C4.1 Further Functional Analysis

Stuart White

University of Oxford
Michaelmas Term 2019*

I'm very grateful to the previous lecturer of this course, David Seifert, who developed the previous version of these notes, and allowed me to use them. Please send comments, corrections, clarifications etc to me at `stuart.white@maths.ox.ac.uk`.

1 Introduction

Functional analysis is the study of normed (or, more generally, of topological) vector spaces and the continuous linear maps between them. This course builds on what is covered in introductory courses on functional analysis, such as B4 Functional Analysis I and II taken by third-year students at Oxford, and in particular we will extend the theory of normed vector spaces and bounded linear operators developed there. Some functional analysts are primarily interested in what might be called geometric properties of Banach spaces, others in properties of operators acting on these spaces. In fact, the two strands are connected and we will deal with elements of both. Functional analysis makes connections right across mathematics; there are many applications, for instance to differential equations, probability, mathematical physics, numerical analysis, as well as strong connections to topics in pure mathematics: the Feldholm index we begin in Section 12 provides the analytic side of the celebrated Atiyah-Singer index theorem, various approximation properties in geometric group theory are studied using tools from functional analysis, and there are strong links between Banach spaces and subjects like metric geometry and descriptive set theory.

Our emphasis here will nevertheless be mainly on the abstract theory, both to avoid excessive overlap with other courses and to keep the prerequisites to a minimum. We will illustrate the abstract theory by considering various specific examples, both in the lectures and especially in the problem sheets.

There are many good books on functional analysis. Among those particularly relevant to this course are the following:

[1] B. Bollobas, *Linear Analysis: An Introductory Course*, CUP, 1999.

[2] H. Brezis, *Functional Analysis, Sobolev Spaces and PDEs*, Springer, 2011.

[3] N.L. Carothers, *A Short Course on Banach Space Theory*, CUP, 2004.

---

*Version of November 11, 2019

[4] J. Conway, *A Course in Functional Analysis*, Springer, 2007.

[5] M. Fabian et al., *Funct. Analysis and Infinite-Dim. Geometry*, Springer, 2001.

[6] R.E. Megginson, *An Introduction to Banach Space Theory*, Springer, 1998.

[7] W. Rudin, *Functional Analysis*, McGraw-Hill, 1991.

[8] A.E. Taylor and D.C. Lay, *Introduction to Functional Analysis*, Wiley, 1980.

Perhaps the most useful of these is [5]. If your college library doesn't already own a copy, you might consider asking it to buy one.

Prerequisites

- Basics of metric spaces, particularly aspects relating to completeness.

- Basics of topological spaces, closures and interiors, and in particular compactness.[1]

- Fundamentals of linear algebra, bases, quotient spaces.

- Normed spaces and Banach spaces. Definitions and fundamental examples. Familiarity with Hilbert spaces, and their fundamental properties.[2]

- Operators between normed spaces, Continuity and boundedness, Completeness of $\mathcal{B}(X,Y)$ when Y is complete.

- The Baire category theorem.[3]

- Open mapping theorem, closed graph theorem, and inverse mapping theorem.[4]

- Measure theory. We use this in Section 8 and a couple of associated examples (and an example on the last example sheet). You'll want enough background for this section, but it's otherwise not essential.

There is a preliminary reading document on the course homepage, prepared by Richard Haydon[5] which contains more details on this background material, including many complete proofs.

---

[1] We will use Tychonoff's theorem that the product of compact spaces is compact as a black box in the course. This is equivalent to the axiom of choice, and not in my list of basic facts about topological spaces.

[2] Most of this course will focus on the structure of Banach spaces and the operators between them, but it'll be useful to contrast the behaviour with known results for Hilbert spaces, such as the projection theorem: there is an orthogonal projection onto a closed subspace of Hilbert space.

[3] If you've not seen this before, this shouldn't be a problem. We will state it, and use it once in Section 2 in order to show that Hamel bases on Banach spaces are necessarily uncountable. Baire's category theorem is also often used to deduce the open mapping theorem and closed graph theorem. But this isn't necessary for the course.

[4] Ideally you've seen these before, but for our course the statements will suffice. I invite you to deduce the open mapping theorem from Baire's category theorem in Sheet 0, and we will see the equivalence of these three classical theorems in sheet 2 (and they're also equivalent to the uniform boundedness principle). One can also prove these theorems directly; indeed as noted in [3, Theorem 5.11 and the discussion which follows] — a good source for a direct proof of uniform boundedness — this is how Banach and Steinhaus most likely first proved the uniform boundedness principle.

[5] Another former lecturer for this course

2 Normed vector spaces

Let X be a vector space over the field $\mathbb{F}$. Throughout this course $\mathbb{F}$ will be either $\mathbb{R}$ or $\mathbb{C}$. We will specify the field when it matters; when we don't it is to be understood that the vector space is either real or complex. If several normed vector spaces are introduced at the same time, they will always be over the same field. Unless the possibility that $X = \{0\}$ is explicitly mentioned we assume that $X \neq \{0\}$.

Recall that a *norm* on X is a map $\|\cdot\| \colon X \to [0,\infty)$ such that

- $\|x\| = 0$ if and only if $x = 0$;

- $\|\lambda x\| = |\lambda| \|x\|$ for all $\lambda \in \mathbb{F}$, $x \in X$;

- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

In particular, any norm $\|\cdot\|$ induces a metric $d \colon X \times X \to [0,\infty)$ given by $d(x,y) = \|x - y\|$, $x, y \in X$. We call a vector space equipped with a norm a *normed vector space*. A normed vector space X is said to be a *Banach space* if the metric space (X, d) is complete. Recall that a complete subspace of a normed vector space is closed and that a closed subspace of a Banach space is complete.

As always in mathematics, we should have a range of examples to hand. For now one should have the first three of the following examples to hand.

- n-dimensional Euclidian space ℓ_n^2, and the various equivalent norms we could put on these spaces to obtain ℓ_n^p, for $1 \leq p \leq \infty$

- The standard spaces of infinite sequences ℓ^2, ℓ^1 and ℓ^∞, c_0, and the more general spaces ℓ^p for $1 \leq p \leq \infty$.

- $C(K)$ for a compact Hausdorff K (think $K = [0,1]$);

- $L^p(\Omega, \Sigma, \mu)$ for $1 \leq p \leq \infty$, and a measure space (Ω, Σ, μ).[6]

Our notation for open and closed balls will be as follows. For $x_0 \in X$ and $r > 0$, we let
$$B_X(x_0, r) = \big\{ x \in X : \|x - x_0\| \leq r \big\}$$
denote the *closed ball*, and we let
$$B_X^\circ(x_0, r) = \big\{ x \in X : \|x - x_0\| < r \big\}$$
denote the *open ball*. For brevity we let $B_X(r) = B_X(0, r)$ and $B_X^\circ(r) = B_X^\circ(0, r)$. We also write $B_X = B_X(1)$ and $B_X^\circ = B_X^\circ(1)$ for the unit balls, and we denote the *unit sphere* $\{ x \in X : \|x\| = 1 \}$ of X by S_X.

Recall that two norms $\|\cdot\|$ and $\|\|\cdot\|\|$ on X are said to be *equivalent* if there exist constants $c, C > 0$ such that

$$c\|x\| \leq \|\|x\|\| \leq C\|x\|, \quad x \in X.$$

We already know that if X is finite-dimensional then all norms are equivalent, and in particular X is complete with respect to the metric induced by any norm. In

---

[6]We will look at these spaces in Section 8. It's not our purpose to give a course on measure theory, so we'll keep the amount of measure theory prerequistes to a minimum.

fact, if $\dim X = n$ for $n \in \mathbb{N}$ then X is isomorphic to $\mathbb{F}^n$ with any particular norm.[7] Here and in what follows *isomorphic* means linearly homeomorphic. If two normed space X and Y are isomorphic we occasionally write $X \simeq Y$, and if they are isometrically isomorphic we write $X \cong Y$. For infinite-dimensional vector spaces it is no longer true that all norms are equivalent (and even two complete norms may be non-equivalent; see Problem Sheet 1),

On the subject of dimension, recall that $\dim X$ is the cardinality of any linearly independent spanning set of X. Such a set is called a *Hamel basis* for X.[8] Thus a subset B of X is a Hamel basis if and only if every $x \in X$ can be written uniquely in the form

$$x = \sum_{k=1}^{n} \lambda_k x_k$$

for some $n \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and $x_1, \ldots, x_n \in B$. In the general setting, the existence of Hamel bases relies on the axiom of choice, in the form of Zorn's lemma (often the way that the axiom of choice is used in practise).

A partial ordered set (poset) is a set $\mathcal{P}$ together with an order $\leq$ which is reflexive, transitive and antisymmetric.[9] The example of most relevance to us, will be a family of subsets of a given set, ordered by inclusion. An element x in a poset $\mathcal{P}$ is *maximal* if $x \leq y$ implies that $y = x$.[10] A *chain* $\mathcal{C}$ in a poset $\mathcal{P}$ is a subset of $\mathcal{P}$ which is totally ordered, i.e. any two elements are comparable.

Theorem 2.1 (Zorn's Lemma). *Let X be a non-empty poset such that every chain has an upper bound. Then X has a maximal element.*

Remark 2.2. Zorn's Lemma is equivalent to the Axiom of Choice; see for instance B1.2 Set Theory and C1.4 Axiomatic Set Theory for details. Note that there is no claim about uniqueness of the maximal element. Not all mathematicians accept the Axiom of Choice, so it is good practice to be aware of which results depend on it and which don't.[11]

Proposition 2.3. *Every non-zero vector space X admits a Hamel basis.*

PROOF. Let $\mathcal{P}$ be the collection of all linearly independent subsets of X ordered by inclusion (this is certainly non-empty) and suppose that $\mathcal{C} \subseteq \mathcal{P}$ is a chain. We claim that $A = \bigcup_{C \in \mathcal{C}} C$ is linearly independent, which requires us to check that any finite collection of vectors in A forms a linearly independent set. So suppose that $\{x_1, \ldots, x_n\} \subseteq A$ for some $n \in \mathbb{N}$. Then there exist $C_k \in \mathcal{C}$ such that $x_k \in C_k$,

---

[7]This is essentially proved by choosing a basis, and using this to show that X is isomorphic to ℓ_n^1; see [5, Proposition 1.22], for example.

[8]In the context of linear algebra, a Hamel basis would simply be called a basis. But in the study of functional analysis, such bases are not so useful, and we prefer to reserve the term basis for things like orthogonal bases in a Hilbert space, where we can write every element as a norm convergent infinite linear combination of the basis. We will return briefly to the topic of Schauder bases, the appropriate notion of a basis for a Banach space in Section 13.

[9]if $x, y \in \mathcal{P}$ have $x \leq y$ and $y \leq x$, then $x = y$.

[10]Note the terminology maximal, rather than maximum. The latter would suggest uniqueness, and in general maximal elements need not be unique.

[11]With that said, the axiom of choice is not really controversial within functional analysis and normally assumed without comment. Fundamental results in the field, such as the Hahn-Banach theorem and the Banach-Alaoglu Theorem

$1 \leq k \leq n$. Since $\mathcal{C}$ is a chain it is easy to see that there exists $m \in \{1, \ldots, n\}$ such that $\bigcup_{k=1}^{n} C_k = C_m$. Since $C_m \in \mathcal{P}$ the vectors $x_1, \ldots, x_n$ are linearly independent.

By Zorn's lemma $\mathcal{P}$ has a maximal element, say B. Let $Y = \operatorname{span} B$ and suppose that $x \in X$. If $x \notin Y$ then the set $B' = B \cup \{x\}$ is linearly independent. Moreover, $B \subseteq B'$ and hence by maximality $B = B'$, which is a contradiction. Thus $Y = X$ and hence B is the required Hamel basis. $\qquad \square$

Remark 2.4. The proof shows slightly more, namely that for every linearly independent subset A of X there exists a Hamel basis B for X such that $A \subseteq B$. To see this replace $\mathcal{P}$ by those linearly independent subsets of X containing A.

The main use we will have for Hamel bases is producing unbounded linear maps; one of the reasons that they're not so useful elsewhere in functional analysis is that (in the case of complete spaces) they're always too large. Precisely, as a consequence of the Baire category theorem, any Hamel basis for an infinite dimensional Banach space is necessarily uncountable.

Theorem 2.5 (Baire's Category Theorem). *Let (X, d) be a complete metric space[12] and suppose that U_n, $n \in \mathbb{N}$, are dense open subsets of X. Then $\bigcap_{n \geq 1} U_n$ is also dense in X.*

Proposition 2.6. *Let X be an infinite-dimensional Banach space. Then any Hamel basis for X must be uncountable.*

PROOF. Suppose, for the sake of a contradiction, that there exists a countable Hamel basis $B = \{x_n : n \geq 1\}$ for X and let $F_n = \operatorname{span}\{x_k : 1 \leq k \leq n\}$, $n \geq 1$. Then each of the spaces F_n is finite-dimensional and hence complete, and in particular each F_n is closed in X. Let $U_n = X \setminus F_n$, $n \geq 1$. Then each U_n is open in X and moreover

$$\bigcap_{n \geq 1} U_n = X \setminus \bigcup_{n \geq 1} F_n = \emptyset.$$

By the Baire Category Theorem there exists $n \geq 1$ such that U_n is not dense in X, which is equivalent to saying that F_n has non-empty interior. Suppose that $x \in X$ and $\varepsilon > 0$ are such that $B_X^\circ(x, \varepsilon) \subseteq F_n$. Since F_n is a vector space and in particular closed under translations, it follows that $B_X^\circ(\varepsilon) \subseteq F_n$, and hence $X \subseteq F_n$. In particular, we have $\dim X \leq n$, which is a contradiction. $\qquad \square$

Remark 2.7. This result shows that even if X is a separable Banach space it cannot have a countable Hamel basis. We will see in Section 13 that a more appropriate notion of basis in the context of Banach spaces is that of a *Schauder basis*.

Example 2.8. Let X be the space of all polynomials $x \colon [0, 1] \to \mathbb{F}$ with coefficients in $\mathbb{F}$, endowed with the supremum norm $\|x\|_\infty = \sup_{0 \leq t \leq 1} |x(t)|$, $x \in X$. Then X is a subspace of the Banach space $C([0, 1])$ of all scalar-valued continuous functions

---

[12]The Baire category theorem also works for locally compact Hausdorff spaces; see [7, 2.2], and provides a notion of typical in the setting of *Baire spaces*, i.e. those spaces for which the Baire category theorem holds. That is we can view a property as holding *generically* if it holds on a set containing a dense countable intersection of open sets. The point is that countable intersections of generic properties remain generic — this way of thinking heads towards descriptive set theory, and gives us a framework for discussing, for example, typical properties of representations.

defined on $[0, 1]$. But X is not closed, for otherwise it would be complete and by Proposition 2.6 this cannot be the case, since the set $\{x_n : n \geq 0\}$, where $x_n(t) = t^n$ for $n \geq 0$ and $0 \leq t \leq 1$, is a countable Hamel basis for X. In fact, we know from the Weierstrass Approximation Theorem that X is dense in $C([0, 1])$.

Recall that any vector space X has an associated *algebraic dual space* X' consisting of all linear functionals $f \colon X \to \mathbb{F}$. If X is a normed vector space it is natural to restrict oneself to the class of *continuous*, or equivalently *bounded*, linear functionals f, which are those satisfying $|f(x)| \leq C\|x\|$ for some $C > 0$ and all $x \in X$. We write X^* for the *(topological) dual space* consisting of all bounded linear functionals $f \colon X \to \mathbb{F}$. Recall that a linear functional $f \in X'$ is bounded if and only if it is continuous at 0, and that this in turn is equivalent to the *kernel* $\mathrm{Ker}(f) = \{x \in X : f(x) = 0\}$ being a closed subspace of X.[13] If X is finite-dimensional then any linear functional on X is automatically bounded and hence $X' = X^*$. The situation is different when X is infinite-dimensional.

Proposition 2.9. *Suppose that X is an infinite-dimensional normed vector space. Then there exists an unbounded linear functional $f \colon X \to \mathbb{F}$.*

PROOF. Let $A = \{x_n : n \in \mathbb{N}\}$ be a linearly independent subset of X. By Proposition 2.3 and Remark 2.4 there exists a Hamel basis B for X such that $A \subseteq B$. Define the functional f by $f(x_n) = n\|x_n\|$, $n \geq 1$, and $f(x) = 0$ for all $x \in B \setminus A$ and extend f linearly to X. Then $f \in X'$ but f is unbounded. $\qquad\square$

Example 2.10. Let X be as in Example 2.8. Then the linear functional f on X given by $f(x) = x'(1)$, $x \in X$, is easily seen to be unbounded, for instance because $f(x_n) = n$ while $\|x_n\|_\infty = 1$ for all $n \geq 1$.[14]

The conclusion of Proposition 2.9 is more interesting in the case where X is complete, because on such a space any unbounded functional must have non-closed graph. This follows from another important consequence of the Baire Category Theorem.[15] Recall that, given two normed vector spaces X and Y we write $\mathcal{B}(X, Y)$ for the space of bounded linear operators $T \colon X \to Y$. Here a linear operator is said to be *bounded* if there exists $C > 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in X$. The infimum over all such $C > 0$ is the *norm* of T, denoted by $\|T\|$. Recall that

$$\|T\| = \sup_{x \in B_X} \|Tx\| = \sup_{x \in B_X^\circ} \|Tx\| = \sup_{x \in S_X} \|Tx\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|},$$

and that $\mathcal{B}(X, Y)$ is complete if and only if Y is complete. Note also that $X^* = \mathcal{B}(X, \mathbb{F})$, which in particular is always complete. For future reference, recall that

---

[13]These make good exercises for reviewing your prior functional analysis courses.

[14]You'll notice that this explicit example is on a non-complete space, and might be wanting an explicit example on a Banach space. The existence of discontinuous linear functionals on Banach spaces relies on the axiom of choice, and it is consistent with ZF set theory without AC that all linear functionals on Banach spaces are complete. Having said that, finding functional analysts that don't subscribe to the axiom of choice is also not straightforward.

[15]The closed graph theorem is more usually deduced from the open mapping theorem, and this from the closed graph theorem. But in fact, Banach's original proof of the open mapping theorem didn't use Baire's category theorem; we will see this in the successive approximations lemma in Section 4.

$\mathcal{B}(X) = \mathcal{B}(X, X)$. Given a linear operator $T\colon X \to Y$ between two normed vector spaces, we denote the *graph* of T by $G_T = \{(x, y) \in X \times Y : Tx = y\}$. For any $1 \le p \le \infty$ we may endow the cartesian product $X \times Y$ with the norm given by

$$\|(x, y)\|_p = \begin{cases} \left(\|x\|^p + \|y\|^p\right)^{1/p}, & 1 \le p < \infty, \\ \max\left\{\|x\|, \|y\|\right\}, & p = \infty, \end{cases}$$

for all $(x, y) \in X \times Y$. It is straightforward to see that all of these norms are equivalent, and in fact the topology they define is the product topology on $X \times Y$. Moreover, if X and Y are Banach spaces then so is $X \times Y$ with respect to any of the norms $\|\cdot\|_p$, $1 \le p \le \infty$.

Theorem 2.11 (Closed Graph Theorem). *Suppose that X and Y are Banach spaces and let $T\colon X \to Y$ be a linear operator. Then $T \in \mathcal{B}(X, Y)$ if and only if G_T is closed in $X \times Y$.*

Remark 2.12. If X is as in Example 2.8 and $Y = C([0, 1])$, then the operator $T\colon X \to Y$ defined by $Tx = x'$, $x \in X$, is unbounded but has closed graph. Hence the completeness assumption on X in the Closed Graph Theorem cannot be omitted. We will see on Problem Sheet 1 that completeness of Y cannot be dropped either.

Remark 2.13. As a complete aside — which you should feel free to ignore — closed graphs are also particularly useful when we deal with *densely defined* unbounded linear operators — such as operators of differentiation on spaces like $C(\mathbb{T})$ or $L^2(\mathbb{T})$ defined where this makes sense — as arise regularly in applications of functional analysis to partial differential equations. In the absence of boundedness, having a closed graph is the next best thing.[16]

3 Direct sums and complemented subspaces

If Y and Z are subspaces of a vector space X, then the sum $X_0 = Y + Z$ is a subspace of X and there is a surjective linear map $T\colon Y \times Z \to X_0$ given by $T(y, z) = y + z$ for all $(y, z) \in Y \times Z$. Recall that this linear map is injective if and only if $Y \cap Z = \{0\}$, in which case we write $X_0 = Y \oplus Z$ and call X_0 the *(algebraic) direct sum* of Y and Z. If X is a normed vector space then the map T is continuous[17] but not necessarily an isomorphism. If the inverse T^{-1} of T is also continuous we write (once again) $X_0 = Y \oplus Z$ and say that X_0 is the *topological direct sum* of Y and Z.

Given a vector space X and a subspace Y of X, we say that Y is *algebraically complemented* in X if there exists a further subspace Z of X such that $X = Y \oplus Z$ as an algebraic direct sum. In this case Z is said to be an *algebraic complement* of Y. It is easy to see that the space Y is algebraically complemented if and only if there exists a projection $P\colon X \to X$ such that $\operatorname{Ran} P = Y$. Recall that a *projection* is a linear map satisfying $P^2 = P$ and that $\operatorname{Ran} P = \{Px : x \in X\}$ is the *range* of P. It follows from Zorn's Lemma that every subspace is algebraically complemented in just the same way as in finite dimensional linear algebra; see Problem Sheet 1.

---

[16]I'm talking about densely defined operators here, not everywhere defined operators, so this isn't contradicting the closed graph theorem.

[17]We give $Y \times Z$ the product topology, and equip it with one of the equivalent norms $\|\cdot\|_p$ as at the end of the previous section.

Similarly, if X is a normed vector space we say that a subspace Y of is *(topologically) complemented* in X if there exists a further subspace Z of X such that $X = Y \oplus Z$ as a topological direct sum. We call Z a *(topological) complement* of Y. In this course, since we're doing functional analysis and not just linear algebra, by a 'complemented subspace' we will always mean a topologically complemented subspace.

Proposition 3.1. *Let X be a normed vector space and suppose that Y and Z are subspaces of X such that $X = Y \oplus Z$ algebraically. Then X is the topological direct sum of Y and Z if and only if the map $P\colon X \to X$ given by $P(y+z) = y$ for $y \in Y$, $z \in Z$ is bounded. In particular, a subspace Y of X is complemented if and only if there exists a bounded projection on X whose range is Y.*

PROOF. Let us endow the product $Y \times Z$ with the ∞-norm. Then the map $T\colon Y \times Z \to X$ introduced above satisfies $\|T\| \le 2$ and its inverse is given by

$$T^{-1}x = \big(Px, (I-P)x\big), \quad x \in X.$$

Thus if P is bounded then so is T^{-1} and in fact $\|T^{-1}\| \le 1 + \|P\|$, while if T^{-1} is bounded then so is P and $\|P\| \le \|T^{-1}\|$. This proves the first part of the result, and the second part follows easily. $\square$

Note that any topologically complemented subspace is closed, as is any topological complement. 'Which closed subspaces of a Banach space are complemented?' is a fundamental question in Banach space theory. In the Hilbert space setting, by the Projection Theorem *any*, closed subspace of a Hilbert space is topologically complemented, even by a projection of norm 1.[18] As we will see right at the end of the course, for the classical sequence spaces $X = c_0$ or $X = \ell^p$ (with $1 \le p < \infty$, every infinite dimensional complemented subspace of X is isomorphic to X.

Finally we note that for Banach spaces, an algebraic direct sum of closed subspaces is automatically a topological direct sum. This is a consequence of the following important result, which we recall from an earlier course.

Theorem 3.2 (Inverse Mapping Theorem). *Let X and Y be Banach spaces and suppose that $T \in \mathcal{B}(X, Y)$ is a bijection. Then T is an isomorphism.*

Theorem 3.3. *Let X be a Banach space and suppose that Y and Z are closed subspaces of X such that $X = Y \oplus Z$ algebraically. Then $X = Y \oplus Z$ topologically.*
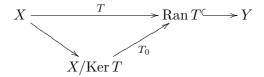
PROOF. If we endow the product $Y \times Z$ with any of the p-norms, then $Y \times Z$ is a Banach space and hence the map $T\colon Y \times Z \to X$ given by $T(y, z) = y + z$, $y \in Y$, $z \in Z$, is a bounded linear bijection between two Banach spaces. By the Inverse Mapping Theorem T is an isomorphism, so $X = Y \oplus Z$ topologically. $\square$

4 Quotient spaces and quotient operators

Given a vector space X and a subspace Y of X we may consider the cosets

$$x + Y = \{x + y : y \in Y\}, \quad x \in X.$$

These form a vector space with zero element $Y = 0 + Y$, addition given by $(x_1 + Y) + (x_2 + Y) = (x_1 + x_2) + Y$, $x_1, x_2 \in X$, and scalar multiplication $\lambda(x + Y) = \lambda x + Y$, $\lambda \in \mathbb{F}$, $x \in X$. Note that two cosets $x_1 + Y$ and $x_2 + Y$ coincide if and only if $x_1 - x_2 \in Y$. We call this space the *quotient space* and denote it by X/Y. Recall that if Z is an algebraic complement of Y in the vector space X,[19] then X/Y is isomorphic to Z as a vector space. We also have a canonical factorisation of linear maps from the first isomorphism theorem. Given a linear map $T : X \to Y$, we get a well defined linear bijection $T_0 : X/\operatorname{Ker}(T) \to \operatorname{Ran} T$ given by $T_0(x + \operatorname{Ker}T) = T(x)$. Then T factorises as

This section aims to develop the analogous theory for normed vector spaces.

If X is a normed vector space we may define a map $\|\cdot\| \colon X/Y \to [0, \infty)$ by

$$\|x + Y\| = \operatorname{dist}(x, Y) = \inf\{\|x + y\| : y \in Y\}, \quad x \in X.$$

But note that $\|x + Y\| = 0$ need not imply that $x + Y = Y$, which is to say $x \in Y$. Instead it only implies that x lies in the closure of Y.

Proposition 4.1. *Let X be a normed vector space and suppose that Y is a closed subspace of X. Then the map $\|\cdot\| \colon X/Y \to [0, \infty)$ given by $\|x + Y\| = \operatorname{dist}(x, Y)$, $x \in X$, defines a norm on X/Y. Moreover, if X is complete then so is X/Y.*

PROOF. It is clear from the above remarks that $\|x + Y\| = 0$ if and only if $x \in Y$. Moreover, it is easy to see that for $\lambda \in \mathbb{F}$ and $x \in X$ we have $\|\lambda x + Y\| = |\lambda| \|x + Y\|$. If $x_1, x_2 \in X$, then for any $y_1, y_2 \in Y$

$$\|x_1 + x_2 + Y\| \le \|x_1 + y_1 + x_2 + y_2\| \le \|x_1 + y_1\| + \|x_2 + y_2\|,$$

so taking the infimum over $y_1, y_2 \in Y$ shows that $\|x_1 + x_2 + Y\| \le \|x_1 + Y\| + \|x_2 + Y\|$. Thus $\|\cdot\|$ defines a norm on X/Y. Now suppose that X is complete, and recall that a normed vector space is complete if and only if every absolutely convergent series is convergent. Suppose that $x_n \in X$, $n \ge 1$, are such that $\sum_{n=1}^{\infty} \|x_n + Y\| < \infty$. For each $n \ge 1$ let $y_n \in Y$ be such that $\|x_n + y_n\| \le \|x_n + Y\| + 2^{-n}$. Then

---

[18]This is not true in a Banach space setting, and we will see some examples in example sheet 2. Moreover a theorem of Lindenstrauss and Tzafriri, beyond the scope of this course, shows that in fact Hilbert spaces are the only spaces all of whose closed subspaces are complemented. Precisely, if X is a Banach space, such that every closed subspace is complemented, then X is isomorphic to a Hilbert space.

[19]such exists by Zorn's lemma, as shown on exercise sheet 1.

$\sum_{n=1}^{\infty} \|x_n + y_n\| < \infty$, so by completeness of X the series $\sum_{n=1}^{\infty}(x_n + y_n)$ converges in norm to some $z \in X$. But now

$$\left\| \sum_{n=1}^{N}(x_n + Y) - (z + Y) \right\| \le \left\| \sum_{n=1}^{N}(x_n + y_n) - z \right\| \to 0, \quad N \to \infty.$$

Hence the norm $\|\cdot\|$ turns X/Y into a Banach space. $\qquad\square$

When X is a normed vector space and Y is a closed subspace we take the quotient space X/Y to be endowed with the norm $\|\cdot\|$ from Proposition 4.1, and we call this the *quotient norm*.

Remark 4.2. (a) If X is a normed vector space and Y a closed subspace, then we may consider the map $\pi\colon X \to X/Y$ given by $\pi(x) = x + Y$, $x \in X$. It is clear that π is a bounded linear operator and that $\pi(B_X^\circ) = B_{X/Y}^\circ$. In fact, π is an open map and the quotient topology it induces on X/Y is precisely the quotient norm topology.

(b) We can generalise the construction of quotient spaces. Indeed, given a normed vector space X and a subspace Y the map $p\colon X \to [0, \infty)$ defined by $p(x) = \mathrm{dist}(x, Y)$ is a *seminorm*, that is to say it satisfies the axioms for a norm except that $p(x) = 0$ need not imply $x = 0$. Given an arbitrary seminorm p on a vector space X we may consider the subspace $Y = \{x \in X : p(x) = 0\}$ of X and endow X/Y with the norm $\|x + Y\| = p(x)$, $x \in X$. Conversely, given a subspace Y of X and a norm on X/Y we may define a seminorm on X by $p(x) = \|x + Y\|$, $x \in X$.

Example 4.3. (a) If X is a Hilbert space and Y is a closed subspace, then by the projection theorem $X = Y \oplus Y^\perp$. Let P denote the orthogonal projection onto Y. Then, given $x \in X$ we have $x + Y = x - Px + Y$ and $\|x + Y\| = \|x - Px\|$. Thus the map $T\colon X/Y \to Y^\perp$ given by $T(x + Y) = x - Px$, $x \in X$, is a well-defined isometric isomorphism, and hence $X/Y \cong Y^\perp$.

(b) As an example of the construction mentioned in part (b) of Remark 4.2 let X be the space of all integrable functions over some measure space (Ω, Σ, μ). If the seminorm p is given by $p(x) = \int_\Omega |x| \, \mathrm{d}\mu$ and if $Y = \{x \in X : p(x) = 0\}$, then X/Y is precisely $L^1(\Omega, \Sigma, \mu)$.

Given vector spaces X and Y and a linear operator $T\colon X \to Y$, the First Isomorphism Theorem tells us that T induces a well-defined linear bijection $T_0\colon X/\operatorname{Ker} T \to \operatorname{Ran} T$ by $T_0(x + \operatorname{Ker} T) = Tx$, $x \in X$. We are interested in the topological version.

Lemma 4.4. *Let X and Y be normed vector spaces and suppose that $T \in \mathcal{B}(X, Y)$. Then the operator $T_0\colon X/\operatorname{Ker} T \to \operatorname{Ran} T$ given by $T_0(x + \operatorname{Ker} T) = Tx$, $x \in X$, is bounded and in fact $\|T_0\| = \|T\|$.*

PROOF. Given $x \in X$ and $z \in \operatorname{Ker} T$ we have

$$\|T_0(x + \operatorname{Ker} T)\| = \|T(x + z)\| \le \|T\| \|x + z\|,$$

and taking the infimum over $z \in \operatorname{Ker} T$ shows that T_0 is bounded with $\|T_0\| \le \|T\|$. On the other hand,

$$\|Tx\| = \|T_0(x + \operatorname{Ker} T)\| \le \|T_0\| \|x + \operatorname{Ker} T\| \le \|T_0\| \|x\|, \quad x \in X,$$
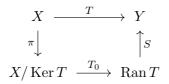
and hence $\|T\| \le \|T_0\|$, as required. $\qquad\square$

Given normed vector spaces X and Y and an operator $T \in \mathcal{B}(X, Y)$ we say that T is a *quotient operator* (an *isometric quotient operator*) if T is surjective and the map T_0 considered in Lemma 4.4 is an (isometric) isomorphism. Our goal is to characterise these.

Example 4.5. (a) If X is a normed vector space and Y is a closed subspace, then the map $\pi \colon X \to X/Y$ given by $\pi(x) = x + Y$ is an isometric quotient operator. Indeed, π_0 is the identity operator on X/Y. We call π the *canonical quotient operator*.

(b) Suppose that $\| \cdot \|$ and $\vertiii{\cdot}$ are two norms on a vector space X such that $\vertiii{x} \leq C\|x\|$ for some $C > 0$ and all $x \in X$, and let T be the identity operator from $(X, \| \cdot \|)$ to $(X, \vertiii{\cdot})$. Then T is a quotient operator if and only if it is an isomorphism, that is to say if and only if the two norms are equivalent.

Remark 4.6. In some cases it is useful to think of a not necessarily surjective operator $T \colon X \to Y$ in terms of its so-called *canonical factorisation*. Indeed, if we write $\pi \colon X \to X/\operatorname{Ker} T$ for the canonical quotient operator and $S \colon \operatorname{Ran} T \to Y$ for the usual embedding, then $T = S \circ T_0 \circ \pi$, where $T_0 \colon X/\operatorname{Ker} T \to \operatorname{Ran} T$ is as above. Hence the following diagram commutes:

$$
\begin{array}{ccc}
X & \xrightarrow{\ \ T\ \ } & Y \\
\pi \big\downarrow & & \big\uparrow S \\
X/\operatorname{Ker} T & \xrightarrow{\ T_0\ } & \operatorname{Ran} T
\end{array}
$$

Theorem 4.7. *Let X and Y be normed vector spaces and suppose that $T \in \mathcal{B}(X, Y)$. Then the following are equivalent:*

(a) *T is a quotient operator;*

(b) *There exists $M > 0$ such that for every $y \in Y$ there exists $x \in X$ with $Tx = y$ and $\|x\| \leq M\|y\|$;*

(c) *There exists $r > 0$ such that $B_Y^\circ(r) \subseteq T(B_X^\circ)$;*

(d) *$T(B_X^\circ)$ has non-empty interior;*

(e) *T is an open map.*

Moreover, T is an isometric quotient operator if and only if $T(B_X^\circ) = B_Y^\circ$.

PROOF. We begin by showing carefully that (a) $\Longleftrightarrow$ (c). Suppose that T is a quotient operator. Then T is surjective and there exists $r > 0$ such that $\|Tx\| \geq r\|x + \operatorname{Ker} T\|$ for all $x \in X$. Let $y \in B_Y^\circ(r)$. Then there exists $x_0 \in X$ such that $Tx_0 = y$ and $\|y\| \geq r\|x_0 + \operatorname{Ker} T\|$. Thus $\|x_0 + \operatorname{Ker} T\| < 1$ so there exists $z_0 \in \operatorname{Ker} T$ such that for $x = x_0 + z_0$ we have $x \in B_X^\circ$ and $Tx = y$. Hence $B_Y^\circ(r) \subseteq T(B_X^\circ)$. Now suppose conversely that $B_Y^\circ(r) \subseteq T(B_X^\circ)$. Then T is clearly surjective. Let $x \in X$ and let $y = Tx$. For $r_0 \in (0, r)$ there exists $z \in B_X^\circ$ such that for $x_0 = r_0^{-1}\|y\|z$ we have $Tx_0 = y$. Hence $r_0\|x + \operatorname{Ker} T\| \leq r_0\|x_0\| \leq \|Tx\|$, so T is a quotient operator. In fact, since $r_0 \in (0, r)$ was arbitrary it follows that $\|Tx\| \geq r\|x + \operatorname{Ker} T\|$, $x \in X$.

It is straightforward to see that (a) $\Longrightarrow$ (b), and in fact if $\|Tx\| \geq r\|x + \operatorname{Ker} T\|$ for all $x \in X$ then (b) holds for any $M > r^{-1}$. Moreover, (b) $\Longrightarrow$ (c) with $r = M^{-1}$, and (c) $\Longrightarrow$ (d) is trivial. If (d) holds and $T(B_X^\circ)$ contains $B_Y^\circ(Tx_0, r)$ for some

$x_0 \in B_X^\circ$ and some $r > 0$, then by symmetry $T(B_X^\circ)$ also contains $B_Y^\circ(-Tx_0, r)$ and hence by convexity it also contains

$$B_Y^\circ(r) = \frac{1}{2} B_Y^\circ(Tx_0, r) + \frac{1}{2} B_Y^\circ(-Tx_0, r).$$

Suppose that $U \subseteq X$ is open and that $y \in T(U)$. Then $y = Tx$ for some $x \in U$. Since U is open there exists $\varepsilon > 0$ such that $B_X^\circ(x, \varepsilon) = x + B_X^\circ(\varepsilon) \subseteq U$. Then by linearity of T we see that $B_Y^\circ(y, r\varepsilon) \subseteq T(B_X^\circ(x, \varepsilon)) \subseteq T(U)$, so $T(U)$ is open and (d) $\implies$ (e). Finally, if (e) holds then $T(B_X^\circ)$ is open and it certainly contains the origin. Hence (e) $\implies$ (c).

For the final statement, note that if T is an isometric quotient operator then the proof of (a) $\implies$ (c) shows that $B_Y^\circ \subseteq T(B_X^\circ)$. Moreover, $\|T\| = \|T_0\| = 1$ and hence $T(B_X^\circ) \subseteq B_Y^\circ$. Conversely, if $T(B_X^\circ) = B_Y^\circ$ then the proof of (c) $\implies$ (a) shows that $\|Tx\| \geq \|x + \operatorname{Ker} T\|$ for all $x \in X$. We also have $\|T\| = 1$ and hence $\|T_0\| = 1$. Thus $\|Tx\| = \|x + \operatorname{Ker} T\|$ for all $x \in X$ and T is an isometric quotient operator. $\qquad \square$

In the Banach space setting we can once again say more.

Theorem 4.8 (Open Mapping Theorem). *Let X and Y be Banach spaces and suppose that $T \in \mathcal{B}(X, Y)$ is a surjection. Then T is an open map.*

By Theorem 4.7 this tells us that if X and Y are Banach spaces then an operator $T \in \mathcal{B}(X, Y)$ is a quotient operator if and only if it is surjective.[20] In order to improve Theorem 4.7 even further we require the following lemma which is used in the proof of the Open Mapping Theorem.

Lemma 4.9 (Successive Approximations Lemma). *Let X be a Banach space, Y a normed vector space and $T \in \mathcal{B}(X, Y)$. Suppose there exist $\varepsilon \in (0, 1)$ and $M > 0$ such that $\operatorname{dist}(y, T(B_X^\circ(M))) < \varepsilon$ for all $y \in B_Y^\circ$. Then $B_Y^\circ \subseteq T(B_X^\circ(M(1 - \varepsilon)^{-1}))$. Furthermore, if $T(B_X^\circ(M))$ contains a dense subset of B_Y°, then $B_Y^\circ \subseteq T(B_X^\circ(M))$. In either case, T is a quotient operator and Y is complete.*

PROOF. Let $y \in B_Y^\circ$. We recursively define sequences $(x_n)_{n=1}^\infty$ in X and $(y_n)_{n=1}^\infty$ in Y as follows. Set $y_1 = y$ and let $x_1 \in B_X^\circ(M)$ be such that $\|Tx_1 - y_1\| < \varepsilon$. Supposing we have $x_n \in X$ and $y_n \in Y$ such that $\|y_n\| < \varepsilon^{n-1}$, $\|x_n\| < M\varepsilon^{n-1}$ and $\|Tx_n - y_n\| < \varepsilon^n$, we set $y_{n+1} = y_n - Tx_n$. Since $\varepsilon^{-n}\|y_{n+1}\| < 1$ there exists $x_{n+1}' \in B_X^\circ(M)$ such that $\|Tx_{n+1}' - \varepsilon^{-n}y_{n+1}\| < \varepsilon$. If we let $x_{n+1} = \varepsilon^n x_{n+1}'$ then $\|x_{n+1}\| < M\varepsilon^n$ and we may continue inductively. Since $\sum_{n=1}^\infty \|x_n\| < \infty$ and X is complete, the series $\sum_{n=1}^\infty x_n$ converges to some $x \in X$ satisfying

$$\|x\| \leq \sum_{n=1}^\infty \|x_n\| < \frac{M}{1 - \varepsilon}.$$

Moreover,

$$\left\| y - \sum_{k=1}^n Tx_k \right\| = \|y_{n+1}\| < \varepsilon^n \to 0, \quad n \to \infty.$$

---

[20]So, in the Banach space setting, we have the same result as in linear algebra: a surjective morphism $T : X \to Y$ (i.e. bounded linear map) induces an isomorphism (in the category of Banach spaces with bounded linear maps) between the quotient space $X/\operatorname{Ker} T$ and Y.

By continuity of T we obtain that $Tx = y$, which proves the first claim. If $T(B_X^\circ(M))$ contains a dense subset of B_Y°, then $B_Y^\circ \subseteq T(B_X^\circ(M)) + B_Y^\circ(\varepsilon)$ and hence $B_Y^\circ(1-\varepsilon) \subseteq T(B_X^\circ(M))$ for all $\varepsilon \in (0,1)$. It follows that

$$B_Y^\circ = \bigcup_{\varepsilon \in (0,1)} B_Y^\circ(1-\varepsilon) \subseteq T(B_X^\circ(M)),$$

as required. In either case Theorem 4.7 shows that T is a quotient operator. Thus Y is isomorphic to the Banach space $X/\operatorname{Ker} T$ and hence Y itself is complete. $\qquad\square$

Remark 4.10. If in Theorem 4.7 we assume X to be complete, then by Lemma 4.9 we may weaken the conditions in (c) and (d). For instance, in (c) it would be sufficient to require that the closure of $T(B_X)$ contains $B_Y^\circ(r)$ for some $r > 0$, or indeed that there exist $\varepsilon \in (0,1)$ and $r > 0$ such that $B_Y^\circ(r) \subseteq T(B_X^\circ) + B_Y^\circ(\varepsilon)$. In the first case we need to recall that by continuity of T the closures of $T(B_X^\circ)$ and $T(B_X)$ coincide. Furthermore, a necessary and sufficient condition for T to be an isometric quotient operator is now that the closure of $T(B_X)$ equals B_Y.

5   The Hahn-Banach Theorems

The Hahn-Banach theorems refer to a range[21] of theorems in functional analysis concerned with extension of functionals, and separation of points and sets.

- **Extension** Given a linear functional g defined on a subspace Y of a vector space X, when can we extend g to a functional defined on X? Asked like this, the answer is always[22], so a more useful version of this question imposes some control on g, and asks for extensions retaining control. For example (and probably the version of Hahn-Banach you're familiar with already), if X is a normed space, we can extend *bounded* linear functionals from subspaces without increasing the norm (see Remarks 5.4 and 5.7).

- **Separation** Given disjoint subsets A and B of a vector space X, when we can we separate these by hyperplanes, i.e. in the case of real vector spaces find a linear functional f on X and constant c such that $f(a) < c$ for all $a \in A$, while $f(b) > c$ for all $b \in B$. Drawing some pictures in 2 dimensions should convince you that at the very least you'll need some convexity (and for the version I've stated both A and B to be closed). Is convexity enough?

The version of control we will use for extension in this course is sublinearity. Given a vector space X, a map $p\colon X \to \mathbb{R}$ is said to be a *sublinear functional* if it satisfies the following two properties:

- $p(\lambda x) = \lambda p(x)$ for all $\lambda > 0$ and all $x \in X$;
- $p(x + y) \le p(x) + p(y)$ for all $x, y \in X$.

---

[21]We will see further versions in Section 9, and in the exercises. Annoyingly in the literature it is quite normal to say 'by Hahn-Banach' to refer to any version, or some corollary thereof. At least in this course we will try and do better.

[22]Using Hamel bases and the axiom of choice.

Note that any sublinear functional satisfies $p(0) = 0$. Any seminorm, and in particular any norm, is a sublinear functional. Moreover, if $\mathbb{F} = \mathbb{R}$ then any linear functional on X is also a sublinear functional.

An important example of a sublinear functional is the *Minkowski functional* of a convex set.[23] Suppose that C is a convex subset of X which is *absorbing* in the sense that for each $x \in X$ there exists $\lambda > 0$ such that $\lambda x \in C$. Then for each $x \in X$ the set $\{\lambda \in \mathbb{R} : \lambda x \in C\}$ is an interval containing 0. The *Minkowski functional p_C* of C is defined by

$$p_C(x) = \inf\left\{\lambda > 0 : \lambda^{-1} x \in C\right\}, \quad x \in X.$$

The set C is said to be *symmetric* (or sometimes *balanced* when $\mathbb{F} = \mathbb{C}$) if $x \in C$ implies $\lambda x \in C$ for all $\lambda \in \mathbb{F}$ with $|\lambda| = 1$. A set of the form $\{\lambda x : \lambda > 0\}$ with $x \in X \setminus \{0\}$ is called an *infinite ray*.

Lemma 5.1. *Suppose that C is a convex absorbing subset of a vector space X. Then p_C is a sublinear functional. If C is symmetric, then p_C is a seminorm. If C in addition contains no infinite rays, then p_C is a norm.*

PROOF. Let $x \in X$. If $x = 0$ then $\lambda^{-1} x \in C$ for all $\lambda > 0$ and hence $p_C(x) = 0$. If $x \neq 0$ and $\mu > 0$ is such that $\mu x \in C$, then by convexity $\lambda^{-1} x \in C$ for all $\lambda \geq \mu^{-1}$. It follows that $p_C(x)$ is well-defined and that $0 \leq p_C(x) \leq \mu^{-1}$. To prove subadditivity, let $x, y \in X$ and define $S_x = \{\lambda > 0 : \lambda^{-1} x \in C\}$ and $S_y = \{\lambda > 0 : \lambda^{-1} y \in C\}$. For $\lambda \in S_x$, $\mu \in S_y$ we have by convexity

$$\frac{x + y}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu}\lambda^{-1}x + \frac{\mu}{\lambda + \mu}\mu^{-1}y \in C,$$

and hence $\lambda + \mu \geq p_C(x+y)$. Taking the infimum over $\lambda \in S_x$ and $\mu \in S_y$ shows that p_C is subadditive. If $\lambda > 0$ and $x \in X$, then $\{\mu > 0 : \mu^{-1}\lambda x \in C\} = \{\lambda\mu : \mu^{-1}x \in C\}$, and hence $p_C(\lambda x) = \lambda p_C(x)$. If C is symmetric, a similar argument shows that $p_C(\lambda x) = |\lambda| p_C(x)$ for all $\lambda \in \mathbb{F}$ and $x \in X$. Suppose finally that $p_C(x) = 0$. Then $\lambda^{-1} x \in C$ for all $\lambda > 0$, so either $x = 0$ or C contains an infinite ray. Thus if C is symmetric and contains no infinite rays then p_C is a norm on X. $\qquad\square$

Remark 5.2. If X is a normed vector space and C is a convex absorbing subset of X, then $C = \{x \in X : p_C(x) < 1\}$ if C is open and $C = \{x \in X : p_C(x) \leq 1\}$ if C is closed. Moreover, if $B_X^\circ \subseteq C \subseteq B_X$ then $p_C(x) = \|x\|$ for all $x \in X$.

When we first consider norms, we often draw pictures of the unit ball in $\mathbb{R}^2$ with respecct to the p-norms. What we're doing with these Minkowski functionals is not dissimilar; one is specifying a ball and producing a corresponding sublinear functional.[24]

Theorem 5.3 (Hahn-Banach Extension Theorem, real case). *Let X be a real vector space and let Y be a subspace of X. Suppose that p is a sublinear functional on X and that $g \in Y'$ is such that $g(y) \leq p(y)$ for all $y \in Y$. Then there exists $f \in X'$ such that $f|_Y = g$ and $f(x) \leq p(x)$ for all $x \in X$.*

---

[23]This starts to expose the deep connection between Hahn-Banach extension type theorems and convexity. This goes a lot further than we'll see in the course.

[24]We'll return to this point on example sheet 2.

Remark 5.4. If p is a (semi)norm then the assumption on g in fact implies that $|g(y)| \leq p(y)$ for all $y \in Y$. Thus Theorem 5.3 contains the standard version of the Hahn-Banach Theorem allowing us to extend a bounded linear functional on a (possibly non-separable) normed vector space without increasing its norm.

Lemma 5.5. *Let X be a real vector space and let Y be a proper subspace of X. Suppose that p is a sublinear functional on X and that $g \in Y'$ is such that $g(y) \leq p(y)$ for all $y \in Y$. Suppose moreover that $x_0 \in X \setminus Y$ and let Z be the linear span of $Y \cup \{x_0\}$. Then there exists $f \in Z'$ such that $f|_Y = g$ and $f(z) \leq p(z)$ for all $z \in Z$.*

PROOF. Every $z \in Z$ can be uniquely expressed in the form $z = y + \lambda x_0$ with $y \in Y$ and $\lambda \in \mathbb{R}$. This forces f to be of the form $f(y + \lambda x_0) = g(y) + c\lambda$ for some $c \in \mathbb{R}$, which remains to be fixed. Now the condition $f(z) \leq p(z)$ for all $z \in Z$ is equivalent to $g(y) + c\lambda \leq p(y + \lambda x_0)$ for all $y \in Y$ and $\lambda \in \mathbb{R}$. For $\lambda = 0$ the condition is true for all $y \in Y$ by assumption, and distinguishing the cases $\lambda \gtrless 0$ it is straightforward to see that our condition is equivalent to having

$$g(y_1) - p(y_1 - x_0) \leq c \leq p(y_2 + x_0) - g(y_2), \quad y_1, y_2 \in Y.$$

We need to show that, for $y_1, y_2 \in Y$,

$$g(y_1 + y_2) \leq p(y_1 - x_0) + p(y_2 + x_0).$$

But this follows immediately from the assumption on g and subadditivity of p. Thus

$$\sup \big\{ g(y) - p(y - x_0) : y \in Y \big\} \leq \inf \big\{ p(y + x_0) - g(y) : y \in Y \big\},$$

and we may choose c to be any number between these two quantities.[25] $\qquad \square$

If the space X in Theorem 5.3 is a separable normed vector space and p is the norm, then we may this lemma repeatedly to extend g by one dimension at a time, thus defining f on a dense subset of X, and then we may extend to the whole of X using continuity. In the non-separable case we use[26] Zorn's Lemma.

PROOF OF THEOREM 5.3: We say that a real valued function f is a *g-extension* if its domain $D(f)$ is a subspace of X containing Y, and f is linear, with $f|_Y = g$ and $f(z) \leq p(z)$ for all $z \in D(f)$. Let $\mathcal{P}$ be the collection of all such g-extensions (noting that this is non-empty as g is a g-extension) equipped with the partial order $f_1 \precsim f_2$ if and only if $D(f_1) \subset D(f_2)$ and $f_2|_{D(f_1)} = f_1$.[27] Given any chain $\{f_i : i \in \mathcal{C}\}$ in $\mathcal{P}$, note that $\bigcup_{i \in \mathcal{C}} D(f_i)$ is a subspace of X (containing Y) and we can define a linear map $f : D(f) \to \mathbb{R}$ by $f(x) = f_i(x)$ for $x \in Dom(f_i)$. This f is a g-extension[28] so provides an upper bound for the chain.

Therefore Zorn's Lemma there is a maximal g-extension f. If $D(f) \neq X$, then choose $x_0 \in X \setminus D(f)$, and extend f by Lemma 5.5 to some $f_1 \in \mathcal{P}$ defined on $\mathrm{Span}(D(f) \cup \{x_0\})$, contradicting maximality. $\qquad \square$

---

[25]So the extension will be unique when these two quantities are equal. We'll investigate how to characterise uniqueness of extension on example sheet 3.

[26]Strictly speaking there's a bit of logical overkill here. Hahn-Banach is weaker than the axiom of choice, but strong enough to imply the existence of non-Lebesgue measurable sets and the Banach-Tarski paradox. The exact logical statement equivalent to Hahn-Banach is a little fiddly though.

[27]Note that $f_1 \precsim f_2$ is exacty the relation that the graph of f_1 is a subset of the graph of f_2.

[28]Can you justify all these claims?

If $\mathbb{F} = \mathbb{C}$ and functionals are complex-valued and complex-linear, then the sublinear functional appearing in Theorem 5.3 needs to be replaced by a seminorm. Surprisingly the following result was obtained around 10 years after the real version.

Theorem 5.6 (Hahn-Banach Extension Theorem, complex case). *Let X be a complex vector space and let Y be a subspace of X. Suppose that p is a seminorm on X and that $g \in Y'$ is such that $|g(y)| \leq p(y)$ for all $y \in Y$. Then there exists $f \in X'$ such that $f|_Y = g$ and $|f(x)| \leq p(x)$ for all $x \in X$.*

PROOF. Observe first that we may regard a complex vector space as a real vector space with the same operations, and that the assignment $f \mapsto \operatorname{Re} f$ sending a complex-linear functional f to the real-linear functional $\operatorname{Re} f$ given by $(\operatorname{Re} f)(x) = \operatorname{Re} f(x)$, $x \in X$, is a bijection. Indeed, it is clear that $\operatorname{Re} f$ is real-linear, and if $\operatorname{Re} f = \operatorname{Re} g$ for two complex-linear functionals f and g then, for $x \in X$,

$$\operatorname{Im} f(x) = -\operatorname{Re}(if(x)) = -\operatorname{Re}(f(ix)) = -\operatorname{Re}(g(ix)) = -\operatorname{Re}(ig(x)) = \operatorname{Im} g(x),$$

and hence $f = g$, so the assignment is injective. On the other hand, if g is a real-linear functional then it is easy to verify that the functional f given by $f(x) = g(x) - ig(ix)$ for $x \in X$ is complex-linear and satisfies $\operatorname{Re} f = g$, so the assignment is surjective. Our next observation is that $\operatorname{Re} f(x) \leq p(x)$ for all $x \in X$ if and only if $|f(x)| \leq p(x)$ for all $x \in X$. Indeed, one implication is trivial and for the other we note that if $\operatorname{Re} f(x) \leq p(x)$ for all $x \in X$ then for some $\theta \in [0, 2\pi)$ depending on $x \in X$ we have

$$|f(x)| = e^{-i\theta} f(x) = \operatorname{Re} f(e^{-i\theta} x) \leq p(e^{-i\theta} x) = p(x), \quad x \in X.$$

Thus given $g \colon Y \to \mathbb{C}$ as in the statement of the theorem, we may apply Theorem 5.3 to find a real-linear functional $f_0 \colon X \to \mathbb{R}$ such that $f_0|_Y = \operatorname{Re} g$ and $f_0(x) \leq p(x)$ for all $x \in X$. Now set $f(x) = f_0(x) - if_0(ix)$, $x \in X$. Then $f \in X'$ and $|f(x)| \leq p(x)$, $x \in X$. Furthermore, we have $\operatorname{Re} f(y) = \operatorname{Re} g(y)$ for all $y \in Y$, so arguing as before we see that $f|_Y = g$, as required. $\square$

Remark 5.7. The above proof also shows that if X is a complex normed vector space and $f \in X^*$, then the real-linear functional $\operatorname{Re} f$ is bounded with $\|\operatorname{Re} f\| = \|f\|$. Note too that Theorem 5.6 contains the usual Hahn-Banach extension theorem for bounded linear functionals on complex normed spaces.

In the next result we collect some of the standard consequences of the Hahn-Banach Theorem. Here and in what follows, given a subset M of a normed vector space X we denote the *annihilator* of M in X^* by

$$M^\circ = \big\{ f \in X^* : f(x) = 0 \text{ for all } x \in M \big\},$$

and given a subset N of X^* we let

$$N_\circ = \big\{ x \in X : f(x) = 0 \text{ for all } f \in N \big\}$$

be the *annihilator* of N in X. It is clear that annihilators are closed subspaces.

Corollary 5.8. *Let X be a normed vector space.*
  (a) *For each $x_0 \in X$ there exists $f \in S_{X^*}$ such that $f(x_0) = \|x_0\|$. In particular, $\|x\| = \max\{|f(x)| : f \in S_{X^*}\}$ for all $x \in X$.*

(b) *If Y is a subspace of X and $x_0 \in X$, then there exists $f \in Y^\circ$ such that $\|f\| \leq 1$ and $f(x_0) = \operatorname{dist}(x_0, Y)$. In particular, the closure of Y coincides with $(Y^\circ)_\circ$ and Y is dense in X if and only if $Y^\circ = \{0\}$.*

PROOF. For the first part of (a) we may take any $f \in S_{X^*}$ if $x_0 = 0$, and otherwise it suffices to apply the Hahn-Banach Theorem to the linear functional g defined on $\operatorname{span}\{x_0\}$ by $g(\lambda x_0) = \lambda \|x_0\|$, $\lambda \in \mathbb{F}$. The second part then follows easily. For part (b) we may consider the seminorm $p(x) = \operatorname{dist}(x, Y)$, $x \in X$, and $g \colon \operatorname{span}\{x_0\} \to \mathbb{F}$ given by $g(\lambda x_0) = \lambda \operatorname{dist}(x_0, Y)$, $\lambda \in \mathbb{F}$. By the Hahn-Banach Theorem there exists a linear functional $f \colon X \to \mathbb{F}$ such that $f(x_0) = g(x_0) = \operatorname{dist}(x_0, Y)$ and $|f(x)| \leq p(x) \leq \|x\|$ for all $x \in X$. In particular, $f \in X^*$ with $\|f\| \leq 1$ and $f \in Y^\circ$ since $|f(x)| \leq p(x) = 0$ for all $x \in Y$. The remaining statements follow straightforwardly from the fact that $\operatorname{dist}(x, Y) = 0$ if and only if x lies in the closure of Y. □

Let X be a normed vector space and suppose that $C \subseteq X$ and $x_0 \in X \setminus C$. We say that x_0 and C are *strictly separated* if there exists $f \in X^*$ such that $\operatorname{Re} f(x_0) > \operatorname{Re} f(x)$ for all $x \in C$, and that they are *uniformly separated* if there exists $f \in X^*$ such that

$$\operatorname{Re} f(x_0) > \sup\{\operatorname{Re} f(x) : x \in C\}.$$

If $\mathbb{F} = \mathbb{R}$ the real parts are redundant.

Theorem 5.9 (Hahn-Banach Separation Theorem). *Let X be a normed vector space and suppose that C is a non-empty convex subset of X and that $x_0 \in X \setminus C$.*

  (a) *If C is open, then x_0 and C are strictly separated.*
  (b) *If C is closed, then x_0 and C are uniformly separated.*

PROOF. (a) Fix $y_0 \in C$ and let $z_0 = x_0 - y_0$ and $C_0 = C - y_0$. Since C is open we have $B_X^\circ(\varepsilon) \subseteq C_0$ for some $\varepsilon > 0$. In particular, the set C_0 is absorbing so the Minkowski functional $p = p_{C_0}$ is a well-defined sublinear functional on X. Note also that $p(x) \leq \varepsilon^{-1}\|x\|$ for all $x \in X$. Consider the functional $g : \operatorname{span}\{z_0\} \to \mathbb{F}$ given by $g(\lambda z_0) = \lambda$, $\lambda \in \mathbb{F}$. Since $z_0 \notin C_0$ we have $p(z_0) \geq 1 = g(z_0)$. Suppose first that $\mathbb{F} = \mathbb{R}$. Then for $\lambda \geq 0$ we have $p(\lambda z_0) = \lambda p(z_0) \geq g(\lambda z_0)$, while for $\lambda < 0$ we have $p(\lambda z_0) \geq 0 > g(\lambda z_0)$. By the Hahn-Banach Extension Theorem there exists a linear functional f on X such that $f(\lambda z_0) = \lambda$ for all $\lambda \in \mathbb{R}$ and $f(x) \leq p(x) \leq \varepsilon^{-1}\|x\|$ for all $x \in X$, so $f \in X^*$. Let $x \in C$. Then there exists $\delta > 0$ such that $x + \delta z_0 \in C$, so $p(x + \delta z_0 - y_0) \leq 1$ and hence

$$f(x) + \delta = f(x + \delta z_0 - x_0) + f(x_0) \leq p(x + \delta z_0 - y_0) - 1 + f(x_0) \leq f(x_0),$$

giving $f(x) < f(x_0)$. If $\mathbb{F} = \mathbb{C}$ we find, by considering X as a real vector space and proceeding as above, a bounded real-linear functional f_0 on X such that $f_0(x) < f_0(x_0)$ for all $x \in C$. As in the proof of Theorem 5.6 we now take $f \in X^*$ to be given by $f(x) = f_0(x) - if_0(ix)$, $x \in X$, so that $f_0 = \operatorname{Re} f$.

(b) If C is closed and $x_0 \notin C$ then there exists $\varepsilon > 0$ such that $x_0 \notin C_\varepsilon = C + B_X^\circ(\varepsilon)$. Since C_ε is open and convex we may apply part (a) to find $f \in X^*$ such that $\operatorname{Re} f(x) < \operatorname{Re} f(x_0)$ for all $x \in C_\varepsilon$. Let $z_0 \in X$ be such that $f(z_0) = 1$ and let $\delta > 0$ be such that $\delta\|z_0\| < \varepsilon$. Then for all $x \in C$ we have $x + \delta z_0 \in C_\varepsilon$ and hence

$$\operatorname{Re} f(x) = \operatorname{Re} f(x + \delta z_0) - \delta < \operatorname{Re} f(x_0) - \delta,$$

which gives the result. □

Recall that, given a bounded linear operator $T\colon X \to Y$ between two normed vector spaces, the *(topological) dual operator* $T^*\colon Y^* \to X^*$ of T is given by $(T^*f)(x) = f(Tx)$ for $f \in Y^*$, $x \in X$. We know that $T^* \in \mathcal{B}(Y^*, X^*)$, and by a standard application of the Hahn-Banach Theorem we have $\|T^*\| = \|T\|$. Recall also that T is said to be an *isomorphic embedding* if there exists a constant $r > 0$ such that $\|Tx\| \geq r\|x\|$ for all $x \in X$. This is equivalent to saying that T maps isomorphically onto its range.

Theorem 5.10. *Let X and Y be normed spaces and suppose that $T \in \mathcal{B}(X, Y)$.*

  (a) *T is an isomorphic embedding (an isometry) if and only if T^* is a quotient operator (an isometric quotient operator).*

  (b) *If T is a quotient operator (an isometric quotient operator) then T^* is an isomorphic embedding (an isometry), and if X is complete the converse holds.*

PROOF. (a) Suppose that T is an isomorphic embedding, so that there exists $r > 0$ such that $\|Tx\| \geq r\|x\|$ for all $x \in X$. Let $Z = \operatorname{Ran} T$. Given $g \in X^*$ we may define $h \in Z'$ by setting $h(Tx) = g(x)$, $x \in X$. This is well-defined by injectivity of T, and moreover $|h(Tx)| \leq \|g\|\|x\| \leq M\|g\|\|Tx\|$, $x \in X$, where $M = r^{-1}$. Thus $h \in Z^*$ and by the Hahn-Banach Theorem there exists $f \in Y^*$ such that $f|_Z = h$ and $\|f\| = \|h\| \leq M\|g\|$. Thus $T^*f = g$ and, by Theorem 4.7, T^* is a quotient operator. If T is an isometry we may take $r = M = 1$ and hence $\|T^*f\| \geq \|f\| \geq \|f + \operatorname{Ker} T^*\|$. Since $\|(T^*)_0\| = \|T^*\| = \|T\| = 1$ it follows that T^* is an isometric quotient operator.

Conversely, suppose that T^* is a quotient operator. Then T^* is surjective and there exists $r > 0$ such that $\|T^*f\| \geq r\|f + \operatorname{Ker} T^*\|$ for all $f \in Y^*$. Hence by Theorem 4.7 there exists $M > 0$ such that for every $g \in X^*$ there exists $f \in Y^*$ with $T^*f = g$ and $\|f\| \leq M\|g\|$, and as observed in the proof of that result we may take any $M > r^{-1}$. Now by Corollary 5.8, given $x \in X$, there exists $g \in S_{X^*}$ such that $g(x) = \|x\|$. Choose $f \in Y^*$ so that $T^*f = g$ and $\|f\| \leq M$. Then

$$\|x\| = g(x) = f(Tx) \leq \|f\|\|Tx\| \leq M\|Tx\|, \tag{5.1}$$

which shows that T is an isomorphic embedding. If T^* is an isometric quotient operator we may choose $r = 1$ and then, for every $x \in X$, (5.1) holds for all $M > 1$. In particular, $\|Tx\| \geq \|x\|$ for all $x \in X$. Since $\|T\| = \|T^*\| = 1$, T is an isometry.

(b) If T is quotient operator, then T is surjective and there exists $r > 0$ such that $\|Tx\| \geq r\|x + \operatorname{Ker} T\|$ for all $x \in X$. By Theorem 4.7 there exists $M > 0$ such that for every $y \in Y$ there exists $x \in X$ with $Tx = y$ and $\|x\| \leq M\|y\|$ and once again any $M > r^{-1}$ works. Given $y \in Y$, let $x \in X$ be as described. Then

$$|f(y)| = |T^*f(x)| \leq \|T^*f\|\|x\| \leq M\|T^*f\|\|y\|,$$

and hence $\|f\| \leq M\|T^*f\|$ for all $f \in Y^*$, so T^* is an isomorphic embedding. If T is an isometric quotient operator we may take $r = 1$ and then, for every $f \in Y^*$, we have $\|f\| \leq M\|T^*f\|$ for all $M > 1$. Hence $\|T^*f\| \geq \|f\|$ for all $f \in Y^*$. Since $\|T^*\| = \|T\| = \|T_0\| = 1$, T^* is an isometry.

The final statement requires the Hahn-Banach Separation Theorem and the Successive Approximations Lemma; see Problem Sheet 2. $\qquad\square$

Remark 5.11. In the setting of Theorem 5.10 we always have $\operatorname{Ran} T^* \subseteq (\operatorname{Ker} T)^\circ$. If T is surjective and $g \in (\operatorname{Ker} T)^\circ$ the map $f(Tx) = g(x)$, $x \in X$, is a well-defined linear functional on Y. If T is a quotient operator an argument similar to the one used in part (b) above shows that $f \in Y^*$. Since $g = T^* f$ we have that $\operatorname{Ran} T^* = (\operatorname{Ker} T)^\circ$ in this case. In particular, T^* has closed range. Recall that we always have $(\operatorname{Ran} T^*)_\circ = \operatorname{Ker} T$ and $(\operatorname{Ran} T)^\circ = \operatorname{Ker} T^*$. In particular, by Corollary 5.8 the closure of $\operatorname{Ran} T$ coincides with $(\operatorname{Ker} T^*)_\circ$, but in general even the closure of $\operatorname{Ran} T^*$ is contained in $(\operatorname{Ker} T)^\circ$ as a proper subset. We shall return to this issue in later sections.

Corollary 5.12. *Let X be a normed vector space and Y a closed subspace of X. Then $Y^* \cong X^*/Y^\circ$ and $(X/Y)^* \cong Y^\circ$.*

PROOF. Let $S\colon Y \to X$ denote the inclusion operator and let $\pi\colon X \to X/Y$ denote the canonical quotient map. Then S is an isometry and π is an isometric quotient operator. By Theorem 5.10, $S^*\colon X^* \to Y^*$ is an isometric quotient operator and $\pi^*\colon (X/Y)^* \to X^*$ is an isometry. But $\operatorname{Ker} S^* = (\operatorname{Ran} S)^\circ = Y^\circ$, and hence $Y^* \cong X^*/Y^\circ$. By Remark 5.11 we have $\operatorname{Ran} \pi^* = (\operatorname{Ker} \pi)^\circ = Y^\circ$, so $(X/Y)^* \cong Y^\circ$. $\qquad\square$

6 Biduals and reflexivity

Given a normed vector space X, the dual space X^* is a Banach space and in particular has a dual space $X^{**} = (X^*)^*$ of its own, the so-called *bidual* of X. Recall that there always exists a well-defined linear map $J_X\colon X \to X^{**}$ given by $(J_X x)(f) = f(x)$ for $x \in X$ and $f \in X^*$. By Corollary 5.8 we have

$$\|J_X(x)\| = \max\{|f(x)| : f \in S_{X^*}\} = \|x\|, \quad x \in X,$$

so J_X is an isometry. The space X is said to be *reflexive* if J_X is surjective.

Given any metric space X, a pair (Y, J) is said to be a *completion* of X if Y is a complete metric space and $J\colon X \to Y$ is an isometry whose range is dense in Y. Every metric space has a completion, and this is usually proved by considering a quotient of the space of all Cauchy sequences. In the case of normed vector spaces we obtain this result with very little effort by a different argument.

Proposition 6.1. *Every normed vector space X has a completion (Y, J) such that Y is a Banach space and J is linear.*

PROOF. Let Y be the closure of $J_X(X)$ in X^{**} and consider the map $J\colon X \to Y$ given by $J(x) = J_X(x)$, $x \in X$. Then Y is a closed subspace of the Banach space X^{**} and therefore itself is complete, and the range of J is dense in Y by construction. $\quad\square$

Let us recall briefly some examples of classical Banach spaces. We write ℓ^∞ for the space of all bounded scalar-valued sequences $x = (x_n)_{n \geq 1}$, and we endow this space with the supremum norm given by $\|x\|_\infty = \sup_{n \geq 1} |x_n|$. We write c for the subspace of ℓ^∞ given by sequences x such that $\lim_{n \to \infty} x_n$ exists, and we let c_0 denote the subspace of sequences converging to zero. With the supremum norm both of these are closed subspaces of ℓ^∞ and hence themselves Banach spaces.

For $1 \leq p < \infty$ we let ℓ^p denote the space of scalar-valued sequences x for which $\sum_{n=1}^{\infty} |x_n|^p < \infty$, endowed with the norm

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

We denote by e_n, $n \geq 1$, the sequence $(\delta_{n,k})_{k \geq 1}$ and we let $c_{00} = \text{span}\{e_n : n \geq 1\}$ be the space of finitely supported sequences. Then c_{00} is dense in c_0 and in ℓ^p for $1 \leq p < \infty$, but not in c or in ℓ^∞.

If $1 \leq p \leq \infty$ we say that q is the *Hölder conjugate* of p if $1 \leq q \leq \infty$ and $p^{-1} + q^{-1} = 1$. You will have shown previously that the dual of ℓ^p for $1 \leq p < \infty$ is isometrically isomorphic to ℓ^q, where q is the Hölder conjugate of p, via the map $\Phi_p \colon \ell^q \to (\ell^p)^*$ given by

$$(\Phi_p y)(x) = \sum_{n=1}^{\infty} x_n y_n, \quad x \in \ell^p, \ y \in \ell^q.$$

In particular, $(\ell^1)^* \cong \ell^\infty$. Similarly, the duals of c and c_0 are both isomorphically isometric to ℓ^1 by the maps $\Phi \colon \ell^1 \to c^*$ and $\Psi \colon \ell^1 \to c_0^*$ given by

$$(\Phi y)(x) = y_1 \cdot \lim_{n \to \infty} x_n + \sum_{n=1}^{\infty} x_n y_{n+1}, \quad x \in c, \ y \in \ell^1,$$

and

$$(\Psi y)(x) = \sum_{n=1}^{\infty} x_n y_n, \quad x \in c_0, \ y \in \ell^1,$$

respectively. Note in particular that for $1 < p < \infty$ the space ℓ^p is isometrically isomorphic to its bidual. Let $X = \ell^p$ for $1 < p < \infty$ and let q be the Hölder conjugate of p. Then $J_X = (\Phi_p^{-1})^* \circ \Phi_q$. Indeed, for $x \in X$ and $f \in X^*$ we have $f = \Phi_p(y)$ for some $y \in \ell^q$, and hence

$$\left((\Phi_p^{-1})^*(\Phi_q x) \right)(f) = (\Phi_q x)(y) = \sum_{n=1}^{\infty} x_n y_n = (\Phi_p y)(x) = f(x) = (J_X x)(f),$$

so X is in fact reflexive. Other examples of reflexive spaces include all finite-dimensional spaces, all Hilbert spaces and, as we shall prove in Section 8, all L^p-spaces for $1 < p < \infty$.

Remark 6.2. A construction due to R.C. James (1951) shows that it is possible for a Banach space to be isometrically isomorphic to its bidual and yet non-reflexive.

We also see from the above considerations that the spaces c and c_0 are non-reflexive. Indeed, both spaces are separable but their bidual is isometrically isomorphic to ℓ^∞, which is non-separable. Recall that a normed vector space is said to be *separable* if it contains a countable dense subset. We will see on Problem Sheet 2 that c_0 not only fails to be isometrically isomorphic to its bidual, it is in fact not (isometrically) isomorphic to the dual of *any* normed vector space. Notice, however, that not knowing the dual of ℓ^∞ makes it hard to say anything about reflexivity of ℓ^1 for the moment. Separability turns out to be useful here too.

Lemma 6.3. *Let X be a normed vector space and suppose that X^* is separable. Then X too is separable.*

PROOF. Let $\{f_n : n \geq 1\}$ be a dense subset of S_{X^*} and, for each $n \geq 1$, let $x_n \in B_X$ be such that $|f_n(x_n)| \geq 1/2$. Let $Y = \mathrm{span}\{x_n : n \geq 1\}$. In order to show that X is separable it suffices, by a standard result, to prove that Y is dense in X. If this is not the case, then by Corollary 5.8 there exists $f \in S_{X^*} \cap Y^\circ$. Thus for some $n \geq 1$ we have $\|f - f_n\| < 1/2$ and consequently

$$|f(x_n)| \geq |f_n(x_n)| - \|f - f_n\| > 0,$$

which is a contradiction. Hence X is separable. $\qquad\square$

It follows that ℓ^1 cannot be reflexive, since if it were then its bidual would have to be separable and hence so would its dual. But we know that the dual of ℓ^1 is isometrically isomorphic to ℓ^∞ and in particular non-separable. More generally, the result shows that any separable normed vector with non-separable dual cannot be reflexive. The next result gives another way of seeing why ℓ^1, and indeed many other spaces, cannot be reflexive.

Theorem 6.4. *Let X be a normed vector space. Then X is reflexive if and only if X is complete and X^* is reflexive.*

PROOF. Suppose first that X is reflexive. Then X is isometrically isomorphic to its own bidual, which is a Banach space, and hence X is necessarily complete. Suppose that $\xi \in X^{***}$ and let $f = \xi \circ J_X$, noting that $f \in X^*$. Then given $\phi \in X^{**}$ we have by reflexivity of X that $\phi = J_X(x)$ for some $x \in X$, and hence

$$\xi(\phi) = \xi(J_X x) = f(x) = (J_X x)(f) = (J_{X^*} f)(\phi),$$

so $\xi = J_{X^*}(f)$ and X^* is reflexive.

Conversely, if X is complete and X^* is reflexive, then the image $Y = J_X(X)$ of X under the isometry J_X is complete and hence closed in the bidual X^{**}. Suppose that $\xi \in Y^\circ$. Then $\xi = J_{X^*}(f)$ for some $f \in X^*$ and $f(x) = (J_X x)(f) = \xi(J_X x) = 0$ for all $x \in X$. Hence $f = 0$ and therefore $\xi = 0$, so $Y^\circ = \{0\}$. By Corollary 5.8 we see that Y is dense in X^{**}. Since Y is closed we have $Y = X^{**}$, so X is reflexive. $\quad\square$

Using this result we see again that ℓ^1 cannot be reflexive because it is (isomorphic to) the dual of the non-reflexive space c_0. A similar argument works for ℓ^∞ and indeed for any Banach space which is isomorphic to the dual of a non-reflexive Banach space. Implicit in these statements is the observation that reflexivity is preserved under isomorphism.

Proposition 6.5. *Let X, Y be two normed spaces which are isomorphic. Then X is reflexive if and only if Y is.*
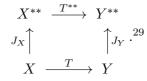
PROOF. Let $T \in \mathcal{B}(X, Y)$. For $x \in X$ and $f \in Y^*$ we have

$$\big((T^{**} \circ J_X)(x)\big)(f) = (J_X x)(T^* f) = f(Tx) = \big((J_Y \circ T)(x)\big)(f),$$

and hence $T^{**} \circ J_X = J_Y \circ T$. If $T \colon X \to Y$ is an isomorphism then $T^{-1}T$ and TT^{-1} are the identity operators on X and Y, respectively, and taking duals we see that T^* too is an isomorphism. Similarly $T^{**} \colon X^{**} \to Y^{**}$ is an isomorphism. Thus J_X is surjective if and only if J_Y is surjective. $\qquad\square$

Remark 6.6. (a) An alternative, but much less direct, way of seeing that the dual operator of any isomorphism is an isomorphism is to appeal to Theorem 5.10.

(b) The first part of the above proof establishes a general fact, namely that $T^{**} \circ J_X = J_Y \circ T$ whenever X and Y are normed vector spaces and $T \in \mathcal{B}(X, Y)$. Thus the following diagram commutes:

$$
\begin{array}{ccc}
X^{**} & \xrightarrow{\;T^{**}\;} & Y^{**} \\
J_X \big\uparrow & & \big\uparrow J_Y \\
X & \xrightarrow{\;\;T\;\;} & Y
\end{array} \cdot{}^{29}
$$

7 Convexity and smoothness of norms

Let X be a normed vector space. If $x, y \in S_X$ then $\frac{1}{2}\|x + y\| \leq 1$, and equality is possible even when $x \neq y$. We say that X (or its norm) is *strictly convex* if whenever $x, y \in S_X$ are distinct then $\frac{1}{2}\|x + y\| < 1$. Let $\delta_X : [0, 2] \to \mathbb{R}$ be given by

$$
\delta_X(\varepsilon) = \inf\left\{ 1 - \frac{\|x + y\|}{2} : x, y \in S_X, \|x - y\| \geq \varepsilon \right\}, \quad 0 \leq \varepsilon \leq 2.
$$

We call δ_X the *modulus of convexity* of X (or its norm). Note that $\frac{1}{2}\|x + y\| \leq 1 - \delta_X(\varepsilon)$ whenever $x, y \in S_X$ with $\|x - y\| \geq \varepsilon$. We say that X (or its norm) is *uniformly convex* if $\delta_X(\varepsilon) > 0$ for all $\varepsilon \in (0, 2]$. Note that any uniformly convex space is strictly convex, and in finite-dimensional spaces the two notions are equivalent. If X is a Hilbert space it follows from the parallelogram law that

$$
\delta_X(\varepsilon) \geq 1 - \left(1 - \frac{\varepsilon^2}{4} \right)^{1/2}, \quad 0 \leq \varepsilon \leq 2,
$$

and in particular every Hilbert space is uniformly convex. We will see on Problem Sheet 3 that every uniformly convex Banach space is reflexive. However, strictly convex Banach spaces are not in general reflexive. An equivalent definition of uniform convexity is that whenever (x_n) and (y_n) are sequences in X such that

$$
\|x_n\| \to 1, \quad \|y_n\| \to 1, \quad \frac{\|x_n + y_n\|}{2} \to 1, \quad n \to \infty,
$$

we have $\|x_n - y_n\| \to 0$ as $n \to \infty$.[30]

Theorem 7.1. *Suppose that X is a uniformly convex Banach space and let $f \in X^* \setminus \{0\}$. Then there exists a unique $x \in S_X$ such that $f(x) = \|f\|$.*

---

[29]This is what we mean by saying that the map J_X is natural. Indeed, in the language of category theory J provides a natural transformation from the identity functor to the functor **.

[30]As is often the case, it'll be this sequential characterisation we want to use in proofs. To see this claim, note that if uniform convexity fails, then there is some $\epsilon > 0$ such that $\delta_X(\epsilon) = 0$. From the definition of the infimum defining $\delta_X(\epsilon)$) we can find sequences $(x_n)_{n=1}^{\infty}, (y_n)_{n=1}^{\infty}$ in the sphere with $\|x_n - y_n\| \geq \epsilon$, and $\|(x_n + y_n)/2\| \to 1$. Conversely, if the sequential condition fails, then passing to subsequences, there is $\epsilon > 0$ and $(x_n)_{n=1}^{\infty}, (y_n)_{n=1}^{\infty}$ in X with $\|(x_n + y_n)/2\| \to 1$, but $\|x_n - y_n\| \geq \epsilon$ for all n. By considering $\tilde{x}_n = x_n/\|x_n\|$ and $\tilde{y}_n = y_n/\|y_n\|$, one gets $\delta_X(\epsilon) = 0$.

PROOF. Let $x_n \in S_X$, $n \geq 1$, be such that $f(x_n) \to \|f\|$ as $n \to \infty$. We show that the sequence (x_n) is Cauchy. By completeness of X and continuity of f the existence part then follows. Suppose not. Then there exist $\varepsilon > 0$ and increasing sequences of integers (n_k), (m_k) such that $\|x_{n_k} - x_{m_k}\| \geq \varepsilon$ for all $k \geq 1$. Let $y_k = \frac{1}{2}(x_{n_k} + x_{m_k})$, $k \geq 1$. Then $\|y_k\| \leq 1$ for all $k \geq 1$ and $f(y_k) \to \|f\|$ as $k \to \infty$. It follows that $\|y_k\| \to 1$ as $k \to \infty$, so by uniform convexity $\|x_{n_k} - x_{m_k}\| \to 0$ as $k \to \infty$, giving the required contradiction. For uniqueness, suppose that $x, y \in S_X$ are two distinct vectors such that $f(x) = f(y) = \|f\|$. By uniform convexity we must have $\frac{1}{2}\|x + y\| < 1$ and hence

$$\|f\| = f\left(\frac{x+y}{2}\right) \leq \|f\| \frac{\|x+y\|}{2} < \|f\|.$$

This contradiction completes the proof. □

Remark 7.2. Note that for the uniqueness part strict convexity was sufficient. In fact, it is not hard to show using Corollary 5.8 that X is strictly convex if and only if every $f \in X^* \setminus \{0\}$ attains its norm in at most one point $x \in S_X$. On the other hand, for the existence part of the theorem reflexivity of X would have been sufficient, as will become clear from Remark 10.2 below.

Thus in a uniformly convex space every functional attains its norm at a unique vector in the unit sphere. We call the vector x above a *norming* vector for f, i.e. $x \in X$ *norms* $f \in X^*$ if $\|x\| = 1$ and $f(x) = \|f\|$.

We now consider the dual question, interchanging the roles of X and X^*. Given $x_0 \in X$, Hahn-Banach (in the guise of Corollary 5.8) shows that for every $x_0 \in X$ there exists a *norming functional*, that is to say a functional $f \in S_{X^*}$ such that $f(x_0) = \|x_0\|$.[31] In general, norming functionals are non-unique even for non-zero vectors x_0, but again we have a geometric condition on X which gives rise to uniqueness.

We say that a real normed vector space X (or its norm) is *smooth at* $x_0 \in X$ if the limit

$$\lim_{h \to 0} \frac{\|x_0 + hx\| - \|x_0\|}{h}$$

exists for all $x \in X$. We say that X (or its norm) is *smooth* if it is smooth at all $x_0 \in X \setminus \{0\}$, or equivalently all $x_0 \in S_X$. Note that no norm can be smooth at 0. Smoothness gives us uniqueness of norming functionals; we record a formally stronger statement for use in the next statement.

Theorem 7.3. *Let X be a real normed vector space such that for $x_0 \in X$ with $\|x_0\| = 1$ and some $p \geq 1$, the limit*

$$\lim_{h \to 0} \frac{\|x_0 + hx\|^p - \|x_0\|^p}{ph}$$

exists for all $x \in X$. Then there is a unique norming functional $f \in S_{X^*}$ with $f(x_0) = \|x_0\|$ given by

$$f(x) = \lim_{h \to 0} \frac{\|x_0 + hx\|^p - \|x_0\|^p}{ph}.$$

---

[31]In contrast without strict convexity norming vectors may not exist. Can you give an example using c_0 and ℓ^1?

In particular if X is smooth, then all non-zero vectors have unique norming functionals.

PROOF. As observed above, existence of norming functionals follows from Corollary 5.8.

Let $f \in S_X$ be a norming functional for x_0 and fix $x \in X$. Note that as f is linear, we can use L'Hopital's rule to compute

$$\lim_{h \to 0} \frac{f(x_0 + hx)^p - f(x_0)^p}{ph} = f(x_0)^{p-1} f(x) = f(x),^{32} \tag{7.1}$$

On the other hand as $\|f\| \leq 1$ and $f(x_0) = \|x_0\|$, we have

$$f(x_0 + hx)^p - f(x_0) \leq \|x_0 + hx\|^p - \|x_0\|^p.$$

So for $h > 0$,

$$\frac{f(x_0 + hx)^p - f(x_0)^p}{ph} \leq \frac{\|x_0 + hx\|^p - \|x_0\|^p}{ph}$$

while for $h > 0$

$$\frac{f(x_0 + hx)^p - f(x_0)^p}{ph} \geq \frac{\|x_0 + hx\|^p - \|x_0\|^p}{ph}.$$

Taking limits as $h \to 0^+$ and as $h \to 0^-$ and using (7.1) gives that

$$f(x) = \frac{f(x_0 + hx)^p - f(x_0)^p}{ph} \leq \frac{\|x_0 + hx\|^p - \|x_0\|^p}{ph}.$$

For the final statement, note that uniqueness of norming functionals for a vector is invariant under non-zero scalar multiplication.[33] □

We will see on Problem Sheet 3 that uniqueness of norming functionals in fact characterises smooth spaces.[34] Moreover, if X is a real normed vector space such that the norm of X^* is strictly convex then the norm of X is smooth, and similarly if the norm of X^* is smooth then the norm of X is strictly convex.

8 Lebesgue spaces

Let (Ω, Σ, μ) be a measure space. For $1 \leq p \leq \infty$ we denote by $L^p(\Omega, \Sigma, \mu)$, or simply $L^p(\Omega)$, the vector space of all (equivalence classes of) μ-measurable functions $x \colon \Omega \to \mathbb{F}$ such that

$$\int_\Omega |x(t)|^p \, \mathrm{d}\mu(t) < \infty$$

---

[32]For this to come out nicely we need $f(x_0) = \|x_0\| = 1$ — which I made a hash of by forgetting to include in the hypothesis in lectures. Sorry.

[33]I retain the direct proof for the case when X is smooth (it's basically the same as above). Given any norming functional $f \in S_{X^*}$ and any $x \in X$ we have

$$hf(x) = f(x_0 + hx) - \|x_0\| \leq \|x_0 + hx\| - \|x_0\|, \quad h \in \mathbb{R}.$$

Dividing by $h \neq 0$ and taking the limit as $h \to 0\pm$ now gives the result.

[34]so given the $p > 1$ version of Theorem 7.3, in fact the norm is smooth.

if $1 \le p < \infty$ and $\operatorname{ess\,sup}_{t \in \Omega} |x(t)| < \infty$ if $p = \infty$. If we endow $L^p(\Omega)$ with the norm

$$\|x\|_p = \left(\int_\Omega |x(t)|^p \, \mathrm{d}\mu(t) \right)^{1/p}$$

for $1 \le p < \infty$ and $\|x\|_p = \operatorname{ess\,sup}_{t \in \Omega} |x(t)|$ if $p = \infty$, then $L^p(\Omega)$ is a Banach space. For $p = 2$ it is a Hilbert space. If $1 \le p \le \infty$ is given we will always take q be the Hölder conjugate of p. We wish to show that for $1 \le p < \infty$ the dual of $L^p(\Omega)$ is isometrically isomorphic to $L^q(\Omega)$. We already know this for $p = 2$, and also when $\Omega = \mathbb{N}$, Σ is the power set of $\mathbb{N}$ and μ is the counting measure, since in this case $L^p(\Omega) = \ell^p$, $1 \le p \le \infty$. A measure space (Ω, Σ, μ) is said to be σ-*finite* if there exists a countable family $\{\Omega_n \in \Sigma : n \ge 1\}$ of measurable subsets of Ω such that $\mu(\Omega_n) < \infty$ for all $n \ge 1$ and $\Omega = \bigcup_{n \ge 1} \Omega_n$. Note that by taking complements and intersections we may assume that the sets Ω_n, $n \ge 1$, are mutually disjoint.

Proposition 8.1. *Let $1 \le p \le \infty$ and, if $p = 1$, assume that the measure space (Ω, Σ, μ) is σ-finite. Then the map $\Phi_p \colon L^q(\Omega) \to L^p(\Omega)^*$ given by*

$$(\Phi_p y)(x) = \int_\Omega x(t) y(t) \, \mathrm{d}\mu(t), \quad x \in L^p(\Omega),\ y \in L^q(\Omega),$$

is a well-defined linear isometry.

PROOF. It follows from Hölder's inequality that the integral defining Φ_p exists and that $|(\Phi_p y)(x)| \le \|x\|_p \|y\|_q$ for $x \in L^p(\Omega)$, $y \in L^q(\Omega)$. Since $\Phi_p y$ is linear and bounded with $\|\Phi_p y\| \le \|y\|_q$ for all $y \in L^q(\Omega)$, we see that Φ_p indeed maps into $L^p(\Omega)^*$. It is clear that Φ_p is a bounded linear operator with $\|\Phi_p\| \le 1$. We need to show that Φ_p is an isometry. Suppose first that $1 < p \le \infty$ and let $y \in L^q(\Omega)$ be such that $\|y\|_q = 1$. If $x(t) = |y(t)|^{q-2}\overline{y(t)}$ for $t \in \Omega$ such that $y(t) \ne 0$ and $x(t) = 0$ otherwise, then $x \in L^p(\Omega)$ with $\|x\|_p = 1$ and $(\Phi_p y)(x) = \|y\|_q^q = 1$, so Φ_p is an isometry. Suppose that $p = 1$ and let $y \in L^\infty(\Omega)$ with $\|y\|_\infty = 1$. Then for every $r \in [0, 1)$ the set $\Omega_r' = \{t \in \Omega : |y(t)| \ge r\}$ lies in Σ and is non-null. By σ-finiteness there exists a measurable subset $\Omega_r \subseteq \Omega_r'$ such that $0 < \mu(\Omega_r) < \infty$. Let

$$x_r(t) = \frac{\overline{y(t)}}{|y(t)|} \mu(\Omega_r)^{-1} \, {}_{\Omega_r}(t), \quad t \in \Omega.$$

Then $x_r \in L^1(\Omega)$ with $\|x_r\|_1 = 1$ and $(\Phi_1 y)(x_r) \ge r$, so Φ_1 too is an isometry. $\square$

It remains to show that the operator Φ_p is surjective when $1 \le p < \infty$ (this is never the case for $p = \infty$ except when the space is finite-dimensional). As in Proposition 8.1 the case $p = 1$ requires σ-finiteness of the measure space; see Problem Sheet 3. We restrict ourselves here to the case $1 < p < \infty$.

Recall that a function $\phi : \mathbb{R} \to \mathbb{R}$ is *convex* if

$$\phi(\lambda x + (1 - \lambda)y) \le \lambda \phi(x) + (1 - \lambda)\phi(y),$$

for $0 < \lambda < 1$ and $x, y \in \mathbb{R}$.[35] It is strictly convex if equality only holds when $x = y$. Recall too that differentiable convex functions have increasing derivative. For use

---

[35]This says that the area above the curve of ϕ is convex.

in Theorem 8.3, if $\phi : \mathbb{R} \to \mathbb{R}$ is convex and differentiable, then the function

$$\psi(s) = \begin{cases} \frac{\phi(s)-\phi(0)}{s}, & s \neq 0 \\ \phi'(0), & s = 0 \end{cases}$$

is continuous[36] and increasing. The example that matters to us is the function $\phi(x) = |a + bx|^p$ for $a, b \in \mathbb{R}$ and $p > 1$. This is strictly convex (for $b \neq 0$) and has $\phi'(x) = pb|a + bx|^{p-1} \operatorname{sgn}(a + bx)$.[37]

Theorem 8.2. *Suppose that $1 < p < \infty$. Then the space $L^p(\Omega)$ is uniformly convex.*

PROOF. Let (x_n) and (y_n) be sequences in $L^p(\Omega)$ such that $\|x_n\|_p, \|y_n\|_p, \frac{1}{2}\|x_n + y_n\|_p \to 1$ as $n \to \infty$. Let $u_n = \frac{1}{2}|x_n + y_n|$ and $v_n = \frac{1}{2}|x_n - y_n|$, $n \geq 1$. Then $\|u_n\|_p \to 1$ as $n \to \infty$ and we need to show that $\|v_n\|_p \to 0$ as $n \to \infty$. Now

$$\|u_n + v_n\|_p^p + \|u_n - v_n\|_p^p = \int_\Omega \frac{||x_n + y_n| + |x_n - y_n||^p}{2^p} + \frac{||x_n + y_n| - |x_n - y_n||^p}{2^p} \, d\mu$$

$$= \int_\Omega \max\{|x_n|, |y_n|\}^p + \min\{|x_n|, |y_n|\}^p \, d\mu$$

$$= \|x_n\|_p^p + \|y_n\|_p^p \to 2, \quad n \to \infty.$$

By passing to subsequences if necessary we may assume that $\|u_n + v_n\|_p \to a$ and $\|u_n - v_n\|_p \to b$ as $n \to \infty$. Then $a^p + b^p = 2$ and

$$2 = 2 \lim_{n \to \infty} \|u_n\|_p \leq \lim_{n \to \infty} \left(\|u_n + v_n\|_p + \|u_n - v_n\|_p \right) = a + b,$$

and hence

$$\frac{a^p + b^p}{2} \leq \left(\frac{a+b}{2} \right)^p.$$

By strict convexity of the function $t \mapsto t^p$ it follows that $a = b = 1$. Since

$$\int_\Omega \frac{||x_n + y_n| + |x_n - y_n||^p}{2^p} \, d\mu \geq \int_\Omega \frac{|x_n + y_n|^p + |x_n - y_n|^p}{2^p} \, d\mu,$$

which is to say $\|u_n + v_n\|_p^p \geq \|u_n\|_p^p + \|v_n\|_p^p$, we obtain that

$$0 \leq \|v_n\|_p^p \leq \|u_n + v_n\|_p^p - \|u_n\|_p^p \to 0, \quad n \to \infty.$$

Thus by a standard subsequence argument the space $L^p(\Omega)$ is uniformly convex. $\square$

From now on we will consider only the case $\mathbb{F} = \mathbb{R}$. However, the main result, Theorem 8.4 below, remains true in the complex case and indeed it is possible to deduce the complex case from the real case by decomposing into real and imaginary parts not only the functionals in question but also the functions themselves.

Theorem 8.3. *Suppose that $1 < p < \infty$. If $x_0 \in L^p(\Omega)$ with $\|x_0\|_p = 1$ then $f = \Phi_p(y_0)$, where $y_0 = |x_0|^{p-1} \operatorname{sgn} x_0$, is the unique element $f \in L^p(\Omega)^*$ with $f(x_0) = 1$ and $\|f\| = 1$.*

---

[36]by differentiability
[37]Here, for $h \in \mathbb{R}$ we let $\operatorname{sgn} h = \pm 1$ according as $h \gtrless 0$, and we let $\operatorname{sgn} 0 = 0$.

PROOF. Certainly $y_0 \in L^q(\Omega)$. By Theorem 7.3 we need to show that

$$(\Phi_p y_0)(x) = \lim_{h \to 0} \frac{\|x_0 + hx\|_p^p - \|x_0\|_p^p}{ph}, \quad x \in L^p(\Omega).$$

For $t \in \mathbb{R}$, set

$$\phi_t(h) = \begin{cases} \frac{|x_0(t) - hx(t)|^p - |x_0(t)|^p}{ph}, & h \neq 0 \\ y_0(t)x(t), & h = 0. \end{cases}$$

As ϕ_t is increasing and continuous, the monotone convergence theorem gives[38]

$$\lim_{h \to 0} \frac{\|x_0 + hx\|_p^p - \|x_0\|_p^p}{ph} = \lim_{h \to 0} \int_\Omega \phi_t(h) d\mu(t) = \int_\Omega y_0(t)x(t) d\mu(t), \qquad (8.1)$$

as required. □

Theorem 8.4. *Suppose that $1 < p < \infty$. Then the map $\Phi_p \colon L^q(\Omega) \to L^p(\Omega)^*$ is an isometric isomorphism. Furthermore, the space $L^p(\Omega)$ is reflexive.*

PROOF. Let $f \in L^p(\Omega)^* \setminus \{0\}$. By Theorem 8.2 the space $L^p(\Omega)$ is uniformly convex, and hence by Theorem 7.1 there exists a (unique) function $x_0 \in L^p(\Omega)$ such that $\|x_0\|_p = 1$ and $f(x_0) = \|f\|$. It follows from Theorem 8.3 that $f \in \operatorname{Ran} \Phi_p$, so Φ_p is surjective and the first part follows from Proposition 8.1. In particular the map $(\Phi_p^{-1})^* \circ \Phi_q \colon L^p(\Omega) \to L^p(\Omega)^{**}$ is an isomorphism. If $x \in L^p(\Omega)$ and $f \in L^p(\Omega)^*$, then $f = \Phi_p(y)$ for some $y \in L^q(\Omega)$ and hence

$$\big((\Phi_p^{-1})^*(\Phi_q x)\big)(f) = (\Phi_q x)(y) = \int_\Omega xy \, d\mu = (\Phi_p y)(x) = f(x) = (J_{L^p(\Omega)} x)(f).$$

Thus $J_{L^p(\Omega)} = (\Phi_p^{-1})^* \circ \Phi_q$, so $J_{L^p(\Omega)}$ is surjective and hence $L^p(\Omega)$ is reflexive. □

9 The weak and weak* topologies

Given a vector space X and a subspace Y of the algebraic dual X', we denote by $\sigma(X, Y)$ the coarsest topology on X for which all of the functionals $f \in Y$ are continuous. An equivalent definition of $\sigma(X, Y)$ is that it is the topology generated by the basic open neighbourhoods

$$\big\{ x \in X : |f_k(x - x_0)| < \varepsilon \text{ for } 1 \le k \le n \big\},$$

where $x_0 \in X$, $n \in \mathbb{N}$, $f_1, \ldots, f_n \in Y$ and $\varepsilon > 0$. If X is a normed vector space, the case $Y = X^*$ is of particular interest, and we call $\sigma(X, X^*)$ the *weak topology* on X. Another important special case is the so-called *weak* topology* $\sigma(X^*, J_X(X))$ on the dual space X^*, which we also denote simply by $\sigma(X^*, X)$. Note that a basic weak*-open neighbourhood has the form

$$\big\{ f \in X^* : |f(x_k) - f_0(x_k)| < \varepsilon \text{ for } 1 \le k \le n \big\},$$

where $f_0 \in X^*$, $n \in \mathbb{N}$, $x_1, \ldots, x_n \in X$ and $\varepsilon > 0$. It is straightforward to verify that both the weak and the weak* topologies are Hausdorff. We certainly have

---

[38] consider separately the limit as $h \to 0^+$ and $h \to 0^-$.

$\sigma(X^*, X) \subseteq \sigma(X^*, X^{**})$. Moreover, since any $f \in X^*$ is norm-continuous on X, and likewise any element of $J_X(X)$ on X^*, the weak and weak* topologies are no finer than the norm topologies on X and X^*, respectively. On finite-dimensional spaces they coincide but, as we shall see in Proposition 9.6 below, in the infinite-dimensional setting the weak and weak* topologies are always strictly coarser than the norm topology. In fact, one of the main reasons for introducing the weak and weak* topologies is that, unlike the norm topology, these coarser topologies give us a rich supply of compact sets even in infinite-dimensional spaces.

Let X be a normed vector space. Given a sequence (x_n) in X it is straightforward to see that the sequence converges in the weak topology to a limit $x \in X$ if and only if $f(x_n) \to f(x)$ as $n \to \infty$ for all $f \in X^*$, which is precisely the notion of weak convergence you may already know in the Hilbert space setting. Similarly, a sequence (f_n) in X^* converges in the weak* topology to a limit $f \in X^*$ if and only if $f_n(x) \to f(x)$ as $n \to \infty$ for all $x \in X$. It follows from the Uniform Boundedness Principle that any weakly convergent sequence is norm-bounded, and if X is complete the same is true of any weak*-convergent sequence in X^*. We say that X has the *Schur property* if every weakly convergent sequence in X is norm-convergent. If $X = c_0$ or $X = \ell^p$ for $1 < p < \infty$ then we see by considering the sequence (e_n) that X does not have the Schur property. The case $p = 1$ is different.

Theorem 9.1. *The space ℓ^1 has the Schur property.*

PROOF. This follows from a 'gliding hump' argument; see Problem Sheet 3. □

Recall that the topology of any metric space can be described in terms of sequences. In general the weak and weak* topologies are not metrisable (see Problem Sheet 3) but we have the following result.

Proposition 9.2. *Let X be a normed vector space.*

(a) *If X is separable then the relative weak* topology on B_{X^*} is metrisable.*
(b) *If X^* is separable then the relative weak topology on B_X is metrisable.*

PROOF. For part (a) let $\{x_n : n \geq 1\}$ be a dense subset of B_X and consider the map

$$d(f, g) = \sum_{n=1}^{\infty} \frac{|f(x_n) - g(x_n)|}{2^n}, \quad f, g \in B_{X^*}.$$

The result in part (b) can be approached analogously or deduced from part (a). The details are left as an exercise; see Problem Sheet 3. □

Remark 9.3. The converse statements in Proposition 9.2 are also true; see Remark 9.10 below for a proof in the case of part (a).

Given a linear map $T \colon X \to Y$ between two vector spaces X and Y, the *algebraic dual operator* $T' \colon Y' \to X'$ of T is given by $(T'f)(x) = f(Tx)$ for $f \in Y'$ and $x \in X$. In particular, if X, Y are normed vector spaces and $f \in Y^*$ then $T'f = T^*f$.

Proposition 9.4. *Let X and Z be vector spaces and suppose that Y is a subspace of X' and W is a subspace of Z'.*

(a) *A functional $f \in X'$ is $\sigma(X, Y)$-continuous if and only if $f \in Y$.*

(b) *A linear map $T\colon X \to Z$ is $\sigma(X,Y)$-to-$\sigma(Z,W)$-continuous if and only if $T'(W) \subseteq Y$.*

PROOF. (a) If $f \in Y$ then f is $\sigma(X,Y)$-continuous by definition of $\sigma(X,Y)$. Suppose that $f \in X'$ is $\sigma(X,Y)$-continuous and let $U = \{x \in X : |f(x)| < 1\}$. Then U is a $\sigma(X,Y)$-open neighbourhood of zero so there exist $n \in \mathbb{N}$, $f_1, \ldots, f_n \in Y$ and $\varepsilon > 0$ such that the basic $\sigma(X,Y)$-open set $V = \{x \in X : |f_k(x)| < \varepsilon \text{ for } 1 \le k \le n\}$ is contained in U. Since $\bigcap_{k=1}^{n} \operatorname{Ker} f_k \subseteq V \subseteq U$ we have by linearity that

$$\bigcap_{k=1}^{n} \operatorname{Ker} f_k \subseteq \operatorname{Ker} f,$$

so $f \in \operatorname{span}\{f_1, \ldots, f_n\} \subseteq Y$ by a result on Problem Sheet 1.

(b) It follows from the definition of $\sigma(Z,W)$ that a linear map $T\colon X \to Z$ is $\sigma(X,Y)$-to-$\sigma(Z,W)$-continuous if and only if $f \circ T\colon X \to \mathbb{F}$ is $\sigma(X,Y)$-continuous for all $f \in W$, which by part (a) is equivalent to having $T'(W) \subseteq Y$, as required. $\quad\square$

Corollary 9.5. *Let X be a normed vector space. Then $\sigma(X^*, X) = \sigma(X^*, X^{**})$ if and only if X is reflexive.*

Recall that a subspace Y of a vector space X is said to have *finite codimension* if $\dim X/Y < \infty$. Typical examples of finite-codimensional subspaces are annihilators of finite-dimensional subspaces. Indeed, if Y is a finite-dimensional subspace of a normed vector space X then by Corollary 5.12 we have $X^*/Y^\circ \cong Y^*$ and hence $\dim X^*/Y^\circ < \infty$. Similarly, if Z is a finite-dimensional subspace of X^*, then by a result on Problem Sheet 1 we have $(Z_\circ)^\circ = Z$ and Corollary 5.12 gives $(X/Z_\circ)^* \cong Z$. In particular, $X/Z_\circ$ must be finite-dimensional. The proof of Proposition 9.4 shows that any basic weakly open neighbourhood of the origin contains a finite-codimensional subspace, and an analogous argument works for the weak* topology.

Proposition 9.6. *Let X be an infinite-dimensional normed vector space. Then S_X is not weakly closed and S_{X^*} is not weak*-closed. In particular, $\sigma(X, X^*)$ and $\sigma(X^*, X)$ are strictly coarser than the respective norm topologies.*

PROOF. We will show that 0 lies in the weak closure of S_X. Indeed, as in the proof of Proposition 9.4 we see that any weakly open neighbourhood U of 0 contains a finite-codimensional subspace Y of the form $Y = \bigcap_{k=1}^{n} \operatorname{Ker} f_k$ for some $f_1, \ldots, f_n \in X^*$. Since X is infinite-dimensional but $\dim X/Y < \infty$, Y must be non-trivial. In particular, $Y \cap S_X = S_Y$ is non-empty and hence so is $U \cap S_X$, as required. The argument for S_{X^*} is completely analogous, and the final statement is then clear. $\quad\square$

Given a family $\{X_\alpha : \alpha \in A\}$ of topological spaces we may view the *product space* $X = \prod_{\alpha \in A} X_\alpha$ as the space of all functions $x\colon A \to \bigcup_{\alpha \in A} X_\alpha$ such that $x(\alpha) \in X_\alpha$ for all $\alpha \in A$. We may endow X with the *product topology*, which is the coarsest topology for which all of the maps $p_\alpha\colon X \to X_\alpha$, $\alpha \in A$, are continuous, where $p_\alpha(x) = x(\alpha)$ for $\alpha \in A, x \in X$. It is not hard to see that if X_α is Hausdorff for each $\alpha \in A$ then so is the product space with the product topology. The corresponding statement for compactness is far less obvious. The following result is proved in C1.3 Analytic Topology; it is equivalent to the Axiom of Choice.

Theorem 9.7 (Tychonoff). *Let $\{X_\alpha : \alpha \in A\}$ be a family of topological spaces and suppose that X_α is compact for each $\alpha \in A$. Then the product space endowed with the product topology is also compact.*

Theorem 9.8 (Banach-Alaoglu). *Let X be a normed vector space. Then B_{X^*} is weak*-compact.*

PROOF. For $x \in X$ let $D_x = \{\lambda \in \mathbb{F} : |\lambda| \le \|x\|\}$, noting that each D_x is compact, and let $K = \prod_{x \in X} D_x$ be endowed with the product topology. Then $B_{X^*} \subseteq K$ and the topology on B_{X^*} induced by the product topology is precisely the weak* topology. Note that K is compact in the product topology by Tychonoff's Theorem. Hence in order to prove that B_{X^*} is weak*-compact it suffices to show that it is closed in K. Now K is simply the set of all functions $f \colon X \to \mathbb{F}$ such that $|f(x)| \le \|x\|$ for all $x \in X$, and B_{X^*} consists precisely of those elements of K which are linear. For $x, y \in X$ and $\lambda \in \mathbb{F}$ let $\Phi_{x,y,\lambda} \colon K \to \mathbb{F}$ be the map defined by

$$\Phi_{x,y,\lambda}(f) = f(x + \lambda y) - f(x) - \lambda f(y), \quad f \in K,$$

and note that $B_{X^*} = \bigcap \{\Phi_{x,y,\lambda}^{-1}(\{0\}) : x, y \in X, \ \lambda \in \mathbb{F}\}$. It follows from the definition of the product topology that the map $\Phi_{x,y,\lambda}$ is continuous on K for every $x, y \in X$ and $\lambda \in \mathbb{F}$, so B_{X^*} is closed in K, as required. $\qquad\square$

Given a compact topological space Ω we write $C(\Omega)$ for the Banach space of scalar-valued continuous functions, endowed with the supremum norm.

Corollary 9.9. *Given any normed vector space X there exists a compact Hausdorff space Ω such that $C(\Omega)$ contains a subspace which is isometrically isomorphic to X.*

Proof. We may take $\Omega = B_{X^*}$ with the subspace topology induced by the weak* topology on X^*. Then Ω is Hausdorff and, by the Banach-Alaoglu Theorem, it is compact. Moreover, the map $T \colon X \to C(\Omega)$ given by $(Tx)(f) = f(x)$ for $x \in X$, $f \in \Omega$, is a well-defined linear operator, and by Corollary 5.8 it is an isometry. $\quad\square$

Remark 9.10. By Proposition 9.2(a) we may take Ω to be a compact *metric* space if X is separable, and in fact, by the so-called *Banach-Mazur Theorem*, we may even take $\Omega = [0, 1]$ in this case. Since $C(\Omega)$ is separable whenever Ω is a compact metric space we see that Corollary 9.9 implies the converse of Proposition 9.2(a).

Theorem 9.11 (Hahn-Banach Separation Theorem, weak* version). *Let X be a normed vector space and suppose that C is a non-empty convex subset of X^* and that $f_0 \in X^* \setminus C$.*

 (a) *If C is weak*-open, then there exists $x \in X$ such that*

$$\mathrm{Re}\, f_0(x) > \mathrm{Re}\, f(x), \quad f \in C.$$

 (b) *If C is weak*-closed, then there exists $x \in X$ such that*

$$\mathrm{Re}\, f_0(x) > \sup\{\mathrm{Re}\, f(x) : f \in C\}.$$

PROOF. (a) Since C is in particular open in the norm topology, Theorem 5.9 implies the existence of a $\phi \in X^{**}$ such that $\mathrm{Re}\,\phi(f_0) > \mathrm{Re}\,\phi(f)$ for all $f \in C$. Let $g_0 \in C$. Then there exists a basic weak*-open neighbourhood U of zero such that $g_0 + U \subseteq C$. Suppose that
$$U = \big\{ f \in X^* : |f(x_k)| < \varepsilon \text{ for } 1 \le k \le n \big\}$$
for some $n \in \mathbb{N}$, $x_1, \ldots, x_n \in X$ and $\varepsilon > 0$. Then in particular
$$\mathrm{Re}\,\phi(f_0) > \mathrm{Re}\,\phi(g_0) + \mathrm{Re}\,\phi(f), \quad f \in U.$$
If $Y = \bigcap_{k=1}^{n} \mathrm{Ker}\, J_X(x_k)$, then $Y \subseteq U$ and because Y is a vector space we must have $Y \subseteq \mathrm{Ker}\,\phi$. It follows that $\phi \in \mathrm{span}\{J_X(x_1), \ldots, J_X(x_n)\} \subseteq J_X(X)$, as required.

(b) Since C is in particular closed in the norm topology, Theorem 5.9 implies the existence of a $\phi \in X^{**}$ such that $\mathrm{Re}\,\phi(f_0) > \sup_{f \in C} \mathrm{Re}\,\phi(f)$. Choosing U to be a basic weak*-open neighbourhood of zero such that $f_0 \notin C + U$, it follows as in part (a) that $\phi \in J_X(X)$. $\qquad\square$

Corollary 9.12. *Let X be a normed vector space and let Y be a subspace of X^*. Then $(Y_\circ)^\circ$ coincides with the weak* closure of Y.*

PROOF. Let C denote the weak* closure of Y, noting that C is a vector space and hence convex. It is clear that $Y \subseteq (Y_\circ)^\circ$ and that
$$(Y_\circ)^\circ = \bigcap \big\{ \mathrm{Ker}\, J_X(x) : x \in Y_\circ \big\}$$
is weak*-closed, so $C \subseteq (Y_\circ)^\circ$. Suppose that $f_0 \in X^* \setminus C$. By Theorem 9.11 there exists $x \in X$ such that $\mathrm{Re}\, f_0(x) > \sup_{f \in C} \mathrm{Re}\, f(x)$. Since C is a vector space this in particular implies that $f(x) = 0$ for all $f \in Y$ and hence $x \in Y_\circ$. Since $f_0(x) \ne 0$ we deduce that $f_0 \notin (Y_\circ)^\circ$, so $(Y_\circ)^\circ \subseteq C$, as required. $\qquad\square$

Theorem 9.13 (Goldstine). *Let X be a normed vector space. Then $J_X(B_X)$ is weak*-dense in $B_{X^{**}}$.*

PROOF. Let C be the weak* closure of $J_X(B_X)$, noting that C is convex. By the Banach-Alaoglu Theorem $B_{X^{**}}$ is weak*-compact. Since the weak* topology is Hausdorff the set $B_{X^{**}}$ is weak*-closed, and hence $C \subseteq B_{X^{**}}$. Suppose that $\phi \in X^{**} \setminus C$. By Theorem 9.11 there exists $f \in X^*$ such that
$$\mathrm{Re}\,\phi(f) > \sup\{\mathrm{Re}\,\psi(f) : \psi \in C\} \ge \sup\{\mathrm{Re}\, f(x) : x \in B_X\} = \|\mathrm{Re}\, f\| = \|f\|.$$
It follows that $\|\phi\| > 1$, so $B_{X^{**}} \subseteq C$ and the result follows. $\qquad\square$

Theorem 9.14. *Let X be a normed vector space. Then X is reflexive if and only if B_X is weakly compact.*

PROOF. Consider the usual map $J_X \colon X \to X^{**}$ given by $(J_X x)(f) = f(x)$ for $f \in X^*$, $x \in X$, and let $Y = J_X(X)$. If U is a basic weakly open subset of X then $J_X(U) = V \cap Y$ for a basic weak*-open subset V of X^{**}, while if V is a basic weak*-open subset of X^{**} then $J_X^{-1}(V \cap Y)$ is a basic weakly open subset of X. It follows that J_X is a homeomorphic embedding from X with the weak topology onto Y with the subspace topology induced by the weak* topology on X^{**}.

Suppose first that X is reflexive. Then $J_X(B_X) = B_{X^{**}}$ and $B_X = J_X^{-1}(B_{X^{**}})$. Since $B_{X^{**}}$ is weak*-compact by the Banach-Alaoglu Theorem, we see that B_X is weakly compact. Conversely, if B_X is weakly compact then $J_X(B_X)$ is weak*-compact and hence weak*-closed in X^{**}, because the weak* topology is Hausdorff. By Goldstine's Theorem $J_X(B_X)$ is weak*-dense in $B_{X^{**}}$ and hence $J_X(B_X) = B_{X^{**}}$, so J_X is surjective and consequently X is reflexive. $\qquad\square$

10 Compactness in normed vector spaces

Let (X, d) be a metric space. Given a subset M of X we define the *diameter* of M as $\operatorname{diam} M = \sup\{d(x, y) : x, y \in M\}$. We say that a subset M of X is *totally bounded* (or *precompact*) if for every $\varepsilon > 0$ there exists a finite cover of M by sets of diameter at most ε. This is equivalent to the existence, for every $\varepsilon > 0$, of a finite set $F \subseteq M$ such that $\operatorname{dist}(x, F) < \varepsilon$ for all $x \in M$, which is to say that $M \subseteq \bigcup_{x \in F} B_X^\circ(x, \varepsilon)$. Such a set F is said to be an *ε-net* for M. The set M is said to be *relatively compact* if its closure in X is compact. Any relatively compact subset is totally bounded and it is easy to see that M is totally bounded or relatively compact if and only if its closure is. We say that X is *sequentially compact* if every sequence with terms in X has a subsequence which converges to an element of X.

Theorem 10.1. *Let (X, d) be a metric space. Then the following are equivalent:*

(a) *X is compact;*

(b) *X is complete and totally bounded;*

(c) *X is sequentially compact.*

PROOF. It is a standard fact in point-set topology that a metric space is compact if and only if it is sequentially compact and that any compact space is complete. Moreover, any compact space is totally bounded. A standard diagonal selection argument shows that in a totally bounded set any sequence has a Cauchy subsequence, and in fact the converse holds as well. Hence if a space is totally bounded and complete then it must be sequentially compact. $\qquad\square$

Remark 10.2. Suppose that X is a normed vector space such that X^* is separable. Then B_X with the relative weak topology is a metric space by Proposition 9.2, so by Theorem 10.1 it is compact if and only if it is sequentially compact. It follows from Theorem 9.14 that X is reflexive if and only if every bounded sequence has a weakly convergent subsequence. In fact, this statement is true even without the assumption that X^* is separable, and this follows from the so-called *Eberlein-Šmulian Theorem*, which states that weak compactness is equivalent to weak sequential compactness for subsets of arbitrary normed vector spaces. The Eberlein-Šmulian Theorem is not part of this course, but note that the proof of one implication relies a diagonal selection argument similar to the one commonly used to show that every bounded sequence in a Hilbert space has a weakly convergent subsequence.

Theorem 10.1 has the following important consequence.

Corollary 10.3. *Let (X, d) be a complete metric space and let $M \subseteq X$. Then M is relatively compact if and only if it is totally bounded.*

Suppose now that X is a normed vector space. If X is finite-dimensional and $M \subseteq X$ then by the Heine-Borel Theorem M is compact if and only if it is closed and bounded, and hence M is totally bounded if and only if it is bounded.

Proposition 10.4. *Let X be a normed vector space and let $M \subseteq X$. Then M is totally bounded if and only if M is bounded and for every $\varepsilon > 0$ there exists a finite-dimensional subspace Y of X such that $\mathrm{dist}(x, Y) < \varepsilon$ for all $x \in M$.*

PROOF. Suppose first that M is totally bounded. Then M is bounded and, given $\varepsilon > 0$, we may let $F \subseteq M$ be a finite ε-net for M. Letting $Y = \mathrm{span}\, F$ we have $\dim Y \leq |F|$ and $\mathrm{dist}(x, Y) < \varepsilon$ for all $x \in M$, as required. Conversely, suppose that M is bounded and, given $\varepsilon > 0$, let Y be a finite-dimensional subspace Y of X such that $\mathrm{dist}(x, Y) < \varepsilon/2$ for all $x \in M$. Let $r > 0$ be such that $M \subseteq B_X(r)$ and note that the set $K = B_Y(r + \varepsilon/2)$ is compact by the Heine-Borel Theorem. Let F be a finite $\frac{\varepsilon}{2}$-net for K. Then given $x \in M$ we may find $y \in Y$ such that $\|x - y\| < \varepsilon/2$. Since $y \in K$ there exists $z \in F$ such that $\|y - z\| < \varepsilon/2$ and hence $\|x - z\| < \varepsilon$. Thus M is totally bounded. $\qquad\square$

Lemma 10.5 (F. Riesz). *Let X be a normed vector space and suppose that Y is a proper closed subspace of X. Then for any $\delta \in (0, 1)$ there exists $x \in S_X$ such that $\mathrm{dist}(x, Y) \geq 1 - \delta$.*

PROOF. By Corollary 5.8 there exists $f \in S_{X^*} \cap Y^\circ$. Given $\delta \in (0, 1)$ we may find $x \in S_X$ such that $|f(x)| > 1 - \delta$, and then for $y \in Y$ we have

$$\|x - y\| \geq |f(x - y)| = |f(x)| > 1 - \delta.$$

Thus $\mathrm{dist}(x, Y) \geq 1 - \delta$, as required. $\qquad\square$

Remark 10.6. The conclusion of Lemma 10.5 in general becomes false if we allow $\delta = 0$. However, using Theorem 7.1 we may show the result remains valid even with $\delta = 0$ if X is uniformly convex, and by a simple application of the Heine-Borel Theorem the same is true when Y is finite-dimensional.

Corollary 10.7. *Let X be a normed vector space. Then B_X is totally bounded if and only if X is finite-dimensional.*

PROOF. If B_X is totally bounded and $\varepsilon \in (0, 1)$ then by Proposition 10.4 there exists a finite-dimensional subspace Y of X such that $\mathrm{dist}(x, Y) < \varepsilon$ for all $x \in B_X$. If $Y \subsetneq X$ then by Riesz's Lemma applied with $\delta = 1 - \varepsilon$ there exists $x \in B_X$ such that $\mathrm{dist}(x, Y) \geq \varepsilon$, a contradiction. Hence $Y = X$, so X is finite-dimensional. Conversely, if X is finite-dimensional then by the Heine-Borel Theorem B_X is compact and in particular totally bounded. $\qquad\square$

We now turn to characterisations of totally bounded subsets, or equivalently of relatively compact subsets, in two classical Banach spaces. Given a compact topological space Ω and a subset M of $C(\Omega)$, we say that M is *equicontinuous* if for every $\varepsilon > 0$ and $t \in \Omega$ there exists an open neighbourhood U of t in Ω such that $|x(s) - x(t)| < \varepsilon$ for all $x \in M$ and all $s \in U$.

Theorem 10.8 (Arzelà-Ascoli). *Let Ω be a compact topological space and suppose that M is a subset of $C(\Omega)$. Then M is relatively compact if and only if it is bounded and equicontinuous.*

PROOF. Suppose that M is relatively compact. Then M is totally bounded, in particular bounded, and given $\varepsilon > 0$ we may find a finite $\frac{\varepsilon}{3}$-net $F \subseteq M$ for M. Suppose that $t \in \Omega$. For each $x \in F$ there exists an open neighbourhood U_x of t in Ω such that $|x(s) - x(t)| < \varepsilon/3$ for all $s \in U_x$. Let $U = \bigcap_{x \in F} U_x$, which is another open neighbourhood of t in Ω. Given $x \in M$ we may find $y \in F$ such that $\|x - y\|_\infty < \varepsilon/3$ and, for $s \in U$, we have

$$|x(s) - x(t)| \leq |x(s) - y(s)| + |y(s) - y(t)| + |y(t) - x(t)| < \varepsilon.$$

Thus M is equicontinuous.

Conversely, suppose that M is bounded and equicontinuous, and let $\varepsilon > 0$. Since M is equicontinuous there exists, for each $t \in \Omega$, an open neighbourhood U_t of t in Ω such that $|x(s) - x(t)| < \varepsilon/3$ for all $x \in M$ and all $s \in U_t$. Then $\{U_t : t \in \Omega\}$ is an open cover of Ω. By compactness of Ω we may select a finite subcover $\{U_{t_k} : 1 \leq k \leq n\}$. Let $T\colon C(\Omega) \to (\mathbb{F}^n, \|\cdot\|_\infty)$ be given by

$$Tx = \big(x(t_1), \ldots, x(t_n)\big), \quad x \in C(\Omega),$$

and let $S = T(M)$. Then by boundedness of M the set S is a bounded, and therefore totally bounded, subset of $(\mathbb{F}^n, \|\cdot\|_\infty)$. Let $F \subseteq M$ be a finite set such that $T(F)$ is an $\frac{\varepsilon}{3}$-net for S and suppose that $x \in M$. Then there exists $y \in F$ such that $|x(t_k) - y(t_k)| < \varepsilon/3$ for $1 \leq k \leq n$. Given $t \in \Omega$ we have $t \in U_{t_k}$ for some $k \in \{1, \ldots, n\}$ and hence

$$|x(t) - y(t)| \leq |x(t) - x(t_k)| + |x(t_k) - y(t_k)| + |y(t_k) - y(t)| < \varepsilon.$$

Thus M is totally bounded and hence, by Corollary 10.3, relatively compact. $\square$

Example 10.9. Given $k \in L^1(0,1)$, consider the the set

$$M = \left\{ t \mapsto \int_0^t k(s) x(s) \, \mathrm{d}s : x \in C([0,1]), \ \|x\|_\infty \leq 1 \right\}.$$

It can be shown that for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\int_I |k(s)| \, \mathrm{d}s < \varepsilon$ for all intervals $I \subseteq [0,1]$ of length less than δ. Hence M is a bounded equicontinuous subset of $C([0,1])$. By the Arzelà-Ascoli Theorem M is relatively compact in $C([0,1])$.

Theorem 10.10 (Kolmogorov-Riesz-Fréchet). *Let $1 \leq p < \infty$ and $n \geq 1$, and suppose that $M \subseteq L^p(\mathbb{R}^n)$. Then M is relatively compact if and only if M is bounded,*

$$\int_{|t| \geq R} |x(t)|^p \, \mathrm{d}t \to 0 \ \text{as } R \to \infty \quad \text{and} \quad \int_{\mathbb{R}^n} |x(s+t) - x(t)|^p \, \mathrm{d}t \to 0 \ \text{as } |s| \to 0,$$

uniformly over $x \in M$.

PROOF. The proof uses the Arzelà-Ascoli Theorem together with a mollification argument. We omit the details; see for instance Brezis, Section 4.5. $\square$

11 Compact operators

Given two normed vector spaces X, Y and a linear operator $T\colon X \to Y$, we say that T is a *compact operator* if the set $T(B_X)$ is relatively compact in Y. Thus T is compact if and only if for every bounded sequence (x_n) in X the sequence (Tx_n) in Y has a convergent subsequence. We write $\mathcal{K}(X,Y)$ for the set of compact linear operators $T\colon X \to Y$ and we let $\mathcal{K}(X) = \mathcal{K}(X,X)$. Notice that if T is compact then the closure of $T(B_X)$ is in particular bounded, so T is bounded and hence $\mathcal{K}(X,Y) \subseteq \mathcal{B}(X,Y)$.

Example 11.1. (a) If X, Y are normed vector spaces and $T \in \mathcal{B}(X,Y)$ has *finite rank*, which is to say that $\dim \operatorname{Ran} T < \infty$, then $T \in \mathcal{K}(X,Y)$. Indeed, $T(B_X)$ is a bounded and hence relatively compact subset of the finite-dimensional space $\operatorname{Ran} T$.

(b) If X is a normed vector space then by Corollary 10.7 the identity operator on X is compact if and only if X is finite-dimensional.

(c) Let $X = C([0,1])$ and $k \in L^1(0,1)$. Then by Example 10.9 the integral operator

$$(Tx)(t) = \int_0^t k(s)x(s)\, \mathrm{d}s, \quad x \in X, \ 0 \le t \le 1,$$

is compact. Integral operators are important in the theory of differential equations.

Proposition 11.2. *Let X and Y be normed vector spaces.*

(a) *The set $\mathcal{K}(X,Y)$ is a subspace of $\mathcal{B}(X,Y)$, and it is closed if Y is complete.*

(b) *If $T \in \mathcal{K}(X,Y)$ and $R \in \mathcal{B}(Y,Z)$, $S \in \mathcal{B}(W,X)$, where W and Z are normed vector spaces, then $RTS \in \mathcal{K}(W,Z)$.*

PROOF. (a) Suppose that $S, T \in \mathcal{K}(X,Y)$ and that $\lambda \in \mathbb{F}$. Let L and M denote the closures or $S(B_X)$ and $T(B_X)$, respectively. Then L and M are compact and so is the set $K = L + |\lambda| M$. Since $(S + \lambda T)(B_X) \subseteq K$, we see that $S + \lambda T \in \mathcal{K}(X,Y)$, so $\mathcal{K}(X,Y)$ is a subspace of $\mathcal{B}(X,Y)$. It remains to show that it is closed if Y is complete. Suppose that $T \in \mathcal{B}(X,Y)$ lies in the closure of $\mathcal{K}(X,Y)$, and let $\varepsilon > 0$. Then there exists $S \in \mathcal{K}(X,Y)$ such that $\|S - T\| < \varepsilon/3$. Since S is compact there exists a finite set $F \subseteq B_X$ such that $S(F)$ is an $\frac{\varepsilon}{3}$-net for $S(B_X)$. Given $x \in B_X$ we may find $y \in F$ such that $\|Sx - Sy\| < \varepsilon/3$, and hence

$$\|Tx - Ty\| \le \|T - S\|\|x\| + \|Sx - Sy\| + \|S - T\|\|y\| < \varepsilon.$$

It follows that $T(B_X)$ is totally bounded, so by completeness of Y and Corollary 10.3 it is relatively compact and hence $T \in \mathcal{K}(X,Y)$. Thus $\mathcal{K}(X,Y)$ is closed in $\mathcal{B}(X,Y)$.

(b) Let $r = \|S\|$ and $B = B_X(r)$, and note that $S(B_W) \subseteq B$. Letting K denote the closure of $T(B)$ we see that K is a compact set containing $TS(B_W)$, and hence $RTS(B_W)$ is contained in the compact set $R(K)$, so $RTS \in \mathcal{K}(X,Y)$. $\qquad\square$

Corollary 11.3. *Let X be a normed vector space and Y a Banach space, and let $T \in \mathcal{B}(X,Y)$. Suppose there exist finite-rank operators $T_n \in \mathcal{B}(X,Y)$, $n \ge 1$, such that $\|T_n - T\| \to 0$ as $n \to \infty$. Then $T \in \mathcal{K}(X,Y)$.*

Remark 11.4. P. Enflo (1973) gave an example of a separable reflexive Banach space X for which there exists a compact operator $T \in \mathcal{K}(X)$ such that T cannot be approximated by finite-rank operators. However, we will see on Problem Sheet 4 that the converse of Corollary 11.3 does hold on many spaces.

Lemma 11.5. *Let X be a normed vector space and let Y be a closed finite-codimensional subspace of X. Then for every $\varepsilon > 0$ there exists a finite-dimensional subspace Z of X such that every $x \in X$ can be expressed as $x = y + z$ with $y \in Y$, $z \in Z$ and $\|z\| \le (1 + \varepsilon)\|x\|$.*

PROOF. Let $0 < \delta \le \varepsilon(2 + \varepsilon)^{-1}$. Since $\dim X/Y < \infty$ the open unit ball $B^\circ_{X/Y}$ is totally bounded. Let $\pi\colon X \to X/Y$ be the canonical quotient map. Then $\pi(B^\circ_X) = B^\circ_{X/Y}$ and hence there exist $z_1, \dots, z_n \in B^\circ_X$ such that $\{\pi(z_k) : 1 \le k \le n\}$ is a δ-net for $B^\circ_{X/Y}$. Let $Z = \mathrm{span}\{z_k : 1 \le k \le n\}$. Then Z is complete and Lemma 4.9 applied to the restricted operator $\pi|_Z\colon Z \to X/Y$ shows that $B^\circ_{X/Y} \subseteq \pi(B^\circ_Z((1-\delta)^{-1}))$. Let $x \in X$. If $x = 0$ we may take $y = z = 0$, so we assume that $x \ne 0$. Then the vector $x_0 = \|x\|^{-1}(1 + \delta)^{-1}x$ satisfies $\|x_0\| < 1$, so there exists $z_0 \in B^\circ_Z((1-\delta)^{-1})$ such that $\pi(z_0) = \pi(x_0)$. Let $z = (1+\delta)\|x\|z_0$. Then $x = y + z$ for some $y \in Y$ and by our choice of δ we have $\|z\| < (1 + \varepsilon)\|x\|$. $\qquad\square$

Theorem 11.6. *Let X and Y be normed vector spaces and suppose that $T \in \mathcal{B}(X, Y)$. Consider the following statements:*

(a) *T is a compact operator;*

(b) *For every $\varepsilon > 0$ there exists a finite-dimensional subspace Z of Y such that $\|\pi_Z \circ T\| \le \varepsilon$, where $\pi_Z\colon Y \to Y/Z$ denotes the canonical quotient operator;*

(c) *For every $\varepsilon > 0$ there exists a closed finite-codimensional subspace W of X such that $\|T|_W\| \le \varepsilon$.*

Then (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c), and if Y is complete then (c) $\Longrightarrow$ (a).

PROOF. If T is compact, then $T(B_X)$ is totally bounded, so by Proposition 10.4 there exists, for every $\varepsilon > 0$, a finite-dimensional subspace Z of Y such that $\|\pi_Z(Tx)\| = \mathrm{dist}(Tx, Z) < \varepsilon$ for all $x \in B_X$, and hence $\|\pi_Z \circ T\| \le \varepsilon$, so (a) $\Longrightarrow$ (b).

If (b) holds and $\varepsilon > 0$, let Z be a finite-dimensional subspace of Y such that $\|\pi_Z \circ T\| \le \varepsilon/3$. Then Z° is closed and has finite codimension in Y^*, so by Lemma 11.5 there exists a finite-dimensional subspace V of Y^* such that every $f \in Y^*$ can be expressed as $f = g + h$ with $g \in Z^\circ$, $h \in V$ and $\|h\| \le 2\|f\|$. It follows that $\|g\| \le 3\|f\|$. Let $W = T^*(V)_\circ$, noting that W is closed and has finite codimension in X. Suppose that $x \in W$ and $f \in B_{Y^*}$. We may write $f = g + h$ as above, and then for any $z \in Z$ we have

$$|f(Tx)| = |T^*(g + h)(x)| = |g(Tx)| = |g(Tx + z)| \le 3\|Tx + z\|.$$

Taking the infimum over all $z \in Z$ we see using Corollary 5.8 that $\|Tx\| \le \varepsilon\|x\|$, and hence $\|T|_W\| \le \varepsilon$. Thus (b) $\Longrightarrow$ (c).

Finally, suppose that (c) holds and assume that Y is complete. Given $\varepsilon > 0$, let W be a closed finite-codimensional subspace of X such that $\|T|_W\| < \varepsilon/3$. By Lemma 11.5 there exists a finite-dimensional subspace V of X such that every $x \in X$ can be expressed as $x = v + w$ with $v \in V$, $w \in W$ and $\|v\| \le 2\|x\|$. It follows that

$\|w\| \leq 3\|x\|$. Let $Z = T(V)$, noting that Z is a finite-dimensional subspace of Y. Given $x \in B_X$, we may write $x = v + w$ as above, and then

$$\text{dist}(Tx, Z) \leq \|Tx - Tv\| \leq \|T|_W\|\|w\| < \varepsilon.$$

Hence the bounded set $T(B_X)$ is totally bounded by Proposition 10.4. By completeness of Y and Corollary 10.3 we see that T is a compact operator, so (c) $\Longrightarrow$ (a). $\square$

Theorem 11.7 (Schauder). *Let X and Y be normed vector spaces, and suppose that $T \in \mathcal{B}(X, Y)$. If T is compact then so is T^*, and if Y is complete then the converse also holds.*

PROOF. Suppose that T is compact and let $\varepsilon > 0$. By Theorem 11.6 there exists a finite-dimensional subspace Z of Y such that $\|\pi_Z \circ T\| \leq \varepsilon$. Then $W = Z^\circ$ is a closed finite-codimensional subspace of Y^*. Let $f \in W$, $x \in X$ and $z \in Z$. Then

$$|T^* f(x)| = |f(Tx + z)| \leq \|f\|\|Tx + z\|,$$

and taking the infimum over all $z \in Z$ we obtain that $|T^* f(x)| \leq \varepsilon \|f\|\|x\|$. Hence $\|T^*|_W\| \leq \varepsilon$, so by completeness of X^* and Theorem 11.6 we see that T^* is compact.

Conversely, suppose that T^* is compact and that Y is complete. Let $\varepsilon > 0$. By Theorem 11.6 there exists a finite-dimensional subspace Z of X^* such that $\|\pi_Z \circ T^*\| \leq \varepsilon$. Then $W = Z_\circ$ is closed and of finite codimension in X. Let $x \in W$, $f \in B_{Y^*}$ and $g \in Z$. Then

$$|f(Tx)| = |(T^* f + g)(x)| \leq \|T^* f + g\|\|x\|,$$

and taking the infimum over all $g \in Z$ gives $|f(Tx)| \leq \varepsilon \|x\|$. It follows from Corollary 5.8 that $\|T|_W\| \leq \varepsilon$ and hence T is compact by Theorem 11.6. $\square$

Remark 11.8. An alternative way to prove the converse statement is to observe that if T^* is compact then by the first part the operator $T^{**} \colon X^{**} \to Y^{**}$ is again compact. Now $T^{**} \circ J_X = J_Y \circ T$, so $J_Y(T(B_X)) \subseteq T^{**}(B_{X^{**}})$, which is totally bounded. Since J_Y is an isometry it follows that $T(B_X)$ is totally bounded and hence, by completeness of Y and Corollary 10.3, $T(B_X)$ is relatively compact.

12 Fredholm theory

Let X and Y be Banach spaces and suppose that $T \in \mathcal{B}(X, Y)$. One is often interested in finding solutions $x \in X$ of an equation of the form

$$Tx = y, \tag{12.1}$$

where $y \in Y$ is a given vector. The problem has a solution if and only if $y \in \text{Ran}\, T$, and we know from Remark 5.11 that the closure of $\text{Ran}\, T$ coincides with $(\text{Ker}\, T^*)_\circ$. In particular, if $\text{Ran}\, T$ is closed then we have a criterion for our problem to have a solution, namely that $f(y) = 0$ for all $f \in \text{Ker}\, T^*$.

Theorem 12.1 (Closed Range Theorem). *Let X and Y be Banach spaces and suppose that $T \in \mathcal{B}(X, Y)$. Then $\text{Ran}\, T$ is closed if and only if $\text{Ran}\, T^*$ is closed.*

PROOF. Define the operators $Q\colon X \to \operatorname{Ran} T$ and $S\colon \operatorname{Ran} T \to Y$ by $Qx = Tx$, $x \in X$, and $Sy = y$, $y \in \operatorname{Ran} T$, so that $T = S \circ Q$. If $\operatorname{Ran} T$ is closed, then by the Open Mapping Theorem Q is a quotient operator. Conversely, if Q is a quotient operator then $\operatorname{Ran} T$ is isomorphic to the Banach space $X/\operatorname{Ker} T$, so $\operatorname{Ran} T$ is complete and therefore closed. Thus $\operatorname{Ran} T$ is closed if and only if Q is a quotient operator. By Theorem 5.10 the latter is equivalent to Q^* being an isomorphic embedding. But $\operatorname{Ker} Q^* = (\operatorname{Ran} Q)^\circ = \{0\}$, so Q^* maps bijectively onto its range. If Q^* is an isomorphic embedding, then $\operatorname{Ran} Q^*$ is closed, while if $\operatorname{Ran} Q^*$ is closed then by the Inverse Mapping Theorem Q^* is an isomorphic embedding. Thus Q^* is an isomorphic embedding if and only if $\operatorname{Ran} Q^*$ is closed, and therefore $\operatorname{Ran} T$ is closed if and only if $\operatorname{Ran} Q^*$ is closed. Note that the operator S is an isomorphic embedding, so by Theorem 5.10 its dual S^* is a quotient operator and in particular surjective. Since $T^* = Q^* \circ S^*$ we see that $\operatorname{Ran} T^* = \operatorname{Ran} Q^*$, so the result is proved. $\qquad\square$

Remark 12.2. It follows from a result on Problem Sheet 3 that $\operatorname{Ran} T$ is norm-closed if and only if it is weakly closed. The above proof shows that if $\operatorname{Ran} T$ is closed then Q is a quotient operator, and hence by Remark 5.11 we see that $\operatorname{Ran} Q^* = (\operatorname{Ker} Q)^\circ$. But $\operatorname{Ran} Q^* = \operatorname{Ran} T^*$ and $\operatorname{Ker} Q = \operatorname{Ker} T$, so we obtain a further equivalent condition for closedness of $\operatorname{Ran} T$ and $\operatorname{Ran} T^*$, namely that $\operatorname{Ran} T^* = (\operatorname{Ker} T)^\circ$. Thus by Corollary 9.12 we see that $\operatorname{Ran} T^*$ is norm-closed if and only if it is weak*-closed.

If X and Y are Banach spaces and $T \in \mathcal{B}(X, Y)$ we say that T is a *Fredholm operator* if $\operatorname{Ker} T$ is finite-dimensional and $\operatorname{Ran} T$ has finite codimension in Y. We know from Problem Sheet 2 that any such operator must have closed range. If T is Fredholm we define the *index* of T as

$$\operatorname{ind} T = \dim \operatorname{Ker} T - \dim Y/\operatorname{Ran} T.$$

Example 12.3. (a) Any invertible operator between two Banach spaces is Fredholm with index zero.
  (b) Let $X = \ell^1$ and define the left-shift $T \in \mathcal{B}(X)$ by $Tx = (x_2, x_3 \dots)$ for $x = (x_n) \in X$. Then T^k for each integer $k \geq 0$ is Fredholm with $\operatorname{ind} T^k = k$.
  (c) If X and Y are finite-dimensional normed vector spaces, then every linear operator $T\colon X \to Y$ is Fredholm and moreover, by the Rank-Nullity Theorem,

$$\operatorname{ind} T = \dim \operatorname{Ker} T - \dim Y + \dim \operatorname{Ran} T = \dim X - \dim Y.$$

Proposition 12.4. *Let X and Y be Banach spaces and suppose that $T \in \mathcal{B}(X, Y)$. Then T is Fredholm if and only if T^* is Fredholm, and if both operators are Fredholm then $\operatorname{ind} T + \operatorname{ind} T^* = 0$.*

PROOF. Exercise; see Problem Sheet 4. $\qquad\square$

Theorem 12.5. *Let X and Y be Banach spaces and suppose that $T \in \mathcal{B}(X, Y)$. If T is Fredholm then there exist a closed finite-codimensional subspace V of X and a finite-dimensional subspace W of Y such that $X = \operatorname{Ker} T \oplus V$ and $Y = \operatorname{Ran} T \oplus W$ topologically, and moreover $T|_V$ maps V isomorphically onto $\operatorname{Ran} T$. Conversely, if there exist closed finite-codimensional subspaces V of X and Z of Y such that $T|_V$ maps V isomorphically onto Z, then T is Fredholm and*

$$\operatorname{ind} T = \dim X/V - \dim Y/Z.$$

PROOF. Suppose that T is Fredholm. The existence of V follows from a result on Problem Sheet 2 and the existence of W is also clear. Then $T|_V$ maps V bijectively onto the Banach space $\operatorname{Ran} T$, so by the Inverse Mapping Theorem it does so isomorphically. Suppose now that V and Z are as described. Then $\operatorname{Ker} T \cap V = \{0\}$, so $\operatorname{Ker} T$ is finite-dimensional, and $Z \subseteq \operatorname{Ran} T$, so $Y/\operatorname{Ran} T$ is finite-dimensional. Hence T is Fredholm. We may find a finite-dimensional subspace U of X such that $X = \operatorname{Ker} T \oplus V \oplus U$. Then $\operatorname{Ran} T = T(V \oplus U) = Z \oplus T(U)$ and $\dim T(U) = \dim U$ since $T|_{V \oplus U}$ is injective. Thus

$$\dim \operatorname{Ker} T = \dim X/(V \oplus U) = \dim X/V - \dim U$$

and

$$\dim Y/\operatorname{Ran} T = \dim Y/(Z \oplus T(U)) = \dim Y/Z - \dim U,$$

and the result follows. $\qquad\square$

From the point of view of solving the equation (12.1) Fredholm operators are particularly nice because they lead to criteria involving only finitely many conditions, both for existence and uniqueness of solutions. Indeed, we have that $\operatorname{Ran} T = (\operatorname{Ker} T^*)_\circ$ and Corollary 5.12 shows that $\operatorname{Ker} T^* = (\operatorname{Ran} T)^\circ \cong (Y/\operatorname{Ran} T)^*$. In particular, $\operatorname{Ker} T^*$ is finite-dimensional. Thus given $y \in Y$ equation (12.1) has a solution $x \in X$ if and only if $f_k(y) = 0$, $1 \leq k \leq n$, where $\{f_k : 1 \leq k \leq n\}$ is a basis for $\operatorname{Ker} T^*$. Moreover, there exists a closed subspace V of X such that $X = \operatorname{Ker} T \oplus V$, and the solution is unique subject to $x \in V$. Since V is closed we have $V = (V^\circ)_\circ$, and Corollary 5.12 shows that $V^\circ \cong (X/V)^*$. Thus $\dim V^\circ = \dim X/V = \dim \operatorname{Ker} T$, so V° is finite-dimensional and the solution $x \in X$ to (12.1) is unique subject to $g_k(x) = 0$, $1 \leq k \leq m$, where $\{g_k : 1 \leq k \leq m\}$ is a basis for V°.

Suppose that X and Y are Banach spaces. If $T \in \mathcal{B}(X, Y)$ is an isomorphism and if $S \in \mathcal{B}(X, Y)$ is such that $\|S\| < \|T^{-1}\|^{-1}$, then as $n \to \infty$ the partial sums

$$\sum_{k=0}^{n} (-1)^k (T^{-1}S)^k T^{-1}, \quad n \geq 0,$$

converge (absolutely) in the norm of $\mathcal{B}(Y, X)$ to the inverse of $T + S$, so $T + S$ is also an isomorphism. In particular, isomorphisms form an open subset of $\mathcal{B}(X, Y)$.

Theorem 12.6. *Let X and Y be Banach spaces. Then the set of Fredholm operators is an open subset of $\mathcal{B}(X, Y)$, and moreover the index map is locally constant, and in particular continuous, on the set of all Fredholm operators.*

PROOF. Let $T \in \mathcal{B}(X, Y)$ be a Fredholm operator. By Theorem 12.5 there exist a closed finite-codimensional subspace V of X and a finite-dimensional subspace W of Y such that $X = \operatorname{Ker} T \oplus V$ and $Y = \operatorname{Ran} T \oplus W$ topologically, and $T|_V$ is an isomorphic embedding. Consider the space $V \times W$ endowed with the ∞-norm and, given $S \in \mathcal{B}(X, Y)$, let $Q_S : V \times W \to Y$ be given by $Q_S(x, y) = Tx + Sx + y$ for $x \in V$, $y \in W$. Then Q_0 is an isomorphism and $\|Q_0 - Q_S\| = \|S|_V\| \leq \|S\|$, $S \in \mathcal{B}(X, Y)$. Thus for $\|S\| < \|Q_0^{-1}\|^{-1}$ the operator Q_S is again an isomorphism. In particular, if we let $Z = (T + S)(V)$ then $(T + S)|_V$ maps V isomorphically onto Z. Moreover, by bijectivity of Q_S we have $Z \oplus W = Y$, so Z has the same finite

codimension in Y as $\operatorname{Ran} T$. It follows from Theorem 12.5 that $T + S$ is Fredholm and that

$$\operatorname{ind}(T + S) = \dim X/V - \dim Y/Z = \dim \operatorname{Ker} T - \dim Y/\operatorname{Ran} T = \operatorname{ind} T,$$

which completes the proof. $\qquad\square$

Remark 12.7. Note that by the above proof we have $\operatorname{ind}(T + S) = \operatorname{ind} T$ and $\dim Y/\operatorname{Ran}(T + S) \leq \dim Y/\operatorname{Ran} T$ for S of sufficiently small norm. It follows that for such operators S we also have $\dim \operatorname{Ker}(T + S) \leq \dim \operatorname{Ker} T$.

Theorem 12.8. *Let X and Y be Banach spaces and suppose that $T, S \in \mathcal{B}(X, Y)$ with T Fredholm and S compact. Then $T + S$ is Fredholm and $\operatorname{ind}(T + S) = \operatorname{ind} T$.*

PROOF. By Theorem 12.5 there exists a closed finite-codimensional subspace V of X such that $T|_V$ is an isomorphic embedding. Since S is compact and restricts to an isomorphic embedding on $V \cap \operatorname{Ker}(T + S)$ this space must be finite-dimensional, and hence $\dim \operatorname{Ker}(T + S) < \infty$. Note that S^* is compact by Schauder's Theorem and T^* is Fredholm by Proposition 12.4, so the above argument shows that $\operatorname{Ker}(T^* + S^*)$ is also finite-dimensional. Now let $r > 0$ be such that $\|Tx\| \geq r\|x\|$ for all $x \in V$, and let $\varepsilon = r/2$. By Theorem 11.6 we may find a closed finite-codimensional subspace W of X such that $\|S|_W\| \leq \varepsilon$. Then $V \cap W$ is closed and of finite codimension in X, and for $x \in V \cap W$ we have $\|(T + S)x\| \geq \varepsilon\|x\|$. Thus $(T + S)(V \cap W)$ is a Banach space which has finite codimension in $\operatorname{Ran}(T + S)$. Since closed finite-codimensional subspaces are complemented it follows that $\operatorname{Ran}(T + S)$ is complete and therefore closed. Thus $\operatorname{Ran}(T + S) = \operatorname{Ker}(T^* + S^*)_\circ$, so $\operatorname{Ran}(T + S)$ has finite codimension in Y and hence $T + S$ is Fredholm. For $0 \leq t \leq 1$ let $Q_t = T + tS$. Then each Q_t is Fredholm and Theorem 12.6 implies that the function $\psi \colon [0, 1] \to \mathbb{Z}$ given by $\psi(t) = \operatorname{ind} Q_t$ is continuous. By the Intermediate Value Theorem ψ must be constant, and in particular $\operatorname{ind}(T + S) = \psi(1) = \psi(0) = \operatorname{ind} T$, as required. $\qquad\square$

Corollary 12.9 (Fredholm Alternative). *Let X be a Banach space and suppose that $T \in \mathcal{K}(X)$ and that $\lambda \in \mathbb{F} \setminus \{0\}$. Then $\lambda - T$ is injective if and only if it is surjective, and*

$$\operatorname{Ran}(\lambda - T) = \operatorname{Ker}(\lambda - T^*)_\circ.$$

Furthermore, $\dim \operatorname{Ker}(\lambda - T^*) = \dim \operatorname{Ker}(\lambda - T)$ and both numbers are finite.

PROOF. For $\lambda \neq 0$ the operator λI is Fredholm with index zero, so by Theorem 12.8 the operator $\lambda - T$ is also Fredholm and $\operatorname{ind}(\lambda - T) = 0$. In particular, $\lambda - T$ has closed range, and it is injective if and only if it is surjective. This proves the first two claims. By Corollary 5.12 we have

$$\dim(X/\operatorname{Ran}(\lambda - T))^* = \dim \operatorname{Ran}(\lambda - T)^\circ = \dim \operatorname{Ker}(\lambda - T^*),$$

and hence $\dim \operatorname{Ker}(\lambda - T^*) = \dim X/\operatorname{Ran}(\lambda - T) = \dim \operatorname{Ker}(\lambda - T)$. $\qquad\square$

Recall that if X is a complex Banach space and $T \in \mathcal{B}(X)$, then the *spectrum*

$$\sigma(T) = \big\{\lambda \in \mathbb{C} : \lambda - T \text{ is not an isomorphism}\big\}$$

of T is a non-empty compact subset and the *spectral radius* $r(T) = \sup_{\lambda \in \sigma(T)} |\lambda|$ satisfies $r(T) = \lim_{n \to \infty} \|T^n\|^{1/n}$. Let $\sigma_p(T)$ denote the *point spectrum* of T, that is to say the set of eigenvalues of T. If X is finite-dimensional then $\sigma(T) = \sigma_p(T)$, but in general this is not true. However, for compact operators we at least have $\sigma(T) \setminus \{0\} \subseteq \sigma_p(T)$, as we now show.

Theorem 12.10. *Let X be a complex Banach space and suppose that $T \in \mathcal{K}(X)$. Then $\sigma(T)$ is at most countably infinite and $\sigma(T) \setminus \{0\}$ consists of isolated points which are eigenvalues with finite-dimensional eigenspaces. In particular, if $\sigma(T)$ is infinite then its unique limit point is zero. Furthermore, $0 \in \sigma(T)$ whenever X is infinite-dimensional.*

PROOF. For $\lambda \in \mathbb{C} \setminus \{0\}$ it follows from Corollary 12.9 and the Inverse Mapping Theorem that $\lambda \in \sigma(T)$ if and only if λ is an eigenvalue of T with finite-dimensional eigenspace. Fix $\lambda \in \mathbb{C} \setminus \{0\}$ and let $Y_n = \mathrm{Ran}(\lambda - T)^n$, $n \geq 0$, and $Y = \bigcap_{n=0}^{\infty} Y_n$. Note that $Y_{n+1} \subseteq Y_n$ and that Y_n has finite codimension in X for all $n \geq 0$. Thus each Y_n, $n \geq 0$, is closed and hence so is Y. We now show that $\lambda - T$ maps Y onto itself. It is straightforward to see that $(\lambda - T)(Y) \subseteq Y$. To prove the converse inclusion let $y \in Y$. For each $n \geq 0$ there exists $x_n \in Y_n$ such that $(\lambda - T)x_n = y$. Moreover, since $\mathrm{Ker}(\lambda - T)$ is finite-dimensional there exists $k \geq 0$ such that $\mathrm{Ker}(\lambda - T) \cap Y_n = \mathrm{Ker}(\lambda - T) \cap Y_k$ for all $n \geq k$. Thus

$$x_k - x_n \in \mathrm{Ker}(\lambda - T) \cap Y_k \subseteq Y_n, \quad n \geq k,$$

which implies that $x_k \in Y$ and hence $Y \subseteq (\lambda - T)(Y)$, as required. Now $T|_Y \in \mathcal{K}(Y)$, so Corollary 12.9 shows that $\mathrm{Ker}(\lambda - T)|_Y = \{0\}$. By Remark 12.7 there exists $\varepsilon > 0$ such that $\mathrm{Ker}(\mu - T)|_Y = \{0\}$ for $|\lambda - \mu| < \varepsilon$. If $\mu \neq \lambda$ and $x \in \mathrm{Ker}(\mu - T)$ then

$$x = (\lambda - \mu)^{-n}(\lambda - T)^n x \in Y_n, \quad n \geq 0,$$

so $x \in Y$. Thus for $0 < |\mu - \lambda| < \varepsilon$ we have $\mathrm{Ker}(\mu - T) \subseteq \mathrm{Ker}(\mu - T)|_Y = \{0\}$ and hence $\mu \notin \sigma(T) \setminus \{0\}$. It follows that $\sigma(T) \setminus \{0\}$ consists of isolated points. Let $\sigma_n(T) = \{\lambda \in \sigma(T) : |\lambda| \geq n^{-1}\}$, $n \geq 1$, and recall that any infinite compact set contains a limit point. Thus $\sigma_n(T)$ is either empty or finite for each $n \geq 1$ and $\sigma(T) \setminus \{0\} = \bigcup_{n \geq 1} \sigma_n(T)$ is at most countably infinite. Moreover, if $\sigma(T)$ is infinite then since the points in $\sigma(T) \setminus \{0\}$ are isolated zero must be the unique limit point of $\sigma(T)$. Note finally that if T is surjective then by the Open Mapping Theorem there exists $r > 0$ such that $B_X^\circ(r) \subseteq T(B_X^\circ)$ and hence by compactness of T the unit ball B_X is compact, so X must be finite-dimensional. Thus $0 \in \sigma(T)$ whenever X is infinite-dimensional. $\square$

Remark 12.11. There is an alternative way to proceed after proving that $\sigma(T) \setminus \{0\}$ consists of eigenvalues corresponding to finite-dimensional eigenspaces. Indeed, one can show by means of a subsequence argument that the subspace of X spanned by all eigenvectors of T corresponding to eigenvalues in $\{\lambda \in \mathbb{C} : |\lambda| \geq r\}$ is finite-dimensional for each $r > 0$; see Bollobas, Chapter 13, Theorem 4.

Let X be a complex Hilbert space with inner product $(\cdot, \cdot)$, and let $T \in \mathcal{B}(X)$. Recall that the *adjoint operator* $T^\star \in \mathcal{B}(X)$ of T is defined by the identity $(Tx, y) =$

$(x, T^\star y)$, $x, y \in X$. By the Riesz Representation Theorem there exists a conjugate-linear isometric surjection $\Phi_X : X \to X^*$ such that $(\Phi_X y)(x) = (x, y)$ for all $x, y \in X$. The adjoint of T is related to the dual of T through the relation $T^* \circ \Phi_X = \Phi_X \circ T^\star$. Indeed, for $x, y \in X$ we have

$$(T^*(\Phi_X x))(y) = (\Phi_X x)(Ty) = (Ty, x) = (y, T^\star x) = (\Phi_X(T^\star x))(y).$$

Thus the following diagram commutes:

$$
\begin{array}{ccc}
X^* & \xrightarrow{\;T^*\;} & X^* \\
\Phi_X \big\uparrow \big\downarrow \Phi_X^{-1} & & \Phi_X \big\uparrow \big\downarrow \Phi_X^{-1} \\
X & \xrightarrow{\;T^\star\;} & X
\end{array}
$$

Recall that an operator T is said to be *self-adjoint* if $T^\star = T$. We say that a set $\{x_\alpha : \alpha \in A\}$ is an *orthonormal basis* for X if $\|x_\alpha\| = 1$ for all $\alpha \in A$, $(x_\alpha, x_\beta) = 0$ whenever $\alpha, \beta \in A$ are distinct and $\mathrm{span}\{x_\alpha : \alpha \in A\}$ is dense in X.

Theorem 12.12 (Spectral Theorem). *Let X be an infinite-dimensional complex Hilbert space and suppose that $T \in \mathcal{K}(X)$ is self-adjoint. Then X admits an orthonormal basis consisting of eigenvectors of T. Moreover there exist a sequence $(\lambda_n)_{n=1}^N$ of non-zero real numbers, where $N \in \mathbb{N} \cup \{\infty\}$, such that $\lambda_n \to 0$ as $n \to \infty$ when the sequence is infinite, and furthermore there exists a sequence of orthogonal finite-rank projections $(P_n)_{n=1}^N$ such that $P_m P_n = 0$ for $m \neq n$ and*

$$T = \sum_{n=1}^N \lambda_n P_n,$$

where the series converges in the norm of $\mathcal{B}(X)$ when $N = \infty$.

PROOF. Recall that $\sigma(T) \subseteq \mathbb{R}$ when T is self-adjoint, and that eigenvectors corresponding to distinct eigenvalues are orthogonal. Hence existence of the sequences (λ_n) and (P_n) follows from Theorem 12.10. If $N = \infty$, the series $\sum_{n=1}^N \lambda_n P_n$ is Cauchy and hence convergent in $\mathcal{B}(X)$. Let $S = \sum_{n=1}^N \lambda_n P_n$. Then $Sx = Tx$ for any $x \in X$ which is a linear combination of eigenvectors of T. Hence the result will follow once we have shown that X admits an orthonormal basis consisting of such eigenvectors. Let Y be the closed linear span of all eigenvectors of T. If we let B_n be an orthonormal basis for $\mathrm{Ker}(\lambda_n - T)$ and if we use Zorn's Lemma to obtain an orthonormal basis B_0 for $\mathrm{Ker}\, T$, then the set $\bigcup_{n=0}^N B_n$ is an orthonormal basis for Y consisting of eigenvectors of T. Let $Z = Y^\perp$. Since Y is T-invariant and T is self-adjoint the space Z is also T-invariant, and moreover $T|_Z$ is a compact self-adjoint operator on Z. Theorem 12.10 implies that $\sigma(T|_Z) = \{0\}$, because any eigenvalue of $T|_Z$ would also be an eigenvalue of T. By self-adjointness we deduce that $\|T|_Z\| = r(T|_Z) = 0$, so $Z \subseteq \mathrm{Ker}\, T \subseteq Y = Z^\perp$. Hence $Z = \{0\}$, so $X = Y$. $\square$

Remark 12.13. The Spectral Theorem can be extended to the case of compact *normal* operators, that is to say compact operators T such that $T^\star T = TT^\star$.

13 Schauder bases

Given a Banach space X, a set $\{x_n : n \geq 1\}$ in X is said to be a *Schauder basis* for X if every $x \in X$ admits a unique representation as a norm-convergent series

$$x = \sum_{n=1}^{\infty} \lambda_n x_n \tag{13.1}$$

with $\lambda_n \in \mathbb{F}$, $n \geq 1$. Note that a Schauder basis necessarily forms a linearly independent set and that any Banach space which admits a Schauder basis must be separable. As usual in the context of bases, a Schauder basis is strictly speaking an *ordered* set, and a permutation of a Schauder basis need not be a Schauder basis. If $\{x_n : n \geq 1\}$ is a Schauder basis for X we may consider the linear maps $P_n \colon X \to X$, $n \geq 1$, given by

$$P_n x = \sum_{k=1}^{n} \lambda_k x_k$$

when $x \in X$ has the representation in (13.1). Then $P_n^2 = P_n$ for all $n \geq 1$ and $\|P_n x - x\| \to 0$ as $n \to \infty$ for all $x \in X$. We may also consider the linear functionals $f_n \in X'$, $n \geq 1$, given by $f_n(x) = \lambda_n$ when $x \in X$ is as in (13.1). The maps P_n, $n \geq 1$, are called the *basis projections* associated with the Schauder basis $\{x_n : n \geq 1\}$ and the functionals f_n, $n \geq 1$, are the associated *basis functionals*, sometimes also referred to as *coordinate functionals*.

Theorem 13.1. *Let X be a Banach space and suppose that $\{x_n : n \geq 1\}$ is a Schauder basis for X. Then the basis projections P_n, $n \geq 1$, and the basis functionals f_n, $n \geq 1$, are all bounded, and in fact there exists $M \geq 1$ such that $\|P_n\| \leq M$ and $\|f_n\| \leq 2M\|x_n\|^{-1}$ for all $n \geq 1$.*

PROOF. Let the map $\||\cdot\|| \colon X \to \mathbb{R}_+$ be defined by

$$\||x\|| = \sup\{\|P_n x\| : n \geq 1\}, \quad x \in X.$$

We will see on Problem Sheet 4 that $\||\cdot\||$ is a complete norm on X. Note also that since $x = \lim_{n \to \infty} P_n x$ for all $x \in X$ we have that

$$\|x\| = \lim_{n \to \infty} \|P_n x\| \leq \||x\||, \quad x \in X.$$

Thus the identity map from $(X, \||\cdot\||)$ to $(X, \|\cdot\|)$ is a continuous bijection, and it follows from the Inverse Mapping Theorem that $\||\cdot\||$ is equivalent to $\|\cdot\|$, so there exists $M \geq 1$ such that $\||x\|| \leq M\|x\|$ for all $x \in X$. Note finally that

$$\|P_n x\| \leq \||x\|| \leq M\|x\|, \quad x \in X, \ n \geq 1,$$

and hence $\|P_n\| \leq M$ for all $n \geq 1$. Since $f_n(x)x_n = (P_n - P_{n-1})x$ for all $x \in X$ and $n \geq 1$ (with P_0 taken to be the zero operator), the final claim follows at once. $\quad\square$

By Remark 11.4 and a result on Problem Sheet 4 there exists a separable Banach space which does not admit a Schauder basis. However, most classical Banach spaces which are separable do admit a Schauder basis.

Example 13.2. (a) If X is a separable Hilbert space then any orthonormal basis for X is a Schauder basis for X.

(b) If $X = \ell^p$ for $1 \leq p < \infty$ or $X = c_0$, then the set $\{e_n : n \geq 1\}$ is a Schauder basis for X. The basis functionals are given by $f_n(x) = x_n$ for $x \in X$, $n \geq 1$. If $X = c$ we may add $e_0 = (1, 1, 1, \dots)$ to the above basis to get a Schauder basis $\{e_n : n \geq 0\}$ for X. The basis functional f_0 corresponding to e_0 is $f_0(x) = \lim_{n \to \infty} x_n$, $x \in X$.

(c) Let $X = L^p(0, 1)$, where $1 \leq p < \infty$. Given $n \geq 1$ we may uniquely express n in the form $n = 2^k + j$ with $k \geq 0$ and $0 \leq j \leq 2^k - 1$, and then we may take x_n to be the function satisfying $x_n(t) = 1$ for $j2^{-k} < t < (2j+1)2^{-k-1}$ and $x_n(t) = -1$ for $(2j+1)2^{-k-1} < t < (j+1)2^{-k}$. Together with the constant function $x_0(t) = 1$, $0 < t < 1$, we obtain a Schauder basis $\{x_n : n \geq 0\}$ for X. This is the *Haar basis*.

(d) Let $X = C([0, 1])$ and, for $n \geq 0$, let x_n be as in (e). Now define $y_n \in X$ by $y_0 = x_0$ and

$$y_n(t) = 2^{n-1} \int_0^t x_{n-1}(s) \, ds, \quad 0 \leq t \leq 1, \ n \geq 1.$$

Then $\{y_n : n \geq 0\}$ is a Schauder basis for X. It is known as *Schauder's basis*.

14 Subspaces of classical sequence spaces

The material covered in this section is not examinable. Given a Banach space X, a subset $S = \{y_n : n \geq 1\}$ of X is said to be *basic* if it is a Schauder basis for its closed linear span $Y = \overline{\operatorname{span}} \, S$. In this case there exist, by Theorem 13.1, basis projections $P_n \in \mathcal{B}(Y)$ and associated basis functionals $f_n \in Y^*$, $n \geq 1$. We also know that $\sup\{\|P_n\| : n \geq 1\} < \infty$, and we call the quantity $M = \sup\{\|P_n\| : n \geq 1\}$ the *basis constant* of S. Then $\|f_n\| \leq 2M\|y_n\|^{-1}$, $n \geq 1$. Note that the basis constant M of any basic set necessarily satisfies $M \geq 1$.

Theorem 14.1. *Let X be a Banach space and suppose that $\{y_n : n \geq 1\} \subseteq S_X$ is a basic set with basis constant M. Suppose furthermore that $z_n \in X$, $n \geq 1$, are such that*

$$\sum_{n=1}^{\infty} \|y_n - z_n\| < \frac{1}{2M}.$$

Then $Y = \overline{\operatorname{span}}\{y_n : n \geq 1\}$ is isomorphic to $Z = \overline{\operatorname{span}}\{z_n : n \geq 1\}$, and moreover Y is complemented in X if and only if Z is.

PROOF. Let $g_n \in X^*$ be Hahn-Banach extensions of the basis functionals $f_n \in Y^*$, so that $\|g_n\| = \|f_n\| \leq 2M$, $n \geq 1$, and let

$$Sx = \sum_{n=1}^{\infty} g_n(x)(y_n - z_n), \quad x \in X.$$

Then by our assumption S is a well-defined element of $\mathcal{B}(X)$ with $\|S\| < 1$. Hence the operator $T = I - S \in \mathcal{B}(X)$ is an isomorphism which satisfies $Ty_n = z_n$, $n \geq 1$, and it follows that $T(Y) = Z$. Suppose that $P \in \mathcal{B}(X)$ satisfies $P^2 = P$. If $\operatorname{Ran} P = Y$ then we consider $Q = TPT^{-1} \in \mathcal{B}(X)$ which satisfies $Q^2 = Q$ and $\operatorname{Ran} Q = Z$, and if $\operatorname{Ran} P = Z$ then we consider $Q = T^{-1}PT \in \mathcal{B}(X)$ which satisfies $Q^2 = Q$ and $\operatorname{Ran} Q = Z$. Hence Y is complemented if and only if Z is. $\qquad\square$

Proposition 14.2. *Let $X = \ell^p$ for $1 \le p < \infty$ or $X = c_0$.*

(a) *If $S = \{y_n : n \ge 1\} \subseteq S_X$ is disjointly supported, then S is a basic set with basis constant $M = 1$ and furthermore the space $Y = \overline{\operatorname{span}}\, S$ is isomorphic to X and complemented in X.*

(b) *If Z is an infinite-dimensional subspace of X then there exists a disjointly supported set $\{y_n : n \ge 1\} \subseteq S_X$ and vectors $z_n \in Z$, $n \ge 1$, such that*

$$\sum_{n=1}^{\infty} \|y_n - z_n\| < \frac{1}{2}.$$

PROOF. The proof of part (a) is straightforward, and the proof of part (b) uses the prototype of a 'gliding hump' argument. $\qquad\square$

Theorem 14.3. *Let $X = \ell^p$ for $1 \le p < \infty$ or $X = c_0$, and suppose that Z is a closed infinite-dimensional subspace of X. Then Z contains a complemented subspace which is isomorphic to X.*

PROOF. This follows from Theorem 14.1 and Proposition 14.2. $\qquad\square$

Corollary 14.4. (a) *Let $X = \ell^1$ or $X = c_0$. Then every infinite-dimensional subspace of X is non-reflexive.*

(b) *Every closed infinite-dimensional subspace of ℓ^1 has non-separable dual.*

We end with a striking result about subspaces of classical sequence spaces.

Theorem 14.5 (Pełczyński). *Let $X = \ell^p$ for $1 \le p < \infty$ or $X = c_0$. Then every infinite-dimensional complemented subspace of X is isomorphic to X.*

PROOF. By considering a partition of $\mathbb{N}$ into countably many infinite subsets we see that

$$X \cong \left(\bigoplus_{n=1}^{\infty} X \right)_p,$$

where if $X = c_0$ we let $p = \infty$ and consider X-valued sequences which converge to zero. If Z is an infinite-dimensional complemented subspace of X then $X = Y \oplus Z$ topologically for some closed subspace Y of X and by Theorem 14.3 we have $Z = Z_0 \oplus Z_1$ topologically, where $Z_1 \simeq X$. Thus $X = Z_0 \oplus Z_1 \oplus Y = Z_1 \oplus Y \oplus Z_0$, all direct sums being topological, and hence

$$Z = Z_0 \oplus Z_1 \simeq Z_0 \oplus X \simeq Z_0 \oplus \left(\bigoplus_{n=1}^{\infty} Z_1 \oplus Y \oplus Z_0 \right)_p \simeq \left(\bigoplus_{n=1}^{\infty} Z_0 \oplus Z_1 \oplus Y \right)_p \simeq X,$$

as required. $\qquad\square$