TOPIC 1

Normed vector spaces

Throughout, vector spaces are over the field $\mathbb{F}$, which is $\mathbb{R}$ or $\mathbb{C}$, unless otherwise stated.

First, we consider bases in a space of continuous functions. This will motivate using (countably) infinite linear combinations. To interpret these, we need some kind of convergence; a norm on a vector space allows us to do this. We ask when linear maps of normed vector spaces are continuous, and when two normed vector spaces are "the same".

1. The space $C(I)$

Denote the interval $[0,1]$ by I. Denote by $C(I)$ the space of continuous $\mathbb{F}$-valued functions on I. This is a vector space.

QUESTION 1.1. What is the dimension of $C(I)$?

DEFINITION 1.2. Let S be a subset of a vector space V. The span of S, written $\operatorname{span} S$, is the set of all finite linear combinations of elements of S. A Hamel basis for V is a linearly independent set S such that $\operatorname{span} S = V$. A subspace U of V is finite-dimensional if $U = \operatorname{span} S$ for some finite subset S of V.

EXERCISE 1.3. Let I be the interval $[0,1]$. For $c \in I$, define the function $f_c : I \to \mathbb{F}$ by

$$f_c(t) = \max\{0, 1 - 10\,|t - c|\} \quad \forall t \in I.$$

Suppose that if $\frac{1}{10} \leq c_1 < \cdots < c_n \leq \frac{9}{10}$, $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ and $\sum_{j=1}^n \lambda_j f_{c_j} = 0$. Deduce that $\lambda_1 = \cdots = \lambda_n = 0$; that is, the set of functions $\{f_c : c \in [\frac{1}{10}, \frac{9}{10}]\}$ is linearly independent in $C(I)$ (when we consider finite linear combinations).

EXERCISE 1.4. Given sets A and B, write A^B for the set of all functions from B to A, and $|A|$ for the cardinality of A. You may assume that $\left|A^B\right| = |A|^{|B|}$.

(a) Show that the map

$$(a_n)_{n\in\mathbb{N}} \longmapsto \Big(t \mapsto \sum_{n\in\mathbb{N}} \frac{a_n}{n^2} \cos(2\pi i n t)\Big)$$

is a one-to-one mapping from $I^{\mathbb{N}}$ into $C(I)$.

(b) Show that the restriction map

$$f \mapsto f|_{I\cap\mathbb{Q}}$$

is a one-to-one mapping from $C(I)$ into $\mathbb{F}^{I\cap\mathbb{Q}}$.

(c) Hence find $|C(I)|$.

EXERCISE 1.5. Show that every vector space V over an arbitrary field $\mathbb{F}$ has a Hamel basis. What relations, if any, connect the cardinality of the basis and that cardinality of the vector space. Hint: first consider the collection of linearly independent subsets of V, ordered by inclusion.

A Hamel basis of $C(I)$ has large cardinality; and finding one uses the Axiom of Choice. However, in $C(I)$ we can use Fourier methods to write an arbitrary function f in the form

$$f(t) = at + \sum_{n \in \mathbb{Z}} a_n e^{2\pi i n t} \quad \forall t \in I.$$

The infinite sum may be interpreted using Abel or Cesáro means. So to calculate effectively, we must consider infinite linear combinations. Infinite sums are interpreted using a limit process, so we need to know when a sequence converges. This requires a *topology.* Normed vector spaces provide this.

2. Normed vector spaces

DEFINITION 1.6. A norm on a vector space V is a map $\|\cdot\| : V \to [0, \infty)$ such that

(a) $\|v\| \geq 0$ for all $v \in V$, and equality holds if and only if $v = 0$;
(b) $\|\lambda v\| = |\lambda| \, \|v\|$ for all $v \in V$ and $\lambda \in \mathbb{F}$;
(c) $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in V$.

DEFINITION 1.7. Let $\|\cdot\|$ be a norm on a vector space V. The *associated metric* is the distance function $d : V \times V \to [0, \infty)$ given by

$$d(u, v) = \|u - v\| \quad \forall u, v \in V.$$

The *associated topology* on V is the collection of all the open subsets of V.

LEMMA 1.8. *Let $(V, \|\cdot\|)$ be a normed vector space. Then the following hold:*

(a) $v_n \to v$ in V if and only if $v_n - v \to 0$ in V;
(b) $v_n \to 0$ in V if and only if $\|v_n\| \to 0$ in $\mathbb{R}$;
(c) if $v_n \to v$ in V, then $\|v_n\| \to \|v\|$ in $\mathbb{R}$;
(d) if $v_n \to v$ and $w_n \to w$ in V as $n \to \infty$, then also $v_n + w_n \to v + w$ in V;
(e) if $\lambda_n \to \lambda$ in $\mathbb{F}$ and $v_n \to v$ in V as $n \to \infty$, then also $\lambda_n v_n \to \lambda v$ in V.

PROOF. We omit this. □

The last two items of this lemma mean that addition and scalar multiplication are *jointly continuous.*

DEFINITION 1.9. Let $(V, \|\cdot\|)$ be a normed vector space. We say that V is *separable* if V contains a countable dense subset. We say that V is *complete* if every Cauchy sequence in V has a limit in V. A complete normed vector space is known as a *Banach space.*

EXERCISE 1.10. Let $(V, \|\cdot\|)$ be a normed vector space. Prove that the following are equivalent:

(a) V is complete, that is, every Cauchy sequence in V converges;
(b) every Cauchy sequence in V has a convergent subsequence;
(c) if $x_n \in V$ and $\sum_{n=1}^{\infty} \|x_n\| < \infty$, then the sequence (y_N), given by

$$y_N = \sum_{n=1}^{N} x_n,$$

converges;
(d) if $x_n \in V$ and $\sum_{n=1}^{\infty} \|x_n\| < \infty$, then the sequence (y_N), given by $y_N = \sum_{n=1}^{N} x_n$, converges; further, if y denotes its limit, then

$$\|y - y_N\| \le \sum_{n=N+1}^{\infty} \|x_n\| .$$

QUESTION 1.11. When are two normed vector spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ "the same"?

We would like V and W to be the same as vector spaces, and to have to same topology (i.e., the same convergent sequences) or the same metric (i.e., the same norm). So we would like to have an isomorphism $T : V \to W$ such that T is either a homeomorphism or an isometry. When we say that T is an *isomorphism*, we mean a linear (vector space) isomorphism; when we call T a *homeomorphism* we mean that T and T^{-1} are continuous; and when we say T is an *isometry*, we mean that

$$d_W(Tu, Tv) := \|Tu - Tv\|_W = \|T(u - v)\|_W = \|u - v\|_V =: d_V(u, v) \quad \forall u, v \in V,$$

or equivalently, that $\|Tv\|_W = \|v\|_V$ for all $v \in V$.

QUESTION 1.12. When is a map $T : V \to W$ between normed vector spaces continuous?

LEMMA 1.13. *Suppose that $T : V \to W$ is a linear map between normed vector spaces. Then the following are equivalent:*

(a) T is continuous at one *point in V;*
(b) T is continuous at all *points in V;*
(c) there is a positive constant C such that

$$\|Tv\|_W \le C \|v\|_V \quad \forall v \in V.$$

PROOF. It is obvious that (b) implies (a).

To show that (a) implies (b), suppose that T is continuous at v' in V. Take v'' in V and a sequence (v_n) such that $v_n \to v''$. Then $v_n - v'' + v' \to v'$, so by (a), $T(v_n - v'' + v') \to T(v')$, and hence $T(v_n) \to T(v'')$.

Now we show that (c) holds if and only if T is continuous at 0. Suppose that (c) holds, and that $v_n \to 0$, then $\|v_n\|_V \to 0$, and

$$\|Tv_n\|_W \le C \|v_n\|_V \to 0,$$

so $Tv_n \to 0$. Suppose that (c) does not hold, then for all $n \in \mathbb{Z}^+$ there exists $v_n \in V$ such that $\|Tv_n\|_W > n\|v_n\|_V$. Clearly $v_n \neq 0$. Now let

$$v'_n = \frac{1}{n\|v_n\|_V} v_n.$$

Then $\|v'_n\|_V = 1/n$ and $v_n \to 0$, but $\|Tv'_n\|_W > 1$ and $Tv'_n \not\to 0$. □

COROLLARY 1.14. *Suppose that $T : V \to W$ is an isomorphism of normed vector spaces. Then T is a homeomorphism if and only if there exist positive constants C_1 and C_2 such that*

$$C_1\|v\|_V \leq \|Tv\|_W \leq C_2\|v\|_V \quad \forall v \in V.$$

EXERCISE 1.15. Let $(V, \|\cdot\|)$ be a normed vector space. A *completion* of V is a complete normed vector space W and an isometric isomorphism T from V to a dense subspace TV of W. Prove that:

(a) at least one completion of V exists;
(b) if W_1 and W_2 are two completions of V, then W_1 and W_2 are isometrically isomorphic.

EXERCISE 1.16. Suppose that V and W are normed linear spaces, and that $T : V \to W$ is a linear isomorphism such that

$$C_1\|v\|_V \leq \|Tv\|_W \leq C_2\|v\|_V \quad \forall v \in V,$$

for some positive constants C_1 and C_2. Show that W is complete if and only if V is complete.

EXERCISE 1.17. Suppose that V and W are normed linear spaces, that V_0 and W_0 are dense subspaces of V and W, and that W is complete. Suppose that $T_0 : V_0 \to W_0$ is linear and that there is a constant C such that $\|Tv\|_W \leq C\|v\|_V$ for all $v \in V_0$.

Show that there is a linear map $T : V \to W$ such that $T|_{V_0} = T_0$ and $\|Tv\|_W \leq C\|v\|_V$ for all $v \in V$. Show also that if T' has the same properties as T, then $T = T'$.

EXERCISE 1.18. Suppose that V and W are complete normed linear spaces, and that V_0 and W_0 are dense subspaces of V and W respectively. Suppose that T_0 is an isometric isomorphism from V_0 onto W_0.

Using the result of the previous exercise, or otherwise, show that there is an isomorphism T from V onto W such that $T|_{V_0} = T_0$ and $\|Tv\|_W = \|v\|_V$ for all $v \in V$. Show also that if T' has the same properties as T, then $T = T'$.

TOPIC 2

ℓ^p spaces

Let S be a set. We will define spaces $\ell^p(S)$, where $1 \le p \le \infty$, and show that they are complete normed vector spaces. We will determine when they are separable; this is a necessary condition to have a countable basis.

1. Preliminaries

In the vector space $\mathbb{F}^n$, we define

$$\|u\|_p := \begin{cases} \left(\sum_{j=1}^n |u_j|^p\right)^{1/p} & \text{if } 1 \le p < \infty; \\ \max\{|u_j| : j = 1, \dots, n\} & \text{if } p = \infty. \end{cases}$$

EXERCISE 2.1. Show that $\lim_{p\to\infty} \|u\|_p = \|u\|_\infty$ for all $u \in \mathbb{F}^n$.

EXERCISE 2.2. Sketch the unit balls of the ℓ^p norm in $\mathbb{R}^2$ when $p = 1, \frac{4}{3}, 2, 4, \infty$. What linear maps of $\mathbb{R}^2$ are isometries for the ℓ^p norm? What linear maps of $\mathbb{F}^n$ are isometries?

EXERCISE 2.3. Suppose that $p, q \in (1, \infty)$ and $1/p + 1/q = 1$. Show that

$$st \le \frac{s^p}{p} + \frac{t^q}{q} \quad \forall s, t \in \mathbb{R}^+,$$

and equality holds if and only if $s^p = t^q$.

Deduce that, if $u, v \in \mathbb{F}^n$ and $\|u\|_p = \|v\|_q = 1$, then

$$\Big|\sum_{j=1}^n u_j \bar{v}_j\Big| \le 1,$$

and equality holds if and only if $u_j\,|u_j|^{p-1} = v_j\,|v_j|^{q-1}$ for all $j \in \{1, \dots, n\}$.

Deduce that, if $u, v \in \mathbb{F}^n$, then

$$\Big|\sum_{j=1}^n u_j \bar{v}_j\Big| \le \|u\|_p \,\|v\|_q\,,$$

(this is known as Hölder's inequality) and for all $u \in \mathbb{F}^n$, there exists $v \in \mathbb{F}^n$ such that $\|v\|_q = 1$ and

$$\Big|\sum_{j=1}^n u_j \bar{v}_j\Big| = \|u\|_p$$

(this is known as the converse of Hölder's inequality).

EXERCISE 2.4. Using the result of the previous exercise, or otherwise, show that

$$\|u+v\|_p \leq \|u\|_p + \|u\|_p \quad \forall u, v \in \mathbb{F}^n.$$

Deduce that $\|\cdot\|_p$ is a norm on $\mathbb{F}^n$.

2. Infinite sums

DEFINITION 2.5. Let S be a set, and let $\mathcal{F}(S)$ denote the collection of all finite subsets of S.

(a) Given $f : S \to [0, \infty)$, we define

$$\sum_S f = \sum_{s \in S} f(s) := \sup_{S_0 \in \mathcal{F}(S)} \sum_{s \in S_0} f(s).$$

(b) Given $f : S \to \mathbb{F}$ such that $\sum_S |f| < \infty$, we define

$$\sum_S f = \sum_{s \in S} f(s) := \sum_{s \in S} a(s) - \sum_{s \in S} b(s) + i \sum_{s \in S} c(s) - i \sum_{s \in S} d(s),$$

where

$$a(s) := \max\{\operatorname{Re} f(s), 0\}, \qquad b(s) := \max\{-\operatorname{Re} f(s), 0\},$$
$$c(s) := \max\{\operatorname{Im} f(s), 0\}, \qquad d(s) := \max\{-\operatorname{Im} f(s), 0\},$$

so that $f = a - b + ic - id$. Note that $0 \leq a \leq |f|$, and similarly for b, c and d, so $\sum_S f$ is well-defined.

We make the following definition.

DEFINITION 2.6. Suppose that $1 \leq p \leq \infty$. For a function $f : S \to \mathbb{F}$, we define

$$\|f\|_p = \begin{cases} \left(\sum_S |f|^p\right)^{1/p} & \text{if } p < \infty \\ \sup_S |f| & \text{if } p = \infty. \end{cases}$$

This expression may be infinite. Further, we define

$$\ell^p(S) := \{f : S \to \mathbb{F} : \|f\|_p < \infty\}.$$

Note that $\left(\sum_S |f|^p\right)^{1/p} = \sup_{S_0 \in \mathcal{F}(S)} \left(\sum_S |f|^p\right)^{1/p}$ and $\sup_S |f| = \sup_{S_0 \in \mathcal{F}(S)} \max_{S_0} |f|$.

EXERCISE 2.7. Show that if $f, g \in \ell^1(S)$ and $\lambda \in \mathbb{F}$, then

$$\sum_S (\lambda f + g) = \lambda \sum_S f + \sum_S g$$

and $|\sum_S f| \leq \|f\|_1$.

THEOREM 2.8. *The space $\ell^1(S)$ is a complete normed vector space.*

PROOF. It is easy to see that $\ell^1(S)$ is a normed vector space. We show only that it is closed under addition and that $\|f+g\|_1 \leq \|f\|_1 + \|g\|_1$.

Given $f, g \in \ell^1(S)$, we see that

$$\sum_{s\in S_0} |(f+g)(s)| \leq \sum_{s\in S_0} |f(s)| + \sum_{s\in S_0} |g(s)| \leq \|f\|_1 + \|g\|_1$$

for all $S_0 \in \mathcal{F}(S)$. Hence

$$\sum_{s\in S} |(f+g)(s)| = \sup_{S_0\in\mathcal{F}(S)} \sum_{s\in S_0} |(f+g)(s)| \leq \sup_{S_0\in\mathcal{F}(S)} (\|f\|_1 + \|g\|_1) = \|f\|_1 + \|g\|_1 .$$

This implies both that $f+g \in \ell^1(S)$ and that $\|f+g\|_1 \leq \|f\|_1 + \|g\|_1$.

Now we show that $\ell^1(S)$ is complete. To do this, we use the result of Exercise 1.10 that it is sufficient to show that if $f_n \in \ell^1(S)$ and $\sum_{n=1}^\infty \|f_n\|_1 < \infty$, then there exists $f \in \ell^1(S)$ such that

$$\left\| \sum_{n=1}^N f_n - f \right\|_1 \to 0$$

as $N \to \infty$.

Take $s \in S$. Clearly $|f(s)| \leq \|f\|_1$, and so

$$\sum_{n=1}^\infty |f_n(s)| \leq \sum_{n=1}^\infty \|f_n\|_1 < \infty.$$

This implies that $\sum_{n=1}^\infty f_n(s)$ converges; call this $f(s)$. For each $s \in S$,

$$\left| \sum_{n=1}^N f_n(s) - f(s) \right| = \left| \sum_{n=N+1}^\infty f_n(s) \right| \leq \sum_{n=N+1}^\infty |f_n(s)| ,$$

and so for any $S_0 \in \mathcal{F}(S)$,

$$\begin{aligned} \sum_{s\in S_0} \left| \sum_{n=1}^N f_n(s) - f(s) \right| &\leq \sum_{s\in S_0} \sum_{n=N+1}^\infty |f_n(s)| \\ &= \sum_{n=N+1}^\infty \sum_{s\in S_0} |f_n(s)| \leq \sum_{n=N+1}^\infty \|f_n\|_1 . \end{aligned}$$

Hence

$$\begin{aligned} \left\| \sum_{n=1}^N f_n(s) - f(s) \right\|_1 &= \sup_{S_0\in\mathcal{F}(S)} \sum_{s\in S_0} \left| \sum_{n=1}^N f_n(s) - f(s) \right| \\ &\leq \sup_{S_0\in\mathcal{F}(S)} \sum_{n=N+1}^\infty \|f_n\|_1 = \sum_{n=N+1}^\infty \|f_n\|_1 \to 0 \end{aligned}$$

as $N \to \infty$, as required. □

EXERCISE 2.9. Show that $\ell^p(S)$ is a complete normed vector space.

EXERCISE 2.10. State and prove versions of Hölder's inequality and its converse for spaces $\ell^p(S)$ and $\ell^q(S)$ when S is an infinite set.

DEFINITION 2.11. Let S be a set. We define $c_c(S)$, sometimes written $c_{00}(S)$, and $c_0(S)$ as follows:

$$c_c(S) := \{f : S \to \mathbb{F} : \mathrm{supp}(f) \text{ is finite}\},$$
$$c_0(S) := \{f : S \to \mathbb{F} : f \to 0 \text{ at infinity}\}.$$

Here $\mathrm{supp}(f)$ denotes the *support of* f, that is, the set $\{s \in S : f(s) \neq 0\}$, and $f \to 0$ *at infinity* in S means that $\{s \in S : |f(s)| > 1/n\}$ is finite for each $n \in \mathbb{Z}^+$. We endow $c_0(S)$ with the norm $\|\cdot\|_\infty$.

EXERCISE 2.12. Show that $c_c(S)$ is dense in $\ell^p(S)$ if $1 \leq p < \infty$ or if S is finite and $1 \leq p \leq \infty$. Show also that $c_c(S)$ is dense in $c_0(S)$. Under what conditions on p and S is $\ell^p(S)$ separable? Under what conditions on S $c_0(S)$ separable?

EXERCISE 2.13. Show that, if $1 \leq p \leq q \leq \infty$, then $\ell^1(S) \subseteq \ell^p(S) \subseteq \ell^q(S) \subseteq \ell^\infty(S)$, and $\|f\|_q \leq \|f\|_p$ whenever $f \in \ell^p(S)$. When are the set inclusions strict? When is the norm inequality an equality or a norm equivalence?

EXERCISE 2.14. Suppose that $1 \leq p < r < q \leq \infty$, and that $1/r = (1-\theta)/p + \theta/q$, where $0 < \theta < 1$. If $f \in \ell^p(S)$, then $f \in \ell^r(S)$ and $f \in \ell^q(S)$ by the preceding exercise. Show that

$$\|f\|_r \leq \|f\|_p^{1-\theta} \|f\|_q^\theta .$$

EXERCISE 2.15. Find the isometric isomorphisms of $\ell^p(S)$.

EXERCISE 2.16. Suppose that T is an isometry of $\mathbb{R}^2$ with the ℓ^p norm, that is,

$$\|Tu - Tv\|_p = \|u - v\|_p \quad \forall u, v \in \mathbb{R}^2.$$

Is T necessarily linear? If not, is T "nearly linear" in some sense?

TOPIC 3

Hilbert spaces

We define Hilbert spaces, which are the "nicest" infinite-dimensional vector spaces, and look at some of their properties. We also consider some examples.

1. Inner product spaces

DEFINITION 3.1. An inner product space is a vector space together with a mapping $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ satisfying the following conditions:

(a) $\langle u, v \rangle = \langle v, u \rangle^{-}$ for all $u, v \in V$;
(b) $\langle \lambda u + v, w \rangle = \lambda \langle u, w \rangle + \langle v, w \rangle$ for all $u, v, w \in V$ and $\lambda \in \mathbb{F}$;
(c) $\langle u, u \rangle > 0$ for all $u \in V \setminus \{0\}$.

The associated norm $\|\cdot\|$ on V is the map $v \mapsto \langle v, v \rangle^{1/2}$.

From the definition that $\langle u, u \rangle = 0$ if and only if $u = 0$ in V, and

$$\langle u, \lambda v + w \rangle = \bar{\lambda} \langle u, v \rangle + \langle u, w \rangle \quad \forall u, v, w \in V, \quad \forall \lambda \in \mathbb{F}.$$

Note that if $\mathbb{F} = \mathbb{R}$ then conjugation is irrelevant. It is easy to check that the associated norm really is a norm. An important fact is the Cauchy–Schwarz inequality.

EXERCISE 3.2 (Cauchy–Schwarz inequality). Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Prove that

$$|\langle u, v \rangle| \leq \langle u, u \rangle^{1/2} \langle v, v \rangle^{1/2} \quad \forall u, v \in V.$$

Deduce that $\langle u + v, u + v \rangle^{1/2} \leq \langle u, u \rangle^{1/2} + \langle v, v \rangle^{1/2}$, and hence show that the associated norm is a norm.

LEMMA 3.3. *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space, and fix $v, w \in V$. If*

$$\langle u, v \rangle = \langle u, w \rangle \quad \forall u \in V,$$

then $v = w$.

PROOF. The hypothesis implies that $\langle u, v - w \rangle = 0$ for all $u \in V$. Take $u = v - w$. □

LEMMA 3.4. *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. The inner product is jointly continuous on $V \times V$.*

Proof. We have to show that if $u_n \to u$ and $v_n \to v$, then $\langle u_n, v_n\rangle \to \langle u, v\rangle$. This is a consequence of the following inequalities and limits:

$$\begin{aligned} |\langle u_n, v_n\rangle - \langle u, v\rangle| &= |\langle u_n - u, v_n - v\rangle + \langle u_n - u, v\rangle + \langle u, v_n - v\rangle| \\ &\leq |\langle u_n - u, v_n - v\rangle| + |\langle u_n - u, v\rangle| + |\langle u, v_n - v\rangle| \\ &\leq \|u_n - u\| \, \|v_n - v\| + \|u_n - u\| \, \|v\| + \|u\| \, \|v_n - v\| \\ &\to 0 \end{aligned}$$

as $n \to \infty$. □

Examples 3.5. The following are inner product spaces, when equipped with their usual inner products: $\mathbb{F}^n$; $\ell^2(S)$, where S is a set; and $L^2(M)$, where M is a measure space. The spaces $L^2([0,1])$ and $L^2(\mathbb{R})$ and $\ell^2(\mathbb{N})$ and $\ell^2(\mathbb{Z})$ are particularly useful examples.

We know how the inner products on $\mathbb{R}^n$ and $\mathbb{C}^n$ behave, but let us check that the inner product is well-defined on $\ell^2(S)$. The standard inner product is

$$\langle f, g\rangle = \sum_S f(s)\,\bar{g}(s);$$

this is well-defined because

$$\sum_S |f(s)|\,|g(s)| \leq \Big(\sum_S |f(s)|^2\Big)^{1/2} \Big(\sum_S |g(s)|^2\Big)^{1/2} < \infty.$$

The required properties of $\langle\cdot,\cdot\rangle$ may be deduced easily from the corresponding properties of summation.

The case of $L^2(M)$ is similar, but requires integration rather than summation.

Question 3.6. Suppose that $(V, \|\cdot\|)$ is a normed vector space. If we know that V is an inner product space, can we find the inner product? And if we don't know whether V is an inner product space or not, can we decide whether it is?

Lemma 3.7 (Polarisation). *If $(V, \langle\cdot,\cdot\rangle)$ is an inner product space over $\mathbb{R}$, then*

$$\langle u, v\rangle = \frac{1}{4}\Big(\|u+v\|^2 - \|u-v\|^2\Big) \quad \forall u, v \in V;$$

if $(V, \langle\cdot,\cdot\rangle)$ is an inner product space over $\mathbb{C}$, then

$$\langle u, v\rangle = \frac{1}{4}\Big(\sum_{k=1}^{4} i^k \left\|u + i^k v\right\|^2\Big) \quad \forall u, v \in V.$$

Proof. We consider the real case only:

$$\|u+v\|^2 - \|u-v\|^2 = \langle u+v, u+v\rangle - \langle u-v, u-v\rangle = 2\langle u, v\rangle + 2\langle v, u\rangle = 4\langle u, v\rangle$$

for all $u, v \in V$. □

Notice that if V is a vector space over $\mathbb{C}$, then it is also a vector space over $\mathbb{R}$, by "restriction of scalars". The lemma above shows that the real part of a complex inner product on a complex vector space is a real inner product on the corresponding real vector space. What can we say about the imaginary part of the inner product? And given a real vector space, when does it arise as a complex vector space with restricted scalars?

LEMMA 3.8. *Suppose that $(V, \|\cdot\|)$ is a real normed vector space and that Appolonius' theorem holds, that is,*

$$\|u+v\|^2 + \|u-v\|^2 = 2\big(\|u\|^2 + \|v\|^2\big) \quad \forall u, v \in V.$$

Then the function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ given by

$$\langle u, v \rangle = \frac{1}{4}\big(\|u+v\|^2 - \|u-v\|^2\big) \quad \forall u, v \in V$$

is an inner product on V.

PROOF. Clearly $\langle u, u \rangle \geq 0$, and equality holds if and only if $u = 0$. It is also evident that $\langle u, v \rangle = \langle v, u \rangle$ for all $u, v \in V$. It remains to show that $\langle u+v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for all $u, v, w \in V$ and that $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$ for all $v, w \in V$ and $\lambda \in \mathbb{R}$.

Note that

$$\begin{aligned}
&4\big(\langle u+v, w \rangle - \langle u, w \rangle - \langle v, w \rangle\big) \\
&= \|u+v+w\|^2 - \|u+v-w\|^2 - \|u+w\|^2 + \|u-w\|^2 - \|v+w\|^2 + \|v-w\|^2 \\
&= \|u+v+w\|^2 - \|u+v-w\|^2 - (\|u+w\|^2 + \|v+w\|^2) + (\|u-w\|^2 + \|v-w\|^2) \\
&= \|u+v+w\|^2 - \|u+v-w\|^2 - \tfrac{1}{2}(\|u+v+2w\|^2 + \|u-v\|^2) \\
&\qquad + \tfrac{1}{2}(\|u+v-2w\|^2 + \|u-v\|^2) \\
&= \|u+v+w\|^2 - \tfrac{1}{2}(\|u+v+2w\|^2 + \|u+v\|^2) \\
&\qquad + \tfrac{1}{2}(\|u+v-2w\|^2 + \|u+v\|^2) - \|u+v-w\|^2 \\
&= \|u+v+w\|^2 - (\|u+v+w\|^2 + \|w\|^2) \\
&\qquad + (\|u+v-w\|^2 + \|w\|^2) - \|u+v-w\|^2 \\
&= 0,
\end{aligned}$$

so $\langle u+v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for all $u, v, w \in V$.

By induction, this implies that $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$ for all $\lambda \in \mathbb{N}$ and all $v, w \in V$; it then follows from Lemma 3.3 that the same formula holds for all $\lambda \in \mathbb{Q}$ and all $v, w \in V$. Finally, a continuity argument implies that the last equality holds for all $\lambda \in \mathbb{R}$. □

EXERCISE 3.9. Prove the complex polarisation identity. State and prove a criterion for a complex normed vector space to admit a complex inner product that defines the norm, analogous to Lemma 3.8.

DEFINITION 3.10. A Hilbert space is a *complete* inner product space.

EXAMPLES 3.11. The following are Hilbert spaces, when equipped with their usual inner products: $\mathbb{F}^n$; $\ell^2(S)$, where S is a set; and $L^2(M)$, where M is a measure space.

We showed that $\ell^p(S)$ is complete in Topic 2; in particular, $\ell^2(S)$ is complete. The case of $L^2(M)$ is similar, but requires integration rather than summation; the key fact is that pointwise limits of measurable functions are measurable (this is not true for continuous or Riemann integrable functions, which is why we need to use Lebesgue integration).

2. Orthonormality and approximation

DEFINITION 3.12. Vectors u, v in an inner product space $(V, \langle\cdot,\cdot\rangle)$ are said to be *orthogonal* if $\langle u, v\rangle = 0$. We often write $u \perp v$ to indicate this. Similarly, if $u \in V$ and $S \subseteq V$, then we write $u \perp S$ if $u \perp v$ for all $v \in S$, and if $S, S' \subseteq V$, then the expression $S \perp S'$ is defined similarly. Finally, $S^\perp$ denotes the set of all vectors v such that $v \perp S$.

EXERCISE 3.13. Suppose that A is a subset of an inner product space $(V, \langle\cdot,\cdot\rangle)$ and that $S = \operatorname{span} A$. (The span uses finite linear combinations only.) Show that

$$A^\perp = S^\perp = (\bar{S})^\perp \quad \text{and} \quad (A^\perp)^\perp = \bar{S}.$$

DEFINITION 3.14. A subset A of an inner product space $(V, \langle\cdot,\cdot\rangle)$ is *orthonormal* if different elements are orthogonal and they are normalised, that is, for all $a, b \in A$,

$$\langle a, b\rangle = \begin{cases} 1 & \text{if } a = b; \\ 0 & \text{if } a \neq b. \end{cases}$$

If A is a finite orthonormal set, that is, we may write $A = \{e_1, \dots, e_k\}$, then the *orthogonal projection onto the linear span S of A* is defined by

$$P_S(v) = \langle v, e_1\rangle\, e_1 + \cdots + \langle v, e_k\rangle\, e_k.$$

This definition would be somewhat problematic if A were infinite, as we would need to deal with an infinite number of terms.

EXERCISE 3.15. Let A be a finite orthonormal set in an inner product space $(V, \langle\cdot,\cdot\rangle)$ with span S. Show that $P_S P_S = P_S$, that $\langle P_S(u), v\rangle = \langle u, P_S(v)\rangle$, and that $P_S(v)$ is the unique element w of S such that $v - w \perp S$.

DEFINITION 3.16 (Gram–Schmidt orthogonalisation process). Given finitely many linearly independent vectors $v_1, \dots, v_n$ in an inner product space $(V, \langle\cdot,\cdot\rangle)$, we may produce an orthonormal set as follows: first, take $w_1 = v_1$ and then set $e_1 = \|w_1\|^{-1} w_1$. Next, take $w_2 = v_2 - P_{\operatorname{span}\{e_1\}}(v_2)$ and then set $e_2 = \|w_2\|^{-1} w_2$. Recursively, once $e_1, \dots, e_{k-1}$ are defined, take $w_k = v_k - P_{\operatorname{span}\{e_1,\dots,e_{k-1}\}}(v_k)$ and then set $e_k = \|w_k\|^{-1} w_k$.

In this construction, the vectors w_k are necessarily nonzero since the vectors $v_1, \dots, v_k$ are linearly independent.

EXERCISE 3.17. Let $\{v_1, \dots, v_n\}$ be a basis of $\mathbb{F}^n$. Define the Gram (or Gramian) matrix G by $G_{ij} = \langle v_i, v_j\rangle$. Show that under the change of basis represented by a matrix P, the Gram matrix changes to P^*GP. Hence find a triangular matrix Q such that $G = Q^*Q$.

LEMMA 3.18. *Suppose that U is a finite-dimensional subspace of an inner product space $(V, \langle\cdot,\cdot\rangle)$. Then U is closed.*

PROOF. Suppose that $v_n \in U$ and $v_n \to v$ in V. We need to show that $v \in U$. Take an orthonormal basis $\{e_1, \dots, e_k\}$ for U (this is possible since U is finite-dimensional). We may write v_n in terms of the basis:

$$v_n = \lambda_n^1 e_1 + \dots + \lambda_n^k e_k,$$

where $\lambda_n^j = \langle v_n, e_j\rangle$, and we define $\lambda^j = \langle v, e_j\rangle$. Now $\lambda^j - \lambda_n^j = \langle v - v_n, e_j\rangle$, so

$$\left|\lambda^j - \lambda_n^j\right| = |\langle v - v_n, e_j\rangle| \le \|v_n - v\| \to 0,$$

as $n \to \infty$, and thus $\lambda_n^j \to \lambda^j$. Now

$$\Big\|v_n - \sum_{j=1}^k \lambda^j e_j\Big\| = \Big\|\sum_{j=1}^k \lambda_n^j e_j - \sum_{j=1}^k \lambda^j e_j\Big\| = \Big\|\sum_{j=1}^k (\lambda_n^j - \lambda^j) e_j\Big\| \le \sum_{j=1}^k \left|\lambda_n^j - \lambda^j\right| \|e_j\| \to 0$$

as $n \to \infty$, that is, $v_n \to \sum_{j=1}^k \lambda^j e_j$ However, $v_n \to v$ by hypothesis, and limits are unique, so $v = \sum_{j=1}^k \lambda^j e_j$. This shows that $v \in U$, as required. □

DEFINITION 3.19. A subset S of a vector space V is convex if $tv + (1-t)w \in S$ whenever $v, w \in S$ and $t \in [0, 1]$.

LEMMA 3.20. *Suppose that S is a closed convex subset of a Hilbert space $(V, \langle\cdot,\cdot\rangle)$ and that $u \in V$. Then there exists a unique vector $v \in S$ such that*

$$\|u - v\| = \min\{\|u - w\| : w \in S\}.$$

PROOF. Write d for $\inf\{\|u - w\| : w \in S\}$; this is called the *distance from u to S*. By definition, for all $n \in \mathbb{Z}^+$ we can find vectors $w_n \in S$ such that

$$\|u - w_n\| \le d + \frac{1}{n}.$$

Take $m, n \in \mathbb{Z}^+$. Since S is convex, $\frac{1}{2}(w_n + w_m) \in S$, and so

$$\left\|u - \tfrac{1}{2}(w_n + w_m)\right\| \ge d.$$

Now, by Apollonius' theorem,

$$\|u - w_m\|^2 + \|u - w_n\|^2 = 2\left(\left\|u - \tfrac{1}{2}(w_m + w_n)\right\|^2 + \left\|\tfrac{1}{2}(w_m - w_n)\right\|^2\right),$$

and so

$$\begin{aligned}\left\|\tfrac{1}{2}(w_m - w_n)\right\|^2 &= \frac{1}{2}\Big(\|u - w_m\|^2 + \|u - w_n\|^2\Big) - \left\|u - \tfrac{1}{2}(x_m + x_n)\right\|^2 \\ &\le \frac{1}{2}\Big(d + \frac{1}{m}\Big)^2 + \frac{1}{2}\Big(d + \frac{1}{n}\Big)^2 - d^2 \\ &= d\Big(\frac{1}{m} + \frac{1}{n}\Big) + \frac{1}{2}\Big(\frac{1}{m^2} + \frac{1}{n^2}\Big).\end{aligned}$$

Thus the sequence (w_n) is Cauchy, and so it converges, to v say; since S is closed, $v \in S$.

Now

$$\|u - v\| \le \|u - w_n\| + \|w_n - v\| \le d + \frac{1}{n} + \|w_n - v\|\,;$$

by letting $n \to \infty$, we see that $\|u - v\| \le d$. But $\|u - v\| \ge d$ because $v \in S$, and hence equality holds.

Uniqueness is easy: if $v, v' \in S$ and $\|u - v\| = \|u - v'\| = d$, then, again by Apollonius' theorem,

$$\left\|u - \tfrac{1}{2}(v + v')\right\|^2 + \left\|\tfrac{1}{2}(v - v')\right\|^2 = \tfrac{1}{2}(\|u - v\|^2 + \|u - v'\|^2) = d^2;$$

now $\left\|u - \frac{1}{2}(v + v')\right\| \ge d$ since $\frac{1}{2}(v + v') \in S$, whence $\left\|\frac{1}{2}(v - v')\right\|^2 \le 0$ so $v = v'$. □

We sometimes call u the *best approximant to v in S*.

Since subspaces of vector spaces are convex, the last lemma enables us to define the orthogonal projection onto a not necessarily finite-dimensional closed subspace of a Hilbert space.

COROLLARY 3.21. *Suppose that S is a closed subspace of a Hilbert space $(V, \langle\cdot,\cdot\rangle)$. For each $u \in V$, write $P_S(u)$ for the unique vector in S such that*

$$\|u - P_S(u)\| = \min\{\|u - w\| : w \in S\},$$

as in Lemma 3.20. Then $P_S(u)$ is the unique vector in S such that $\langle u - P_S(u), w\rangle = 0$ for all $w \in S$. Further, $P_S : V \to S$ is an orthogonal projection, that is, P_S is linear, $P_S^2 = P_S$, and $\langle P_S(u), v\rangle = \langle u, P_S(v)\rangle$ for all $u, v \in V$.

PROOF. Take $w \in S$. By Lemma 3.20,

$$\|u - (P_S(u) + \lambda w)\| \ge \|u - P_S(u)\|$$

for all $\lambda \in \mathbb{F}$, and equality holds if and only if $\lambda = 0$. Hence

$$\langle u - (P_S(u) + \lambda w), u - (P_S(u) + \lambda w)\rangle \ge \langle u - P_S(u), u - P_S(u)\rangle\,,$$

that is,

$$|\lambda|^2 \langle w, w\rangle - 2\,\mathrm{Re}(\lambda \langle w, u - P_S(u)\rangle) \ge 0$$

for all $\lambda \in \mathbb{F}$, and equality holds if and only if $\lambda = 0$. This implies that $\langle w, u - P_S(u)\rangle = 0$.

Further, if $v_1, v_2 \in S$ and $\langle w, u - v_1\rangle = 0 = \langle w, u - v_2\rangle$ for all $w \in S$, then it follows that $\langle w, v_1 - v_2\rangle = 0$ for all $w \in S$, and hence $v_1 = v_2$. Thus $P_S(u)$ is the unique vector in S such that $\langle u - P_S(u), w\rangle = 0$ for all $w \in S$.

We have just shown that each $u \in V$ may be written uniquely as $u_S + u_\perp$, where $u_S \in S$ and $u_\perp \perp S$. Given two vectors $u, u' \in S$ and $\lambda \in \mathbb{F}$, we see that

$$\lambda u + u' = \lambda u_S + u'_S + \lambda u_\perp + u'_\perp,$$

and $\lambda u_S + u'_S \in S$ while $\lambda u_\perp + u'_\perp \perp S$. It follows that $P_S(\lambda u + u') = \lambda P_S(u) + P_S(u')$, that is, P_S is linear. The other properties of P_S follow similarly. □

We observe that if S is finite-dimensional, then the projection defined here coincides with the projection defined earlier; the key is that $v - P_S(v) \in S^\perp$ (see Exercise 3.15).

EXERCISE 3.22. Suppose that S_1 and S_2 are two closed convex subsets of a Hilbert space $(V, \langle\cdot,\cdot\rangle)$. What additional hypotheses, if any, are needed to ensure that there exist points $u_1 \in S_1$ and $u_2 \in S_2$ such that

$$\|u_1 - u_2\| = \inf\{\|v_1 - v_2\| : v_1 \in S_1, v_2 \in S_2\}?$$

3. Bessel's inequality and orthonormal bases

LEMMA 3.23. *Suppose that A is an orthonormal set in an inner product space $(V, \langle\cdot,\cdot\rangle)$. Then Bessel's inequality holds, that is,*

$$\sum_{e\in A} |\langle v, e\rangle|^2 \le \|v\|^2 \quad \forall v \in V. \tag{3.1}$$

PROOF. Suppose first that A is finite, and write S for $\operatorname{span} A$. Write $v = v_S + v_\perp$, where $v_S = P_S(v)$ and $v_\perp \perp S$; then

$$\|v\|^2 = \|v_S\|^2 + \|v_\perp\|^2 \ge \|v_S\|^2 .$$

Further, $v_S = \sum_{e\in A} \langle v, e\rangle\, e$ and so $\|v_S\|^2 = \sum_{e\in A} |\langle v, e\rangle|^2$. The result follows in the case when A is finite.

If A is infinite, then

$$\sum_{e\in A_0} |\langle v, e\rangle|^2 \le \|v\|^2 \quad \forall v \in V$$

for each finite subset A_0 of A. Taking the supremum over finite subsets of A gives the result in the general case. □

Finally we come to our main theorem. Recall that $\operatorname{span} A$ denotes the vector space of all *finite linear combinations* of elements of a subset A of a vector space V.

THEOREM 3.24. *Let A be an orthonormal subset of a Hilbert space $(V, \langle\cdot,\cdot\rangle)$. The following conditions are equivalent:*

(a) $\operatorname{span} A$ *is dense in* V*;*
(b) $A^\perp = \{0\}$*;*
(c) A *is a maximal orthonormal subset of* V*;*
(d) Bessel's inequality (3.1) *is always an equality;*
(e) the mapping $T : V \to \ell^2(A)$*, given by*

$$T(v) = (e \mapsto \langle v, e\rangle),$$

is an isometric isomorphism.

PROOF. We show that (a) and (b) are equivalent. Write S for $\operatorname{span} A$; then Exercise 3.13 shows that

$$A^\perp = S^\perp = (\bar{S})^\perp,$$

so if S is dense, then $A^\perp = \{0\}$. And if S is not dense, then $\bar{S}$ is a proper closed subspace of V; take $v \in V \setminus \bar{S}$, and then $v - P_{\bar{S}}v \in A^\perp \setminus \{0\}$.

Similarly, (c) is equivalent to (b): indeed, if $v \in A^\perp \setminus \{0\}$, then $A \cup \{\|v\|^{-1} v\}$ is a larger orthonormal set than A; conversely, if A' is an orthonormal set and $A' \supset A$, then $A' \setminus A \subset A^\perp \setminus \{0\}$.

If $A^\perp \neq \{0\}$, then take $v \in A^\perp \setminus \{0\}$. For this v, the left hand side of Bessel's inequality is 0 but the right hand side is not.

We now assume that $A^\perp = \{0\}$, and prove that Bessel's inequality is always an equality. Combined with the result of the previous paragraph, this will imply that (b) is equivalent to (d). Take $v \in V$ and $\varepsilon \in \mathbb{R}^+$. If $A^\perp = \{0\}$, then $\operatorname{span} A$ is dense in V, so there exists $v_0 \in \operatorname{span} A$ such that $\|v - v_0\| < \varepsilon$. As v_0 is a finite linear combination of elements of A, there is a finite subset A_0 of A such that $v_0 \in \operatorname{span} A_0$. Write S for $\operatorname{span} A_0$. Now

$$\|v - P_S(v)\| \leq \|v - v_0\| < \varepsilon$$

and

$$\|v - P_S(v)\|^2 + \|P_S(v)\|^2 = \|v\|^2 ,$$

so $\|P_S(v)\|^2 \geq \|v\|^2 - \varepsilon^2$. Furthermore, if $e \in A_0$, then

$$\langle v, e\rangle = \langle (v - P_S(v)), e\rangle + \langle P_S(v), e\rangle = \langle P_S(v), e\rangle$$

and so

$$\sum_{e\in A} \langle v, e\rangle^2 \geq \sum_{e\in A_0} \langle v, e\rangle^2 = \sum_{e\in A_0} \langle P_S(v), e\rangle^2 = \|P_S(v)\|^2 \geq \|v\|^2 - \varepsilon^2.$$

As this holds for an arbitrary ε, it follows that

$$\sum_{e\in A} \langle v, e\rangle^2 \geq \|v\|^2 ;$$

combined with Bessel's inequality, this implies that Bessel's inequality is an equality for v; as v is arbitrary, (d) holds.

Finally, if (d) holds, then (a) holds too, and so for all v in $\operatorname{span} A$, the map

$$v \longmapsto (e \mapsto \langle v, e\rangle)$$

is a linear isometry. This maps isometrically from $\operatorname{span} A$, a dense subspace of V, onto $c_c(A)$, a dense subspace of $\ell^2(A)$. This extends to an isometric isomorphism by Exercise 1.18; this extension is also an isomorphism of inner product spaces by polarisation.

If (e) holds, then Bessel's inequality is an equality by definition: one side is the square of the $\ell^2(A)$ norm and the other is the square of the V norm. □

Definition 3.25. A maximal orthonormal set is also called a *complete orthonormal set* or a *basis.*

Theorem 3.26. *Let $(V, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Then there is a basis in V.*

Proof. It suffices to show that a maximal orthonormal set exists. This can be done using a Zorn's lemma argument. □

Exercise 3.27. If A is an orthonormal subset of a Hilbert space $(V, \langle\cdot,\cdot\rangle)$, how can we interpret $\sum_{e\in A} \lambda(e)e$, when $\lambda \in \ell^2(A)$?

EXAMPLES 3.28. The set of functions $\{t \mapsto e^{2\pi int} : n \in \mathbb{Z}\}$ is orthonormal in $L^2(I)$; alternatively, we may consider the set $\{1, \sqrt{2}\cos(2\pi nt), \sqrt{2}\sin(2\pi nt) : n \in \mathbb{Z}^+\}$. Both these sets are bases. This may be proved in various ways. Essentially all the proofs use the fact that $C(I)$ is a dense subspace of $L^2(I)$, but there are several ways to show that the set of trigonometric polynomials (the span of the complex exponentials, or of the trigonometric functions) is dense in $C(I)$. These include using Sturm–Liouville theory, which is about differential equations; using the Stone–Weierstrass theorem, which is about algebras of functions; and using convolution, which is a technique of harmonic analysis.

Fourier series is a prototype of Hilbert space theory.

The theory of Fourier integrals does not give us a basis for $L^2(\mathbb{R})$, since periodic functions are not square integrable. Various bases for $L^2(\mathbb{R})$ were found in the 1800s, with names such as Hermite functions or Laguerre functions; these were first constructed using differential operators. We will consider the Hermite functions later.

EXERCISE 3.29. Define functions $h_j : I \to \mathbb{R}$, where $j \in \mathbb{N}$, by $h_0 = 1$, and, for $n = 0, 1, 2, \dots$ and $k = 0, 1, \dots, 2^n - 1$,

$$h_{2^n+k}(t) = \begin{cases} 2^{n/2} \operatorname{sgn} \sin(2^n \pi t) & \text{if } k \leq 2^n t \leq k+1 \\ 0 & \text{otherwise,} \end{cases}$$

where sgn is the signum function: $\operatorname{sgn} t = 1$ if $t > 0$, $\operatorname{sgn} t = -1$ if $t < 0$, and $\operatorname{sgn} 0 = 0$.

(a) Sketch the graphs of the functions h_1 and h_6.
(b) Show that $\{h_j : j \in \mathbb{N}\}$ is an orthogonal set in $L^2(I)$ (where $I = [0, 1]$).
(c) Show that every function $f \in C(I)$ may be approximated uniformly by a finite sum of the form $\sum_{j=0}^{J} c_j h_j$.
(d) Deduce that $\{h_j : j \in \mathbb{N}\}$ is an orthonormal basis in $L^2(I)$. You may assume that $C(I)$ is dense in $L^2(I)$.

(This basis is known as the Haar basis, and is used in signal processing.)

REMARK 3.30. Spaces $L^2(M_1)$ and $L^2(M_2)$ can be "the same" in different ways. We say that measure spaces M_1 and M_2 are isomorphic if there is a measure-preserving bijection $\varphi : M_1 \to M_2$. Such a map φ induces an isometric isomorphism $T : L^2(M_2) \to L^2(M_1)$:

$$Tf = f \circ \varphi.$$

Measure space isomorphism is a "coarse" equivalence. For example, $\mathbb{R}^m$ and $\mathbb{R}^n$ are measure-isomorphic for all $m, n \in \mathbb{Z}^+$, and hence the corresponding L^2 spaces are isomorphic as Hilbert spaces. However, no two of $\mathbb{Z}$, $\mathbb{R}$ and I are measure-isomorphic, but the corresponding Hilbert spaces are isometrically isomorphic. Indeed, there is a unique isometric isomorphism equivalence class of infinite-dimensional separable Hilbert spaces. Hence Hilbert space equivalence is an even "coarser" equivalence.

An important research area in modern functional analysis involves studying Hilbert spaces with additional structure so that, for instance, $L^2(\mathbb{R}^m)$ and $L^2(\mathbb{R}^n)$ with the additional structure are not isomorphic.

TOPIC 4

Dual spaces

We discuss linear functionals on normed vector spaces, and consider the examples of the ℓ^p spaces. We then prove the Hahn–Banach theorem, one of the pillars of functional analysis, and look at some applications.

1. Definition of the dual space

In this section, we will deal with several normed vector spaces, and we will add a subscript to the norm to clarify the space in which the norm is being used.

DEFINITION 4.1. Let V be a vector space. A *linear functional on V* is a linear mapping $L : V \to \mathbb{F}$. The set of all linear functionals on V forms a vector space.

DEFINITION 4.2. Let $(V, \|\cdot\|_V)$ be a normed vector space. A *continuous linear functional on V* is a linear functional L on V for which there is a constant C such that

$$|L(v)| \leq C \|v\|_V \quad \forall v \in V. \tag{4.1}$$

The set of all continuous linear functionals on V is denoted V^*, and the norm of $L \in V^*$ is given by

$$\|L\|_{V^*} = \sup\{|L(v)| : v \in V, \|v\|_V \leq 1\}.$$

From (4.1), we see that $\|L\|_{V^*} \leq C$.

EXERCISE 4.3. Let L be a linear functional on a normed vector space $(V, \|\cdot\|_V)$. Show that

$$\sup\{|L(v)| : v \in V, \|v\|_V \leq 1\} = \inf\{C \in [0, \infty) : \text{(4.1) holds}\}$$

and that the infimum is attained.

PROPOSITION 4.4. *Let $(V, \|\cdot\|_V)$ be a normed vector space. The set of all continuous linear functionals on V, endowed with the norm $\|\cdot\|_{V^*}$, is a Banach space.*

PROOF. We omit this proof. □

REMARK 4.5. Let $(V, \|\cdot\|_V)$ be a normed vector space. Since the dual space V^* is a normed vector space, it too has a dual space, written V^{**}. The space V maps continuously into V^{**}; indeed, define $J : V \to V^{**}$ by

$$(Jv)(L) = L(v) \quad \forall L \in V^*.$$

This mapping is obviously linear. Further,

$$\begin{aligned}\|Jv\|_{V^{**}} &= \sup\{|(Jv)(L)| : L \in V^*, \|L\|_{V^*} \le 1\} \\ &= \sup\{|L(v)| : L \in V^*, \|L\|_{V^*} \le 1\} \\ &\le \sup\{\|L\|_{V^*} \|v\|_V : L \in V^*, \|L\|_{V^*} \le 1\} \\ &= \|v\|_V ,\end{aligned}$$

so the mapping is also continuous. But *a priori* there is no reason why J should be injective, surjective, or isometric.

Definition 4.6. A Banach space V is reflexive if the map $J : V \to V^{**}$ is an isometric isomorphism.

Example 4.7. Suppose that S is a set, that $p, q \in [1, \infty]$ and $1/p + 1/q = 1$, and that $g \in \ell^q(S)$. Then the mapping $T(g) : \ell^p(S) \to \mathbb{F}$ given by

$$T(g) : f \mapsto \sum_{s \in S} f(s)g(s)$$

is well-defined and linear, from Topic 2 (in particular, from Hölder's inequality); further

$$|T(g)(f)| \le \|g\|_q \|f\|_p \quad \forall f \in \ell^p(S).$$

Thus the mapping T is a linear mapping from $\ell^q(S)$ into $\ell^p(S)^*$. This mapping is an isometry (by Hölder's inequality and its converse); so in particular it is injective.

Conversely, given a continuous linear functional L on $\ell^p(S)$, we may define $g : S \to \mathbb{F}$ by the formula

$$g(s) = L(\delta_s),$$

where $\delta_s \in \ell^p(S)$ is defined by requiring that $\delta_s(t)$ is equal to 1 if $s = t$ and 0 otherwise. It is evident that $|g(s)| \le \|L\|_{V^*}$. By considering linear combinations of the δ_s where $s \in S_0$, a finite subset of S, and using the converse of Hölder's inequality, we see that

$$\Big(\sum_{s \in S_0} |g(s)|^q\Big)^{1/q} \le \|L\|_{V^*}$$

if $1 \le q < \infty$ or

$$\max_{s \in S_0} |g(s)| \le \|L\|_{V^*}$$

if $q = \infty$, and then taking suprema over such subsets S_0, it follows that

$$\|g\|_q \le \|L\|_{V^*} .$$

Further, by construction, $L(f) = T(g)(f)$ for all $f \in c_c(S)$. If $c_c(S)$ is dense in $\ell^p(S)$, then it follows that $L = T(g)$; but if $c_c(S)$ is not dense in $\ell^p(S)$ (this only happens if $p = \infty$), then it is possible that $L \ne T(g)$.

We will see soon that $T(\ell^1(S))$ is a proper subset of $\ell^\infty(S)^*$ when S is infinite. Thus $\ell^p(S)$ is reflexive if $1 < p < \infty$ or if S is finite, and is not reflexive otherwise.

EXAMPLE 4.8. If H is a Hilbert space, then we may define a mapping $\widetilde{T} : H \to H^*$ by

$$\widetilde{T}u(v) = \langle v, u \rangle \quad \forall v \in H.$$

Because the second variable in the inner product is conjugate-linear, $\widetilde{T}$ is a conjugate-linear mapping of H into H^*. From the Cauchy–Schwarz inequality,

$$\left|\widetilde{T}u(v)\right| \leq \|u\|_H \|v\|_H \quad \forall v \in H,$$

and it follows that

$$\left\|\widetilde{T}u\right\|_{H^*} \leq \|u\|_H ;$$

by taking v equal to u, we see that in fact equality holds, and in particular $\widetilde{T}$ is injective. But it is not *a priori* obvious that $\widetilde{T}$ is surjective. However, by identifying H with $\ell^2(A)$ for some set A, we can see that the mapping $\widetilde{T}u$ corresponds to $T(\bar{u})$, where T is the mapping considered in Example 4.7, and hence $\widetilde{T}$ is also surjective.

EXERCISE 4.9. Suppose that $(V, \|\cdot\|_V)$ is a Banach space. If V is reflexive, must V^* be reflexive? And if V^* is reflexive, must V be reflexive?

2. The Hahn–Banach theorem

In this section, the symbol $\|\cdot\|$ will always refer to the same space, and so we omit subscripts.

THEOREM 4.10. *Let U be a subspace of a normed vector space $(V, \|\cdot\|)$. Suppose that $L : U \to \mathbb{F}$ is a continuous linear mapping with the property that $|L(v)| \leq C\|v\|$ for all $v \in U$. Then there exists a linear mapping $M : V \to \mathbb{F}$ with the properties that $|M(v)| \leq C\|v\|$ for all $v \in V$ and $M|_U = L$.*

PROOF. There are several steps in the proof. First we show that, when the scalars are real, it is always possible to extend a linear functional defined on a subspace to a larger subspace *without increasing the norm*. Then we consider the collection of all possible extensions, and use Zorn's lemma to deduce the result. An extra trick is needed to deal with complex scalars.

Step 1: Enlarging the domain. Suppose that V is a normed vector space over $\mathbb{R}$. If W is a subspace of V and $v \in V \setminus W$, then $W' = W + \mathbb{R}v$ is a larger subspace than W. We show how to extend a linear functional L on W to a linear functional L' on W' without increasing the norm. Suppose that

$$|L(w)| \leq C\|w\| \quad \forall w \in W;$$

we will construct L' such that $|L'(w')| \leq C\|w'\|$ for all $w' \in W'$.

Every element of W' has a unique representation in the form $w + \lambda v$, where $w \in W$ and $\lambda \in \mathbb{R}$. We may therefore take $\mu \in \mathbb{R}$, and define

$$L'(w + \lambda v) = L(w) + \lambda\mu,$$

to obtain a linear functional on W'. Indeed,

$$\begin{aligned} L'(\nu(w+\lambda v)+\nu'(w'+\lambda' v)) &= L'((\nu w+\nu' w')+(\nu\lambda+\nu'\lambda')v) \\ &= L(\nu w+\nu' w')+(\nu\lambda+\nu'\lambda')\mu \\ &= \nu L(w)+\nu' L(w')+\nu\lambda\mu+\nu'\lambda'\mu \\ &= \nu L'(w+\lambda v)+\nu' L'(w'+\lambda' v) \end{aligned}$$

for all $w+\lambda v, w'+\lambda' v \in W'$, showing that L' is linear.

However, we also want the following inequality to hold:

$$|L'(w+\lambda v)| \leq C\,\|w+\lambda v\| \quad \forall w+\lambda v \in W'. \tag{4.2}$$

This is certainly true if $\lambda = 0$, so we may assume that $\lambda \neq 0$. Replacing w by λw, and dividing out a factor of $|\lambda|$ from each side, we see that it suffices to prove that

$$|L'(w+v)| \leq C\,\|w+v\| \quad \forall w \in W.$$

This inequality is equivalent to each of the following inequalities:

$$\begin{gathered} -C\,\|w+v\| \leq L'(w+v) \leq C\,\|w+v\| \\ -C\,\|w+v\| \leq L(w)+\mu \leq C\,\|w+v\| \\ -C\,\|w+v\| - L(w) \leq \mu \leq C\,\|w+v\| - L(w) \end{gathered}$$

for all $w \in W$, and the last inequality is equivalent to

$$\sup\{-C\,\|w+v\| - L(w) : w \in W\} \leq \mu \leq \inf\{C\,\|w+v\| - L(w) : w \in W\}.$$

This in turn is equivalent to each of the inequalities

$$\begin{gathered} -C\,\|w_1+v\| - L(w_1) \leq \mu \leq C\,\|w_2+v\| - L(w_2) \\ -C\,\|w_1+v\| - L(w_1) \leq C\,\|w_2+v\| - L(w_2) \\ L(w_2)-L(w_1) \leq C\,\|w_2+v\| + C\,\|w_1+v\| \\ L(w_2-w_1) \leq C\,\|w_2+v\| + C\,\|w_1+v\| \end{gathered}$$

for all $w_1, w_2 \in W$. Now

$$\begin{aligned} L(w_2-w_1) &\leq |L(w_2-w_1)| \\ &\leq C\,\|w_2-w_1\| \\ &\leq C\,\|(w_2+v)-(w_1+v)\| \\ &\leq C\,\|w_2+v\| + C\,\|w_1+v\| \end{aligned}$$

for all $w_1, w_2 \in W$, which proves the inequality (4.2) that enables us to choose μ.

Step 2: Induction. Now we present the Zorn's lemma induction argument, still supposing that the scalars are real. Recall that we start with a linear functional $L : U \to \mathbb{R}$ satisfying $|L(u)| \leq C \|u\|$ for all $u \in U$. We consider the collection of linear mappings $M_W : W \to \mathbb{R}$, where W is a subspace of V and $U \subseteq W$, with the properties that $M_W|_U = L$ and

$$|M_W(w)| \leq C \|w\| \quad \forall w \in W.$$

We partially order the set of these mappings by decreeing that $M_W \preceq M'_{W'}$ if $W \subseteq W'$ and $M'_{W'}|_W = M_W$. Given a chain of such mappings

$$M_W \preceq M'_{W'} \preceq M''_{W''} \dots,$$

we take $\widetilde{W}$ to be the union of the spaces W, W', W'' and so on, and define $\widetilde{M}$ on $\widetilde{W}$ by $\widetilde{M}(w) = M^{\dagger}_{W^{\dagger}}(w)$ for any $W^{\dagger}$ such that $w \in W^{\dagger}$; this is well-defined by the definition of the partial order $\preceq$. It is clear that $\widetilde{M}_{\widetilde{W}}$ is an upper bound for the chain, so Zorn's lemma applies. If $\widetilde{M}_{\widetilde{W}}$ is a maximal element, then $\widetilde{W} = V$ by Step 1. This establishes the theorem in the case where the scalars are real.

Step 3: Complex scalars. When the scalars are complex, we may proceed in two different ways. Suppose that $|L(u)| \leq C \|u\|$ for all $u \in U$. The simplest approach is to take $L : U \to \mathbb{C}$, and consider $\operatorname{Re} L$, which is a linear functional on U when the scalars are taken to be $\mathbb{R}$; further, $|\operatorname{Re} L(u)| \leq |L(u)| \leq C \|u\|$ for all $u \in U$. We extend $\operatorname{Re} L$ to a real-linear functional $M_{\mathbb{R}} : V \to \mathbb{R}$, and then define

$$M(v) = M_{\mathbb{R}}(v) - iM_{\mathbb{R}}(iv) \quad \forall v \in V.$$

Clearly $M(v + w) = M(v) + M(w)$ and $M(\lambda v) = \lambda M(v)$ for all $v, w \in V$ and all $\lambda \in \mathbb{R}$. Next,

$$M(iv) = M_{\mathbb{R}}(iv) - iM_{\mathbb{R}}(i^2 v) = i(M_{\mathbb{R}}(v) - iM_{\mathbb{R}}(iv)) = iM(v) \quad \forall v \in V.$$

Thus

$$M((\lambda + i\mu)v) = M(\lambda v + i\mu v) = M(\lambda v) + M(i\mu v) = \lambda M(v) + i\mu M(v) = (\lambda + i\mu)M(v)$$

for all $\lambda, \mu \in \mathbb{R}$, so M is complex-linear. Further, if $v \in V$, there exists $\omega \in \mathbb{C}$ such that $|\omega| = 1$ and $\omega M(v) \geq 0$. Hence $M(\omega v) \geq 0$, so $M(v) = \omega^{-1} M(\omega v) = \omega^{-1} M_{\mathbb{R}}(\omega v)$, and

$$|M(v)| = |M_{\mathbb{R}}(\omega v)| \leq C \|\omega v\| = C \|v\|,$$

and if in addition $v \in U$, then

$$M(v) = \omega^{-1} M_{\mathbb{R}}(\omega v) = \omega^{-1} L(\omega v) = L(v),$$

and M has all the required properties.

However, we could also use the one-dimensional extension argument twice to extend $\operatorname{Re} L : W \to \mathbb{R}$ to $L''_{\mathbb{R}} : W + \mathbb{C}v \to \mathbb{R}$, next use the complexification argument to produce $L'' : W + \mathbb{C}v \to \mathbb{C}$ with the required properties, and finally use the Zorn's lemma argument with complex linear mappings. $\square$

3. Consequences of the Hahn–Banach theorem

COROLLARY 4.11. *Suppose that $(V, \|\cdot\|_V)$ is a normed vector space. For all $v \in V$, there exists a continuous linear functional M on V of norm 1 such that $M(v) = \|v\|_V$. Hence the mapping $J : V \to V^{**}$, defined in Remark 4.5, is an isometric injection.*

PROOF. We saw that $\|Jv\|_{V^{**}} \leq \|v\|_V$ for all $v \in V$. It suffices to show that the opposite inequality holds.

Given $v \in V$, the mapping $\lambda v \mapsto \lambda \|v\|_V$ (for all $\lambda \in \mathbb{F}$) is a continuous linear functional on $\mathbb{F}v$, of norm 1. By the Hahn–Banach theorem, there is a continuous linear functional M on V of norm 1 such that $M(v) = \|v\|_V$. Hence

$$\|Jv\|_{V^{**}} = \sup\{|L(v)| : L \in V^*, \|L\|_{V^*} \leq 1\} \geq |M(v)| = \|v\|_V ,$$

as required. □

EXERCISE 4.12. Suppose that H_0 is a closed subspace of a Hilbert space H, and that $L : H_0 \to \mathbb{F}$ is a linear functional on H_0. Find a description of all continuous linear functionals $M : H \to \mathbb{F}$ such that $\|M\| = \|L\|$ and $M|_{H_0} = L$.

COROLLARY 4.13. *There exists a nonzero continuous linear functional M on $\ell^\infty(\mathbb{N})$ such that $M(f) = 0$ for all $f \in c_0(\mathbb{N})$; in particular, $M \notin J\ell^1(\mathbb{N})$.*

PROOF. Denote by $\underline{1}$ the constant sequence $(1, 1, 1, \dots)$, and observe that the map $L : \mathbb{F}\underline{1} + c_0(\mathbb{N}) \to \mathbb{F}$ given by

$$L(f) = \lim_{n\to\infty} f(n) \quad \forall f \in \mathbb{F}\underline{1} + c_0(\mathbb{N}),$$

is a continuous linear functional on this subspace of $\ell^\infty(\mathbb{N})$ of norm 1; indeed,

$$\big|\lim_{n\to\infty} f(n)\big| \leq \sup\{|f(n)| : n \in \mathbb{N}\}.$$

We observe that $L(f) = 0$ for all $f \in C_0(\mathbb{N})$.

We may extend L to a continuous linear functional $M : \ell^\infty \to \mathbb{F}$ of norm 1, with the properties that $M(\underline{1}) = 1$ (so $M \neq 0$) and $M(f) = 0$ for all $f \in c_0(\mathbb{N})$. □

REMARK 4.14. In the last corollary, if we replace $\underline{1}$ by another sequence that is not in $c_0(\mathbb{N})$, such as $(\sin 0, \sin 1, \sin 2, \dots)$, then the argument still goes through. In advanced courses, it is sometimes shown that the space $\ell^\infty(\mathbb{N})^* \setminus i\ell^1(\mathbb{N})$ is generated by its *multiplicative* elements, that is, linear functionals T such that $T(fg) = T(f)T(g)$, and these may all be obtained as limits of subsequences (or more precisely, subnets) of point evaluations $f \mapsto f(n)$.

EXERCISE 4.15. Show that there exists a continuous linear functional L on $\ell^\infty(\mathbb{N})$ such that

$$L((a_n)) = a_0 + \lim_{n\to\infty} a_{2n} - \lim_{n\to\infty} a_{2n+1}$$

if both limits exist. What can be said about $\|L\|_{(\ell^\infty)^*}$?

THEOREM 4.16. *Suppose that U is a closed subspace of a Banach space V. The quotient vector space V/U, endowed with the* quotient norm $\|\cdot\|_q$, *defined by*

$$\|v+U\|_q = \inf\{\|v+u\| : u \in U\}$$

for all $v \in V$, is a Banach space.

PROOF. We need to show that $\|\cdot\|_q$ is a norm on V/U, and that V/U is complete. We show only additivity and that if $\|v+U\|_q = 0$, then $v \in U$, that is, $v+U = U$.

Suppose that $v, w \in V$, and observe that

$$\begin{aligned}
\|(v+U)+(w+U)\|_q &= \|v+w+U\|_q \\
&= \inf\{\|v+w+u\| : u \in U\} \\
&= \inf\{\|v+w+u_1+u_2\| : u_1, u_2 \in U\} \\
&\leq \inf\{\|v+u_1\| + \|w+u_2\| : u_1, u_2 \in U\} \\
&= \inf\{\|v+u_1\| : u_1 \in U\} + \inf\{\|w+u_2\| : u_2 \in U\} \\
&= \|v+U\|_q + \|w+U\|_q\,.
\end{aligned}$$

Suppose now that $\|v+U\|_q = 0$. Then there exist $u_n \in U$ such that $\|v+u_n\| \leq 1/n$. Then

$$\|u_m - u_n\| = \|(v+u_m)-(v+u_n)\| \leq \|v+u_m\| + \|v+u_n\| \leq \frac{1}{m} + \frac{1}{n}\,,$$

and the sequence (u_n) is Cauchy. Since U is a closed subspace of a complete space, U is complete, and so $u_n \to u$, for some $u \in U$. It follows that $\|v-u\| = 0$, that is, $v \in U$, so $v+U = U$. □

DEFINITION 4.17. Suppose that U is a closed subspace of a Banach space V. We define the subspace $U^\perp$ of V^* by

$$U^\perp := \{T \in V^* : T|_U = 0\}.$$

COROLLARY 4.18. *Suppose that U is a closed subspace of a Banach space V. If $T \in V^*$, then T determines a continuous linear functional $T|_U$ on U by restriction. The mapping $T \mapsto T|_U$ from V^* to U^* has kernel $U^\perp$, and induces an isometric isomorphism from $V^*/U^\perp$ onto U^*.*

PROOF. Omitted. □

COROLLARY 4.19. *Suppose that U is a closed subspace of a Banach space V. If $T \in U^\perp$, then T determines a continuous linear functional $\dot{T}$ on V/U by*

$$\dot{T}(v+U) = T(v) \quad \forall v \in V.$$

The mapping $T \mapsto \dot{T}$ is an isometric isomorphism from $U^\perp$ onto $(V/U)^*$.

PROOF. Omitted. □

4. More on dual spaces

EXERCISE 4.20. Let $C^k(I)$ be the vector space of all functions $f : I \to \mathbb{F}$ that are k times continuously differentiable, and set

$$\|f\|_{C^k} = \max\{\left|f^{(j)}(t)\right| : t \in I, 0 \le j \le k\}.$$

(a) Show that $(C^k(I), \|\cdot\|_{C^k})$ is a Banach space, and that the map $f \mapsto f^{(j)}(t_0)$ is a continuous linear functional on this space when $t_0 \in I$ and $0 \le j \le k$.
(b) Find a sequence of functions $g_n \in C(I)$ such that

$$f(t_0) = \lim_{n\to\infty} \int_I f(t)\, g_n(t)\, dt \quad \forall f \in C^k(I).$$

(c) Can you find a sequence of functions $g_n \in C(I)$ such that

$$f^{(j)}(t_0) = \lim_{n\to\infty} \int_I f(t)\, g_n(t)\, dt \quad \forall f \in C^k(I).$$

(d) Can you describe the set of all linear functionals on $(C^k(I), \|\cdot\|_{C^k})$?

EXERCISE 4.21. Suppose that $(V, \|\cdot\|)$ is a normed vector space, that U is a subspace of V, and that $U^\perp$ is the subset of V^* of all linear functionals that annihilate U.

(a) Show that $U^\perp$ is closed in the weak-star topology, that is, if $L \in V^*$ and $L_\alpha(v) \to L(v)$ for all $v \in V$ for some net $(L_\alpha)_{\alpha\in\mathcal{A}}$ in $U^\perp$, then $L \in U^\perp$.
(b) Conversely, if W is a weak-star closed subspace of V^*, is W a dual space?

EXERCISE 4.22. Suppose that $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ are normed vector spaces, and that $: U \to V$ is a continuous linear mapping. Define $T^\mathsf{T} : V^* \to U^*$ by the formula

$$(T^\mathsf{T} L)(u) = L(Tu) \quad \forall u \in V,\ \forall L \in V^*.$$

(a) Show that T^T is a continuous linear mapping from V^* to U^*.
(b) Compute $\left\|T^\mathsf{T}\right\|$ in terms of $\|T\|$.

TOPIC 5

Weak topologies and compactness

We show that in infinite-dimensional vector spaces, closed bounded sets need not be compact, and we vary the topology to make it easier for a set to be compact.

1. Weak topologies

Suppose that V is a vector space. One way to define a topology on V is to require that certain functions f on V are continuous. If these functions are $\mathbb{F}$-valued, then we require that the sets $f^{-1}(J)$ are open whenever J is an open subset of $\mathbb{F}$. But every open subset of $\mathbb{F}$ is a union of open balls (these are intervals when $\mathbb{F} = \mathbb{R}$), and $f^{-1}(\bigcup_\alpha J_\alpha) = \bigcup_\alpha f^{-1}(J_\alpha)$, so if we require that $f^{-1}(J)$ is open whenever J is an open ball, then it will follow that $f^{-1}(J)$ is open whenever J is an open subset of $\mathbb{F}$. We use this fact when the functions f are linear.

Definition 5.1. Suppose that $(V, \|\cdot\|)$ is a normed vector space, and that S is a subset of V^*. The $\sigma(V, S)$ topology on V is the topology for which an open base consists of all the sets

$$\{v \in V : L_1(v) \in J_1, L_2(v) \in J_2, \ldots, L_n(v) \in J_n\},$$

where $n \in N$, while $L_j \in S$ and J_j is an open ball in V for all $j \in \{1, 2, \ldots, n\}$. Equivalently, a subbase consists of the sets $\{v \in V : L(v) \in J$, where $L \in S$ and J is an open ball in $\mathbb{F}$.

Observe that this definition only requires that V be a vector space; the norm plays no role, except that we ask that the linear functionals L_j be continuous in the usual (norm) topology. We could also relax this requirement and consider arbitrary linear functionals on an arbitrary vector space V; indeed, the Zariski topology in algebraic geometry is defined in this way (except that the functions used are polynomials and the open sets J are complements of finite sets).

Usually, we consider two particular instances of these topologies.

Definition 5.2. Suppose that $(V, \|\cdot\|)$ is a normed vector space. The *weak topology* on V is the $\sigma(V, V^*)$ topology; if V is a dual space, that is, $V = U^*$ for some Banach space U, then we may view the predual space U as a subspace of the dual space V^* using the injection of U into U^{**} considered in Corollary 4.11), and the weak-star topology is the $\sigma(V, U)$ topology on V.

Lemma 5.3. *The weak and weak-star topologies are coarser than the norm topology, that is, every set that is open in the weak or weak-star topology is open in the norm topology.*

PROOF. It suffices to show that a subbasic set $\{v \in V : L(v) \in J\}$, where $L \in V^*$ and J is an open ball in $\mathbb{F}$, is open in the norm topology.

If $L(v_0) \in J$ and J is open in $\mathbb{F}$, then there exists $\varepsilon \in \mathbb{R}^+$ such that $B_{\mathbb{F}}(L(v_0), \varepsilon) \subseteq J$; now if $\|v - v_0\| < \varepsilon / \|L\|_{V^*}$, then

$$|L(v) - L(v_0)| = |L(v - v_0)| \leq \|L\|_{V^*} \|v - v_0\| < \varepsilon,$$

and so

$$B_V(v_0, \varepsilon / \|L\|_{V^*}) \subseteq \{v \in V : L(v) \in J\},$$

as required. □

LEMMA 5.4. *Suppose that S is a convex subset of V. Then S is weakly closed (closed in the weak topology) if and only if it is closed in the norm topology.*

SKETCH OF THE PROOF. If S is closed in the weak topology, then it is automatically closed in the norm topology. So we may assume that S is closed in the norm topology and show that it is closed in the weak topology.

Given any point v in $V \setminus S$, we can use the Hahn–Banach theorem to construct a continuous linear functional M on V such that $M(v) \notin L(S)^-$. It then follows that

$$S = \bigcap_{L \in V^*} \{v \in V : L(v) \in L(S)^-\}.$$

For each $L \in V^*$, the set $L(S)^-$ is closed in $\mathbb{F}$, and the set $\{v \in V : L(v) \in L(S)^-\}$ is closed in the weak topology; it follows that S is closed. □

EXERCISE 5.5. Suppose that $(V, \|\cdot\|)$ is a normed vector space and that W is a subspace of V^*. Show that the weak topology $\sigma(V, W)$ is the coarsest topology for which all the linear functionals in W are continuous, in the sense that $L^{-1}(S)$ is open in V for all open subsets S of $\mathbb{F}$ and all $L \in W$.

EXERCISE 5.6. Suppose that $(V, \|\cdot\|)$ is a normed vector space and that E is a subset of V^*. Show that the weak topology $\sigma(V, E)$ is the same as $\sigma(V, \operatorname{span} E)$. Is this the same as $\sigma(V, (\operatorname{span} E)^-)$?

2. The Banach–Alaoglu Theorem

THEOREM 5.7. *The closed unit ball $\bar{B}$ in a dual Banach space $(V, \|\cdot\|)$ is compact in the weak-star topology.*

PROOF. Let U be a Banach space such that $V = U^*$. Let $\mathcal{F}$ be the collection of all functions $f : U \to \mathbb{F}$ that satisfy the condition

$$|f(u)| \leq \|u\|_U \quad \forall u \in U.$$

We proceed to identify each function f in $\mathcal{F}$ with the element $(f(u))_{u \in U}$ of the product space $\prod_{u \in U} \bar{B}_{\mathbb{F}}(0, \|u\|_U)$; by Tychonov's theorem, this space is compact in the product topology, since each $\bar{B}_{\mathbb{F}}(0, \|u\|_U)$ is compact in $\mathbb{F}$, and so $\mathcal{F}$ is compact with the topology of pointwise convergence on U.

Each element of $\bar{B}$ is an element of $\mathcal{F}$. It therefore suffices to show that the subset $\bar{B}$ of $\mathcal{F}$ is closed.

But if $f_\alpha \to f$ and each f_α is linear, then

$$f(\lambda u + \mu u') = \lim_\alpha f_\alpha(\lambda u + \mu u') = \lim_\alpha \lambda f_\alpha(u) + \mu f_\alpha(u') = \lambda f(u) + \mu f(u'),$$

so f is linear. □

3. Exercises

EXERCISE 5.8. Suppose that $(V, \|\cdot\|)$ is a Banach space. Show that the unit ball of V is dense in the unit ball of V^{**} in the weak-star topology on V^{**} (that is, the $\sigma(V^{**}, V^*)$ topology).

EXERCISE 5.9. Suppose that $(V, \|\cdot\|)$ is a Banach space. If V is reflexive, is V^* reflexive? If V^* is reflexive, is V reflexive?

TOPIC 6

Linear operators

We study linear mappings between Banach spaces in greater depth.

1. Linear operators

DEFINITION 6.1. Suppose that U and V are normed vector spaces. A *bounded* or *continuous linear operator* $T : U \to V$ is a linear mapping T from U to V for which there exists a constant C such that

$$\|Tu\|_V \leq C \|u\|_U \quad \forall u \in V. \tag{6.1}$$

LEMMA 6.2. *Suppose that $T : U \to V$ is a continuous linear operator. Then*

$$\sup\{\|Tu\|_V : u \in U, \|u\|_U \leq 1\} = \min\{C : (6.1) \textit{ holds}\}. \tag{6.2}$$

PROOF. Clearly, if (6.1) holds, then $C \geq 0$, and if $C' \geq C$, then

$$\|Tu\|_V \leq C' \|u\|_U \quad \forall u \in V.$$

Hence the set of C for which (6.1) holds is an interval, of the form $[C_0, +\infty)$ or $(C_0, +\infty)$, where $C_0 \geq 0$. In either case,

$$\|Tu\|_V \leq (C_0 + \varepsilon) \|u\|_U \quad \forall u \in V$$

for all $\varepsilon \in \mathbb{R}^+$. Exchanging the order of the quantifiers, and letting $\varepsilon \to 0$, we see that

$$\|Tu\|_V \leq C_0 \|u\|_U \quad \forall u \in V.$$

This shows that the infimum is attained, and justifies the use of minimum.

Since

$$\|Tu\|_V \leq C_0 \|u\|_U \quad \forall u \in V,$$

by specialising to u of norm at most 1 we obtain $\|Tu\|_V \leq C_0$, and

$$\sup\{\|Tu\|_V : u \in U, \|u\|_U \leq 1\} \leq C_0.$$

Conversely, for all $u \in U \setminus \{0\}$,

$$\|Tu\|_V = \left\|T(\|u\|_U^{-1} u)\right\|_V \|u\|_U \leq \sup\{\|Tu'\|_V : u' \in U, \|u'\|_U \leq 1\} \|u\|_U ,$$

so $C_0 \leq \sup\{\|Tu\|_V : u \in U, \|u\|_U \leq 1\}$. □

DEFINITION 6.3. The (operator) norm of a linear operator $T : U \to V$ between normed vector spaces $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ is the nonnegative number on either side of equality (6.2).

LEMMA 6.4. *Suppose that $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ are normed vector spaces. Then the set $B(U,V)$ of all continuous linear operators $T : U \to V$, with the norm defined above, is a normed vector space. If V is a Banach space, so is $B(U,V)$.*

PROOF. We omit the proof. It is standard that linear maps between vector spaces form another vector space, so the crucial issue is whether the norm as defined is a norm, and when $B(U,V)$ is complete. □

2. Special types of linear operators

DEFINITION 6.5. Suppose that $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ are normed vector spaces. A linear operator $T : U \to V$ is *compact* if the image of the unit ball in U is precompact in V, that is, $T(B_U(0,1))^-$ is a compact subset of V.

EXAMPLE 6.6. Given a set S and $m \in \ell^\infty(S)$, we define the linear operator $T_m : \ell^p(S) \to \ell^p(S)$ (where $1 \le p \le \infty$) by

$$[T_m f](s) = m(s)\, f(s) \quad \forall s \in S;$$

that is, T_m multiplies by m. It is easy to check that T_m is linear and that

$$\|T_m f\|_p \le \|m\|_\infty \|f\|_p \quad \forall f \in \ell^p(S),$$

so T_m is bounded. If, in addition, m has finite support, then the image of the unit ball in $\ell^p(S)$ under T_m is bounded and contained in a set of functions with finite support, so in a finite-dimensional subspace of $\ell^p(S)$, and T_m is compact.

EXERCISE 6.7. Let $T_m : \ell^p(S) \to \ell^p(S)$ be as in the previous example. Show that T_m is compact if and only if $m \in c_0(S)$, and that if T is compact then T_m may be approximated in operator norm by operators of finite rank.

EXERCISE 6.8. Suppose that $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ are normed vector spaces. Show that T is compact if $T \in B(U,V)$ is of finite rank. Show that, if $T_n \in B(U,V)$ is compact for all $n \in \mathbb{N}$ and $\|T_n - T\| \to 0$ as $n \to \infty$, then T is compact.

The question of whether a general compact operator can be approximated by operators of finite rank is delicate.

EXAMPLE 6.9. Let $T_m : \ell^p(S) \to \ell^p(S)$ be as in Example 6.6. For $t \in S$, let $\delta_t : S \to \{0,1\}$ be the characteristic function of $\{t\}$, that is, $\delta_t(t) = 1$ and $\delta_t(s) = 0$ for all $s \in S \setminus \{t\}$. Evidently, $T_m \delta_t = m(t)\delta_t$, so δ_t is an eigenfunction for T_m with eigenvalue $m(t)$. We may rephrase the result of Exercise 6.7 as saying that T_m is compact if and only if, for any $\varepsilon \in \mathbb{R}^+$, T_m has finitely many eigenfunctions of absolute value greater than ε.

DEFINITION 6.10. Suppose that $(U, \langle\cdot,\cdot\rangle_U)$ and $(V, \langle\cdot,\cdot\rangle_V)$ are Hilbert spaces and that $T \in B(U,V)$. We define *the adjoint $T^* : V \to U$ of T* by

$$\langle T^* v, u\rangle_U = \langle v, Tu\rangle_V \quad \forall u \in U,\ \forall v \in V.$$

LEMMA 6.11. *The map T^* defined above is an element of $B(V,U)$, and $\|T^*\| = \|T\|$. The map $T \mapsto T^*$ is conjugate linear.*

PROOF. We omit much of this proof, as it is standard that the adjoint of a linear map between vector spaces is another linear map; the issue is showing that T^* is bounded and finding its norm. The key is to show that

$$\|T\| = \sup\{|\langle v, Tu \rangle| : u \in U, \|u\|_U \le 1, v \in V, \|v\|_V \le 1\}, \tag{6.3}$$

which follows from the Cauchy–Schwarz inequality. □

DEFINITION 6.12. A linear operator T on a Hilbert space V is *self-adjoint* if $T = T^*$.

EXAMPLE 6.13. For a function $m \in L^\infty(M)$, define $T_m : L^2(M) \to L^2(M)$ by

$$T_m f(t) = m(t)\, f(t) \quad \forall t \in M.$$

The adjoint of T_m is $T_{\bar{m}}$, and so T_m is self-adjoint if m is real-valued.

If $M = I$, the unit interval, and $m(t) = t$ for all $t \in I$, then the operator T_m is self-adjoint. However this operator does not have any eigenvectors or eigenvalues. Indeed, if $T_m f = \lambda f$ for some function $f \in L^2(I)$ and $\lambda \in \mathbb{C}$, then

$$(m(t) - \lambda) f(t) = 0 \quad \forall t \in I$$

and so $f = 0$ almost everywhere. Thus, in contrast to the finite-dimensional case, in infinite-dimensional spaces, being self-adjoint does not ensure that an operator has eigenvectors and eigenvalues. However, T_m is a limit of operators with eigenvalues and eigenvectors. Indeed, define the step functions $m_k : I \to \mathbb{R}$ by $m_k(t) = \lfloor kt \rfloor / k$ for all $t \in I$, where $k \in \mathbb{Z}^+$. Then T_{m_k} is self-adjoint, and has eigenvalues j/k, where $0 \le j < k$, and each eigenspace is infinite-dimensional; further, $T_m = \lim_{k\to\infty} T_{m_k}$ in the operator norm.

It is also worth pointing out that the supremum $\sup\{\|T_m u\|_V : u \in L^2(I), \|u\|_2 \le 1\}$ is not attained for this example.

PROPOSITION 6.14. *Suppose that* $(V, \langle \cdot, \cdot \rangle)$ *is a Hilbert space and that* $T \in B(V, V)$ *is self-adjoint. Then*

$$\|T\| = \sup\{|\langle Tv, v \rangle| : v \in V, \|v\| \le 1\}.$$

PROOF. Define the number M_T as follows:

$$M_T = \sup\{|\langle Tv, v \rangle| : v \in V, \|v\| \le 1\},$$

so that $|\langle Tv, v \rangle| \le M_T \|v\|^2$ for all $v \in V$; note that

$$\|T\| = \sup\{|\langle Tv, w \rangle| : v \in V, \|v\| \le 1, w \in V, \|v\| \le 1\}.$$

so clearly

$$M_T \le \|T\|.$$

Now we show that $\|Tv\| \le M_T \|v\|$, which implies that $\|T\| \le M_T$ and proves the proposition. Clearly this inequality holds if $Tv = 0$, so we may suppose that $Tv \ne 0$. If

$\lambda \in \mathbb{R}^+$, then $\langle T^2 v, \lambda v\rangle = \lambda \langle Tv, Tv\rangle$ since T is self-adjoint, whence

$$\begin{aligned}
&\langle T(Tv+\lambda v), (Tv+\lambda v)\rangle - \langle T(Tv-\lambda v), (Tv-\lambda v)\rangle \\
&\quad = \left\langle T^2 v, Tv\right\rangle + \langle T\lambda v, Tv\rangle + \left\langle T^2 v, \lambda v\right\rangle + \langle T\lambda v, \lambda v\rangle \\
&\qquad - \left\langle T^2 v, Tv\right\rangle + \langle T\lambda v, Tv\rangle + \left\langle T^2 v, \lambda v\right\rangle - \langle T\lambda v, \lambda v\rangle \\
&\quad = 4\lambda \left\|Tv\right\|^2 .
\end{aligned}$$

Hence

$$\begin{aligned}
4\lambda \left\|Tv\right\|^2 &= \langle T(Tv+\lambda v), (Tv+\lambda v)\rangle - \langle T(Tv-\lambda v), (Tv-\lambda v)\rangle \\
&\le M_T \left\|Tv+\lambda v\right\|^2 + M_T \left\|Tv-\lambda v\right\|^2 \\
&= M_T(\langle Tv+\lambda v, Tv+\lambda v\rangle + \langle Tv-\lambda v, Tv-\lambda v\rangle) \\
&= M_T(2\langle Tv, Tv\rangle + 2\lambda^2 \langle v, v\rangle) \\
&= 2M_T(\left\|Tv\right\|^2 + \lambda^2 \left\|v\right\|^2),
\end{aligned}$$

and so

$$2\lambda \left\|Tv\right\|^2 \le M_T(\left\|Tv\right\|^2 + \lambda^2 \left\|v\right\|^2).$$

Take $\lambda = \|Tv\| \,/\, \|v\|$; then

$$2\left\|Tv\right\|^3 / \left\|v\right\| \le 2M_T \left\|Tv\right\|^2 ,$$

and so

$$\|Tv\| \le M_T \|v\| ,$$

as required. □

3. Compact operators on a Hilbert space

An important feature of compact operators in Hilbert spaces is that they attain their norms.

PROPOSITION 6.15. *Suppose that $(V, \langle\cdot,\cdot\rangle)$ is a Hilbert space and that $T \in B(V,V)$ is compact and self-adjoint. Then there exists $v \in V$ such that $\|v\| = 1$ and $\|Tv\| = \|T\|$. Moreover, we may choose v so that $Tv = \lambda v$, where $\lambda \in \mathbb{R}$ and $|\lambda| = \|T\|$.*

PROOF. We may and shall suppose that $T \ne 0$. From Proposition 6.14, we may find a sequence $(v_n)_{n\in\mathbb{N}}$ in V such that $\|v\| = 1$ and $|\langle Tv_n, v_n\rangle| \to \|T\|$ as $n \to \infty$. Since T is self-adjoint, $\langle Tv, v\rangle = \langle v, Tv\rangle = \langle Tv, v\rangle^-$ for any $v \in V$, and so in particular each $\langle Tv_n, v_n\rangle$ is a real number. By passing to a subsequence if necessary, we may assume that $\langle Tv_n, v_n\rangle \to \lambda$ and $\operatorname{sgn}\langle Tv_n, v_n\rangle = \operatorname{sgn}\lambda$, where $\lambda = \pm\|T\|$.

Now

$$\begin{aligned}
0 &\le \langle Tv_n - \lambda v_n, Tv_n - \lambda v_n\rangle \\
&= \langle Tv_n, Tv_n\rangle - 2\lambda\langle Tv_n, v_n\rangle + \lambda^2 \langle v_n, v_n\rangle \\
&\le 2\left\|T\right\|^2 - 2\left\|T\right\| |\langle Tv_n, v_n\rangle| \to 0
\end{aligned}$$

as $n \to \infty$, and hence $Tv_n - \lambda v_n \to 0$ in V.

Because T is compact, the bounded sequence Tv_n admits a convergent subsequence; by again passing to a subsequence if necessary, we may suppose that $(Tv_n)_{n\in\mathbb{N}}$ converges, and hence $(\lambda v_n)_{n\in\mathbb{N}}$ converges, to λv say. Now $Tv = \lim_n Tv_n = \lim_n \lambda v_n = \lambda v$. □

THEOREM 6.16. *Suppose that $(V, \langle\cdot,\cdot\rangle)$ is a Hilbert space and that $T \in B(V,V)$ is compact and self-adjoint. Then there exist a set $\mathcal{N}$, equal to $\mathbb{Z}^+$ or $\{1,\dots,n\}$ for some positive integer n or $\emptyset$, and an orthonormal set $\{v_n : n \in \mathcal{N}\}$ of eigenvectors v_n with corresponding nonzero eigenvalues λ_n, such that $|\lambda_1| \geq |\lambda_2| \geq \dots$ and $\lambda_n \to 0$ as $n \to \infty$ if $\mathcal{N} = \mathbb{Z}^+$. Further $T = 0$ on $\{v_n : n \in \mathcal{N}\}^\perp$.*

PROOF. We may suppose that $T \neq 0$, and argue by induction.

By the previous proposition, there exists $v_1 \in V$ such that $\|v_1\| = 1$ and $Tv_1 = \lambda_1 v_1$, where $|\lambda| = \|T\|$.

Now we take $V_1 = \{v_1\}^\perp$, and observe that if $v \in V_1$, then

$$\langle Tv, v_1\rangle = \langle v, Tv_1\rangle = \lambda_1 \langle v, v_1\rangle = 0,$$

and so $T(V_1) \subseteq V_1$. We apply the proposition to $T|_{V_1}$, the restriction of T to V_1, and find $v_2 \in V_1$ such that $\|v_2\| = 1$ and $Tv_2 = \lambda_2 v_2$, where $|\lambda_2| = \|T|_{V_1}\|$. Clearly $|\lambda_1| \geq |\lambda_2|$.

Now we take $V_2 = \{v_1, v_2\}^\perp$, and observe that if $v \in V_2$, then

$$\langle Tv, v_j\rangle = \langle v, Tv_j\rangle = \lambda_j \langle v, v_j\rangle = 0$$

when $j \in \{1,2\}$, and so $T(V_2) \subseteq V_2$. We apply the proposition to $T|_{V_2}$, the restriction of T to V_2, and find $v_3 \in V_2$ such that $\|v_3\| = 1$ and $Tv_3 = \lambda_3 v_3$, where $|\lambda_3| = \|T|_{V_2}\|$. Clearly $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$.

Either this process continues inductively then stops after n steps because $T|_{V_n} = 0$, or it continues infinitely. In the latter case, the set $\{\lambda_1 v_1, \lambda_2 v_2, \dots\}$ is contained in a compact set, and this is only possible if for all $\varepsilon \in \mathbb{R}^+$, there are only finitely many λ_n such that $|\lambda_n| \geq \varepsilon$, that is, $\lambda_n \to 0$. □

COROLLARY 6.17. *Suppose that $(U, \langle\cdot,\cdot\rangle_U)$ and $(V, \langle\cdot,\cdot\rangle_V)$ are Hilbert spaces and that $T \in B(U,V)$ is compact. Then there exist a set $\mathcal{N}$, equal to $\mathbb{Z}^+$ or $\{1,\dots,n\}$ for some positive integer n or $\emptyset$, and orthonormal sets $\{u_n : n \in \mathcal{N}\}$ in U and $\{v_n : n \in \mathcal{N}\}$ in V, such that $Tu_n = \mu_n v_n$ for all $n \in \mathcal{N}$, where $\mu_1 \geq \mu_2 \geq \dots > 0$ and $\mu_n \to 0$ as $n \to \infty$ if $\mathcal{N} = \mathbb{Z}^+$. Further $T = 0$ on $\{u_n : n \in \mathcal{N}\}^\perp$.*

PROOF. The operator T^*T is *positive*, that is, $\langle T^*Tu, u\rangle \geq 0$ for all $u \in U$. Hence, when we apply the previous theorem to T^*T, the nonzero eigenvalues of T^*T produced by the theorem are all positive, and for the corresponding eigenvectors u_n we may write

$T^*Tu_n = \mu_n^2 u_n$, where $\mu_n > 0$. Take $v_n = \mu_n^{-1} Tu_n$. Now

$$\begin{aligned}\langle v_m, v_n\rangle &= \mu_m^{-1}\mu_n^{-1}\langle Tu_m, Tu_n\rangle \\ &= \mu_m^{-1}\mu_n^{-1}\langle T^*Tu_m, u_n\rangle \\ &= \mu_m^{-1}\mu_n^{-1}\mu_m^2\langle u_m, u_n\rangle \\ &= \begin{cases} 1 & \text{if } m = n \\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

and so the v_n also form an orthonormal set. □

The decomposition in the corollary is called the *singular value decomposition of T*.

EXERCISE 6.18. Suppose that $(U, \langle\cdot,\cdot\rangle_U)$ and $(V, \langle\cdot,\cdot\rangle_V)$ are Hilbert spaces, and that $T : U \to V$ is a linear map. Show that T^* is compact if T is compact. Is the converse true?

4. The open mapping theorem

THEOREM 6.19. *Suppose that $T : U \to V$ is a continuous linear operator between Banach spaces $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$. If T is surjective, then T is an open mapping, that is, the image of an open set under T is open.*

PROOF. This proof uses the Baire category theorem.

Since T is surjective and linear,

$$V = TU = T\bigcup_{n\in\mathbb{N}} B_U(0,n) = \bigcup_{n\in\mathbb{N}} TB_U(0,n) = \bigcup_{n\in\mathbb{N}} nTB_U(0,1),$$

and *a fortiori*

$$V = \bigcup_{n\in\mathbb{N}} (nTB_U(0,1))^-.$$

By the Baire category theorem, there exists $n \in \mathbb{N}$ such that $(nTB_U(0,1))^-$ has nonempty interior. Since multiplication by nonzero scalars is a homeomorphism, it follows that $(TB_U(0,1))^-$ has nonempty interior, that is,

$$(TB_U(0,1))^- \supseteq B_V(v_0, r_0)$$

for some $v_0 \in V$ and $r_0 \in \mathbb{R}^+$. In particular, $TB_U(0,1) \cap B_V(v_0,r_0)$ is dense in $B_V(v_0,r_0)$.

If $v \in B_V(0,r_0)$, then $v + v_0 \in B_V(v_0,r_0)$, and so there exists a sequence $(u_n)_{n\in\mathbb{N}}$ in $B_U(0,1)$ such that $Tu_n \to v + v_0$ as $n \to \infty$; in particular, there is a sequence $(u'_n)_{n\in\mathbb{N}}$ such that $Tu_n \to v_0$. Now $u_n - u'_n \in B_U(0,2)$ and $T(u_n - u'_n) \to v + v_0 - v_0 = v$ as $n \to \infty$, and thus

$$(TB(0,2))^- \supseteq B_V(0,r_0).$$

I claim that, given $v \in B_V(0, r_0)$, there exists a sequence $(u_n)_{n\in\mathbb{N}}$ in $B_U(0, 2)$ such that

$$v = \sum_{n\in\mathbb{N}} 2^{-n} T u_n = \sum_{n\in\mathbb{N}} T(2^{-n} u_n) = T\Big(\sum_{n\in\mathbb{N}} 2^{-n} u_n\Big).$$

Since

$$\Big\|\sum_{n\in\mathbb{N}} 2^{-n} u_n\Big\|_U \le \sum_{n\in\mathbb{N}} \|2^{-n} u_n\|_U \le \sum_{n\in\mathbb{N}} 2^{1-n} = 4,$$

this claim implies that $v \in TB_U(0, 4)$, and so $TB_U(0, 4) \supseteq B_V(0, r_0)$. By linearity, it follows that

$$TB_U(u_0, r) = Tu_0 + TB_U(0, r) \supseteq Tu_0 + B_V(0, rr_0/4),$$

and hence that T sends open sets to open sets. Indeed, if u_0 is a point in an open set E in U, then $B_U(u_0, r) \supseteq E$ for a suitably small r_0, and so the typical point Tu_0 in TE lies inside a ball $B_V(Tu_0, rr_0/4)$ in V that is a subset of TE.

To prove the claim, we choose the points $u_n \in B_U(0, 2)$ recursively. First, take u_0 so that $\|v - Tu_0\|_V \le r_0/2$. Once we have chosen u_0, u_1, ..., u_k in $B_U(0, 2)$ such that

$$\Big\|v - \sum_{n=0}^{k} 2^{-n} T u_n\Big\|_V < 2^{-k-1} r_0,$$

$2^{k+1}(v - \sum_{n=0}^{k} 2^{-n} T u_n) \in B_V(0, r_0)$, so we may choose $u_{k+1} \in B_U(0, 2)$ such that

$$\Big\|2^{k+1}(v - \sum_{n=0}^{k} 2^{-n} T u_n) - Tu_{k+1}\Big\|_V < \frac{r_0}{2},$$

and then

$$\Big\|v - \sum_{n=0}^{k+1} 2^{-n} T u_n\Big\| = \Big\|\Big(v - \sum_{n=0}^{k} 2^{-n} T u_n\Big) - 2^{-k-1} T u_{k+1}\Big\| < 2^{-k-2} r_0.$$

Once the u_n have been chosen, passing to the limit as $k \to \infty$ proves the claim, and hence the theorem. □

COROLLARY 6.20. *Suppose that $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ are Banach spaces, and that $T : U \to V$ is a continuous linear operator. If T is bijective, then T^{-1} is also continuous.*

PROOF. This follows easily from the theorem. □

In the next exercises and examples, given an (integrable) functions g on the unit interval I, we write $\hat{g}$ for its Fourier transform, which is a function on $\mathbb{Z}$:

$$\hat{g}(n) = \int_I g(t) e^{-2\pi i n t}\, dt \quad \forall n \in \mathbb{Z}.$$

We study the properties of the linear operator $g \mapsto \hat{g}$ from $L^1(I)$ to $c_0(\mathbb{Z})$.

EXERCISE 6.21. Show that

$$\|\hat{g}\|_\infty \leq \|g\|_1 \quad \forall g \in L^1(I).$$

Deduce that $\hat{f} \in c_0(\mathbb{Z})$ for all $f \in L^1(I)$. You may assume that every function $f \in L^1(I)$ may be written as $\lim_{k\to\infty} f_k$, where each $f_k \in L^2(I)$ and $\hat{f}_k \in \ell^2(\mathbb{Z}) \subseteq c_0(\mathbb{Z})$.

EXERCISE 6.22. Define the *Dini kernel* $D_N : I \to \mathbb{C}$ by

$$D_N(t) = \sum_{|n|\leq N} e^{2\pi int} \quad \forall t \in I.$$

(a) Show that

$$D_N(t) = \frac{\sin((2N+1)\pi t)}{\sin(\pi t)}.$$

(b) Show that there is a positive constant C such that

$$\int_I |D_N(t)|\, dt \geq C \log(N+1) \quad \forall N \in \mathbb{N}.$$

EXAMPLE 6.23. The continuous linear mapping $g \mapsto \hat{g}$ from $L^1(I)$ to $c_0(\mathbb{Z})$ is injective (from Fourier series), but not surjective. Indeed, if it were surjective, then by the corollary to the open mapping theorem, the inverse mapping would be continuous, that is, there would be a constant C such that

$$\|f\|_1 \leq C \|\hat{f}\|_\infty \quad \forall f \in L^1(I).$$

However, the functions D_N studied above show that no such constant can exist, and so the Fourier transformation $g \mapsto \hat{g}$ cannot be surjective.

EXERCISE 6.24. By using the open mapping theorem, show that there exists a function $f \in C(I)$ such that $f(0) = f(1)$ and $\sum_{n\in\mathbb{Z}} |\hat{f}(n)| = \infty$.

5. The closed graph theorem

The graph $G(T)$ of a linear operator $T : U \to V$ is the set of all pairs $(u, v) \in U \times V$ such that $v = Tu$. It is clear that it is a subspace of $U \times V$.

THEOREM 6.25. *Suppose that $T : U \to V$ is a continuous linear operator between Banach spaces $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$. Then T is continuous if and only if $G(T)$ is a closed subspace of $U \times V$.*

PROOF. Write P_1 and P_2 for the projections of $U \times V$ onto the factors U and V respectively. A sequence of pairs $((u_n, v_n))_{n\in\mathbb{N}}$ in $U \times V$ converges if and only if the sequences $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ converge in U and V, so the projections P_1 and P_2 are continuous. If $(u_n, v_n) \in G(T)$, and $(u_n, v_n) \to (u, v)$ in $U \times V$ as $n \to \infty$, then $u_n \to u$, and $v_n = Tu_n \to Tu$ if T is continuous. But $v_n \to v$, and so $v = Tu$, proving that $(u, v) \in G(T)$ and $G(T)$ is closed.

Conversely, suppose that $G(T)$ is closed. The linear mapping $P_1 : U \times V \to U$ is continuous, and so its restriction $P_1|_{G(T)}$ to $G(T)$ is also continuous. Further, $P_1|_{G(T)}$ is birjective, since $P_1(u, Tu) = u$ for all $u \in U$. Then $P_1|_{G(T)}{}^{-1} : U \to G(T)$ is also continuous, by the corollary to the open mapping theorem. Since $T = P_2 P_1|_{G(T)}{}^{-1}$, it follows that T is continuous. □

EXERCISE 6.26. Suppose that $m \in \ell^\infty(\mathbb{Z})$, and suppose that for all $f \in L^p(I)$, there exists $T_m f \in L^p(I)$ such that $(T_m f)\widehat{\ }(n) = m(n)\hat{f}(n)$ for all $n \in \mathbb{Z}$. Show that the mapping T_m is continuous and linear.

6. The uniform boundedness principle

THEOREM 6.27. *Suppose that $(V, \|\cdot\|_V)$ is a Banach space and $(W, \|\cdot\|_W)W$ is a normed vector space. Suppose also that $T_n : V \to W$ is a continuous linear operator for all $n \in \mathbb{N}$, such that $\sup_{n\in\mathbb{N}} \|T_n v\|_W < \infty$ for all $v \in V$. Then*

$$\sup_{n\in\mathbb{N}} \|T_n\|_{B(V,W)} < \infty.$$

PROOF. For any positive real number r, define $E_r = \{v \in V : \sup_{n\in\mathbb{N}} \|T_n v\|_W \leq r\}$. Then

$$E_r = \bigcap_{n\in\mathbb{N}} \{v \in V : \|T_n v\|_W \leq r\},$$

and so E_r is closed, as each set $\{v \in V : \|T_n v\|_W \leq r\}$ is closed. Indeed, if $v_j \in V$ and $v_j \to v$ as $j \to \infty$, then

$$\|T_n v\|_W = \lim_j \|T_n v_j\|_W \leq r.$$

By definition and hypothesis, $V = \bigcup_{r\in\mathbb{Z}^+} E_r$. By the Baire category theorem, at least one of the sets E_r must not be nowhere dense, that is, at least one E_r must contain an open set, and hence contain an open ball $B(v_0, s)$, where $v_0 \in V$ and $s \in \mathbb{R}^+$. □

EXERCISE 6.28. Suppose that $(V, \|\cdot\|_V)$ is a Banach space and $(W, \|\cdot\|_W)$ is a normed vector space. Suppose also that $T_n : V \to W$ is a continuous linear operator for each $n \in \mathbb{N}$, and that $\lim T_n v$ exists for all $v \in V$. Prove that

$$\sup_{n\in\mathbb{N}} \|T_n\|_{B(V,W)} < \infty.$$

EXERCISE 6.29.

7. Further problems

EXERCISE 6.30. Suppose that $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ are normed vector spaces, and that $T : U \to V$ is a linear map. Show that T^T is compact if T is compact. Is the converse true?

TOPIC 7

Approximation in spaces of continuous functions

In this chapter, we give some conditions for a subspace A of the space $C(X)$ of continuous functions on a topological space X to be dense in $C(X)$. This type of result is used in approximation theory, as well as to show that the span of an orthogonal set of continuous functions $\mathcal{A}$ is dense in $C(X)$ and so also in $L^2(X)$, and hence to establish that $\mathcal{A}$ is a basis for $L^2(X)$.

1. Weierstrass' approximation theorem

THEOREM 7.1. *Every continuous function on the interval I may be approximated uniformly by polynomials.*

PROOF. Suppose that $f \in C(I)$. Choose $\varepsilon \in \mathbb{R}^+$. We must find a polynomial p such that $|f(t) - p(t)| < \varepsilon$ for all $t \in I$.

Extend f to a continuous function on $\mathbb{R}$, supported in $[-1, 2]$ (for instance, extend f linearly in $[-1, 0]$ and in $[1, 2]$), still denoted f.

Define the Weierstrass kernel $w_\delta : \mathbb{R} \to \mathbb{R}$, where $\delta \in \mathbb{R}^+$, by

$$w_\delta(t) = \frac{1}{\sqrt{2\pi\delta}} e^{-t^2/2\delta} \quad \forall t \in \mathbb{R},$$

and define the convolution $w_\delta * f : \mathbb{R} \to \mathbb{F}$ by

$$w_\delta * f(t) = \int_{\mathbb{R}} w_\delta(t-s)\, f(s)\, ds = \int_{\mathbb{R}} w_\delta(s)\, f(t-s)\, ds \quad \forall t \in \mathbb{R}.$$

We will argue that $w_\delta * f$ approximates f (as $\delta \to 0$) and then, by taking finitely many terms of the power series expansion of $w_\delta(t-s)$ in powers of $t-s$, show that $w_\delta * f(t)$ may be approximated by a polynomial in t.

Since f is continuous and compactly supported, it is uniformly continuous, which means that there exists $\eta \in \mathbb{R}^+$ such that $|f(t-s) - f(t)| < \frac{1}{4}\varepsilon$ for all $t \in \mathbb{R}$ provided that $|s| < \eta$. Further, it is not hard to see that

$$\int_{\mathbb{R}} w_\delta(s)\, ds = 1$$

$$\int_{\mathbb{R}\setminus[-\eta,\eta]} w_\delta(s)\, ds \to 0+ \quad \text{as } \delta \to 0+,$$

irrespective of the value of η, so in particular for η chosen above, we may take $\delta \in \mathbb{R}^+$ such that

$$\int_{\mathbb{R}\setminus[-\eta,\eta]} w_\delta(s)\,ds < \frac{\varepsilon}{4\,\|f\|_\infty}\,.$$

Now for this δ and η,

$$\begin{aligned} w_\delta * f(t) - f(t) &= \int_{\mathbb{R}} w_\delta(s)\,[f(t-s)-f(t)]\,ds \\ &= \int_{\mathbb{R}\setminus[-\eta,\eta]} w_\delta(s)\,[f(t-s)-f(t)]\,ds + \int_{[-\eta,\eta]} w_\delta(s)\,[f(t-s)-f(t)]\,ds \end{aligned}$$

for all $t \in \mathbb{R}$, and hence

$$\begin{aligned} |w_\delta * f(t) - f(t)| &\le \int_{\mathbb{R}\setminus[-\eta,\eta]} w_\delta(s)\,|f(t-s)-f(t)|\,ds + \int_{[-\eta,\eta]} w_\delta(s)\,|f(t-s)-f(t)|\,ds \\ &\le \int_{\mathbb{R}\setminus[-\eta,\eta]} w_\delta(s)\,(2\,\|f\|_\infty)\,ds + \int_{[-\eta,\eta]} w_\delta(s)\,(\tfrac{1}{4}\varepsilon)\,ds \\ &\le 2\,\|f\|_\infty \int_{\mathbb{R}\setminus[-\eta,\eta]} w_\delta(s)\,ds + \tfrac{1}{4}\varepsilon \int_{\mathbb{R}} w_\delta(s)\,ds \\ &\le \tfrac{1}{4}\varepsilon + \tfrac{1}{4}\varepsilon = \tfrac{1}{2}\varepsilon \end{aligned}$$

for all $t \in \mathbb{R}$.

Now we show that $w_\delta * f$ may be approximated by a polynomial to within $\varepsilon/2$ on I. This is not possible on all $\mathbb{R}$, since f is zero outside $[-1,2]$ and any polynomial that is bounded outside $[-1,2]$ must be constant.

Observe that

$$\begin{aligned} w_\delta * f(t) &= \int_{\mathbb{R}} w_\delta(t-s)\,f(s)\,ds \\ &= \frac{1}{\sqrt{2\pi\delta}} \int_{\mathbb{R}} e^{-(t-s)^2/2\delta}\,f(s)\,ds \\ &= \frac{1}{\sqrt{2\pi\delta}} \int_{\mathbb{R}} \sum_{n=0}^{\infty} \frac{[-(t-s)^2/2\delta]^n}{n!}\,f(s)\,ds. \end{aligned}$$

If $t \in I$ and $s \in [-1,2]$, then $|t-s| \le 2$. Choose $N \in \mathbb{N}$ such that

$$\sum_{n=N+1}^{\infty} \frac{[2/\delta]^n}{n!} < \frac{\varepsilon\sqrt{2\pi\delta}}{2\,\|f\|_1}\,,$$

and then

$$
\begin{aligned}
\left| \frac{1}{\sqrt{2\pi\delta}} \int_{\mathbb{R}} \sum_{n=N+1}^{\infty} \frac{[-(t-s)^2/2\delta]^n}{n!} f(s)\, ds \right| &\le \frac{1}{\sqrt{2\pi\delta}} \int_{\mathbb{R}} \left| \sum_{n=N+1}^{\infty} \frac{[-(t-s)^2/2\delta]^n}{n!} \right| |f(s)|\, ds \\
&\le \frac{1}{\sqrt{2\pi\delta}} \int_{\mathbb{R}} \sum_{n=N+1}^{\infty} \frac{|[-(t-s)^2/2\delta]^n|}{n!} |f(s)|\, ds \\
&\le \frac{1}{\sqrt{2\pi\delta}} \int_{\mathbb{R}} \sum_{n=N+1}^{\infty} \frac{[2/\delta]^n}{n!} |f(s)|\, ds \\
&\le \frac{1}{\sqrt{2\pi\delta}} \int_{\mathbb{R}} \frac{\varepsilon\sqrt{2\pi\delta}}{2\left\| f \right\|_1} |f(s)|\, ds \\
&= \tfrac{1}{2}\varepsilon
\end{aligned}
$$

for all $t \in I$, as required. □

2. The Müntz–Szász Theorem

According to the Weierstrass theorem, every continuous function f on I may be approximated uniformly by polynomials. However, more is true: suppose, for instance, that $f \in C(I)$, and define $g \in C(I)$ by $g(s) = f(s^{1/2}$ for all $s \in I$. We may approximate g by polynomials, and since

$$
\max\{|g(s) - p(s)| : s \in I\} = \max\{\left|g(s^2) - p(s^2)\right| : s \in I\} = \max\{\left|f(t) - p(t^2)\right| : t \in I\},
$$

we may also approximate f by polynomials involving even powers only. How many powers do we need to be able to approximate continuous functions? This question is answered by the next result.

THEOREM 7.2. *Suppose that $\Lambda \subseteq \mathbb{N}$. Every continuous function on the interval I may be approximated uniformly by polynomials involving powers from Λ if and only if*

(a) $0 \in \Lambda$;

(b) $\displaystyle\sum_{n \in \Lambda \setminus \{0\}} \frac{1}{n} = +\infty.$

SKETCH OF THE PROOF. Write $\mathbb{C}^+$ for the right half plane $\{x + iy : x \in \mathbb{R}^+, y \in \mathbb{R}\}$ in $\mathbb{C}$, and A for the closure in $C(I)$ of $\operatorname{span}\{t \mapsto t^n : n \in \Lambda\}$. We need the condition $0 \in \Lambda$ to ensure that there are functions in A that do not vanish at 0.

For every continuous linear functional L on $C(I)$, we associate a function $h_L : \mathbb{C}^+ \to \mathbb{C}$ by the formula

$$
h_L(z) = L(t \mapsto t^z) \quad \forall z \in \mathbb{C}^+
$$

It is clear that $\|h_L\|_\infty \le \|L\|_{C(I)^*}$, and that $L(f) = 0$ for all $f \in A$ if and only if $h_L(n) = 0$ for all $n \in \Lambda$.

From complex analysis, much is known about the space $H^\infty(\mathbb{C}^+)$ of bounded holomorphic functions on $\mathbb{C}^+$. In particular, suppose that $\{0\} \subseteq \Lambda \subseteq \mathbb{N}$. Then there exists $h \in H^\infty(\mathbb{C}^+)$ such that $h(n) = 0$ for all $n \in \Lambda \subseteq \mathbb{N}$ if and only if condition (b) above holds. This result is key to the proof of the theorem.

On the one hand, if A is not dense in $C(I)$, then by the Hahn–Banach theorem, there exists a linear functional L on $C(I)$ such that $L|_A = 0$ but $L \neq 0$. By Weierstrass' theorem, there exists $n \in \mathbb{N}$ such that $L(t \mapsto t^n) \neq 0$, and $n \notin \Lambda$ by our earlier observation. Hence $h_L \neq 0$, so condition (b) above does not hold.

On the other hand, if condition (b) above does not hold, then the function that is constructed using complex analysis turns out to be of the form h_L for some linear functional L. We can certainly set $L(t \mapsto t^n) = h(n)$, and so define $L(p)$ for all polynomials p in $C(I)$. What is needed is to show that there exists a constant C such that $|L(p)| \leq C \|p\|_\infty$ for all polynomials p; this ensures that we may define $L(f)$ unambiguously as $\lim L(p_n)$ for any sequence of polynomials p_n that converge to f. To do so requires some more complex analysis and we omit the details. □

3. The Stone–Weierstrass theorem

Our final theorem concerns functions on a general compact Hausdorff topological space. This theorem enables us to answer questions about approximation by polynomials on compact sets in $\mathbb{R}^n$, such as spheres and rectangles, and more. That is quite a generalisation!

THEOREM 7.3. *Suppose that X is a compact Hausdorff topological space, and that A is a subset of $C(X)$. Suppose also that*

(a) A is an algebra under pointwise operations;
(b) A contains the constant function 1*;*
(c) A is closed under (pointwise) complex conjugation;
(d) A separates points, that is, given two distinct points x and y in X, there exists $f \in A$ such that $f(x) \neq f(y)$.

Then A is dense in $C(X)$, so every continuous function may be approximated uniformly by elements of A.

PROOF. If A satisfies the conditions above, so does $\bar{A}$. Hence we may and shall suppose that A is also closed, and show that $A = C(X)$.

If $\mathbb{F} = \mathbb{C}$, then the real part and the imaginary part of every function in A lie in A, and if A satisfies the conditions above, the so does $A_\mathbb{R}$, the (real) subspace of A of real-valued functions, provided that we restrict to real scalars. Further, if $A_\mathbb{R} = C(X, \mathbb{R})$, then $A = C(X, \mathbb{C})$. Hence we may and shall suppose that $\mathbb{F} = \mathbb{R}$.

Step 1: some special polynomials on I. By the argument given just before the statement of this theorem, it is possible to approximate $t \mapsto t$ uniformly on I by (real) polynomials in even powers of t, and so it is possible to approximate $t \mapsto |t|$ uniformly on $[-1, 1]$ by even polynomials. In other words, for each $\varepsilon \in \mathbb{R}^+$, there exists an even polynomial p_ε

such that

$$|p_\varepsilon(t) - |t|| < \varepsilon \quad \forall t \in I.$$

Step 2: A is closed under taking absolute values. Suppose that $f \in A$ and $\|f\|_\infty \leq 1$. Since A is an algebra, $p_\varepsilon \circ f \in A$. Further,

$$\begin{aligned}\max\{|p_\varepsilon \circ f(x) - |f(x)|| : x \in X\} &= \max\{|p_\varepsilon(f(x)) - |f(x)|| : x \in X\} \\ &\leq \max\{|p_\varepsilon(t) - |t|| : t \in [-1, 1]\} < \varepsilon.\end{aligned}$$

Thus we may approximate $|f|$ by polynomials in f. Since A is closed, we deduce that $|f| \in A$. For an arbitrary $f \in A \setminus \{0\}$, let $c = \|f\|_\infty^{-1}$. We see that $cf \in A$ and $\|cf\|_\infty = 1$, whence $|cf| \in A$, and so $|f| \in A$.

Step 3: A is closed under taking maxima and minima of two functions. Suppose that $f, g \in A$. Observe that

$$\max\{f(x), g(x)\} = \frac{1}{2}\big(f(x) + g(x) + |f(x) - g(x)|\big) \quad \forall x \in X,$$

and so $\max\{f, g\} \in A$. Similarly, $\min\{f, g\} \in A$.

Step 4: A contains "bump functions" at every point. Suppose that K is a compact subset of X and $p \in X \setminus K$. Because A is a vector space containing constants, and A separates points, for each $q \in K$, there exists a function $f_q \in A$ such that $f_q(p) = 1$ and $f_q(q) = -1$; then the set $\{x \in X : f_q(x) < 0\}$ is open and contains q. The various sets $\{x \in X : f_q(x) < 0\}$, where the index q ranges over K, form an open cover of K, and since K is compact, there exists a finite subcover, indexed by q_1, q_2, …, q_n, say. Set

$$b_p = \max\{\min\{f_{q_1}, f_{q_2}, , f_{q_n}, 1\}, 0\}.$$

Then $b_p \in A$, and $b_p(p) = 1$ while $b_p(x) = 0$ for all $x \in K$. Further, $0 \leq b_p \leq 1$.

Step 5: $A = C(X)$. Take $f \in C(X)$ and $\varepsilon \in \mathbb{R}^+$. We shall approximate f by a function in A to within ε. Since A is closed, this implies that $A = C(X)$, as required.

Because A contains constants, it suffices to approximate $f + \lambda 1$, for some real λ. Thus we may assume that $f \geq 0$. We will construct $g \in A$ such that $f - \varepsilon < g < f + \varepsilon$.

For each p in X, define $K = \{x \in X : f(x) \leq f(p) - \varepsilon\}$, and consider the bump function b_p constructed in Step 4. If $x \in K$, then $f(p)b_p(x) = 0 < f(x) + \varepsilon$, while if $x \notin K$, then

$$f(p)b_p(x) \leq f(p) < f(x) + \varepsilon,$$

by definition of K, and so

$$f(p)b_p(x) < f(x) + \varepsilon \quad \forall x \in X.$$

Finally, the function $x \mapsto f(p)b_p(x) - f(x)$ is continuous and takes the value 0 when $x = p$, and so the set $\{x \in X : f(p)b_p(x) - f(x) > -\varepsilon\}$ is open and contains p. Thus the various sets $\{x \in X : f(p)b_p(x) - f(x) > -\varepsilon\}$, as the index p ranges over X, form an open cover of X, and since K is compact, there exists a finite subcover, indexed by p_1, p_2, …, p_m, say. Set

$$g = \max\{f(p_1)b_{p_1}, f(p_2)b_{p_2}, \ldots, f(p_m)b_{p_m}\}.$$

Then $g \in A$, and on the one hand, $g < f + \varepsilon$ since each $f(p)b_p < f + \varepsilon$, while on the other hand, each $x \in X$ belongs to at least one of the sets $\{x \in X : f(p_j)b_{p_j}(x) - f(x) > -\varepsilon\}$, and hence $g(x) \geq f(p_j)b_{p_j}(x) > f(x) - \varepsilon$. □

This theorem has many important consequences, many of which were already known before it was proved, but with a variety of different proofs, so it allows us to present a unified approach.

COROLLARY 7.4. *The set of trigonometric polynomials* $t \mapsto \sum_{n\in\mathbb{Z}} a_n e^{2\pi int}$ *(where only finitely many* a_n *are nonzero) is dense in the set of continuous* 1*-periodic functions on* $\mathbb{R}$.

PROOF. Let T denote the unit circle in $\mathbb{C}$. By the Stone–Weierstrass theorem, the algebra of all polynomials in the usual coordinates x and y on $\mathbb{C}$ is dense in $C(T)$. However, every polynomial in x and y may also be written as a polynomial in z and $\bar{z}$, and $z\bar{z} = 1$ on T so all monomials of the form $z^j \bar{z}^k$ may be simplified to z^{j-k} or $\bar{z}^{k-j}$, depending on whether $j \geq k$ or $k \geq j$. We conclude that the set of all polynomials of the form $\sum_{n\in\mathbb{Z}} a_n z^n$ (where only finitely many a_n are nonzero) is dense in $C(T)$.

Now $E : t \mapsto e^{2\pi it}$ is a continuous surjection from $\mathbb{R}$ to T, so $f \mapsto f \circ E$ maps continuous functions on T to continuous 1-periodic functions on $\mathbb{R}$. It is easy to see that this map is a bijection, and that the trigonometric polynomial $t \mapsto \sum_{n\in\mathbb{Z}} a_n e^{2\pi int}$ corresponds to the polynomial $\sum_{n\in\mathbb{Z}} a_n z^n$ on T. □

COROLLARY 7.5. *The Fourier transformation* $f \mapsto (\hat{f}(n))_{n\in\mathbb{Z}}$ *is injective on* $L^1(I)$.

PROOF. If $\hat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then

$$\int_I f(t)\, p(t)\, dt = 0$$

for all trigonometric polynomials p. Since the set of trigonometric polynomials is dense in $C(T)$, it follows that

$$\int_I f(t)\, g(t)\, dt = 0$$

for all $g \in C(T)$.

From the theory of measure and integration, we know that for each $f \in L^1(I)$ and $\varepsilon \in \mathbb{R}^+$, there exists $f_1 \in C(T)$ such that $\|f - f_1\|_1 < \varepsilon$. It follows that

$$\begin{aligned}
\left|\int_I f_1(t)\, g(t)\, dt\right| &= \left|\int_I f(t)\, g(t)\, dt - \int_I [f - f_1](t)\, g(t)\, dt\right| \\
&\leq \int_I |[f - f_1](t)|\, |g(t)|\ dt \\
&\leq \left(\int_I |[f - f_1](t)|\ dt\right) \|g\|_\infty \\
&< \varepsilon \, \|g\|_\infty \,.
\end{aligned}$$

It follows that $\|f_1\|_1 < \varepsilon$, and then that $\|f\|_1 < 2\varepsilon$; since ε is arbitrary, we deduce that $f = 0$ almost everywhere. $\square$

APPENDIX A

Notes on topology and nets

You should know the following definitions and facts from Analysis. Suppose that (M, d) is a metric space. First, for p in M and r in $\mathbb{R}^+$, the *open ball* $B(p, r)$ is the subset of M given by

$$B(p, r) = \{q \in M : d(q, p) < r\}.$$

We say that a subset S of M is *open* if S is empty, or if S may be written as a union of open balls; equivalently, given any point $p \in S$, there exists $r \in \mathbb{R}^+$ such that $B(p, r) \subseteq S$. A subset S of M is *closed* if its complement $M \setminus S$ is open.

Any union of open sets is open and any intersection of closed sets is closed. In any topological space M, the *closure* of a subset S is the smallest closed subset of M that contains S; it is the intersection of all closed subsets of M that contain S. A subset S of M is *dense* if its closure is M itself.

In metric spaces, topological concepts such as continuity may be expressed using open sets or sequences. For instance, a function f is continuous provided that $f^{-1}(U)$ is open whenever U is open, or equivalently, provided that $f(x_n) \to f(x)$ whenever $x_n \to x$. Similarly, K is a compact set if every cover of K by open sets has a finite subcover, or equivalently if every sequence in K has a convergent subsequence. To solve problems about metric spaces, sometimes one point of view is more convenient, and sometimes the other is better; we can choose which one to use.

Unfortunately, sequences are not always effective in general topological spaces. We give some examples of sequences “behaving badly”, and explain how to replace “bad” sequences by “good” nets.

1. Topologies, bases and subbases

Suppose that M is a set and $\mathcal{T}$ is a topology on M, that is, a collection of subsets of M that contains $\emptyset$ and M and is closed under finite intersections and arbitrary unions. Recall from Analysis that a *base* for $\mathcal{T}$ is a subcollection $\mathcal{B}$ of $\mathcal{T}$ such that every element of $\mathcal{T}$ is a union of elements of $\mathcal{B}$. For example, in a metric space, the open balls, together with the empty set and the whole set, form a base, since every open set is a union of open balls. A *subbase* for $\mathcal{T}$ is a subcollection $\mathcal{B}$ of $\mathcal{T}$ such that the collection of all finite intersections of elements of $\mathcal{B}$ is a base for $\mathcal{T}$.

We may specify a topology by specifying either a subbase or a base. For example, if we specify that $\mathcal{B}$ is a base, then we are implicitly specifying that $\mathcal{T}$ is the collection of all unions of elements of $\mathcal{B}$.

A set may have many topologies. A topology $\mathcal{T}_1$ is *finer* than a topology $\mathcal{T}_2$, or $\mathcal{T}_2$ is *coarser* than $\mathcal{T}_1$, if $\mathcal{T}_2 \subseteq \mathcal{T}_1$.

Let M be a set. Then $\{\emptyset, M\}$ is the coarsest possible topology on M; the set containing $\emptyset$ and all subsets of M whose complements are finite is another coarse topology. At the other extreme, the power set of M is the finest possible topology. Note that if $\mathcal{T}_2 \subseteq \mathcal{T}_1$, then a subset N of M that is compact for $\mathcal{T}_1$ is automatically compact for $\mathcal{T}_2$.

2. The Baire category theorem

The Baire category theorem is an important result about metric spaces. It is applied to normed vector spaces to prove several of the main theorems.

Theorem A.1 (Baire). *Suppose that $(U_n)_{n\in\mathbb{N}}$ is a countable collection of open dense subsets of a topological space M that is a complete metric space or a locally compact Hausdorff space. Then $\bigcap_{n\in\mathbb{N}} U_n$ is also dense.*

An alternative version of the theorem states that if $(F_n)_{n\in\mathbb{N}}$ is a countable collection of closed sets in M, and $\bigcap_{n\in\mathbb{N}} F_n = M$, then at least one of the F_n must have nonempty interior, that is, much contain a (nonempty) closed set.

3. Problems with sequences

Example A.2. It can be inconvenient to use sequences in the study of functions of several variables, for instance, when we would like to consider expressions such as

$$f_{(M,N)}(x,y) = \sum_{|m|\le M, |n|\le N} c_{(m,n)} e^{2\pi i(mx+ny)}.$$

It would be nice to be able to write expressions such as

$$f(x,y) = \lim_{(M,N)} f_{(M,N)}(x,y),$$

but we need to clarify what we mean by this limit. We could number the points in $\mathbb{N}\times\mathbb{N}$ using one copy of $\mathbb{N}$, but this would obscure the natural role of the indices M and N.

Example A.3. Let I be the standard interval $[0,1]$. Consider the space $S = \{0,1\}^I$ of all functions from I to the discrete two-point set $\{0,1\}$. We give S the topology in which the open sets are the empty set, together with the arbitrary unions of the sets

$$\{f \in X : f(t_1) = v_1, f(t_2) = v_2, \dots, f(t_n) = v_n\},$$

where $\{t_1, t_2, \dots, t_n\}$ is a finite subset of I and each of the values v_1, v_2, …, v_n is either 0 or 1.

Suppose that $g, g_n \in S$ and $g_n \to g$ in S. If $t_0 \in I$, then $U = \{f \in X : f(t_0) = g(t_0)\}$ is an open set containing g, and since $g_n \to g$, then $g_n \in U$ for sufficiently large n, that is, $g_n(t_0) = g(t_0)$. Since t_0 is arbitrary, it follows that if $g_n \to g$, then $g_n \to g$ pointwise.

Tychonov's theorem shows that S is compact.

However, consider the function h_n in S given by taking $h_n(t)$ to be the nth entry b_n in the binary expansion

$$t = \frac{b_1}{2} + \frac{b_2}{4} + \frac{b_3}{8} + \dots$$

of t; if t admits two different binary expansions, then we take the expansion in which all b_n are 0 when n is sufficiently large. Suppose that the sequence (h_n) has a convergent subsequence (h_{n_k}). Then $h_{n_k}(t)$ converges for all $t \in I$. But there is a real number t whose nth binary coefficient b_n is given by

$$b_n = \begin{cases} 1 & \text{if } n \in \{n_2, n_4, n_6, n_8, \dots\} \\ 0 & \text{otherwise.} \end{cases}$$

and the sequence $(h_{n_k}(t))$ does not converge.

Thus compactness cannot be described in terms of convergent subsequences.

One way to get around these problems is to use *nets*.

4. Definition of a net

Definition A.4. A *directed set* is a set $\mathcal{A}$ with a reflexive and transitive order relation $\preceq$ with the additional property that for all finite subsets F of $\mathcal{A}$, there exists $\mu \in \mathcal{A}$ such that $\alpha \preceq \mu$ for all $\alpha \in F$. The element μ is called an upper bound for F.

We write either $\alpha \preceq \beta$ or $\beta \succeq \alpha$.

Example A.5. The set $\mathbb{N}$, with the usual order, is a directed set. The upper bound of a finite set may be taken to be the maximum element.

Example A.6. The set $\mathbb{Z}$, with the usual order, is a directed set. The upper bound of a finite set may be taken to be the maximum element.

Example A.7. The set $\mathbb{Z}$, with the opposite order, that is, $m \preceq n$ if and only if $m \geq n$, is a directed set. The upper bound of a finite set may be taken to be the minimum element.

Example A.8. The set $\mathbb{N}^2$, with the product ordering, namely, $(M, N) \preceq (M', N')$ provided that $M \leq M'$ and $N \leq N'$, is a directed set. The upper bound of a set $\{(M_1, N_1), (M_2, N_2), \dots, (M_n, N_n)\}$ may be taken to be $(\max\{M_j\}, \max\{N_j\})$.

Example A.9. Let S be any set. Then the set $\mathcal{F}$ of all finite subsets of S ordered by set inclusion is a directed set. The upper bound of the finite subsets $F_1, F_2, \dots, F_n$ may be taken to be their union $\bigcup_{j=1}^n F_j$.

We can omit the condition that the sets are finite in the previous example; the power set of S, ordered by set inclusion, is also a directed set.

Example A.10. Let p be a point in a topological space M. Then the set $\mathcal{N}$ of all open subsets of S that contain p is a directed set, with the ordering given by reverse set inclusion, that is, $U \preceq V$ when $U \supseteq V$. The upper bound of the subsets $U_1, U_2, \dots, U_n$ may be taken to be their intersection $\bigcap_{j=1}^n U_j$.

EXERCISE A.11. Is the power set of M, ordered by reverse set inclusion, a directed set?

DEFINITION A.12. A *net* in a set S is a function from a directed set $\mathcal{A}$ into S; we write x_α rather than $x(\alpha)$ and $(x_\alpha)_{\alpha\in\mathcal{A}}$ rather than $x : \mathcal{A} \to S$. We say that a net $(x_\alpha)_{\alpha\in\mathcal{A}}$ in a topological space M converges to x in M, and we write $x_\alpha \to x$ or $\lim_\alpha x_\alpha = x$, if for each open set U containing x, there exists $\gamma \in \mathcal{A}$ (possibly depending on U) such that $x_\alpha \in U$ whenever $\alpha \succeq \gamma$. We often write (x_α) rather than $(x_\alpha)_{\alpha\in\mathcal{A}}$.

LEMMA A.13. *A subset S of a topological space M is closed if and only if the limit of every net of elements of S that converges in M is an element of S.*

PROOF. Suppose that S is closed, and $y \notin S$. Then $M \setminus S$ is an open set U containing y. If (x_α) is a convergent net in S, then no x_α is in $M \setminus S$, so $x_\alpha \not\to y$.

Conversely suppose that S is not closed. If, for each y in $M \setminus S$, we could find an open set U_y such that $y \in U_y \subseteq M \setminus S$, then we could write

$$M \setminus S = \bigcup_{y\in M\setminus S} U_y,$$

and $M \setminus S$ would be open; this is not possible. Consequently, there exists $y_0 \in M \setminus S$ such that every open set U containing y_0 meets S. Write $\mathcal{N}$ for the collection of open sets U containing y_0, ordered by reverse set inclusion. For each U in $\mathcal{N}$, choose $x_U \in U \cap S$. Then $(x_U)_{U\in\mathcal{N}}$ is a net in S, and further $x_U \to y_0 \notin S$. □

EXAMPLE A.14. A net indexed by $\mathbb{N}$, with the usual order, is just a sequence. The net converges precisely when the sequence converges.

EXAMPLE A.15. A net indexed by $\mathbb{N}^2$, with the product ordering, is a "doubly indexed sequence", and, for instance, $\lim_{(M,N)} f_{(M,N)} = f$ provided that we may make $f_{(M,N)}$ close to f by taking both M and N large. If $\lim_{(M,N)} f_{(M,N)} = f$, then it follows that $\lim_{k\to\infty} f_{(m(k),n(k))} = f$ for any increasing functions m and n from $\mathbb{N}$ to $\mathbb{N}$.

EXAMPLE A.16. Let $\mathcal{F}$ be the collection of all finite subsets of $\mathbb{N}$, ordered by inclusion. For a real or complex sequence (a_n) and $F \in \mathcal{F}$, define $s_F = \sum_{n\in F} a_n$. Then $\lim s_F = s$ precisely when the series $\sum a_n$ converges unconditionally to s.

The topology on a space M determines the convergent nets in M, by Definition A.12. Conversely, if we know the convergent nets in M, then we can determine the topology, for instance, by using Lemma A.13 to determine the closed sets.

5. Subnets

DEFINITION A.17. Suppose that $\mathcal{A}$ and $\mathcal{B}$ are directed sets and that $(\alpha_\beta)_{\beta\in\mathcal{B}}$ is a net in $\mathcal{A}$ such that $\alpha_{\beta'} \succeq \alpha_\beta$ whenever $\beta' \succeq \beta$ and, for all $\gamma \in \mathcal{A}$, there exists $\beta \in \mathcal{B}$ such that $\alpha_\beta \succeq \gamma$. Then $(x_{\alpha_\beta})_{\beta\in\mathcal{B}}$ is a *subnet* of the net $(x_\alpha)_{\alpha\in\mathcal{A}}$.

A subsequence is a subnet, but the converse is not true: $(1, 2, 2, 3, 3, 3, \dots)$ is a subnet of $(1, 2, 3, \dots)$ but is not a subsequence.

DEFINITION A.18. We say that x is a *limit point* of the *set* $\{x_\alpha : \alpha \in \mathcal{A}\}$ if for all open sets U containing x, there is $\gamma \in \mathcal{A}$, which may depend on U, such that $x_\gamma \in U$.

We say that x is a *limit point* of the *net* $(x_\alpha)_{\alpha\in\mathcal{A}}$ if for all open sets U containing x, and all $\gamma \in \mathcal{A}$, there is $\alpha \in \mathcal{A}$, which may depend on γ and U, such that $\alpha \succeq \gamma$ and $x_\alpha \in U$.

EXAMPLE A.19. Suppose that $a_n = \tan^{-1}(n)$ for all $n \in \mathbb{Z}$. Then both ± 1 are limit points of the set $\{a_N : n \in \mathbb{Z}\}$, but only 1 is a limit point of the net $(a_n)_{N\in\mathbb{Z}}$ when $\mathbb{Z}$ is given its standard order.

We prove one sample result on subnets.

LEMMA A.20. *The point x is a limit point of the net $(x_\alpha)_{\alpha\in\mathcal{A}}$ if and only if there is a subnet of $(x_\alpha)_{\alpha\in\mathcal{A}}$ that converges to x.*

PROOF. Suppose that x is a limit point of a net $(x_\alpha)_{\alpha\in\mathcal{A}}$. Given $\gamma \in \mathcal{A}$ and an open set U containing x, take $\alpha \in \mathcal{A}$ such that $\alpha \succeq \gamma$ and $x_\alpha \in U$, and write α as $\alpha_{(\gamma,U)}$. Write $\mathcal{N}$ for the collection of open sets containing x ordered by reverse set inclusion; then $\mathcal{A}\times\mathcal{N}$, with the product ordering, is a also a directed set, and $(x_{\alpha_{(\gamma,U)}})_{(\gamma,U)\in\mathcal{A}\times\mathcal{N}}$ is a subnet of the net $(x_\alpha)_{\alpha\in\mathcal{A}}$ that converges to x.

Conversely, suppose that there is a subnet $(x_{\alpha_\beta})_{\beta\in\mathcal{B}}$ of $(x_\alpha)_{\alpha\in\mathcal{A}}$ that converges to x. Then for each $\gamma \in \mathcal{A}$ and open set U containing x, there exists $\beta_1 \in \mathcal{B}$ such that $x_{\alpha_\beta} \in U$ whenever $\beta \succeq \beta_1$, and there exists $\beta_2 \in \mathcal{A}$ such that $\alpha_\beta \succeq \gamma$ whenever $\beta \succeq \beta_2$. Take an upper bound β for $\{\beta_1, \beta_2\}$. Then $\alpha_\beta \succeq \gamma$ and $x_{\alpha_\beta} \in U$, and so x is a limit point of the net $(x_\alpha)_{\alpha\in\mathcal{A}}$. □

THEOREM A.21. *A subset S of a topological space M is compact if and only if every net in S has a subnet that converges to a point of S.*

PROOF. We omit this proof. It uses the Axiom of Choice. □

EXAMPLE A.22. Consider again the space $\{0,1\}^I$ of $\{0,1\}$-valued functions on the interval I, and the sequence h_n defined in Example A.3. There there is a subnet of this sequence that converges pointwise on I. It does not seem to be possible to describe this subnet simply and explicitly.

APPENDIX B

Notes on measure and integration

We use without proof some ideas from the theory of measure and integration.

1. Measure theory

A *σ-algebra* $\mathcal{M}$ of subsets of a set S is a collection of subsets of S that contains $\emptyset$ and S, and is closed under taking complements and countable unions and intersections. The elements of $\mathcal{M}$ are called *measurable sets*.

A (positive) measure μ is a mapping from $\mathcal{M}$ to $[0, +\infty]$ that satisfies the conditions that $\mu(\emptyset) = 0$, and

$$\mu\Big(\bigcup_{n\in\mathbb{N}} E_n\Big) = \sum_{n\in\mathbb{N}} \mu(E_n)$$

whenever $E_n \in \mathcal{M}$ for all $n \in \mathbb{N}$ and the sets E_n are pairwise disjoint. The total mass of a positive measure on S is $\mu(S)$.

If S is a topological space, then the Borel σ-algebra is the smallest σ-algebra that contains all open sets. It is a theorem that "the smallest σ-algebra makes sense.

In many examples, such as $\mathbb{R}^n$, there are simple objects, such as rectangles with sides parallel to the axes, on which we have a natural notion of measure, and then we can extend this measure to the Borel σ-algebra.

2. Measurable functions

A function $f : S \to \mathbb{F}$ is measurable if $f^{-1}(B) \in \mathcal{M}$ for all balls B in $\mathbb{F}$. If f is complex-valued, then it is measurable if and only if its real and imaginary parts are measurable, and the theory is usually developed with a focus on real-valued functions.

There are various theorems about measurable functions. For example, these form an algebra under pointwise operations, the supremum and infimum of a sequence of measurable functions is measurable, and the pointwise limit of a pointwise convergent sequence of measurable functions is measurable. These properties make measurable functions much more versatile that Riemann integrable functions.

3. Integrals

Given a measure space, that is, a triple $(S, \mathcal{M}, \mu)$ as above, we may define the integral of a nonnegative measurable function $f : S \to [0, +\infty]$ by the formula

$$\int_S f\,d\mu = \sup\{\sum_{n\in\mathbb{N}} a_n\mu(E_n) : \sum_{n\in\mathbb{N}} a_n 1_{E_n} \leq f\},$$

where $a_n \in [0, +\infty]$ and $E_n \in \mathcal{M}$ for all $n \in \mathbb{N}$.

The integral of a general measurable function is then defined by breaking f up into four parts: $f = a - b + ic - id$, where

$$a(s) := \max\{\operatorname{Re} f(s), 0\}, \qquad b(s) := \max\{-\operatorname{Re} f(s), 0\},$$
$$c(s) := \max\{\operatorname{Im} f(s), 0\}, \qquad d(s) := \max\{-\operatorname{Im} f(s), 0\},$$

and then setting

$$\int_S f\,d\mu = \int_S a\,d\mu - \int_S b\,d\mu + i\int_S c\,d\mu - i\int_S d\,d\mu,$$

subject to the condition that the four summands on the right hand side are finite. If two measurable functions differ on a set of measure 0, then they have the same integral.

We define the Lebesgue spaces $L^p(S, \mathcal{M}, \mu)$, usually abbreviated to $L^p(S, \mu)$ or just $L^p(S)$, to be the spaces of measurable spaces for which

$$\left(\int_S |f|^p\,d\mu\right)^{1/p} < \infty,$$

and we identify functions that are equal except on a set of measure 0. The expression on the left hand side of the inequality above is then the L^p-norm of f (or more precisely, of its equivalence class). The definition of $L^\infty(S)$ is similar.

When the measure space arises from a topological space, then it can be shown that continuous functions with compact support are dense in the spaces $L^p(S)$, except perhaps when $p = \infty$.

4. Complex measures

In general, there are many measures defined on the same σ-algebra of subsets of a space S. A complex measure is a linear combination (with coefficients ± 1 and $\pm i$ of four positive measures on S, each of which has finite total mass. These assign a measure to a general set that is a complex number.

When the measure space arises from a topological space, then it can be shown that

$$\int_S d\mu$$

may be fined for all complex measures and all bounded continuous functions f. The Riesz representation theorem states that the linear functionals on $C_0(S)$ (the space of continuous functions on S that vanish at infinity) coincide with the maps $f \mapsto \int_S d\mu$, where μ is an arbitrary complex measure.