$, is that it is complete. The field of complex numbers, $\mathbb{C}$, is also complete.

Example 1.20 *Let K be a compact metric space. Then the space $C(K)$, with the norm defined in example 1.11, is complete.*

Proof: Let (f_n) be a Cauchy sequence in the space $C(K)$. Let $x \in K$. Then for any natural numbers $m, n \in \mathbb{N}$ we have the inequality

$$|f_m(x) - f_n(x)| \leq \sup\{|f_m(x) - f_n(x)| \mid x \in K\} = \|f_m - f_n\|$$

Hence the sequence $(f_n(x))$ is a Cauchy sequence of numbers in the field $\mathbb{F}$. By completeness of the field $\mathbb{F}$ the sequence $(f_n(x))$ converges to some limit, $f(x)$.[2]

Let $\varepsilon > 0$ be a real number. Then the sequence of functions (f_n) is a Cauchy sequence so we can find a natural number N such that $\|f_m - f_n\| < \frac{\varepsilon}{2}$ whenever $m, n \geq N$. Let $x \in X$, and let $m, n \geq N$. Then $|f_m(x) - f_n(x)| < \frac{\varepsilon}{2}$ and so $|f_m(x) - f(x)| < \varepsilon$ if we let $n \to \infty$.

Therefore $\|f_m - f\| < \varepsilon$ whenever $m \geq N$ and the function f is the limit of the sequence (f_n). All that remains is to show that the function f is continuous.

Again, let $\varepsilon > 0$. Let $x \in X$. Then we can find a neighbourhood $U \ni x$ and a natural number N such that $|f_N(y) - f_N(x)| < \frac{\varepsilon}{3}$ whenever $y \in U$ and $\|f_N - f\| < \frac{\varepsilon}{3}$. For each point $y \in U$ we therefore have the inequality

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

The function f is therefore continuous. □

The above example illustrates a pattern which holds for many proofs of completeness. The pattern is roughly as follows.

- Consider a Cauchy sequence in the metric space we are looking at. Produce a point which looks like it should be the limit of the sequence.
- Prove that the point produced is the limit of the Cauchy sequence we are considering.
- If necessary, prove that the point produced belongs to the space we are looking at.

Example 1.21 *The space $\mathbb{F}^k$ can be equipped with a norm defined by the formula*

$$\|(\alpha^{(1)}, \ldots, \alpha^{(k)})\| = |\alpha^{(1)}| + \cdots + |\alpha^{(k)}| \qquad \alpha_i \in \mathbb{F}.$$

If we give the space $\mathbb{F}^k$ this norm, it is a Banach space.

Proof: The proof that the given formula defines a norm is similar to example 1.12.

Let (v_n) be a Cauchy sequence in the space $\mathbb{F}^k$. Write

$$v_n = (\alpha_n^{(1)}, \ldots, \alpha_n^{(k)}).$$

[2] Recall that $\mathbb{F}$ denotes either $\mathbb{R}$ or $\mathbb{C}$

Let $j \in \{1, \ldots, k\}$. Observe that for all $m, n \in \mathbb{N}$, we have $|\alpha_m^{(j)} - \alpha_n^{(j)}| \leq \|v_m - v_n\|$. Hence the sequence $(\alpha_n^{(j)})_{n=1}^{\infty}$ is a Cauchy sequence in the field $\mathbb{F}$. By completeness of the field $\mathbb{F}$, this sequence converges to some limit, $\alpha^{(j)}$. Let

$$v = (\alpha^{(1)}, \ldots, \alpha^{(k)}) \in \mathbb{F}^k.$$

We claim that $\|v_n - v\| \to 0$ as $n \to \infty$. To see this,
Now, observe

$$\lim_{n\to\infty} \|v_n - v\| = \lim_{n\to\infty} |\alpha_n^{(1)} - \alpha^{(1)}| + \cdots + \lim_{n\to\infty} |\alpha_n^{(k)} - \alpha^{(k)}| = 0.$$

So the sequence (v_n) has limit v. We see that the space $\mathbb{F}^k$ is complete. □

We shall see later on that any finite-dimensional normed vector space is a Banach space. Some further examples of Banach spaces include:

- The space l^1 of all sequences (α_n) in the field $\mathbb{F}$ such that the sum $\sum_{n=1}^{\infty} |\alpha_n|$ converges. As noted earlier, we give this space the norm defined by the formula $\|(\alpha_n)\| = \sum_{n=1}^{\infty} |\alpha_n|$.
- The space l^∞ of all bounded sequences (α_n) in the field $\mathbb{F}$, with the norm defined by the formula $\|(\alpha_n)\| = \sup\{|\alpha_n| \mid n \in \mathbb{N}\}$.
- The subspace c_0 of l^∞ consisting of all sequences (α_n) in the field $\mathbb{F}$ such that $\alpha_n \to 0$ as $n \to \infty$.

The relevant proofs are left as exercises.

Example 1.22 *Recall that we can define a norm on the vector space* $C[-1, 1]$ *by the formula*

$$\|f\| = \int_{-1}^{1} |f(x)| \, dx$$

Let $n \in \mathbb{N}$. *Define* $f_n \in C[-1, 1]$ *by*

$$f_n(t) = \begin{cases} (1-t)^n & t \leq 0 \\ 1 & t \geq 0 \end{cases}$$

Let

$$g(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

But

$$\|f_n - g\| = \int_{-1}^{0} (1-t)^n \, dt = \frac{1}{n+1} \to 0$$

as $n \to \infty$. *So the sequence* (f_n) *converges to* g *with respect to the above norm. In particular, the sequence* (f_n) *is Cauchy.*

But g *is not continuous, and, clearly, if* $f \in C[-1, 1]$, *then* $\|f - g\| > 0$. *So no function in* $C[-1, 1]$ *can be the limit of the Cauchy sequence* (f_n) *for our given norm.*

Thus $C[-1, 1]$ *is* not *a Banach space for this norm.*

We have to be careful here; we have an alternative norm on $C[-1,1]$ defined by

$$\|f\|' = \sup\{|f(t)| \mid t \in [-1,1]\}$$

for which, as we saw above, $C[-1,1]$ *is* a Banaach space.

1.4 Closures and Completions

The notion of *completion* is a way of passing from normed vector spaces to Banach spaces.

Definition 1.23 *Let V be a normed vector space. Let A be a subset of V. Then we define the* closure *of A, $\overline{A}$, to be the set of all elements $v \in V$ that are norm-limits of sequences of the space A.*

We call the subset $A \subseteq V$ closed *if $\overline{A} = A$.*

Example 1.24 *Let V be a normed vector space, and $x \in V$. Then the one-point set $\{x\}$ is closed.*

Example 1.25 *Let V be a normed vector space. Let $x \in V$ and $\delta > 0$. Then we define the* open ball

$$B(x,\delta) = \{v \in V \mid \|v - x\| < \delta\}.$$

The closure is the closed ball

$$B(x,\delta) = \{v \in V \mid \|v - x\| \leq \delta\}.$$

Note that when we need to be more precise, we write $B_V(x,\delta)$ to denote the open ball in the space V at the point x with radius δ. This more precise notation is sometimes needed when we are considering more than one normed vector space in a problem.

Proposition 1.26 *If U is a vector subspace of a normed vector space V, then the closure $\overline{U}$ is also a vector subspace.*

Proof: Observe $0 \in U \subseteq \overline{U}$, so $\overline{U} \neq \emptyset$.

Let $v, w \in \overline{U}$ and $\alpha, \beta \in \mathbb{F}$. Let (v_n) and (w_n) be sequences in the space U converging to v and w respectively.

Then certainly, for each $n \in \mathbb{N}$, we have $\alpha v_n + \beta w_n \in U$, and

$$\lim_{n\to\infty} \|\alpha v_n + \beta w_n - (\alpha v + \beta w)\| \leq |\alpha| \lim_{n\to\infty} \|v_n - v\| + |\beta| \lim_{n\to\infty} \|w_n - w\| = 0.$$

Hence $\alpha v + \beta w \in \overline{U}$, and the closure $\overline{U}$ is a vector subspace as required. □

We leave the proof of the following result as a straightforward exercise.

Proposition 1.27 *Let V be a Banach space. Then a vector subspace of V is a Banach space if and only if it is closed.* □

Definition 1.28 *Let V be a normed vector space. We call a subspace U* dense *in V if $\overline{U} = V$.*

The notion of a subspace being dense can be rephrased as follows.

Proposition 1.29 *Let V be a normed vector space, and let $U \subseteq V$ be a subspace. Then U is dense in V if and only if for all $v \in V$ and $\varepsilon > 0$ there exists $u \in U$ such that $\|u - v\| < \varepsilon$.* □

Example 1.30 *Let U be the subspace of l^1 consisting of all sequences (α_n) for which there exists $N \in \mathbb{N}$ such that $\alpha_n = 0$ whenever $n \geq N$.*

Let $v = (\beta_n) \in l^1$. Let $\varepsilon > 0$. The series $\sum_{n=1}^{\infty} |\beta_n|$ converges to the norm $\|v\|$. Hence, looking at partial sums, we have $N \in \mathbb{N}$ such that

$$\left| \sum_{n=1}^{N} |\beta_n| - \|v\| \right| < \varepsilon.$$

Let

$$\alpha_n = \begin{cases} \beta_n & n \leq N \\ 0 & n > N \end{cases}.$$

Then $u = (\alpha_n) \in U$, and $\|u - v\| < \varepsilon$. We conclude that U is a dense subset of l^1.

Definition 1.31 *Let V and W be normed vector spaces. We call a linear map $T: V \to W$ such that $\|T(v)\| = \|v\|$ for all $v \in V$ a* linear isometry*.*

Proposition 1.32 *Let $T: V \to W$ be a linear isometry. Then T is injective.*

Proof: Let $v \in V$, and suppose $T(v) = 0$. Then certainly $\|T(v)\| = 0$.

Hence we also have $\|v\| = 0$, and $v = 0$. We have shown that $\ker T = \{0\}$, which means that T is injective. □

We call a surjective linear isometry an *isometric isomorphism*. If there is an isometric isomorphism between two normed vector spaces, they have identical structures.

Definition 1.33 *Let V be a normed vector space. Then we call a Banach space, $\hat{V}$, a* completion *of V if there is a linear isometry $i: V \to \hat{V}$ such that the image $i[V]$ is dense in $\hat{V}$.*

Note that if a Banach space V has a dense vector subspace U, then V is a completion of U.

Theorem 1.34 *Any normed vector space has a completion.*

We defer the proof of this statement to chapter 3, where we obtain it as a straightforward corollary of the Hahn-Banach theorem.

The following result tells us that completions are to all intents and purposes unique.

Proposition 1.35 *Let U be a normed vector space. Let V_1 and V_2 be completions of U. Then there is an isometric isomorphism $T: V_1 \to V_2$.*

Proof: Let $i_1: U \to V_1$ and $i_2: U \to V_2$ be linear isometries with dense images. We certainly have a linear isometry $S: i_1[U] \to i_2[U]$ defined by the formula

$$S(u) = i_2(i_1^{-1})(u).$$

Let $v \in V_1$. Then, since $i_1[U]$ is dense in V_1, we have a sequence (v_n) in $i_1[U]$ converging in norm to v. Try to define a map $T: V_1 \to V_2$ by writing

$$T(v) = \lim_{n\to\infty} S(v_n).$$

Let (v'_n) be another sequence in $i_1[U]$ converging to v. Then $\lim_{n\to\infty} \|v'_n - v_n\| = 0$.

Hence, since S is a linear isometry, we have $\lim_{n\to\infty} \|S(v'_n) - S(v_n)\| = 0$. Therefore

$$\lim_{n\to\infty} S(u'_n) = \lim_{n\to\infty} S(u_n)$$

and we conclude the map T is well-defined.

Let $u, v \in V_1$. Let $\alpha, \beta \in \mathbb{F}$, and let (u_n) and (v_n) be sequences in $i_1[U]$ converging to u and v respectively. Then

$$\lim_{n\to\infty} S(\alpha u_n + \beta v_n) = \alpha \lim_{n\to\infty} S(u_n) + \beta \lim_{n\to\infty} S(v_n).$$

so

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$$

and the map T is linear.

Finally, let $v \in V_1$, and let (v_n) be a sequence in $i_1[U]$ converging to v. Then the norm is a continuous function on a normed vector space by proposition 1.8, so

$$\|T(v)\| = \|\lim_{n\to\infty} S(v_n)\| = \lim_{n\to\infty} \|S(v_n)\| = \lim_{n\to\infty} \|v_n\| = \|v\|.$$

Thus the map T is a linear isometry. It remains to prove surjectivity.

Let $w \in V_2$. Let (u_n) be a sequence in U such that $i_2(u_n) \to w$ as $n \to \infty$. The sequence $i_2(u_n)$ is certainly a Cauchy sequence such the maps i_2^{-1} and i_1 are isometries. So we have a norm limit $v = \lim_{n\to\infty} i_1(u_n)$.

But by construction, $T(v) = w$, so the map T is surjective, and we are done. □

It therefore makes sense for us to talk about *the* completion of a normed vector space, which is unique up to isometric isomorphism. We will abuse notation slightly, and assume that if we have a normed vector space V, then $V \subseteq \hat{V}$. If we do this, then the space V is a dense subspace of its completion.

Example 1.36 *Recalle that the vector space $C[a, b]$ can be equipped with a norm defined by the formula*

$$\|f\| = \int_a^b |f(x)| \, dx$$

It is not complete with respect to this norm. We define the Banach space $L^1[a, b]$ to be the completion of $C[a, b]$ for this particular norm.

1.5 Convexity

The following concept will come in handy later on, though we will not use it immediately.

Definition 1.37 *Let V be a real or complex vector space. We call a subset $C \subseteq V$* convex *if for all $x, y \in C$ and $\alpha \in [0, 1]$, we have*

$$\alpha x + (1 - \alpha)y \in C.$$

In other words, for any two points $x, y \in C$, the straight line segment joining x and y is also in C.

Example 1.38 *Let V be a normed vector space, $x \in V$ and $\delta > 0$. Then the open ball $B(x, \delta)$ is convex.*

We need the notion of convexity in a few places in these notes (and in particular in the proof of the open mapping theorem in the next chapter). The proof of the following properties are left as exercises.

Proposition 1.39 *Let V be a normed vector space, and let $C \subseteq V$ be convex. Then the closure $\overline{C}$ is also convex.* □

Proposition 1.40 *Let V and W be normed vector spaces, let $T\colon V \to W$ be a linear map, and let $C \subseteq V$ be convex. Then the image $T[C] \subseteq W$ is convex.* □

Chapter 2

Linear Maps and Continuity

2.1 Open Sets and Continuity

Let X be a metric space, $x \in X$, and $\delta > 0$. Then, just as in the normed vector space case, we define the open ball

$$B(x, \delta) = \{y \in X \mid d(x, y) < \delta\}.$$

Definition 2.1 *Let X be a metric space. We call a subset $U \subseteq X$* open *if for all $x \in U$ we have $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$.*

Example 2.2 *Let $x \in X$ and $\delta > 0$. Then the open ball $B(x, \delta)$ is itself an open set.*

The next two results are standard results from the theory of metric spaces; the details are left as an exercise.

Proposition 2.3 *Let X be a metric space.*

- *The sets $\emptyset$ and X are open.*
- *The intersection of finitely many open sets is open.*
- *The union of an arbitrary collection of open sets is open.*

$\square$

Theorem 2.4 *Let X and Y be metric spaces. A map $f: X \to Y$ is continuous if and only if for any open set $U \subseteq W$, the inverse image $f^{-1}[U] \subseteq V$ is open.* $\square$

The notions of closure and closed subsets also work for metric spaces as well as for normed spaces. To be precise, given a metric space X and a subset $A \subseteq X$, we define the closure of A

$$\overline{A} = \{x \in V \mid B(x, \varepsilon) \cap A \neq \emptyset \text{ for all } \varepsilon > 0\}.$$

We call A *closed* if and only if $\overline{A} = A$. We leave the proof of the following as an exercise.

Proposition 2.5 *Let X be a metric space. Then a subset $A \subseteq X$ is closed if and only if the complement $X \backslash A$ is open.* □

The next result follows immediately from the above and proposition 2.3

Corollary 2.6 *Let X be a metric space.*

- *The sets $\emptyset$ and X are closed.*
- *The union of finitely many closed sets is closed.*
- *The intersection of an arbitrary collection of closed sets is closed.*

□

2.2 Bounded Linear Maps

In this chapter we investigate continuous linear maps between normed vector spaces.

Definition 2.7 *Let $T: V \to W$ be a linear map between normed vector spaces. We call T* bounded *if there exists $M \geq 0$ such that $\|Tv\| \leq M\|v\|$ for all $v \in V$.*

If required, we can always choose $M > 0$.

Note that the above definition is *not* the same as the image of the map T being a bounded subset of the space W.

Proposition 2.8 *Let $T: V \to W$ be a linear map between normed vector spaces. Then the following are equivalent:*

(i) *T is continuous.*

(ii) *T is continuous at the point* 0.

(iii) *T is bounded.*

Proof:

$(i) \Rightarrow (ii)$: This is trivial.

$(ii) \Rightarrow (iii)$: Let us take $\varepsilon = 1$ in the definition of continuity at the point 0. Then we can find $\delta > 0$ such that $d(v, 0) < \delta$ implies $d(Tv, 0) < 1$, ie: $\|v\| < \delta$ implies $\|Tv\| < 1$.

Take $M = \frac{2}{\delta}$. Let $v \in V \backslash \{0\}$. Then

$$\left\| \frac{\delta v}{2\|v\|} \right\| = \frac{\delta}{2} < \delta$$

Hence $\|T(\delta v/2\|v\|)\| < 1$. By linearity

$$\frac{\delta}{2\|v\|}\|Tv\| < 1 \quad \text{and} \quad \|Tv\| < \frac{2}{\delta}\|v\| = M\|v\|$$

Certainly, if $v = 0$, then $\|Tv\| = 0 \leq M\|v\|$. So the map T is bounded.

$(iii) \Rightarrow (i)$: Let $T: V \to W$ be bounded. Then we can find $M > 0$ such that $\|Tv\| \leq M\|v\|$ for all $v \in V$.

Choose $\varepsilon > 0$. Write $\delta = \varepsilon/M$. Then, for $v, w \in V$ with $d(v, w) < \delta$, we have

$$\|v - w\| < \frac{\varepsilon}{M}$$

and

$$\|Tv - Tw\| = \|T(v - w)\| \leq M\|v - w\| < \frac{M\varepsilon}{M} = \varepsilon$$

We see that the map T is continuous.

□

Definition 2.9 *Let V and W be normed vector spaces, with $V \neq \{0\}$, and let $T: V \to W$ be a bounded linear map. Then we define the* operator norm *of the operator T by the formula*

$$\|T\| = \sup_{v \in V, v \neq 0} \frac{\|Tv\|}{\|v\|}$$

Certainly $\|T\| \geq 0$. By the above definition, we have the inequality

$$\|Tv\| \leq \|T\|\|v\|$$

for all $v \in V$, and the value $\|T\|$ is the *smallest* number where we have such an inequality.

Actaully, our definition breaks down in the case $V = \{0\}$. In this case, we simply define $\|T\| = 0$. The proof of the following is left as an exercise.

Proposition 2.10 *Let $T: V \to W$ be a bounded linear map. Then*

$$\|T\| = \sup_{v \in V, \|v\| \leq 1} \|Tv\|.$$

If $V \neq \{0\}$, we have

$$\|T\| = \sup_{v \in V, \|v\| = 1} \|Tv\|.$$

□

Definition 2.11 *Let V and W be normed vector spaces. Then we write $Hom(V, W)$ to denote the set of all bounded linear maps from V to W.*

The following is straightforward.

Proposition 2.12 *The set $Hom(V, W)$ is a normed vector space. The operations of addition and scalar multiplication defined by the formulae*

$$(S+T)v = Sv + Tv \qquad S, T \in Hom(V, W),\ v \in V$$

and

$$(\alpha T)v = \alpha T(v) \qquad T \in Hom(V, W),\ \alpha \in \mathbb{F},\ v \in V$$

The norm is defined by the above operator norm. □

Example 2.13 *Let $w \in W$. Let $T_w : \mathbb{F} \to W$ be the linear map defined by the formula $T_w(\alpha) = \alpha w$. Then, for any scalar $\alpha \in \mathbb{F}$, we have*

$$\|T_w(\alpha)\| = \|\alpha w\| = |\alpha|\|w\|$$

Hence the linear map T_w is bounded, and $\|T_w\| = \|w\|$.

Note that any linear map $T : \mathbb{F} \to W$ takes the form T_w for some $w \in W$. Hence we can identify the space $Hom(\mathbb{F}, W)$ with the normed vector space W.

Proposition 2.14 *Let V be a normed vector space, and let W be a Banach space. Then the space $Hom(V, W)$ is a Banach space.*

Proof: Let (T_n) be a Cauchy sequence in the space $Hom(V, W)$. Let $v \in V$. Observe, for all $m, n \in \mathbb{N}$, we have

$$\|T_m v - T_n v\| \leq \|T_m - T_n\| \cdot \|v\|.$$

Hence the sequence $(T_n v)$ is a Cauchy sequence in the space W. But the space W is complete, so we have a norm limit

$$Tv = \lim_{n\to\infty} T_n v.$$

Let $\alpha, \beta \in \mathbb{F}$, and $u, v \in V$. Then

$$T(\alpha u+\beta v) = \lim_{n\to\infty} T_n(\alpha u+\beta v) = \alpha \lim_{n\to\infty} T_n(u) + \beta \lim_{n\to\infty} T_n(v) = \alpha T(u) + \beta T(v)$$

and the map T is linear.

Since the norm is a continuous map, the sequence $(\|T_n\|)$ is a Cauchy sequence in $\mathbb{R}$. Since the real numbers are complete, we have a limit

$$M = \lim_{n\to\infty} \|T_n\|.$$

For $v \in V$, we have

$$\|Tv\| = \lim_{n\to\infty} \|T_n(v)\| \leq \lim_{n\to\infty} \|T_n\| \cdot \|v\| = M\|v\|.$$

So the map T is bounded. We claim that T is the norm limit of the Cauchy sequence (T_n) in the space $Hom(V, W)$, which tells us that the space $Hom(V, W)$ is complete, and finishes the proof.

Let $\varepsilon > 0$. Then there is a natural number $N \in \mathbb{N}$ such that $\|T_m - T_n\| < \varepsilon/2$ whenever $m, n \geq N$. Let $v \in V$ and $\|v\| \leq 1$. Then for $m, n \geq N$, we have

$$\|T_m(v) - T_n(v)\| < \frac{\varepsilon}{2}.$$

Taking the limit as $m \to \infty$, we see for $n \geq N$ that

$$\|T(v) - T_n(v)\| \leq \frac{\varepsilon}{2}.$$

Hence

$$\|T - T_n\| = \sup_{v \in V, \|v\| \leq 1} \|Tv - T_n v\| \leq \frac{\varepsilon}{2} < \varepsilon.$$

It follows that T is the norm limit of the sequence (T_n) and we are done. □

2.3 Finite-Dimensional Spaces

We shall see in this section that any two normed vector spaces of the same dimension are linearly isomorphic. As a matter of convenience, when we mention the space $\mathbb{F}^n$ in this section, we equip it with the norm

$$(\alpha_1, \ldots, \alpha_n) = |\alpha_1| + \cdots + |\alpha_n|.$$

Lemma 2.15 *Let V be a normed vector space, and let $T: \mathbb{F}^n \to V$ be a linear map. Then the map T is bounded.*

Proof: Let $\{e_1, \ldots, e_n\}$ be the standard basis for $\mathbb{F}^n$. Let $v \in \mathbb{F}^n$, and write

$$v = \alpha_1 e_1 + \cdots + \alpha_n e_n \qquad \alpha_i \in \mathbb{F}$$

Let $M = \max(\|Te_1\|, \ldots, \|Te_n\|)$, where $\| - \|$ is the norm on V. Then

$$\begin{array}{rcl} \|Tv\| & = & \|\alpha_1 Te_1 + \cdots + \alpha_n Te_n\| \\ & \leq & |\alpha_1| \|Te_1\| + \cdots + |\alpha_n| \|Te_n\| \\ & \leq & |\alpha_1| M + \cdots + |\alpha_n| M \\ & = & M\|v\| \end{array}.$$

So the map T is bounded. □

Recall that if K is a metric space, we call K *compact* if every sequence (x_n) has a convergent subsequence with a limit in the space K. The following result is known as the *Heine-Borel theorem*.

Theorem 2.16 *A subset $K \subseteq \mathbb{F}^n$ is compact if and only if it is closed and bounded.* □

Here, by a metric spae X being bounded, we mean there is a constant $C \geq 0$ such that $d(x, y) \leq C$ for all $x, y \in X$. This notion is different to the notion of a linear map being bounded.

The proof of the following result is left as an exercise.

Proposition 2.17 *Let X and Y be metric spaces, let $K \subseteq X$ be compact, and let $f\colon X \to Y$ be a continuous map. Then the image $f[K] \subseteq Y$ is compact.* □

Lemma 2.18 *Let K be a compact subset of a normed vector space such that $0 \notin K$. Then we can find $m > 0$ such that $\|a\| \geq m$ for all $a \in K$.*

Proof: We have already seen that we have a continuous map $p\colon K \to \mathbb{R}$ defined by the formula $p(a) = \|a\|$. Hence the image $p[K]$ is compact, and therefore closed and bounded.

Let $m = \inf p[K] = \inf\{\|a\| \mid a \in K\}$. Then y is the greatest lower bound of the set $p[K]$. As a lower bound, $\|a\| \geq 0$ for all $a \in K$. Since m is the *greatest* lower bound, there is a sequence of elements, (y_n), in $p[K]$ converging to m. Hence $y \in \overline{p[K]} = p[K]$, and $m = p(a_0) = \|a_0\|$ for some $a_0 \in K$.

But $0 \notin K$, so $a_0 \neq 0$, and $m = \|a_0\| > 0$. □

Theorem 2.19 *Let V be an n-dimensional normed vector space. Then V is linearly homeomorphic to the space $\mathbb{F}^n$.*

Proof: Since the spaces $\mathbb{F}^n$ and V have the same dimension, there is an invertible linear map $T\colon \mathbb{F}^n \to V$. By lemma 2.15, the map T is continuous. We want to show that the inverse T^{-1} is also continuous.

Since the norm is a continuous map, and the map T is continuous, the composition

$$\mathbb{F} \stackrel{T}{\to} V \stackrel{\|-\|}{\to} \mathbb{R}$$

is continuous.

Let $S = \{x \in \mathbb{F}^n \mid \|x\| = 1\}$ be the unit sphere in the space $\mathbb{F}^n$. Then the subset S is closed and bounded, and therefore compact by the Heine-Borel theorem. It follows by lemma 2.18 that we can find $m > 0$ such that $\|Tx\| \geq m$ for all $x \in S$.

Write $M = \frac{1}{m}$. Observe that $\|T^{-1}(0)\| = 0 \leq M\|0\|$. Choose $v \in V \backslash \{0\}$. Then $T^{-1}v \neq 0$, and $\frac{T^{-1}v}{\|T^{-1}v\|} \in S$. Hence

$$T\left(\frac{T^{-1}v}{\|T^{-1}v\|}\right) \geq m$$

and $\|v\| \geq m\|T^{-1}v\|$.

We see that $\|T^{-1}v\| \leq M\|v\|$ for all $v \in V$. Thus the map T^{-1} is bounded, and we are done. □

The above result tells us that a finite-dimensional vector space has the same open sets for any choice of norm. In particular, when looking at properties such as continuity, limits, or which subsets are compact, the choice of norm on a finite-dimensional vector space is irrelevant.

Corollary 2.20 *Let V and W be normed vector spaces, where V is finite-dimensional. Let $T: V \to W$ be a linear map. Then T is continuous.*

Proof: By the above theorem, we have a linear homeomorphism $\phi: \mathbb{F}^n \to V$. By lemma 2.15, the map $T \circ \phi: \mathbb{F}^n \to W$ is continuous.

The inverse of the homeomorphism ϕ is also continuous. Hence the map $T = T \circ \phi \circ \phi^{-1}: V \to W$ is also continuous, and we are done. □

2.4 The Open Mapping Theorem

Definition 2.21 *A mapping $f: V \to W$ between two normed vector spaces is said to be* open *if for every open set $U \subseteq V$, the image $f[U] \subseteq W$ is open.*

Note that this is different to the definition of continuity in terms of open sets, which asserts that a function is continuous if the *inverse image* of an open set is open. So, if we have an open bijective map, f, then the inverse f^{-1} is continuous.

The *open mapping theorem* is the initially surprising statement that any bounded linear map between Banach spaces is open. The fact that the normed vector spaces we have are complete is important.

In order to prove this result, we begin with a technical result from the theory of metric spaces called the *Baire category theorem*. We will not prove the Baire category theorem here; the interested reader can consult the book by Kreyszig, for example, for a proof.

Theorem 2.22 *Let X be a complete metric space. Let $(A_n)_{n=1}^{\infty}$ be a sequence of closed subsets such that $X = \cup_{n=1}^{\infty} A_n$. Then at least one set A_n contains a non-empty open set.* □

Now, before we attack the open mapping theorem, we need two technical lemmas. The proof of the first is left as an exercise.

Lemma 2.23 *Let V and W be normed vector spaces, and let $T: V \to W$ be a linear map. Then T is open if and only if there exists $\delta > 0$ such that*

$$B_W(0, \delta) \subseteq T[B_V(0, 1)] \subseteq W.$$

□

Lemma 2.24 *Let V and W be Banach spaces. Let $T: V \to W$ be a surjective bounded linear map. Then we have $\delta > 0$ such that*

$$B_W(0, \delta) \subseteq \overline{T[B_V(0, 1)]} \subseteq W.$$

Proof: Observe

$$V = \bigcup_{n=1}^{\infty} B_V(0,n)$$

so by surjectivity

$$W = \bigcup_{n=1}^{\infty} T[B_V(0,n)] \subseteq \bigcup_{n=1}^{\infty} \overline{T[B_V(0,n)]}.$$

Hence, by the Baire category theorem, some set $\overline{T(B_V(0,n))}$ contains a non-empty open subset.

Hence for some $y_0 \in W$ and $\varepsilon > 0$, we have

$$B_W(y_0,\varepsilon) \subseteq \overline{T[B_V(0,n)]}.$$

Let $y \in B_W(0,1)$. Then clearly

$$y_0 + \varepsilon y \in B_W(y_0,\varepsilon) \subseteq \overline{T[B_V(0,n)]}$$

and

$$y_0 \in B_W(y_0,\varepsilon) \subseteq \overline{T[B_V(0,n)]}.$$

Now, the open ball $B(0,2n)$ is convex. Hence the closure of the image, $\overline{T[B_V(0,2n)]}$ is also convex.

Observe $y_0 \in \overline{T[B_v(0,n)]}$, so

$$-2y_0 \in 2\overline{T[B_V(0,n)]} = \overline{T[B_V(0,2n)]}$$

and similarly

$$2(y_0 + \varepsilon y) \in \overline{T[B_V(0,2n)]}.$$

Hence, by convexity

$$\varepsilon y = \frac{1}{2}(2(y_0+\varepsilon y) - 2y_0) \in \overline{T[B_V(0,2n)]} = 2n\overline{T[B_V(0,1)]}.$$

So if we set $\delta = \frac{\varepsilon}{2n}$, then $\delta y \in \overline{T[B_V(0,1)]}$, and we are done. □

We are now ready for the main result.

Theorem 2.25 (The Open Mapping Theorem) *Let V and W be Banach spaces, and let $T\colon V \to W$ be a surjective bounded linear map. Then T is open.*

Proof: We show first that

$$\overline{T[B_V(0,1)]} \subseteq 2T(B_V(0,1)).$$

By the above lemma, we have $\delta > 0$ such that $B_W(0,\delta) \subseteq \overline{T[B_V(0,1)]}$. Let $y \in \overline{T[B_V(0,1)]}$. Then we have $x_1 \in B_V(0,1)$ such that $\|y - Tx_1\| < \frac{1}{2}\delta$.

Hence

$$2(y - Tx_1) \in B_W(0, \delta) \subseteq \overline{T[B_V(0, 1)]}.$$

Repeating this argument with $2(y-Tx_1)$ instead of y, we have $x_2 \in B_V(0, 1)$ such that

$$\|2(y - Tx_1) - Tx_2\| < \frac{1}{2}\delta$$

so

$$\|y - Tx_1 - \frac{1}{2}Tx_2\| < \frac{1}{4}\delta.$$

Repeating again (or more formally, working by induction), we have points $x_1, x_2, x_3, \ldots \in B_V(0, 1)$ such that

$$\left\| y - T\left(x_1 + \frac{1}{2}Tx_2 + \cdots + \frac{1}{2^{n-1}}x_n \right) \right\| < \frac{\delta}{2^n}.$$

Since V is complete, and $\|x_n\| < 1$ for all n, the sum $\sum_{n=1}^{\infty} \frac{x_n}{2^{n-1}}$ converges in V, say to a point x. Further,

$$\|x\| = \lim_{n\to\infty} \left\| x_1 + \frac{1}{2}Tx_2 + \cdots + \frac{1}{2^{n-1}}x_n \right\| < 2.$$

Since the map T is continuous, we have

$$\|y - Tx\| = \lim_{n\to\infty} \left\| y - T\left(x_1 + \frac{1}{2}Tx_2 + \cdots + \frac{1}{2^{n-1}}x_n \right) \right\| < \frac{\delta}{2^n} = 0.$$

Thus $y = Tx$, which means $y \in T[B_V(0, 2)]$. Therefore

$$\overline{T[B_V(0, 1)]} \subseteq T[B_V(0, 2)]$$

as we desired.

Now, combining this with the above lemma, we see

$$B_W(0, \frac{\delta}{2}) \subseteq T[B_V(0, 1)].$$

But by lemma 2.23, this means that the map T is open. □

The following consequence of the open mapping theorem is called the Banach isomorphism theorem.

Corollary 2.26 *Let V and W be Banach spaces. Let $T\colon V \to W$ be a bijective bounded linear map. Then T is a linear homeomorphism.* □

2.5 The Closed Graph Theorem

Definition 2.27 *Let A and B be sets, and let $f: A \to B$ be a map. We define the* graph *of f to be the set*

$$Gr(f) = \{(x, f(x)) \mid x \in A\} \subseteq A \times B.$$

Proposition 2.28 *Let X and Y be metric spaces. Let $f: X \to Y$ be a continuous map. Then the graph $Gr(f)$ is a closed subset of $X \times Y$.*

Proof: Let $(x_n, f(x_n))$ be a sequence in $Gr(f)$ that converges to some point $(x, y) \in X \times Y$. Then $x_n \to x$ as $n \to \infty$.

So by continuity, $f(x_n) \to f(x)$ as $n \to \infty$. But by the above, $f(x_n) \to y$ as $n \to \infty$. So, by the uniqueness of limits in metric spaces, $y = f(x)$, which means $(x, y) \in Gr(f)$.

It follows that $Gr(f)$ is a closed subset of $X \times Y$. □

The above result has a converse in the case of Banach spaces. This converse is called the *closed graph theorem.*

Theorem 2.29 *Let V and W be Banach spaces, and let $T: V \to W$ be a linear map with a closed graph. Then T is continuous.*

Proof: Observe that $Gr(T)$ is a closed linear subspace of $V \times W$, and therefore a Banach space. Define a linear map $P: Gr(T) \to V$ by the formula

$$P((x, Tx)) = x \qquad x \in V.$$

The map P is clearly bijective. Observe that

$$P(x, Tx)\| = \|x\| \leq \|(x, Tx)\|.$$

so the map P is continuous. By the Banach isomorphism theorem, the inverse $P^{-1}: V \to Gr(T)$ is also continuous.

Similarly, the map $Q: Gr(T) \to W$ defined by the formula $Q(x, Tx) = Tx$ is continuous. Therefore the composite

$$T = Q \circ P^{-1}: V \to W$$

is also continuous, and we are done. □

Chapter 3

Spaces of Continuous Functions

3.1 Dual Spaces

Definition 3.1 *Let V be a normed vector space over the field $\mathbb{F}$. Then a* linear functional *is a bounded linear map $f: V \to \mathbb{F}$.*

Example 3.2 *Let K be a compact metric space. Consider the normed vector space $C(K)$ from example 1.11. Fix $x \in K$. Then we have a bounded linear map $E_x: C(K) \to \mathbb{F}$ defined by the formula*

$$E_x(f) = f(x) \qquad f \in C(K).$$

Example 3.3 *The integration function*

$$f \mapsto \int_0^1 f(x)\ dx$$

defines a linear functional on the space $L^1[0, 1]$

Definition 3.4 *Let V be a normed vector space. Then we define the* dual space, *V^*, to be the vector space of bounded linear functionals $f: V \to \mathbb{F}$.*

We have $V^* = Hom(V, \mathbb{F})$, which is a normed vector space when it is given the operator norm. By theorem 2.14, since the field $\mathbb{F}$ is complete, the space V^* is a Banach space.

Definition 3.5 *Let V be a normed vector space, and let $W \leq V$ be a subspace. Suppose we have a bounded linear functional $f: W \to \mathbb{F}$. Then a bounded linear functional $F: V \to \mathbb{F}$ is called an* extension of f *if $F(v) = f(v)$ whenever $v \in W$.*

3.2 Zorn's Lemma

Zorn's Lemma is an alternative formulation of the axiom of choice.

Definition 3.6 *A set P is said to be* partially ordered *if there is a relation, writtten $\leq$, on the elements of P such that the following axioms are satisfied:*

- *$x \leq x$ for every element $x \in P$*
- *If $x \leq y$ and $y \leq z$ then $x \leq z$*
- *If $x \leq y$ and $y \leq x$ then $x = y$*

A partially ordered set P is said to be *totally ordered* if for any two elements $x, y \in P$ either $x \leq y$ or $y \leq x$.

Example 3.7 *Let S be any set, and let P be the collection of subsets of S. Then the set P is partially ordered by the relation $\subseteq$ of set inclusion.*

If the set S has more than one element then the set P is not totally ordered.

Example 3.8 *The set of real numbers, $\mathbb{R}$, is a totally ordered set, where the relation $\leq$ has the usual meaning.*

Definition 3.9 *Let P be a partially ordered set. Then an element $x \in P$ is said to be* maximal *if for all elements $y \in P$ the relation $x \leq y$ implies that $x = y$.*

Let Q be a subset of P. Then an element $x \in P$ is said to be an upper bound *for Q if $q \leq x$ for all elements $q \in Q$.*

Example 3.10 *Let P be the collection of subsets of some set S. Then the set S itself is both an upper bound for P and a maximal element of P.*

Example 3.11 *Let B be a bounded subset of the real numbers. Then the supremum of B, α, is an upper bound for B. If $\alpha \in B$ then α is a maximal element of B.*

We are now ready to state *Zorn's lemma* which can be considered a fundamental axiom of set theory.

Zorn's Lemma

Let P be a partially ordered set. Suppose that every totally ordered subset has an upper bound. Then the set P has a maximal element.

The following result provides an example of how Zorn's lemma is used.

Proposition 3.12 *Let X be a vector space, and let Y be a subspace of X. Then we can find a subspace Z such that $X = Y \oplus Z$.*

Proof: Let P be the set of subspaces, A, such that $Y \cap A = \{0\}$. We can define a partial ordering on the set P by writing $A \leq B$ if the space A is contained in the space B.

Suppose that T is a totally ordered subset of P. Form the set

$$M = \bigcup_{A \in T} A$$

Certainly $M \cap Y = \{0\}$. Because the set T is totally ordered, the set M is a subspace of V. Hence we have an element $M \in P$ that is an upper bound for the subset T. By Zorn's lemma the set P must have a maximal element, Z.

We know that $Y \cap Z = \{0\}$. Suppose that $Y \oplus Z \neq X$. Then we can find a vector $x \in X \backslash (Y \oplus Z)$. The vector space $Z \oplus \langle x \rangle$ is an element of the set P, contains the space Z, and is not equal to Z. This fact contradicts the maximality of the space Z.

Therefore $Y \oplus Z = X$ and we are done. □

3.3 The Hahn-Banach Theorem

The following result is called the *Hahn-Banach theorem* for real normed vector spaces. It's proof uses Zorn's lemma.

Theorem 3.13 *Let V be a real normed vector space, and let $W \leq V$ be a subspace. Let $f \in W^*$ be a linear functional. Then there is an extension $F \in V^*$ such that $\|F\| = \|f\|$.*

Proof: The proof proceeds by showing that we can extend the linear functional f one dimension at a time and then applying Zorn's lemma. Suppose without loss of generality that $\|f\| = 1$.

Let Z be a vector space containing W and let $g\colon Z \to \mathbb{R}$ be an extension of the linear functional f such that $\|g\| = 1$.

Suppose that $Z \neq X$. Choose a point $x_1 \in X \backslash Z$ and form the space

$$Z_1 = Z \oplus \langle x_1 \rangle$$

Let $x, y \in Z$. Then:

$$g(x) + g(y) = g(x+y) \leq \|x+y\| \leq \|x - x_1\| + \|x_1 + y\|$$

so we have the inequality

$$g(x) - \|x - x_1\| \leq \|y + x_1\| - g(y)$$

We can therefore find a real number β such that:

$$\sup_{x \in Z}(g(x) - \|x - x_1\|) \leq \beta \leq \inf_{x \in Z}(\|x + x_1\| - g(x))$$

Define a linear functional $g_1 \colon Z_1 \to \mathbb{R}$ by the formula

$$g_1(x + \alpha x_1) = g(x) + \alpha\beta$$

for all vectors $x \in X$ and scalars $\alpha \in \mathbb{R}$. We claim that $g_1(x+\alpha x_1) \leq \|x+\alpha x_1\|$. There are three cases to consider.

- $\alpha = 0$:
$$g_1(x + \alpha x_1) = g(x) \leq \|x\| = \|x + \alpha x_1\|$$

- $\alpha > 0$:
$$g_1(x + \alpha x_1) = \alpha(g(\alpha^{-1}x) + \beta) \leq \alpha\|\alpha^{-1}x + x_1\| = \|x + \alpha x_1\|$$

- $\alpha < 0$:
$$g_1(x + \alpha x_1) = -\alpha(g(-\alpha^{-1}x) + \beta) \leq -\alpha\| - \alpha^{-1}x - x_1\| = \|x + \alpha x_1\|$$

Replacing the vector $x + \alpha x_1$ by $-(x + \alpha x_1)$ we see that $|g_1(x + \alpha x_1)| \leq \|x + \alpha x_1\|$. Thus $\|g_1\| \leq 1$, and g is an extension of g. Therefore we also have $\|g\| \geq \|g_1\| = 1$, so $\|g\| = 1$. This completes the first step of the proof.

Now let P be the set of pairs (Z, g) such that Z is a subspace of V that contains W, and $g \colon Z \to \mathbb{R}$ is an extension of the linear functional f such that $\|g\| = 1$. Define a partial ordering on the set P by writing $(Z, g) \leq (Z', g')$ if the set Z' contains the set Z and the linear functional g' extends the linear functional g.

Let $T = \{(Z_\lambda, g_\lambda) \mid \lambda \in \Lambda\}$ be a totally ordered subset of the set P. Then we can form a vector space

$$Z = \bigcup_{\lambda \in \Lambda} Z_\lambda$$

and a linear functional $g \colon Z \to \mathbb{R}$ defined by the formula $g(x) = g_\lambda(x)$ whenever $x \in Z_\lambda$. The pair (Z, g) is an upper bound for the set T, so by Zorn's lemma the set P has a maximal element, (M, F).

Suppose that $M \neq V$. Then by the first step of this proof we have a pair (M_1, F_1) such that $(M, F) \leq (M_1, F_1)$ and $(M, F) \neq (M_1, F_1)$. This contradicts maximality of the element (M, F).

Therefore $M = V$ and we have an extension $F \colon X \to \mathbb{R}$ such that $\|F\| = 1$. □

The complex version of the Hahn-Banach theorem follows as a simple corollary.

Corollary 3.14 *Let V be a complex normed vector space, and let $W \leq V$ be a subspace. Let $f \in W^*$ be a linear functional. Then there is an extension $F \in V^*$ such that $\|F\| = \|f\|$.*

Proof: The spaces V and W can be considered real vector spaces. We have a real-valued linear functional $g\colon W \to \mathbb{R}$ defined by writing $g(x) = \Re(f(x))$ for all points $x \in W$. By the previous version of the Hahn-Banach theorem there is a real-valued linear functional $G\colon X \to \mathbb{R}$ that extends g, with norm $\|G\| = \|g\|$.

Define a complex-valued linear functional $F\colon X \to \mathbb{C}$ by the formula

$$F(x) = G(x) - iG(ix)$$

Then the functional F is an extension of f, with norm $\|F\| = \|f\|$. $\square$

3.4 The Stone-Weierstrass Theorem

Let X be a compact metric space. Recall that we define $C(X)$ to be the algebra of continuous functions $f\colon X \to \mathbb{F}$. If we want to be careful and specify the real or the complex case, we write this algebra $C_{\mathbb{R}}(X)$ or $C_{\mathbb{C}}(X)$ respectively. Addition and multiplication are defined pointwise.

We have already seen that the algebra $C(X)$ is in fact a Banach space, with the norm

$$\|f\| = \sup\{|f(x)| \mid x \in X\}$$

Thus, $C(X)$ has the structure of a vector space, with pointwise addition and scalar multiplication of functions.

Further, given functions $f, g \in C(X)$, we can define the product fg by writing

$$fg(x) = f(x)g(x) \qquad x \in X.$$

Definition 3.15 *We call a non-empty subset $A \subseteq C(X)$ a* subalgebra *if:*

- $\alpha f + \beta g \in A$ *whenever* $f, g \in A$ *and* $\alpha, \beta \in \mathbb{F}$.
- $fg \in A$ *whenever* $f, g \in A$.

We call a subalgebra $A \subseteq C(X)$ unital *if the constant function,* 1 *(defined by writing $x \mapsto 1$ for all $x \in X$) is an element of A.*

We say that a subalgebra $A \subseteq C(X)$ seperates points *if for any two points $x, y \in X$ with $x \neq y$ there is a function $f \in C(X)$ such that $f(x) \neq f(y)$.*

Thus a subalgebra of $C(X)$ is a linear subspace that is closed under the product.

Lemma 3.16 *Let A be a closed unital subalgebra of the real algebra $C_{\mathbb{R}}(X)$. Let $f, g \in A$. Then the functions* $\max(f, g)$ *and* $\min(f, g)$*, defined by the formulae*

$$\max(f, g)(x) = \max(f(x), g(x)) \qquad \min(f, g)(x) = \min(f(x), g(x))$$

respectively, belong to the algebra A.

Proof: Since $2\min(f,g) = |f-g| + (f+g)$ and $-\min(f,g) = \max(-f,-g)$ it suffices to show that $|f| \in A$ whenever $f \in A$. By multiplying by a constant, let us assume that $\|f\| < 1$.

Let $g(t) = (1-t)^{\frac{1}{2}}$. By looking at the Taylor series of the function $g(t)$, given $\varepsilon > 0$ there is a polynomial $p(t)$ such that

$$|g(t) - p(t)| < \varepsilon$$

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

Now let $t = 1 - f(x)^2$. Then $t \in [0,1]$, and $g(t) = (1-t)^{\frac{1}{2}} = |f(x)|$. We therefore have the inequality

$$|\ |f(x)| - p(1 - f(x)^2)| < \varepsilon$$

for all $x \in X$. Hence $\|\ |f| - p(1-f^2)\| \leq \varepsilon$.

Since the set A is a unital algebra, and p is a polynomial, the function $p(1-f^2)$ belongs to A. Since the set A is closed, it follows that $|f| \in A$ and we are done. □

The following handy result is known as the *Stone-Weierstrass theorem*

Theorem 3.17 *Let $A \subseteq C_{\mathbb{R}}(X)$ be a unital subalgebra that separates points. Then A is a dense subset of $C_{\mathbb{R}}(X)$.*

Proof: Note that if $A \subseteq C_{\mathbb{R}}(X)$ is a unital subalgebra that seperates points, then the closure $\overline{A}$ is a closed unital subalgebra that seperates points. Further, if A is dense, then $\overline{A} = C_{\mathbb{R}}(X)$. Thus it suffices to prove that if A is a closed unital subalgebra that seperates points, then $A = C_{\mathbb{R}}(X)$.

Consider a function $f \in C_{\mathbb{R}}(X)$, and a real number $\varepsilon > 0$. We will find a function $g \in A$ such that $\|f - g\| \leq \varepsilon$.

Since the algebra A is unital and separates points, for any two points $x, y \in X$ with $x \neq y$ we can find a function $g_{x,y} \in A$ such that

$$g_{x,y}(y) = f(y) + \frac{\varepsilon}{2} \qquad g_{x,y}(x) > f(x)$$

For each point $x \in X$ there is an open neighbourhood $U_x \ni x$ such that $g_{x,y}(z) > f(z)$ for all $z \in U_x$. Since the space X is compact, we have a finite cover of the form $\{U_{x_1}, \ldots, U_{x_m}\}$.

Define $g_y = \max(g_{x_1}, \ldots, g_{x_m})$. Then

$$g_y(y) = f(y) + \frac{\varepsilon}{2} \qquad g_y > f$$

and $g_y \in A$ by lemma 3.16. For each point $y \in X$ there is an open neighbourhood $V_y \ni y$ such that $g_y(z) < f(z) + \varepsilon$ for all $z \in V_y$. Since the space X is compact, we have a finite cover of the form $\{V_{y_1}, \ldots, V_{y_n}\}$.

We can define a function $g = \min(g_{y_1}, \ldots, g_{y_n})$. By construction we know that $f \leq g \leq g + \varepsilon$, and so $\|f - g\| \leq \varepsilon$. But lemma 3.16 tells us that $g \in A$ so we are done. □

To formulate the complex version of this result, note that we have an operation $f \mapsto f^*$ on the space $C_{\mathbb{C}}(X)$ defined by writing $f^*(x) = \overline{f(x)}$ for all $x \in X$.

Definition 3.18 *We call a subalgebra $A \subseteq C_{\mathbb{C}}(X)$ a* $*$-subalgebra *if $f^* \in A$ whenever $f \in A$.*

The complex version of the Stone-Weierstrass theorem is then as follows.

Theorem 3.19 *Let $A \subseteq C_{\mathbb{C}}(X)$ be a unital $*$-subalgebra that separates points. Then A is a dense subset of $C_{\mathbb{C}}(X)$.* □

The classical Weierstrass approximation theorem follows as a corollary.

Corollary 3.20 *Let f be a continuous function on the interval $[0,1]$. Let $\varepsilon > 0$. Then there is a polynomial p such that $|f(t) - p(t)|$ for all $t \in [0,1]$.*

Proof: Let A be the set of polynomials on the interval $[0,1]$. Clearly, A is a $*$-subalgebra of the space $C[0,1]$, and $1 \in A$ so A is unital.

Let $x, y \in [0,1]$ with $x \neq y$. We have a polynomial defined by writing $p(t) = t$ for all $t \in [0,1]$; certainly $p(x) \neq p(y)$. So the subalgebra A seperates points.

Hence, given $\varepsilon > 0$ and a continuous function $f \in C[0,1]$, we have a polynomial p such that $\|f - p\| < \varepsilon$. By definition of the norm on $C[0,1]$, this means that $|f(t) - p(t)| < \varepsilon$ for all $t \in [0,1]$. □

Example 3.21 *Let*

$$\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$$

Then by the same argument as above, the set of all polynomials in z is a dense subset of $C_{\mathbb{C}}(\mathbb{T})$.

Chapter 4

Hilbert Spaces

4.1 Inner Products

If norms provide a measure of distance within a vector space, inner products provide a measure of angle. In the following definition, given a scalar $\alpha \in \mathbb{F}$, when we write $\overline{\alpha}$, we mean the complex conjugate. So, if $x, y \in \mathbb{R}$, then $\overline{x + iy} = x - iy$. If $\alpha \in \mathbb{R}$, then obviously $\overline{\alpha} = \alpha$.

Definition 4.1 *Let V be a vector space over the field $\mathbb{F}$. An* inner product *on V is a map*

$$\langle -, - \rangle \colon V \times V \to \mathbb{F}$$

such that:

- *Let $u, v, w \in V$, and $\alpha, \beta \in \mathbb{F}$. Then $\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle$.*
- *Let $u, v \in V$. Then $\langle u, v \rangle = \overline{\langle v, u \rangle}$.*
- *Let $v \in V$. Then $\langle v, v \rangle \in \mathbb{R}^{\geq 0}$, and $\langle v, v \rangle = 0$ if and only if $v = 0$.*

We call a vector space V equipped with an inner product an inner product space.

Example 4.2 *Let $x, y \in \mathbb{C}^n$. Write $x = (u_1, \ldots, u_n)$ and $y = (v_1, \ldots, v_n)$. Then we can define an inner product on the space $\mathbb{C}^n$ by the formula*

$$\langle x, y \rangle = \overline{u_1} v_1 + \cdots + \overline{u_n} v_n.$$

- *Let $x, y, z \in \mathbb{C}^n$, and $\alpha, \beta \in \mathbb{C}$. Write $x = (u_1, \ldots, u_n)$, $y = (v_1, \ldots, v_n)$, and $z = (w_1, \ldots, w_n)$. Then*

$$\begin{array}{rcl} \langle x, \alpha y + \beta z \rangle & = & \overline{u_1}(\alpha v_1 + \beta w_1) + \cdots + \overline{u_n}(\alpha v_n + \beta w_n) \\ & = & \alpha(\overline{u_1} v_1 + \cdots + \overline{u_n} v_n) + \beta(\overline{u_1} w_1 + \cdots + \overline{u_n} w_n) \\ & = & \alpha \langle x, y \rangle + \beta \langle x, z \rangle. \end{array}$$

- *Let* $x = (u_1, \ldots, u_n)$ *and* $y = (v_1, \ldots, v_n) \in \mathbb{F}^n$. *Then*

$$\begin{aligned} \langle x, y \rangle &= \overline{u_1} v_1 + \cdots + \overline{u_n} v_n \\ &= \overline{\overline{v_1} u_1 + \cdots + \overline{v_n} u_n} \\ &= \overline{\langle y, x \rangle}. \end{aligned}$$

- *Let* $x = (u_1, \ldots, u_n) \in \mathbb{F}^n$. *Then*

$$\langle x, x \rangle = |u_1|^2 + \cdots + |u_n|^2.$$

Observe $\langle x, x \rangle \in \mathbb{R}^{\geq 0}$. *Each of the terms* $|u_i|^2$ *is non-negative. So if* $\langle x, x \rangle = 0$, *we must have* $|u_i|^2 = 0$ *for all* i, *meaning* $x = 0$.

We call the above the *standard inner product* on $\mathbb{C}^n$.

Now, suppose we consider elements of the space $\mathbb{C}^n$ to be column vectors. Thus, given $x, y \in \mathbb{C}^n$, we write $x = (u_1, \ldots, u_n)^T$ and $y = (v_1, \ldots, v_n)^T$. Then by definition of matrix multiplication

$$\langle x, y \rangle = \overline{x}^T y.$$

This idea can be a useful way of viewing the standard inner product for calculations.

Definition 4.3 *Let* V *be an inner product space, and let* $v \in V$. *Then we write*

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

We shall see soon that this formula defines a norm, as our notation suggests.

Theorem 4.4 (The Cauchy-Schwarz inequality) *Let* V *be an inner product space. Let* $u, v \in V$. *Then*

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\|.$$

Proof: If $v = 0$, then clearly $\langle u, v \rangle = 0$, and $\|v\| = 0$, so the result holds. So suppose $v \neq 0$. Then $\langle v, v \rangle \neq 0$.

For any scalar $\alpha \in \mathbb{F}$, we have

$$\langle u - \alpha v, u - \alpha v \rangle \geq 0.$$

But

$$\langle u - \alpha v, u - \alpha v \rangle = \langle u, u \rangle - \alpha \langle u, v \rangle \overline{\alpha} \overline{\langle u, v \rangle} + |\alpha|^2 \langle v, v \rangle.$$

Set

$$\alpha = \frac{\overline{\langle u, v \rangle}}{\langle v, v \rangle}.$$

Then we see

$$\langle u, u \rangle - \frac{|\langle u, v \rangle|^2}{\langle v, v \rangle} \geq 0.$$

Rearranging, we see

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle \cdot \langle v, v\rangle.$$

Taking the square root, the desired result follows. □

Corollary 4.5 *Let V be an inner product space. Then $\| - \|$ is a norm on V.*

Proof: Certainly $\|v\| \geq 0$ for all $v \in V$. We simply check the axioms for a norm.

- Let $\alpha \in \mathbb{F}$ and $v \in V$. Then

  $$\|\alpha v\|^2 = \langle \alpha v, \alpha v\rangle = |\alpha|^2 \langle v, v\rangle.$$

  so $\|\alpha v\| = |\alpha| \|v\|$.

- Let $u, v \in V$. Then, by the Cauchy-Schwarz inequality

  $$\begin{array}{rcl} \|u+v\|^2 & = & \|u\|^2 + \langle u, v\rangle + \overline{\langle u, v\rangle} + \|v\|^2 \\ & = & \|u\|^2 + 2\Re(\langle u, v\rangle) + \|v\|^2 \\ & \leq & \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \\ & \leq & \|u\|^2 + 2\|u\| \cdot \|v\| + \|v\|^2 \\ & = & (\|u\| + \|v\|)^2 \end{array}.$$

  Taking square roots, we have the triangle inequality.

- Let $v \in V$. If $v = 0$, then certainly $\|v\| = \sqrt{\langle v, v\rangle} = 0$. Conversely, if $\|v\| = 0$, then $\langle v, v\rangle = 0$, so $v = 0$.

□

Definition 4.6 *Let V be a normed vector space. Then we say the norm satisfies the* parallelogram law *if*

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

for all $u, v \in V$.

The following result is left as an exercise.

Proposition 4.7 *Let V be a vector space over the field $\mathbb{F}$. A norm on the space V comes from an inner product if and only if it satisfies the parallelogram law. Further, the inner product producing the norm is unique.* □

Since an inner product space is also a normed vector space, we can talk about continuity of functions on inner product spaces. The following result tells us that inner products are continuous.

Proposition 4.8 *Let V be an inner product space. Let (u_n) and (v_n) be sequences in V, with norm limits u and v respectively. Then*

$$\lim_{n\to\infty} \langle u_n, v_n\rangle = \langle u, v\rangle.$$

Proof: Observe, by the Cauchy-Schwarz inequality, that

$$\begin{aligned} |\langle u_n, v_n\rangle - \langle u, v\rangle| &= |\langle u_n, v_n\rangle - \langle u, v_n\rangle + \langle u, v_n\rangle - \langle u, v\rangle| \\ &\le |\langle u_n - u, v_n\rangle| + |\langle u, v_n - v\rangle| \\ &\le \|v_n\| \cdot \|u_n - u\| + \|u\| \cdot \|v_n - v\|. \end{aligned}$$

By convergence of the sequences (u_n) and (v_n), we have $\|u_n - u\| \to 0$ and $\|v_n - v\| \to 0$ as $n \to \infty$. By continuity of the norm, $\|v_n\| \to \|v\|$ as $n \to \infty$. Hence

$$\lim_{n\to\infty} |\langle u_n, v_n\rangle - \langle u, v\rangle| = 0$$

and we are done. □

Definition 4.9 *An inner product space that is complete (as a normed vector space) is called a* Hilbert space.

Example 4.10 *Recall that we define l^2 to be the set of sequences(α_n) in the field $\mathbb{F}$ such that*

$$\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty.$$

We claim that the space l^2 is a Hilbert space, with inner product defined by the formula

$$\langle(\alpha_n), (\beta_n)\rangle = \sum_{n=1}^{\infty} \overline{\alpha_n}\beta_n.$$

The first step is to show that the above series converges. Given $(\alpha_n) \in l^2$, let us define

$$\|\alpha\| = \left(\sum_{n=1}^{\infty} |\alpha_n|^2\right)^{\frac{1}{2}}.$$

By example 4.2, the space $\mathbb{F}^N$ is an inner product space, and the Cauchy-Schwarz inequality for $\mathbb{F}^N$ tells us that

$$\sum_{n=1}^{N} |\overline{\alpha_n}\beta_n| = \sum_{n=1}^{N} \overline{|\alpha_n|}|\beta_n| \le \left(\sum_{n=1}^{N} |\alpha_n|^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^{N} |\beta_n|^2\right)^{\frac{1}{2}} \le \|(\alpha_n)\|\|(\beta_n)\|.$$

Hence the series $\sum_{n=1}^{\infty} \overline{\alpha_n}\beta_n$ converges absolutely, and therefore converges.

Therefore the formula

$$\langle(\alpha_n),(\beta_n)\rangle = \sum_{n=1}^{\infty} \overline{\alpha_n}\beta_n$$

is well-defined. We can check it defines an inner product as in example 4.2.

We have associated norm given by the above formula

$$\|\alpha\| = \left(\sum_{n=1}^{\infty} |\alpha_n|^2\right)^{\frac{1}{2}} \qquad (\alpha_n) \in l^2.$$

We saw in the exercises at the end of chapter 1 that the space l^2, with this norm, is complete. Hence the space l^2 is a Hilbert space.

Definition 4.11 *Let V_1 and V_2 be inner product spaces. An* isomorphism *$T\colon V_1 \to V_2$ is then a bijective linear map such that*

$$\langle Tx, Ty\rangle = \langle x, y\rangle$$

for all $x, y \in V_1$.

Observe that any isomorphism between inner product spaces is an isometry. The following definition is similar to the corresponding definition for normed vector spaces.

Definition 4.12 *Let V be an inner product space. We call a Hilbert space $\hat{V}$ a* completion *of V if there is an isomorphism $i\colon V \to W$ where W is a dense subset of $\hat{V}$.*

The first part of the following result is proved similarly to theorem 1.34. The second part is similar to proposition 1.35.

Theorem 4.13 *Any inner product space has a completion. Further, any two completions of the same inner product space are isomorphic* □

Example 4.14 *It is shown in the exercises that the space $C[0,1]$ can be equipped with an inner product defined by the formula*

$$\langle f, g\rangle = \int_0^1 \overline{f(t)}g(t)\ dt.$$

The inner product defines a norm

$$\|f\| := \sqrt{\langle f, f\rangle} = \left(\int_0^1 |f(t)|^2\ dt\right)^{\frac{1}{2}}.$$

Thus the space $C[0,1]$ is not complete by proposition **??**. *However, we can form the completion, which* is *a Hilbert space. We call this Hilbert space $L^2[0,1]$.*

4.2 Orthogonal Complements

Definition 4.15 *Let H be a Hilbert space. Let S be a subset of H. Then we define the* orthogonal complement *of S:*

$$S^{\perp} = \{v \in H \mid \langle x, v\rangle = 0 \text{ for all } x \in S\}$$

It is left as an exercise to show that

$$\|x + v\|^2 = \|x\|^2 + \|v\|^2$$

whenever $x \in S$ and $v \in S^{\perp}$.

Proposition 4.16 *The orthogonal complement $S^{\perp}$ is a closed linear subspace of H.*

Proof: It is left as an exercise to show that $S^{\perp}$ is a linear subspace of H. But we will check it is closed here.

To see this, let $v \in \overline{S^{\perp}}$. Then there is a sequence, (v_n), in $S^{\perp}$, converging in norm to v.

Let $x \in S$. Then $\langle x, v_n\rangle = 0$ for all n. By continuity of the inner product, if we let $n \to \infty$, we see $\langle x, v\rangle = 0$.

It follows that $v \in S^{\perp}$. So $\overline{S^{\perp}} = S^{\perp}$ and we are done. □

The proof of the following result is again left as an exercise.

Proposition 4.17 *Let H be a Hilbert space, and let $V \subseteq H$ be a subspace. Then $(\overline{V})^{\perp} = V^{\perp}$.* □

We call a vector space V the *direct sum* of two subspaces V_1 and V_2 if $V_1 \cap V_2 = \{0\}$, and for every element $v \in v$, we can write $v = v_1 + v_2$ where $v_1 \in V_1$ and $v_2 \in V_2$. We write $V = V_1 \oplus V_2$ when V is the direct sum of V_1 and V_2.

Our main aim in the rest of this section is to prove that given a closed subspace, V, of a Hilbert space H, we have $H = V \oplus V^{\perp}$. We call this result the *projection theorem*.

The following lemma, which is needed in the proof, is sometimes called the *nearest point theorem*.

Lemma 4.18 *Let H be a Hilbert space, let $x \in H$, and let $C \subseteq H$ be a closed convex subset. Set*

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

Then there is a unique point $y \in C$ such that $\|x - y\| = d(x, C)$.

Proof: Let $d = d(x, C)$. Then we have a sequence, (y_n), in C such that $\|x - y_n\| \to d$ as $n \to \infty$. Set $z_n = x - y_n$. Then $\|z_n\| \geq d$ for all n, and $\|x - y_n\| \to d$ as $n \to \infty$.

By the parallelogram law

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

But

$$\|z_n + z_m\|^2 = \|2x - y_m - y_n\|^2 = 4\|x - \frac{1}{2}(y_m + y_n)\|^2$$

and $\frac{1}{2}(y_m + y_n) \in C$ because the set C is convex and $y_m, y_n \in C$. Hence

$$\|y_m - y_n\|^2 \leq 2(\|z_m\|^2 + \|z_n\|^2) - 4d^2 \to 0$$

as $m, n \to \infty$.

It follows that the sequence (y_n) is Cauchy. Hence, as H is a Hilbert space, and so complete, there is a point $y \in H$ such that $y_n \to y$ as $n \to \infty$.

But by hypothesis, the set C is closed. So $y \in C$. Now, by definition of the sequence (y_n), we had $\|x - y_n\| \to d$ as $n \to \infty$. Hence, by continuity of the norm, we have $\|x - y\| = d$, which is exactly what we need.

Uniqueness of the point y is left as an exercise. □

Note in particular that any linear subspace of H is convex. Then given $x \in H$, and a closed linear subspace, V, of H, we have $y \in V$ such that $\|x - y\| = d(x, V)$.

Theorem 4.19 *Let H be a Hilbert space. Let $V \subseteq H$ be a closed linear subspace. Then $H = V \oplus V^\perp$.*

Proof: Firstly, observe that V and $V^\perp$ are both subspaces of H. Let $v \in V \cap V^\perp$. Then $\langle v, v\rangle = 0$, which implies $v = 0$. Thus $V \cap V^\perp = \{0\}$.

Now, let $x \in H$. Let $d = d(x, V)$. By the above lemma, we have a unique point $y \in V$ such that $\|x - y\| = d$. We claim that $v - y \in V^\perp$. If this claim holds, then

$$x = y + (x - y) \qquad y \in V,\ x - y \in V^\perp$$

and we are done.

Suppose $x - y \notin V^\perp$. Then we have a point $v \in V$ such that $\langle x - y, v\rangle = \alpha \neq 0$.

Let $z = x - y$ and let $\lambda \in \mathbb{F}$. Observe

$$\begin{array}{rcl} \|z - \lambda v\|^2 & = & \|z\|^2 - \lambda\langle z, v\rangle - \overline{\lambda\langle z, v\rangle} + |\lambda|^2\|v\|^2 \\ & = & d^2 - \lambda\alpha - \overline{\lambda\alpha} + |\lambda|^2\|v\|^2 \end{array}.$$

Let

$$\lambda = \frac{\overline{\alpha}}{\|v\|^2}.$$

Then $\|z - \lambda v\|^2 = d^2 - 2|\alpha|^2$, so

$$\|z - \lambda v\| = \|x - (y + \lambda v)\| < d.$$

But $y + \lambda v \in V$, and $d = d(x.V) = \inf\{\|x - w\| \mid w \in V\}$. So the above is a contradiction. It follows that $x - y \in V^\perp$.

This observation completes the proof. □

Thus, if H is a Hilbert space, and $V \subseteq H$ is a closed linear subspace, an element $x \in H$ can be written uniquely as a sum $x = v + w$ where $v \in V$ and $w \in V^\perp$. We can therefore define a map $P\colon H \to H$ by writing $P(x) = v$.

We call the map P the *orthogonal projection* onto V. The above theorem guarantees the existence of the map P, which is why we call it the projection theorem.

The following result is left as an exercise.

Proposition 4.20 *The map P is a bounded linear map with the property $P^2 = P$.* □

Corollary 4.21 *Let H be a Hilbert space, and let V be a linear subspace of H. Then $(V^\perp)^\perp = \overline{V}$.*

Proof: Let $v \in V$. Then for any vector $w \in V^\perp$, we have $\langle w, v\rangle = 0$. Hence $v \in (V^\perp)^\perp$. As the space $(V^\perp)^\perp$ is closed by proposition 4.16, it follows that $\overline{V} \subseteq (V^\perp)^\perp$.

Conversely, let $x \in (V^\perp)^\perp$. By proposition 4.17, we have $(\overline{V})^\perp = V^\perp$. Hence, by the projection theorem, we can write

$$H = \overline{V} \oplus V^\perp.$$

Hence $x = v + w$, where $v \in \overline{V} \subseteq (V^\perp)^\perp$, and $w \in V^\perp$. It follows that $w = x - v \in (V^\perp)^\perp$.

But $w \in V^\perp$, and certainly $V^\perp \cap (V^\perp)^\perp = \{0\}$. Therefore $w = 0$, and $x = v \in \overline{V}$.

We have therefore shown that $(V^\perp)^\perp \subseteq \overline{V}$, which completes the proof. □

In particular, if V is a closed subspace of H, then $(V^\perp)^\perp = H$.

Corollary 4.22 *Let H be a Hilbert space, and let V be a linear subspace of H. Then V is dense in H if and only if $V^\perp = \{0\}$.*

Proof: Let V be dense in H. Then $\overline{V} = H$. By proposition 4.17, we have $(\overline{V})^\perp = V^\perp$. Hence, by the projection theorem, we can write

$$H = \overline{V} \oplus V^\perp.$$

It follows that $V^\perp = \{0\}$.

Conversely, suppose $V^\perp = \{0\}$. Then $(V^\perp)^\perp = H$. But by the above corollary, we have $(V^\perp)^\perp = \overline{V}$. Thus $\overline{V} = H$, and we are done. □

4.3 Dual Spaces

Let H be a Hilbert space. Let $v \in H$. Then we can define a map $J_v \colon H \to \mathbb{F}$ by the formula

$$J_v(x) = \langle v, x \rangle.$$

Proposition 4.23 *The map J_v is a bounded linear map, with norm $\|J_v\| = \|v\|$.*

Proof: It is easy to see that the map J_v is linear. By the Cauchy-Schwarz inequality, for any point $x \in H$ we have

$$|J_v(x)| \leq \|v\| \cdot \|x\|$$

which tells us that the operator J_v is bounded, with norm $\|J_v\| \leq \|v\|$.

Observe

$$|J_v(v)| = \langle v, v \rangle = \|v\|^2$$

so $\|J_v\| \geq \|v\|$. We conclude that $\|J_v\| = \|v\|$. □

Thus the map J_v is an element of the dual space, H^*

We call a map $T \colon V \to W$ between two vector spaces over the field $\mathbb{F}$ *conjugate-linear* if

$$T(\alpha u + \beta v) = \overline{\alpha} T(u) + \overline{\beta} T(v)$$

for all $\alpha, \beta \in \mathbb{F}$ and $u, v \in V$. Observe that if $\mathbb{F} = \mathbb{R}$, conjugate-linearity is the same as linearity.

The proof of the following is an exercise.

Proposition 4.24 *Let H be a Hilbert space. Then we have a conjugate-linear isometry $J \colon H \to H^*$ defined by writing $J(v) = J_v$, that is*

$$J(v)(x) = \langle v, x \rangle \qquad v \in H,\ x \in H.$$

□

The following result is called the *Riesz representation theorem*; it tells us that the above map is surjective.

Theorem 4.25 *Let $f \colon H \to \mathbb{F}$ be a bounded linear map. Then we have a unique vector $R_f \in H$ such that*

$$f(x) = \langle R_f, x \rangle$$

for all $x \in H$.

Proof: Since f is a continuous map, $\ker f$ is closed. If $f = 0$, the result is trivial. Otherwise, we can find $u \in H$ such that $f(u) = 1$.

By the projection theorem, we can write

$$u = v + w \qquad v \in \ker f,\ w \in (\ker f)^\perp.$$

Observe that $f(u) = f(w)$, so $f(w) = 1$. More generally, let $x \in H$. Then

$$f(x - f(x)w) = f(x) - f(x)f(w) = 0$$

so $x - f(x)w \in \ker f$. It follows that $\langle w, x - f(x)w \rangle = 0$, so

$$\langle w, x \rangle = f(x)\langle w, w \rangle.$$

Set $R_f = w/\|w\|^2$. Then by the above

$$\langle R_f, x \rangle = f(x)$$

and we have established existence of the vector R_f.

Let $R'_f \in H$ be another vector where $f(x) = \langle R'_f, x \rangle$ for all $x \in H$. Let $x = R_f - R'_f$. Then

$$\langle x, x \rangle = \langle R_f, x \rangle - \langle R'_f, x \rangle = 0.$$

Therefore $x = 0$, which tells us that $R_f = R'_f$ and establishes uniqueness. □

We leave the following two results as exercises.

Proposition 4.26 *We have a conjugate-linear isometry $R \colon H^* \to H$ defined by the formula $f \mapsto R_f$.* □

Proposition 4.27 *The above map R is the inverse of the map $J \colon H \to H^*$.*

4.4 Adjoints

Theorem 4.28 *Let H be a Hilbert space, and let $T \colon H \to H$ be a bounded linear map. Then there is a unique map $T^* H \to H$ such that*

$$\langle x, Ty \rangle = \langle T^* x, y \rangle$$

for all $x, y \in H$.

Proof: Let $x \in H$. Then we have a bounded linear map $y \mapsto \langle x, Ty \rangle$. Hence, by the Riesz representation theorem, we have a unique element $T^* x \in H$ such that

$$\langle T^* x, y \rangle = \langle x, Ty \rangle$$

for all $y \in H$. □

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

Example 4.29 *Let $I \colon H \to H$ be the identity map. Then for all $x, y \in H$, we have*

$$\langle Ix, y \rangle = \langle x, y \rangle = \langle x, Iy \rangle.$$

Hence $I^* = I$.

Some general properties follow immediately from the definition. The precise details of the following two results are left as exercises.

Proposition 4.30 *The map $T^*: H \to H$ is a bounded linear map, with norm $\|T^*\| = \|T\|$.* □

Proposition 4.31 *Let $S, T: H \to H$ be bounded linear maps, and let $\alpha, \beta \in \mathbb{F}$. Then:*

- $(\alpha S + \beta T)^* = \overline{\alpha} S^* + \overline{\beta} T^*$.
- $(S^*)^* = S$.
- $(ST)^* = T^* S^*$.

□

Example 4.32 *Let $A = (a_{ij})$ be an $n \times n$ complex matrix. Then A defines a bounded linear map $\mathbb{C}^n \to \mathbb{C}^n$ by matrix multiplication on a column vector.*

Thus, if we consider column vectors $u, v \in \mathbb{C}^n$, then

$$\langle u, Av \rangle = \overline{u}^T Av = \overline{(\overline{A}^T u)^T} v = \langle \overline{A}^T u, v \rangle.$$

Therefore $A^* = \overline{A}^T$.

Example 4.33 *Define bounded linear operators $R, L: l^2 \to l^2$ by writing*

$$R(a_1, a_2, a_3, \ldots) = (0, a_1, a_2, a_3, \ldots)$$

and

$$L(a_1, a_2, a_3, \ldots) = (a_2, a_3, \ldots)$$

respectively.

Observe

$$\langle L(a_1, a_2, a_3, \ldots), (b_1, b_2, b_3, \ldots) \rangle = a_2 b_1 + a_3 b_2 + \cdots$$

and

$$\langle (a_1, a_2, a_3, \ldots), R(b_1, b_2, b_3, \ldots) \rangle = a_2 b_1 + a_3 b_2 + \cdots$$

We conclude that $R^* = L$. Further, $L^* = (R^*)^* = R$.

Definition 4.34 *Let H and H' be Hilbert spaces. We call a bounded linear operator $U: H \to H'$* unitary *if $U^* U = I_H$ and $UU^* = I_{H'}$.*

Example 4.35 *A unitary $n \times n$ matrix, U, defines a unitary operator $U: \mathbb{C}^n \to \mathbb{C}^n$.*

The proof of the following result is left as an exercise.

Proposition 4.36 *Let $T: H \to H'$ be an invertible bounded linear map. Then the map T is a unitary if and only if*

$$\langle Tu, Tv \rangle = \langle u, v \rangle$$

for all $u, v \in H$. □

Chapter 5

Orthonormal Sets

5.1 Orthonormal Sets and Bases

Let H be a Hilbert space. Then we call two elements $x, y \in H$ *orthogonal* if $\langle x, y = 0$. It is easy to see that if x and y are orthogonal, then

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

The following notion is a generalisation of this idea.

Definition 5.1 *Let H be a Hilbert space. We call a subset $M \subset H$* orthogonal *if $\langle x, y\rangle = 0$ whenever $x, y \in M$ and $x \neq y$.*

An orthonormal set *is an orthogonal set where each element has norm one.*

Thus a subset $M \subseteq H$ is orthonormal if for elements $x, y \in H$, we have

$$\langle x, y\rangle = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}.$$

The following result lets us construct orthonormal sets from orthogonal sets. It is easy to check.

Proposition 5.2 *Let M be an orthogonal set. Let*

$$M' = \{v/\|v\| \mid v \in M\}.$$

Then the set M' is orthonormal. □

Example 5.3 *Let $H = L^2[0, 2\pi]$. Set*

$$f_n(t) = \sin(nt)$$

where $n \in \mathbb{N}$.

Now, let $m, n \in \mathbb{N}$, with $m \neq n$. Without loss of generality, suppose that $n > m$.

Observe

$$\langle f_m(t), f_n(t)\rangle = \int_0^{2\pi} \sin(mt)\sin(nt)\,dt = \frac{1}{2}\int_0^{2\pi} \cos((n-m)t)\,dt - \frac{1}{2}\int_0^{2\pi} \cos((n+m)t)\,dt = 0.$$

Therefore the set $\{f_n \mid n \in \mathbb{N}\}$ *is orthogonal.*
Now, observe

$$\|f\|^2 = \int_0^{2\pi} \sin^2(nt)\,dt = \frac{1}{2}\int_0^{2\pi} 1 - \cos(2nt)\,dt = \pi.$$

Set

$$g_n(t) = \frac{1}{\sqrt{\pi}}\sin(nt).$$

Then the set $\{g_n \mid n \in \mathbb{N}\}$ *is orthonormal.*

Now, we call a sequence of vectors, (e_n), in a Hilbert space H an *orthonormal sequence* if $\|e_n\| = 1$ for all n, and $\langle e_m, e_n\rangle = 0$ if $m \neq n$. Thus, in the above example, we have an orthonormal sequence, (g_n), in the space $L^2[0, 2\pi]$.

The proof of the following is left as an exercise.

Proposition 5.4 *Any orthogonal set of vectors is linearly independent.* □

The following result (or rather, its proof) tells us how to construct an orthonormal sequence out of a linearly independent sequence. It is called the *Gram-Schmidt orthonormalisation process.*

Theorem 5.5 *Let H be a Hilbert space. Let $\{v_n \mid n \in \mathbb{N}\}$ be a linearly independent set of vectors in H. Then we have an orthonormal sequence, (e_n), such that*

$$Span\{e_1, \dots, e_n\} = Span\{v_1, \dots, v_n\}.$$

for all n.

Proof: We set

$$e_1 = \frac{1}{\|v_1\|} v_1.$$

Then $\|e_1\| = 1$, and certainly $Span\{e_1\} = Span\{v_1\}$..

Let

$$w_2 = v_2 - \langle e_1, v_2\rangle e_1$$

and

$$e_2 = \frac{1}{\|w_2\|} w_2.$$

Observe that $\langle e_1, w_2\rangle = 0$, so $\langle e_1, e_2\rangle = 0$. Clearly $\|e_2\| = 1$.

Certainly $e_1, e_2 \in Span\{v_1, v_2\}$. Since the vectors e_1 and e_2 are orthogonal, they are linearly independent, so their span has dimension equal to the span of v_1 and v_2. But this means that $Span\{e_1, e_2\} = Span\{v_1, v_2\}$.

Now, suppose we have an orthonormal set $\{v_1, \dots, v_{n-1}\}$ such that

$$Span\{e_1, \dots, e_{n-1}\} = Span\{v_1, \dots, v_{n-1}\}.$$

Let

$$w_n = v_n - \langle e_1, v_n\rangle e_1 - \dots - \langle e_{n-1}, v_n\rangle.$$

and

$$e_n = \frac{1}{\|w_n\|} w_n.$$

Then as above, the set $\{e_1, \dots, e_n\}$ is orthonormal, and $Span\{e_1, \dots, e_n\} = Span\{v_1, \dots, v_n\}$. The result therefore follows by induction. □

In general, in a vector space V and a subset S, the span of S, $Span(S)$ is the smallest linear subspace of V that contains S. More concretely,

$$Span(S) = \{\alpha_1 v_1 + \dots + \alpha_n v_n \mid \alpha_i \in \mathbb{F}, v_i \in V,\ n \in \mathbb{N}\}.$$

Thus the span of a subset of the set of all finite linear combinations of elements in that set. Note that we only include finite linear combinations, even when the set is infinitely large.

Definition 5.6 *A subset, M, of a normed vector space V is called* total *if the span of M is a dense subset of V, that is to say*

$$\overline{Span(M)} = V.$$

A total orthonormal set of a Hilbert space, H, is called an orthonormal basis *of H.*

Proposition 5.7 *Let H be a Hilbert space, and let $M \subseteq H$. Then M is total if and only if $M^\perp = \{0\}$.*

Proof: Let M be total. Then $\overline{Span(M)} = H$. By proposition 4.17, we know that $\overline{Span(M)}^\perp = Span(M)^\perp$. It is shown in the exercises that $M^\perp = Span(M)^\perp$. Hence

$$M^\perp = H^\perp = \{0\}.$$

Conversely, suppose that $M^\perp = \{0\}$. Then $Span(M)^\perp = \{0\}$, and by corollary 4.21

$$\overline{Span(M)} = (Span(M)^\perp)^\perp = \{0\}^\perp = H.$$

This completes the proof. □

5.2 Series Related to Orthonormal Sequences

The following result is called *Bessel's inequality.*

Proposition 5.8 *Let (e_n) be an orthonormal sequence in a Hilbert space, H. Let $x \in H$. Then the series $\sum_{n=1}^{\infty} |\langle e_n, x\rangle|^2$ converges, and*

$$\sum_{n=1}^{\infty} |\langle e_n, x\rangle|^2 \leq \|x\|^2.$$

Proof: Since $|\langle e_n, x\rangle|^2 \geq 0$ for all n, it suffices to show that for any $N \in \mathbb{N}$, we have

$$\sum_{n=1}^{N} |\langle e_n, x\rangle|^2 \leq \|x\|^2.$$

Set

$$y = \sum_{n=1}^{N} \langle e_n, x\rangle e_n \qquad z = x - y.$$

Then $x = y + z$. Observe

$$\langle z, y\rangle = \langle x - y, y\rangle = \langle x, y\rangle - \langle y, y\rangle.$$

But

$$\langle x, y\rangle = \sum_{n=1}^{N} \langle e_n, x\rangle\langle x, e_n\rangle = \sum_{n=1}^{N} |\langle e_n, x\rangle|^2 = \langle y, y\rangle.$$

We conclude that z and y are orthogonal. Hence $\|x\|^2 = \|y\|^2 + \|z\|^2$. In particular, it follows that

$$\|y\|^2 = \sum_{n=1}^{N} |\langle e_n, x\rangle|^2 \leq \|x\|^2$$

and we are done. □

Lemma 5.9 *Let H be a Hilbert space. Let $a_n \in \mathbb{F}$. Then the series $\sum_{n=1}^{\infty} a_n e_n$ converges in norm in H if and only if the series $\sum_{n=1}^{\infty} |a_n|^2$ converges in $\mathbb{R}$.*

Proof: Let

$$s_N = \sum_{n=1}^{N} a_n e_n \qquad S_N = \sum_{n=1}^{N} |a_n|^2.$$

Then for $M, N \in \mathbb{N}$ with $N > M$ we have

$$\|s_N - s_M\|^2 = \|\sum_{n=M+1}^{N} a_n e_n\|^2 = \sum_{n=M+1}^{N} |a_n|^2 = S_N - S_M.$$

Suppose the series $\sum_{n=1}^{\infty} a_n e_n$ converges. Then the sequence of partial sums, (s_n), is Cauchy. By the above, it follows that the real sequence (S_n) is Cauchy, and therefore converges. Hence the series $\sum_{n=1}^{\infty} |a_n|^2$ converges.

Conversely, suppose the series $\sum_{n=1}^{\infty} |a_n|^2$ converges. Then the sequence of partial sums, (S_n), is Cauchy. By the above, the sequence of partial sums (s_n) is also Cauchy. Since the space H is a Hilbert space, and therefore complete, this means the sequence (s_n) converges, that is to say the series $\sum_{n=1}^{\infty} a_n e_n$ converges. □

Theorem 5.10 *Let H be a Hilbert space, and let $(e_n)_{n=1}^{\infty}$ be an orthonormal basis. Let $x \in H$. Set $a_n = \langle e_n, x\rangle$. Then the series $\sum_{n=1}^{\infty} a_n e_n$ converges in norm, and*

$$x = \sum_{n=1}^{\infty} a_n e_n.$$

Proof: By Bessel's inequality, the series $\sum_{n=1}^{\infty} |a_n|^2$ converges in norm. The above lemma tells us that the series $\sum_{n=1}^{\infty} a_n e_n$ therefore converges.

Let

$$y = \sum_{n=1}^{\infty} a_n e_n \qquad z = x - y.$$

Observe

$$\langle z, y\rangle = \langle x - y, y\rangle = \langle x, y\rangle - \langle y, y\rangle.$$

But by continuity of the inner product

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \langle e_n, x\rangle\langle x, e_n\rangle = \sum_{n=1}^{\infty} |\langle e_n, x\rangle|^2 = \langle y, y\rangle.$$

Therefore $\langle z, y\rangle = 0$. Let $M = Span\{e_n \mid n \in \mathbb{N}\}$. Then $y \in \overline{M}$, so $z \in \overline{M}^{\perp}$. But proposition 5.7 tells us that $\overline{M}^{\perp} = \{0\}$. Therefore $z = x - y = 0$, so

$$x = y = \sum_{n=1}^{\infty} a_n e_n.$$

□

Our last result in this section is called *Parseval's law.*

Corollary 5.11 *Let H be a Hilbert space, and let $(e_n)_{n=1}^{\infty}$ be an orthonormal basis. Then for any element $x \in H$, we have*

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle e_n, x\rangle|^2.$$

Proof: By the above theorem, we have

$$x = \sum_{n=1}^{\infty} \langle e_n, x \rangle e_n.$$

Let

$$s_N = \sum_{n=1}^{N} \langle e_n, x \rangle e_n.$$

Then the sequence (s_N) converges to x. By continuity of the norm, it follows that $\|s_N\|^2 \to \|x\|^2$ as $N \to \infty$.

But

$$\|s_N\|^2 = \sum_{n=1}^{N} |\langle e_n, x \rangle|^2.$$

So if we let $N \to \infty$, we see

$$\|x\|^2 = \lim_{N\to\infty} \|s_N\|^2 = \sum_{n=1}^{\infty} |\langle e_n, x \rangle|^2$$

as required. □

5.3 Examples of Orthonormal Bases

Recall that we define $L^2[0,1]$ to be the completion of the space of continuous functions $C[0,1]$ with respect to the inner product given by the formula

$$\langle f, g \rangle = \int_0^1 \overline{f(x)} g(x)\ dx.$$

Let $a < b$. Then we similarly define $L^2[a,b]$ to be the completion of the space of continuous functions $C[a,b]$ with respect to the inner product given by the formula

$$\langle f, g \rangle = \int_a^b \overline{f(x)} g(x)\ dx.$$

So we have a norm on the space $L^2[a,b]$ defined by writing

$$\|f\|_2 = \sqrt{\int_0^1 |f(x)|^2\ dx}.$$

The following is relevant when we look at Fourier series.

Lemma 5.12 *Let $C_P[a,b] = \{f \in C[a,b] \mid f(a) = f(b)\}$. Then $C_P[a,b]$ is a dense subset of $L^2[a,b]$.*

Proof: It suffices to prove that $C_P[a,b]$ is a dense subset of $C[a,b]$ with respect to the norm $\| - \|_2$. So, let $f \in C[a,b]$. Let $\varepsilon > 0$.

Define

$$g(t) = \{ \begin{array}{ll} f(t) & a \leq t \leq b-\delta \\ (t-a-\delta)f(a)/\delta + (1-(t-a+\delta)/\delta)f(b-\delta) & b-\delta \leq t \leq b \end{array}$$

The above formula comes from performing a linear interpolation between $f(b-\varepsilon)$ and $f(a)$ in the region $[b-\varepsilon, b]$. So g is continuous and $g(a) = g(b)$, so $g \in C_P[a,b]$. Now, let $M = \sup\{|f(t)| \mid t \in [a,b]\}$. Then certainly $|f(t)-g(t)| \leq 2M$ for all $t \in [a,b]$, and $f(t) = g(t)$ if $t \geq b-\delta$.

So

$$\|f-g\|_2^2 = \int_{b-\delta}^{b} |f(t)-g(t)|^2 \; dt \leq \int_{b}^{b-\delta} 4M^2 \; dt = 4M^2\delta$$

so $\|f-g\|_2 \to 0$ as $\delta \to 0$. Thus $C_P[a,b]$ is a dense subset of $C[a,b]$ and we are done. □

Note that we have different norm on the space $C[a,b]$ (or $C_P[a,b]$) defined by writing

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in [a,b]\}.$$

Comparison of these two norms enables us to apply the Stone-Weierstrass theorem to prove when certain subspaces of the spaces $L^2[a,b]$ are total.

Lemma 5.13 *Let M be a dense subset of the Banach space $(C[a,b], \| - \|_\infty)$. Then M is also a dense subset of the Hilbert space $L^2[a,b]$.*

Proof: Let $f \in L^2[a,b]$, and $\varepsilon > 0$. Then we have $g \in C[a,b]$ such that $\|g-f\|_2 < \frac{\varepsilon}{2}$.

By hypothesis, we also have $h \in M$ such that $\|h-g\|_\infty < \frac{\varepsilon^2}{4(b-a)}$. Observe

$$\|h-g\|_2^2 = \int_a^b |h(x)-g(x)| \; dx \leq (b-a)\|h-g\|_\infty$$

so

$$\|h-g\|_2 < \frac{\varepsilon}{2}.$$

It follows that

$$\|h-f\|_2 \leq \|h-g\|_2 + \|g-f\|_2 < \varepsilon$$

and we are done. □

Similarly, by the above, if M is a dense subset of the Banach space $(C_P[a,b], \| - \|_\infty)$, then M is also a dense subset of the Hilbert space $L^2[a,b]$.

We are now ready to apply our theory to Fourier series.

Proposition 5.14 *The set $M = \{\frac{1}{\sqrt{2\pi}}e^{ikx} \mid k \in \mathbb{Z}\}$ is an orthonormal basis of the Hilbert space $L^2[0,2\pi]$.*

Proof: Let A be the set of all functions

$$x \mapsto \sum_{k=-m}^{n} \frac{a_k}{\sqrt{2\pi}} e^{ikx}$$

where $m, n \in \mathbb{N}$ and $a_k \in \mathbb{C}$. Then $A = Span(M)$.

Recall from example 3.21 that the set of all Laurent polynomials is a dense subset of the space $C(\mathbb{T})$, where $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$. We can define an isometric isomorphism $T\colon C_P[0, 2\pi] \to C(\mathbb{T})$ by the formula $T(f)(e^{ix}) = f(x)$. Observe that

$$T\left(\frac{1}{\sqrt{2\pi}} e^{ikx}\right)(z) = \frac{1}{\sqrt{2\pi}} z^k$$

so T sends the set A to the set of all Laurent polynomials.

Therefore A is a dense subset of $C_P[0, 2\pi]$, and so by the above lemma, of $L^2[0, 2\pi]$. Therefore the set M is total. We leave it as an exercise to show that the set M is orthonormal. □

Corollary 5.15 *The sequence* $\frac{1}{\sqrt{2\pi}}$, $\frac{1}{\sqrt{\pi}}\cos x$, $\frac{1}{\sqrt{\pi}}\sin x$, $\frac{1}{\sqrt{\pi}}\cos(2x)$, $\frac{1}{\sqrt{\pi}}\sin(2x)$, $\frac{1}{\sqrt{\pi}}\cos(3x)$, $\frac{1}{\sqrt{\pi}}\sin(3x)$, ... *forms an orthonormal basis of the Hilbert space* $L^2[0, 2\pi]$.

Proof: Let

$$M = \{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos x, \frac{1}{\sqrt{\pi}}\sin x, \frac{1}{\sqrt{\pi}}\cos(2x), \frac{1}{\sqrt{\pi}}\sin(2x), \frac{1}{\sqrt{\pi}}\cos(3x), \frac{1}{\sqrt{\pi}}\sin(3x), \ldots\}.$$

Then we can check that M is orthonormal. By the above, and the identity $e^{i\theta} = \cos\theta + i\sin\theta$, we see that M is also total. □

Combining the above with theorem 5.10 immediately yields the following result.

Corollary 5.16 *Let* $f \in C[0, 2\pi]$. *Set*

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\ dx \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\ dx.$$

Then the series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

converges to the function f *in the space* $L^2[0, 2\pi]$. □

Parseval's Law gives us the following result

Corollary 5.17 *Let $f \in C[0, 2\pi]$. Let a_n and b_n be as in the above result. Then*

$$\int_0^{2\pi} |f(x)|^2 \, dx = \frac{\pi |a_0|^2}{2} + \sum_{n=1}^{\infty} \left(\pi |a_n|^2 + \pi |b_n|^2 \right).$$

□

We now turn our attention to another example. The following is left as an exercise.

Lemma 5.18 *The sequence of polynomials $1, x, x^2, x^3, \ldots$ is total in the Hilbert space $L^2[-1, 1]$.* □

Now, the set $\{1, x, x^2, x^3, \ldots\}$ is also linearly independent. We can therefore apply the Gram-Schmidt orthonormalisation process to obtain an orthonormal set with the same span. This new set will also be total, and so will be an orthonormal basis for $L^2[-1, 1]$.

Proposition 5.19 *Let*

$$P_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} [(t^2 - 1)^n]$$

and

$$e_n(t) = \sqrt{\frac{2n+1}{2}} P_n(t).$$

Then $\{e_0, e_1, e_2, \ldots\}$ is an orthonormal basis of the space $L^2[-1, 1]$.

Proof: Clearly $e_n(t)$ is a polynomial of degree n, so $Span\{e_0, e_1, e_2, \ldots\} = Span\{1, x, x^2, \ldots\}$. Thus it suffices to prove that the set $\{e_0, e_1, e_2, \ldots\}$ is orthonormal.

We begin by showing that $\|P_n\| = \sqrt{2/(2n+1)}$. Let $v(t) = (t^2 - 1)^n$. Observe that $v(1) = v(-1) = 0$, $v'(1) = v'(-1) = 0$, $v''(1) = v''(-1) = 0$, $\ldots$, $v^{(n-1)}(1) = v^{(n-1)}(-1) = 0$, and $v^{(2n)}(t) = (2n)!$. Hence

$$(2^n n!)^2 \|P_n\|^2 = \int_{-1}^{1} v^{(n)}(t) v^{(n)}(t) \, dt.$$

Integrating by parts we see that

$$(2^n n!)^2 \|P_n\|^2 = \left[v^{(n-1)}(t) v^{(n)}(t) \right]_{-1}^{1} - \int_{-1}^{1} v^{(n-1)}(t) v^{(n+1)}(t) \, dt.$$

So by the above,

$$(2^n n!)^2 \|P_n\|^2 = -\int_{-1}^{1} v^{(n-1)}(t) v^{(n+1)}(t) \, dt.$$

Repeating this process with the new integral, we see

$$(2^n n!)^2 \|P_n\|^2 = (-1)^n \int_{-1}^{1} v(t) v^{(2n)}(t)\ dt = (2n)! \int_{-1}^{1} (1-t^2)^n\ dt.$$

Let $t = \sin u$. Then $(1-t^2)^n = \cos^{2n}(u)$, and $dt/du = \cos u$. When $t = \pm 1$, we have $u = \pm\frac{\pi}{2}$. Hence

$$(2^n n!)^2 \|P_n\|^2 = (2n)! \int_{-\pi/2}^{\pi/2} \cos^{2n+1} u\ du = \frac{2^{2n+1}(n!)^2}{2n+1}$$

where the last integral is standard. Alternatively, we could have just used MAPLE for whole calculation.

Anyway, it follows that $\|P_n\| = \sqrt{2/(2n+1)}$ as required. It remains to check that $\langle P_m, P_n \rangle = 0$ when $m \neq n$.

Let $m < n$. Since P_m is a polynomial, it suffices to check that $\langle t^k, P_n \rangle = 0$ whenever $k < n$. To do this, we again use integration by parts.

To be precise

$$2^n n! \langle x^k, P_n \rangle = \int_{-1}^{1} t^k v^{(n)}(t)\ dt = \Big[t^k v^{(n-1)}(t) \Big]_{-1}^{1} - k \int_{-1}^{1} t^{k-1} v^{(n-1)}(t)\ dt$$

that is

$$2^n n! \langle x^k, P_n \rangle = -k \int_{-1}^{1} t^{k-1} v^{(n-1)}(t)\ dt$$

Repeating this process with the new integral, we see

$$2^n n! \langle x^k, P_n \rangle = (-1)^k k! \int_{-1}^{1} v^{(n-k)}(t)\ dt = (-1)^k k! \Big[v^{(n-k-1)} \Big]_{-1}^{1} = 0$$

and we are done. □

The polynomials $P_n(t)$ are called the *Legendre polynomials.*