C4.1 Further Functional Analysis

Stuart White

University of Oxford
Michaelmas Term 2021*

I'm very grateful to David Seifert, who developed the previous version of these notes, and allowed me to use them in MT 2019. Please send comments, corrections, clarifications etc to me at `stuart.white@maths.ox.ac.uk`.

I have made some very small tweaks since the MT 2020 version of these notes, but kept the numbering the same to allow for compatibility with last years lecture videos. I have numbered any such additional results in the form 1.A, 1.B.

0 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.

*Version of November 26, 2021

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

[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. For the first part of the course, all our topologies will come from metrics.
- Basics of topological spaces, closures and interiors, and in particular compactness. You'll also want to be familiar with constructing topological spaces by specifiying a basis of open sets. This becomes most relevant by the time we get to Sections 9, when we start to work with topologies on normed spaces which do not come from the norm, and are not a-priori metrisable. In terms of the part A topology course, this means we will essentially use the first half of it, but we won't need any of the aspects of the classification of surfaces. If you've not studied the first half of part A topology or its equivalent, I really recommend the topological spaces part of Wilson Sutherland's little book 'An Introduction to metric and topological spaces.'[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] We set out some of this background. and in particular our notation, in Section 1
- Operators between normed spaces, Continuity and boundedness, Completeness of $\mathcal{B}(X,Y)$ when Y is complete. Again we'll review some of these and set out our notation in Section 1.
- The Baire category theorem.[3] We will briefly describe Baire's category theorem in Appendix A.
- Open mapping theorem, closed graph theorem, and inverse mapping theorem.[4]. These are discussed a bit in appendix A, where we deduce them from

[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.

[4] Ideally you've seen these before, but for our course the statements will suffice

Baire's category theorem (in a very similar fashion to as done in B.4.2). We will see the equivalence of these three classical theorems in sheet 2 (and they're also equivalent to the uniform boundeness principle).

- 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 (the monotone convergence theorem will be enough), but it's otherwise not essential.

1 Normed linear 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.[5] 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 $X \neq \{0\}$.

Definition 1.1. 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 $d(x, y) = \|x - y\|$, $x, y \in X$.

Definition 1.2. 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.

- n-dimensional Euclidian space ℓ^2_n, and the various equivalent norms we could put on these spaces to obtain ℓ^p_n, 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]

[5] Researchers studying Banach spaces often prefer $\mathbb{F} = \mathbb{R}$, while those studying operator theory prefer $\mathbb{F} = \mathbb{C}$. In the latter case this is because we use the fundamental theorem of algebra, even for matrices, to see that every complex matrix has an eigenvalue. In general, as you may know, the spectrum of a bounded operator on a complex Banach space is necessarily non-empty — on a real Banach space, even $\mathbb{R}^2$, the spectrum might be empty.

[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.

Our notation for open and closed balls will be as follows. For $x_0 \in X$ and $r > 0$, we let

$$B_X(x_0, r) = \{x \in X : \|x - x_0\| \leq r\}$$

denote the *closed ball*, and we let

$$B_X^\circ(x_0, r) = \{x \in X : \|x - x_0\| < r\}$$

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 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).

Recall that any vector space X has an associated *algebraic dual space* X' consisting of all linear functionals $f \colon X \to \mathbb{F}$. Recall that for a linear functional $f \in X'$, the following are equivalent:.[8]

- f is bounded (i.e. $|f(x)| \leq C\|x\|$ for some $C > 0$ and all $x \in X$);
- f is continuous
- f is continuous at 0
- the *kernel* $\mathrm{Ker}(f) = \{x \in X : f(x) = 0\}$ is a closed subspace of X

If X is a normed vector space it is natural to restrict oneself to the class of *continuous*, or equivalently *bounded*, linear functionals f

Definition 1.3. Let X be a normed space. Write X^* for the *(topological) dual space* consisting of all bounded linear functionals $f \colon X \to \mathbb{F}$.

Recall that, whether or not X is complete, X^* is always complete, when given the norm

$$\|f\| = \sup_{x \in X} |f(x)|.$$

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

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

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 $\mathcal{B}(X) = \mathcal{B}(X,X)$. You might well want to look over the basics of the spectrum of a bounded operator for the last part of the course when we start studying compact and Fredholm operators, and in particular that the spectrum is always compact and non-empty (and provided $\mathbb{F} = \mathbb{C}$).

You'll probably be familiar with the dual operator $T^* \in \mathcal{B}(Y^*,X^*)$ associated to $T \in \mathcal{B}(X,Y)$. This will reappear in Section 5. For now check you can define it, and verify that T^* is bounded. What's an estimate for $\|T^*\|$?

2 Hamel bases and unbounded functionals

Definition 2.1. A *Hamel basis* for a vector space X is a linearly independent spanning set for X.[9]

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

$$x = \sum_{k=1}^{n} \lambda_k x_k \tag{2.1}$$

for some $n \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and $x_1, \ldots, x_n \in B$.

In general, 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). We will set this up in a slightly different (though equivalent) fashion to how you might have seen it in a set theory course.

Definition 2.2. A partially ordered set (poset) is a set $\mathcal{P}$ together with a relation $\leq$ which is reflexive, transitive and antisymmetric.[10] An element x in a poset $\mathcal{P}$ is *maximal* if $x \leq y$ implies that $y = x$.[11] 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.

The key example of a poset, is the family of subsets of a given set, ordered by inclusion.

[9]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 14. But note the key difference: the sums in (2.1) are required to be finite even if X is a normed space. Hamel bases take no account of the analytic structure.

[10]antisymmetry means that if $x, y \in \mathcal{P}$ have $x \leq y$ and $y \leq x$, then $x = y$.

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

Axiom 2.3 (Zorn's Lemma). *Let $\mathcal{P}$ be a non-empty poset such that every chain has an upper bound. Then $\mathcal{P}$ has a maximal element.*

Remark 2.4. 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. 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.[12]

Proposition 2.5. *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 $x_k \in C_k$, $1 \leq k \leq n$. 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. $\square$

Remark 2.6. 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 (Baire's Category Theorem; see Theorem A.1 in the Appendix). *Let (X, d) be a complete metric space*[13] *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 2.7. *Let X be an infinite-dimensional Banach space. Then any Hamel basis for X must be uncountable.*

[12] 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 rely on it, and these power the vast applications of functional analysis to other areas of mathematics. However there is a lot of interplay between functional analysis and mathematical logic; the set theory side of this tends to focus on the consequences for functional analysis of other set theoretic axioms, such as the continuum hypothesis.

[13] 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.

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. □

Remark 2.8. This result shows that even if X is a separable Banach space it cannot have a countable Hamel basis. We will see in Section 14 that a more appropriate notion of basis in the context of Banach spaces is that of a *Schauder basis*.

Example 2.9. 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 defined on $[0,1]$. But X is not closed, for otherwise it would be complete and by Proposition 2.7 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 $C([0,1])$.

If X is finite-dimensional then any linear functional on X is automatically bounded and so the algebraic and topological dual spaces agree, i.e. $X' = X^*$. The situation is different when X is infinite-dimensional.

Proposition 2.10. *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 2.5 and Remark 2.6 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. □

Example 2.11. Let X be as in Example 2.9. 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 2.10 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

[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 continuous. Having said that, finding functional analysts that don't subscribe to the axiom of choice is also not straightforward.

Theorem.[15] 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

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

for all $(x,y) \in X \times Y$. I invite you to check that all of these norms are equivalent, and 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 (Closed Graph Theorem; see Theorem A.5 in the appendix.). *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 2.9 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.

Definition 3.1. Let X be a vector space and let Y be a subspace of X.

1. 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.

[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.

[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.

2. If X is additionally a normed space, we say that 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.

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.

In this course, since we're doing functional analysis and not just linear algebra, by a 'complemented subspace' we will mean a topologically complemented subspace unless stated otherwise.

Proposition 3.2. *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 (Inverse Mapping Theorem; see Theorem A.4). *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.*

[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.

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

Let's start out by recalling how quotient spaces of vector spaces work. 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 $T_0 : X/\ker(T) \to \operatorname{Ran} T$ given by $T_0(x + \mathrm{kerT}) = T(x)$.

Then T factorises as

$$\begin{tikzcd}
X \arrow[r, "T"] \arrow[d, "x \mapsto x+\ker T"'] & Y \\
X/\ker T \arrow[r, "T_0"'] & \operatorname{Ran} T \arrow[u, hook]
\end{tikzcd}$$

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

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

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 $y_1, y_2 \in Y$

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

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

so taking the infimum over $y_1, y_2 \in Y$ shows that $\|x_1+x_2+Y\| \leq \|x_1+Y\|+\|x_2+Y\|$. 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 $x_n \in X$, $n \geq 1$, are such that $\sum_{n=1}^{\infty} \|x_n + Y\| < \infty$. For each $n \geq 1$ let $y_n \in Y$ be such that $\|x_n + y_n\| \leq \|x_n + Y\| + 2^{-n}$. Then $\sum_{n=1}^{\infty} \|x_n + y_n\| < \infty$, 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

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

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

Definition 4.2. 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.*

Definition 4.3. 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 $\pi(x) = x + Y$, $x \in X$. This is known as the *canonical quotient operator* from X onto X/Y.

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 topology induced by the quotient norm.

Example 4.4. 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 $X/Y \cong Y^\perp$.

Before turning to the first isomorphism theorem, we briefly take a detour through seminorms.

Definition 4.5. A *seminorm* on a vector space X is a function $p : X \to [0, \infty)$ such that

- $p(\lambda x) = |\lambda| p(x)$ for $x \in X$ and $\lambda \in \mathbb{F}$
- $p(x + y) \leq p(x) + p(y)$ for $x, y \in X$.

Remark 4.6. So a seminorm satisfies all the norm axioms except that $p(x) = 0$ need not imply $x = 0$. If X is a normed space and Y is a subspace of X which is not closed, then the proof of Propositon 4.1 $p(x + Y) = \operatorname{dist}(x, Y)$ defines a seminorm on X/Y.

Remark 4.7. Seminorms can be used to give norms on the quotient by their kernel. 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 $p(x) = \|x + Y\|$, $x \in X$.

If we start with a non-closed subspace Y of a normed space X and define the seminorm $p(x + Y) = \operatorname{dist}(x, Y)$ on X/Y. The zero subspace of this seminorm is precisely $Z = \{x + Y : x \in \overline{Y}\}$ and $(X/Y)/Z$ (with the norm induced above) is isometrically isomorphic to $X/\overline{Y}$.

Example 4.8. As an example of the construction of Remark 4.7, 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 $T: X \to Y$, the First Isomorphism Theorem tells us that T induces a well-defined linear bijection $T_0: X/\operatorname{Ker} T \to \operatorname{Ran} T$ by $T_0(x + \operatorname{Ker} T) = Tx$, $x \in X$. We are interested in the topological version.

Lemma 4.9. *Let X and Y be normed vector spaces and suppose that $T \in \mathcal{B}(X, Y)$. Then the operator $T_0: X/\operatorname{Ker} T \to \operatorname{Ran} T$ 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

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

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

$$\|Tx\| = \|T_0(x + \operatorname{Ker} T)\| \leq \|T_0\|\|x + \operatorname{Ker} T\| \leq \|T_0\|\|x\|, \quad x \in X,$$

and hence $\|T\| \leq \|T_0\|$, as required. □

Remark 4.10. In some cases it is useful to think of a not necessarily surjective operator $T: X \to Y$ in terms of its so-called *canonical factorisation.* Indeed, if we write $\pi: X \to X/\operatorname{Ker} T$ for the canonical quotient operator and $S: \operatorname{Ran} T \to Y$ for the usual embedding, then $T = S \circ T_0 \circ \pi$, where $T_0: X/\operatorname{Ker} T \to \operatorname{Ran} T$ is as above. Hence the following diagram commutes:

$$\begin{tikzcd}
X \arrow[r, "T"] \arrow[d, "\pi"'] & Y \\
X/\operatorname{Ker} T \arrow[r, "T_0"] & \operatorname{Ran} T \arrow[u, "S"']
\end{tikzcd}$$

Note that as a consequence of Lemma 4.9, if T is bounded, then so too is T_0. The maps π and S are always bounded, so the canonical factorisation of a bounded map consists of bounded maps. But this does not mean that $/X \operatorname{Ker} T$ is isomorphic to $\operatorname{Ran} T$ as a normed linear space: the map T_0^{-1} need not be bounded.

Definition 4.11. 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 4.9 is an (isometric) isomorphism.

Observe that if X is a normed vector space and Y is a closed subspace, then the canonical quotient operator $\pi\colon X \to X/Y$ given by $\pi(x) = x + Y$ is an isometric quotient operator. Indeed, π_0 is the identity operator on X/Y.

Example 4.12. Suppose that $\|\cdot\|$ and $|||\cdot|||$ are two norms on a vector space X such that $|||x||| \leq C\|x\|$ for some $C > 0$ and all $x \in X$, and let T be the identity operator from $(X, \|\cdot\|)$ to $(X, |||\cdot|||)$. 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.

Our goal is to characterise quotient operators.

Theorem 4.13. *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 prove 2 cycles of equivalences: (a)$\Rightarrow$(b)$\Rightarrow$(c)$\Rightarrow$(a) and (c)$\Rightarrow$(d)$\Rightarrow$(e)$\Rightarrow$(c).

(a)$\Rightarrow$(b). As T is a quotient operator, T_0^{-1} is bounded. Take $M > \|T_0^{-1}\|$. For $y \in Y$ with $y \neq 0$ (the result is trivial for $y = 0$), we have

$$\inf_{z \in T_0^{-1}(y)} \|z\| = \|T_0^{-1}(y)\| \leq \|T_0^{-1}\|\|y\| < M\|y\|.$$

Therefore there exists $z \in T_0^{-1}(y)$ with $\|z\| < M\|y\|$. Such a z has $T(z) = y$.

(b)$\Rightarrow$(c) is immediate with $r = M^{-1}$.

(c)$\Rightarrow$(a). Fix $r > 0$ with $B_Y^\circ(r) \subseteq T(B_X^\circ)$. Then T is clearly surjective. Fix $0 < r_0 < r$ and a non-zero $y \in Y$. Then there exists $z \in B_X^\circ$ with $Tz = r_0 y/\|y\|$ and so $\|T_0^{-1}(y)\| \leq r_0^{-1}\|y\|$. In particular T_0^{-1} is bounded, and so T is a quotient operator. Note for later reference that on substituting $y = Tx$, and letting $r_0 \to r$, this gives $r\|x + \ker T\| \leq \|Tx\|$.

(c) $\implies$ (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) $\implies$ (b) $\implies$ (c) shows that $B_Y^\circ(r) \subseteq T(B_X^\circ)$ for all $r < 1$ so $B_Y^\circ \subseteq T(B_X^\circ)$. 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 estimate noted in at the end of (c) $\implies$ (a) shows that $\|T_0(x + \operatorname{Ker} T)\| \geq \|x + \operatorname{Ker} T\|$ for all $x \in X$. We also have $\|T\| = 1$ and hence $\|T_0\| = 1$. Thus $\|T_0(x + \operatorname{Ker} T)x\| = \|x + \operatorname{Ker} T\|$ for all $x \in X$ and T is an isometric quotient operator. $\square$

In the Banach space setting, the open mapping theorem allows us to say more.

Theorem (Open Mapping Theorem; see Theorem A.3). *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 4.13 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 4.13 even further we require the following lemma which is used in the proof of the Open Mapping Theorem (as discussed in Appendix A).

Lemma 4.14 (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,

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

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 4.13 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. $\square$

[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.

Remark 4.15. If in Theorem 4.13 we assume X to be complete, then by Lemma 4.14 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 5.7 and 5.10).

- **Separation** Given disjoint subsets A and B of a vector space X, when 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.

Definition 5.1. 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 $x \in X$;
- $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$.

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] First some concepts regarding convex sets in vector spaces.

Definition 5.2. Let X be a vector space and $C \subset X$ convex, i.e. for $x, y \in C$, and $0 < \lambda < 1$, we have $\lambda x + (1 - \lambda)y \in C$. Say that C is:

[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.

[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.

- *absorbing* if for each $x \in X$ there exists $\lambda > 0$ such that $\lambda x \in C$.
- *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.*

Definition 5.3. Let C be a convex absorbing subset of a vector space X. The *Minkowski functional* p_C of C is defined by

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

Lemma 5.4. *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. □

Remark 5.5. 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 respeect 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 5.6 (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 $x \in X$.*

Remark 5.7. 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 5.6 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.

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

Lemma 5.8. *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 $g(y) \le p(y)$ 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 $f(z) \le p(z)$ 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 $f(z) \le p(z)$ for all $z \in Z$ is equivalent to $g(y) + c\lambda \le p(y + \lambda x_0)$ 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

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

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

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

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

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

and we may choose c to be any number between these two quantities.[25] □

If the space X in Theorem 5.6 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 5.6: 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 $f(z) \le p(z)$ 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 $D(f) \ne X$, then choose $x_0 \in X \setminus D(f)$, and extend f by Lemma 5.8 to some $f_1 \in \mathcal{P}$ defined on $\mathrm{Span}(D(f) \cup \{x_0\})$, contradicting maximality. □

If $\mathbb{F} = \mathbb{C}$ and functionals are complex-valued and complex-linear, then the sublinear functional appearing in Theorem 5.6 needs to be replaced by a seminorm. Surprisingly the following result was obtained around 10 years after the real version.

[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 exactly the relation that the graph of f_1 is a subset of the graph of f_2.

[28] Can you justify all these claims?

Theorem 5.9 (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 $|g(y)| \le p(y)$ for all $y \in Y$. Then there exists $f \in X'$ such that $f|_Y = g$ and $|f(x)| \le p(x)$ 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 $\operatorname{Re} f(x) \le p(x)$ for all $x \in X$ if and only if $|f(x)| \le p(x)$ for all $x \in X$. Indeed, one implication is trivial and for the other we note that if $\operatorname{Re} f(x) \le p(x)$ for all $x \in X$ then for some $\theta \in [0, 2\pi)$ depending on $x \in X$ we have

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

Thus given $g\colon Y \to \mathbb{C}$ as in the statement of the theorem, we may apply Theorem 5.6 to find a real-linear functional $f_0\colon X \to \mathbb{R}$ such that $f_0|_Y = \operatorname{Re} g$ and $f_0(x) \le p(x)$ for all $x \in X$. Now set $f(x) = f_0(x) - if_0(ix)$, $x \in X$. Then $f \in X'$ and $|f(x)| \le p(x)$, $x \in X$. 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. □

Remark 5.10. 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 5.9 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. First we introduce topological versions of anhilators as you might have seen in a purely algebraic setting in a linear algebra course.[29]

Definition 5.11. 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.[30] It is clear that annihilators are closed subspaces.

[29] The algebraic anhilator of a subset M of X is the set $\{f \in X' : f(x) = 0, \text{ for all } x \in M\}$ and the algebraic preanhilator defined similarly.

[30] Sometimes called the preannihilator. Be careful to distinguish between the annilhilator N° of N in X^{**} and the annihilator $N_\circ$ of N in X.

Corollary 5.12. *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. □

Note that part (b) is the topological version of the fact that if Y is a subspace of a vector space, then Y is equal to the algebraic preanhilator of its algebraic anhilator.[31] In the topological setting we must work with closed spaces, so note that its saying that if Y is a closed subspace of a normed space, then $Y = (Y^\circ)_\circ$.

We now turn to the separation versions of Hahn Banach: separating points from convex sets. The precise separation statements we can obtain depend on whether we work with open or closed convex sets.

Definition 5.13. 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 in the previous definition. Now Hahn-Banach spearation:

Theorem 5.14 (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

[31] You've probably seen this in finite dimensions; in the Oxford Part A linear algebra course this sort of result is hinted at, but not proved, right at the end of the lecture notes. You should take an (algebraic) complement Z, obtained through the axiom of choice.

$p(\lambda z_0) \geq 0 > g(\lambda z_0)$. 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 5.9 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. □

We now turn to the duality between embeddings and quotients in the setting of normed vector spaces: we are aiming for Corollary 5.18 which is again the correct normed space version of a fact you'll be familiar with in finite dimensional linear algebra. Recall that a bounded linear operator $T : X \to Y$ 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.

Definition 5.15. 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 $f \in Y^*$, $x \in X$.

Recall that the dual operator T^* of T is bounded, i.e. $T^* \in \mathcal{B}(Y^*, X^*)$. Moreover, by a standard application of the Hahn-Banach Theorem we have $\|T^*\| = \|T\|$.

Theorem 5.16. *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 4.13, 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 4.13 there exists $M > 0$ such that for every $g \in X^*$ there exists $f \in Y^*$ with $T^*f = g$ and $\|f\| \le M\|g\|$, and as observed in the proof of that result we may take any $M > r^{-1}$. Now by Corollary 5.12, 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 $\|f\| \le M$. Then

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

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, $\|Tx\| \ge \|x\|$ 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 $\|Tx\| \ge r\|x + \operatorname{Ker} T\|$ for all $x \in X$. By Theorem 4.13 there exists $M > 0$ such that for every $y \in Y$ there exists $x \in X$ with $Tx = y$ and $\|x\| \le M\|y\|$ and once again any $M > r^{-1}$ works. Given $y \in Y$, let $x \in X$ be as described. Then

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

and hence $\|f\| \le M\|T^*f\|$ 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 $\|f\| \le M\|T^*f\|$ for all $M > 1$. Hence $\|T^*f\| \ge \|f\|$ 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. $\square$

Remark 5.17. In the setting of Theorem 5.16 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 5.12 the closure of $\operatorname{Ran} T$ coincides with $(\operatorname{Ker} T^*)_\circ$.

Corollary 5.18. *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 5.16, $S^*\colon X^* \to Y^*$ 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 5.17 we have $\operatorname{Ran} \pi^* = (\operatorname{Ker} \pi)^\circ = Y^\circ$, so $(X/Y)^* \cong Y^\circ$. $\square$

We end this section by picking up the closed range theorem, which will be crucial in our later work on Fredholm operators.

Theorem 5.19 (Closed Range Theorem). *Let X and Y be Banach spaces and suppose that $T \in \mathcal{B}(X, Y)$. Then $\operatorname{Ran} T$ is closed if and only if $\operatorname{Ran} T^*$ 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 5.16 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 5.16 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. $\square$

Returning briefly to Remark 5.17. Let X and Y be normed spaces. Recall that we always have $(\operatorname{Ran} T^*)_\circ = \operatorname{Ker} T$ and $(\operatorname{Ran} T)^\circ = \operatorname{Ker} T^*$. In particular, by Corollary 5.12 the closure of $\operatorname{Ran} T$ coincides with $(\operatorname{Ker} T^*)_\circ$. In general even the closure of $\operatorname{Ran} T^*$ might be strictly contained in $(\operatorname{Ker} T)^\circ$, but in the setting of the closed range theorem we will have equality:[32]

Remark 5.20. For Banach spaces X, Y and $T \in \mathcal{B}(X, Y)$, if $\operatorname{Ran} T$ is closed, then $(\operatorname{Ran} T^*) = (\operatorname{Ker} T)^\circ$.

Proof. Continuing the proof of the closed range theorem, we have that when $\operatorname{Ran} T$ is closed, then Q is a quotient operator. Then, as noted in Remark 5.17 $\operatorname{Ran} Q^* = (\operatorname{Ker} Q)^\circ$. But (as in the proof of the closed range theorem) $\operatorname{Ran} T^* = \operatorname{Ran} Q^*$ and it is iimmediate that $\operatorname{Ker} T = \operatorname{Ker} Q$. $\square$

See also Remark 9.B later.

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 5.12 we have

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

so J_X is an isometry.

Definition 6.1. 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 6.2. *Every normed vector space X has a completion (Y, J) such that Y is a Banach space and J is linear.*

[32] Next year the following remark will be incorporated into the closed range theorem but I'm not keen to change the statement this year.

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. $\square$

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 $c_{00} = \operatorname{span}\{e_n : n \geq 1\}$ 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 may well have previously seen 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

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

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 6.3. 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. Reflexivity says that the natural map J_X implements an isometric isomorphism between X and X^{**}.

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 6.4. *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 $Y = \operatorname{span}\{x_n : n \geq 1\}$. 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 5.12 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. $\square$

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 6.5. *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 5.12 we see that Y is dense in X^{**}. Since Y is closed we have $Y = X^{**}$, so X is reflexive. $\square$

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 6.6. *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. $\square$

Remark 6.7. (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 5.16.

(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:

$$\begin{array}{ccc} X^{**} & \xrightarrow{T^{**}} & Y^{**} \\ {\scriptstyle J_X}\uparrow & & \uparrow{\scriptstyle J_Y} \\ X & \xrightarrow{T} & Y \end{array}.$$[33]

Reflexivity is a 2 from 3 property[34] for Banach spaces; i.e. it is inherited by closed subspaces and quotients, and also passes back from closed subspaces and quotients to the entire space.

Theorem 6.A. *Let X be a Banach space and Y a closed subspace of X. Then X is reflexive if and only if both Y and X/Y are reflexive.*

Proof. See problem sheet 2 (B.7 and C.1). $\square$

7 Convexity and smoothness of norms

Our objective over the next two sections is to show that the Lebesque spaces $L^p(\Omega, \Sigma, \mu)$ are reflexive for $1 < p < \infty$ — an important result in applications (as it shows via the results in Section 9, that the unit ball of these spaces is weakly compact). We will do this through examining norming vectors and functionals, via

[33] This is what is meant by saying that the map J_X is natural. In the part A linear algebra course 'naturality' was described in terms of 'can be defined without reference to a basis', but it in fact naturality has a precise definition in a very general setting. Indeed, in the language of category theory J provides a natural transformation from the identity functor to the functor **.

[34] i.e. if 2 out of the 3 times in a short exact sequence enjoy the property, so too does the third.

geometric properties of the norm / unit ball of a Banach space. We set out the abstract theory here, and return to the Lebesgue spaces in the next section.

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 $x \neq y$.

Definition 7.1. 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 $\delta_X \colon [0,2] \to \mathbb{R}$ 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 $\|x-y\| \geq \varepsilon$.

Definition 7.2. 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. 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 $n \to \infty$.[35]

Theorem 7.3. *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 $f(x) = \|f\|$.*

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

[35] 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 $\delta_X(\epsilon) = 0$.

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 7.4. Note that for the uniqueness part strict convexity was sufficient. In fact, it is not hard to show using Corollary 5.12 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 10.3 below.

We call the vector x above a *norming* vector for f. Precisely:

Definition 7.5. Let X be a normed space and $f \in X^*$. Say that $x \in X$ *norms* $f \in X^*$ if $\|x\| = 1$ and $f(x) = \|f\|$.

Thus in a uniformly convex space every non-zero functional has a unique norming vector.

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 5.12) 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 $f(x_0) = \|x_0\|$.[36] 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.

Definition 7.6. 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 section.

Theorem 7.7. *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

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

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

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

Proof. As observed above, existence of norming functionals follows from Corollary 5.12.

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

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

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

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

For the final statement, note that uniqueness of norming functionals for a vector is invariant under non-zero scalar multiplication.[37] $\square$

We will see on Problem Sheet 3 that uniqueness of norming functionals in fact characterises smooth spaces.[38] 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.

Moreover, in the extension exercises we will learn that every uniformly convex Banach space is reflexive. This is not the case for strictly convex Banach spaces.

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

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

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

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

[37] 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.

[38] so given the $p > 1$ version of Theorem 7.7, in fact the norm is smooth.

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 $\Omega = \bigcup_{n\ge1} \Omega_n$. 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

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

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. □

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 $x, y \in \mathbb{R}$.[39] It is strictly convex if equality only holds when $x = y$. Recall too that differentiable convex functions have increasing derivative. For use in Theorem 8.3, if $\phi : \mathbb{R} \to \mathbb{R}$ is convex and differentiable, then the function

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

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

is continuous[40] 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 $\phi'(x) = pb|a + bx|^{p-1} \operatorname{sgn}(a + bx)$.[41]

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

$$\begin{aligned}\|u_n + v_n\|_p^p + \|u_n - v_n\|_p^p &= \int_\Omega \frac{\big||x_n + y_n| + |x_n - y_n|\big|^p}{2^p} + \frac{\big||x_n + y_n| - |x_n - y_n|\big|^p}{2^p} \,\mathrm{d}\mu \\ &= \int_\Omega \max\big\{|x_n|, |y_n|\big\}^p + \min\big\{|x_n|, |y_n|\big\}^p \,\mathrm{d}\mu \\ &= \|x_n\|_p^p + \|y_n\|_p^p \to 2, \quad n \to \infty.\end{aligned}$$

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

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

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

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

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. □

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 $\|f\| = 1$.*

Proof. Certainly $y_0 \in L^q(\Omega)$. By Theorem 7.7 we need to show that

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

[40] by differentiability

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

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[42]

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

as required.[43] $\square$

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 $f \in L^p(\Omega)^*$ have $\|f\| = 1$. By Theorem 8.2 the space $L^p(\Omega)$ is uniformly convex, and hence by Theorem 7.3 there exists a (unique) function $x_0 \in L^p(\Omega)$ such that $\|x_0\|_p = 1$ and $f(x_0) = 1$. 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

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

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. $\square$

9 The weak and weak* topologies

Definition 9.1. 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 $x_0 \in X$, $n \in \mathbb{N}$, $f_1, \dots, f_n \in Y$ and $\varepsilon > 0$.

If X is a normed vector space, the case $Y = X^*$ is of particular interest.

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

[43] I retain David Seifert's alternative argument for this using the dominated convergence theorem: As the function $G\colon \mathbb{R} \to \mathbb{R}$ given by $G(h) = |h|^p$ is differentiable with derivative $G'(h) = p|h|^{p-1}\operatorname{sgn} h$, for each $t \in \Omega$ and $h \in \mathbb{R}$ we may apply the Mean-Value Theorem to obtain $\theta = \theta(t, h) \in (0, 1)$ such that

$$|x_0(t) + hx(t)|^p = |x_0(t)|^p + p|x_0(t) + \theta hx(t)|^{p-1}hx(t)\operatorname{sgn}(x_0(t) + \theta hx(t)).$$

Note also that for $|h| \le 1$ we have $|x_0 + \theta hx|^{p-1}|x| \le (|x_0| + |x|)^p \in L^1(\Omega)$. Hence the Dominated Convergence Theorem gives

$$\frac{\|x_0 + hx\|_p^p - \|x_0\|_p^p}{ph} = \int_\Omega |x_0 + \theta hx|^{p-1} x \operatorname{sgn}(x_0 + \theta hx)\,\mathrm{d}\mu \to (\Phi_p y_0)(x), \quad h \to 0.$$

Definition 9.2. Let X be a normed space. We call $\sigma(X, X^*)$ the *weak topology* on X. The *weak* topology* on X^* is the 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

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

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 $\sigma(X^*, X) \subseteq \sigma(X^*, X^{**})$. 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 9.9 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 can 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^*.

Definition 9.3. 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 9.4. *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.[44] In general the weak and weak* topologies are not metrisable (see Problem Sheet 3) but we have the following result.

Proposition 9.5. *Let X be a normed vector space.*

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

[44] In the generality of topological spaces it is not the case that the topology is determined by sequential convergence. Indeed sequential compactness is not generally equivalent to compactness. In the setting of topological spaces we sometimes use various generalisations of sequences in order to describe topologies. In the Part C Analytic topology course, filters are used for this. Most functional analysts prefer to use convergence of *nets* to play this role - this is essentially equivalent to working with filters- one then gets results like a topological space X is compact if and only if every net in X has a convergent subnet. The standard reference for this is Kelley's 'General Topology'. Terminology warning: The nets used to generalise convergence sequences are not the ϵ-nets found in the next section!

(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 9.6. The converse statements in Proposition 9.5 are also true; see Remark 9.13 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 9.7. *Let X and Z be vector spaces and suppose that Y is a subspace of X' and W is a subspace of Z'.*

(1) *A functional $f \in X'$ is $\sigma(X,Y)$-continuous if and only if $f \in Y$.*
(2) *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. (1) 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 $n \in \mathbb{N}$, $f_1, \dots, f_n \in Y$ and $\varepsilon > 0$ such that the basic $\sigma(X,Y)$-open set $V = \{x \in X : |f_k(x)| < \varepsilon \text{ for } 1 \leq k \leq n\}$ 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 $f \in \operatorname{span}\{f_1, \dots, f_n\} \subseteq Y$ by a result on Problem Sheet 1.
(2) 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 (1) is equivalent to having $T'(W) \subseteq Y$, as required. □

Note that in particular the previous result says that if X is a normed space then a functional $f \in X'$ is weakly continuous if and only if $f \in X^*$. So the weakly continuous functionals are precisely the continuous functionals. See exercise sheet 3 for an extension to maps.

Corollary 9.8. *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 5.18 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 5.18 gives $(X/Z_\circ)^* \cong Z$. In particular, $X/Z_\circ$ must be finite-dimensional. The proof of Proposition 9.7 shows that any basic weakly open neighbourhood of the origin contains a finite-codimensional subspace, and an analogous argument works for the weak* topology.[45] As such, for infinite dimensional spaces the weak and (weak*)topologies are genuinely weaker than the norm topology.

Proposition 9.9. *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 9.7 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 $f_1, \dots, f_n \in X^*$. 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. $\square$

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.[46]

Theorem 9.10 (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.*

We use Tychonoff's theorem to obtain weak*-compactness of the unit ball.

Theorem 9.11 (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

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

[45] Note that in the case when X is finite dimensional, this finite codimensional subspace could be the zero subspace.

[46] It can also be proved using nets, when the proof is reminiscent of the diagonal argument used to prove that a countable product of compact metric spaces is compact, albeit with a Zorn's lemma maximality argument in place of a diagonal sequence argument.

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. $\square$

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 9.12. *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 5.12 it is an isometry. $\square$

Remark 9.13. By Proposition 9.5(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.[47] Since $C(\Omega)$ is separable whenever Ω is a compact metric space we see that Corollary 9.12 implies the converse of Proposition 9.5(a).

Now we turn to Hahn-Banach for the weak and weak*-topology. In the case of the weak topology, the Hahn-Banach theorem shows that convex sets have the same weak and norm closure.[48]

Theorem 9.A. *Let X be a normed vector space and suppose that C is a non-empty convex subset of X. Then the weak and norm closures of C are equal.*

Proof. See Sheet 3 B.5. $\square$

Theorem 9.14 (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*

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

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

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

[47] One way to obtain this is via the Hausdorff-Alexandroff theorem that every compact metric space is a continuous image of the Cantor set Z (a number of self contained short proofs of this can be found online). Given a continuous surjection $Z \to \Omega$ we obtain an isometric embedding $C(\Omega) \to C(Z)$. So to complete the argument we just need an isometric embedding $T : C(Z) \to C[0, 1]$. We get this by regarding the Cantor set Z as the usual middle thirds Cantor set. Given $f \in C(Z)$, we extend it to Tf on $[0, 1]$ by interpolating over the removed intervals as follows. For each $t \in [0, 1] \setminus Z$, t lies in a unique removed middle third interval (a_t, b_t), where $a_t = \sup\{a \in Z : a < t\}$ and $b_t = \inf\{b \in Z : b > z\}$, and we can write $t = \lambda a_t + (1 - \lambda_t b_t)$ for $0 < \lambda_t < 1$. Then define $(Tf)(t) = \lambda_t f(a_t) + (1 - \lambda_t) f(b_t)$. This is readily checked to give an isomorphic embedding $C(Z) \to C[0, 1]$.

[48] This is perhaps my personal favourite version of the Hahn-Banach theorem; it's certainly the one I've used the most in my own research.

Proof. (a) Since C is in particular open in the norm topology, Theorem 5.14 implies the existence of a $\phi \in X^{**}$ such that $\operatorname{Re}\phi(f_0) > \operatorname{Re}\phi(f)$ 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

$$U = \{f \in X^* : |f(x_k)| < \varepsilon \text{ for } 1 \le k \le n\}$$

for some $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $\varepsilon > 0$. Then in particular

$$\operatorname{Re}\phi(f_0) > \operatorname{Re}\phi(g_0) + \operatorname{Re}\phi(f), \quad f \in U.$$

If $Y = \bigcap_{k=1}^{n} \operatorname{Ker} J_X(x_k)$, then $Y \subseteq U$ and because Y is a vector space we must have $Y \subseteq \operatorname{Ker}\phi$. It follows that $\phi \in \operatorname{span}\{J_X(x_1), \dots, J_X(x_n)\} \subseteq J_X(X)$, as required.

(b) Since C is in particular closed in the norm topology, Theorem 5.14 implies the existence of a $\phi \in X^{**}$ such that $\operatorname{Re}\phi(f_0) > \sup_{f \in C} \operatorname{Re}\phi(f)$. 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 $\phi \in J_X(X)$. $\square$

Corollary 9.15. *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

$$(Y_\circ)^\circ = \bigcap \{ \operatorname{Ker} J_X(x) : x \in Y_\circ \}$$

is weak*-closed, so $C \subseteq (Y_\circ)^\circ$. Suppose that $f_0 \in X^* \setminus C$. By Theorem 9.14 there exists $x \in X$ such that $\operatorname{Re} f_0(x) > \sup_{f \in C} \operatorname{Re} f(x)$. 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 $f_0(x) \neq 0$ we deduce that $f_0 \notin (Y_\circ)^\circ$, so $(Y_\circ)^\circ \subseteq C$, as required. $\square$

In particular for a normed space X, $J_X(X)$ is weak*-dense in X^{**}, as $J_X(X)_\circ = \{0\}$. However this is not particularly useful, as it doesn't allow one to control norms, that is if we approximate operators in X^{**} in the weak*-topology by those in $J_X(X)$, we want to do so with norm control. Goldstine's theorem enables to do this.

Theorem 9.16 (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 9.14 there exists $f \in X^*$ such that

$$\operatorname{Re}\phi(f) > \sup\{\operatorname{Re}\psi(f) : \psi \in C\} \ge \sup\{\operatorname{Re} f(x) : x \in B_X\} = \|\operatorname{Re} f\| = \|f\|.$$

It follows that $\|\phi\| > 1$, so $B_{X^{**}} \subseteq C$ and the result follows. $\square$

Theorem 9.17. *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. $\square$

We can also use the weak*-Hahn Banach separation theorem to extend the characterisations of closed range for adjoint operators.

Proposition 9.B. *Let X, Y be Banach spaces, and $T \in \mathcal{B}(X, Y)$. Then T^* has closed range if and only if $\operatorname{Ran} T^*$ is weak*-closed.*

Proof. If $\operatorname{Ran} T^*$ is weak*-closed it is closed. Conversely if $\operatorname{Ran} T^*$ is closed, then by Remark 5.20, $\operatorname{Ran} T^* = (\operatorname{Ker} T)^\circ$. Since $(\operatorname{Ker} T)$ is a closed subspace, $\operatorname{Ker} T = ((\operatorname{Ker} T)^\circ)_\circ$ by Corollary 5.12, and hence $((\operatorname{Ran} T^*)_\circ)^\circ = \operatorname{Ran} T^*$. Thus $\operatorname{Ran} T^*$ is weak*-closed by Corollary 9.15. $\square$

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* 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.[49] We say that X is *sequentially compact* if every sequence with terms in X has a subsequence which converges to an element of X.

You may well be familiar with the following characterisations of compactness in a metric space setting (which here is found spread between the part A metric spaces and complex analysis course and the part A topology course). It is this result that enables us to use sequence arguments to work with compact metric spaces — this will be the approach we take over the next part of the course.

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.*

[49] This is not the same usage of the word 'net' as in the net approach to filter convergence which most functional analysts use to work with general topological spaces.

Note that for general topological spaces both the implications (a)⇒(c) and (c)⇒(a) fail.[50]

The set M is said to be *relatively compact* (or pre-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. The previous theorem characterises relatively compact subsets of complete metric spaces

Corollary 10.2. *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.*

Remark 10.3. 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 9.5, so by Theorem 10.1 it is compact if and only if it is sequentially compact. It follows from Theorem 9.17 that X is reflexive if and only if every bounded sequence has a weakly convergent subsequence.[51]

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. Our next goal is to show that for infinite dimensional spaces the unit ball is never totally bounded, so is not compact. The key step is Riesz's lemma.

Lemma 10.4 (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* $\operatorname{dist}(x, Y) \geq 1 - \delta$.

Proof. By Corollary 5.12 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 $\operatorname{dist}(x, Y) \geq 1 - \delta$, as required. □

Remark 10.5. The conclusion of Lemma 10.4 in general becomes false if we allow $\delta = 0$. However, using Theorem 7.3 we may show the result remains valid even with $\delta = 0$ if X is a uniformly convex Banach space, and by a simple application of the Heine-Borel Theorem the same is true when Y is finite-dimensional.

Theorem 10.6. *Let X be a normed vector space. Then B_X is totally bounded if and only if X is finite-dimensional.*

[50] An uncountable product of copies of the unit interval is compact (by Tychonoff) but not sequentially compact; the interval $(0, \omega_1)$ equipped with the order topology, where ω_1 is the first uncountable ordinal, is sequentially compact but not compact. A good reference for this sort of stuff is 'Counterexamples in Topology' by Steen and Seebach.

[51] In fact, this statement is true even without the assumption that X^* is separable, and this follows from the *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.

Proof. It remains to show that if X is infinite dimensional then B_X is not totally bounded. So let X be infinite dimensional. We will construct a sequence $(x_n)_{n=1}^{\infty}$ in B_X such that $\|x_n - x_m\| \geq 1/2$ for all $n \neq m$. Choose x_1 arbitrarily in B_X, and suppose $x_1, \ldots, x_n$ have been found. Let $Y = \operatorname{span}(x_1, \ldots, x_n)$ (a proper closed subspace of X) and use Riesz's lemma to find x_{n+1} such that $\operatorname{dist}(x_{n+1}, Y) \geq 1/2$. $\square$

We now turn to the Arzelà-Ascoli theorem which characterises totally bounded subsets, or equivalently of relatively compact subsets, in $C(\Omega)$ for compact spaces Ω.[52]

Definition 10.7. Given a topological space Ω and a subset $\mathcal{F}$ of $C(\Omega)$, we say that $\mathcal{F}$ is *equicontinuous* if for every $\varepsilon > 0$ and $t \in \Omega$ there exists an open neighbourhood U of t in Ω such that $|f(s) - f(t)| < \varepsilon$ for all $f \in \mathcal{F}$ and all $s \in U$.

The following example of equicontinuous functions will be important to us in the proof of Schauder's theorem in the next section.

Example 10.8. Let X be a normed space, and let $\mathcal{F} \subset Y^*$, which we view as a subset of $C(X)$. Then $\mathcal{F}$ is equicontinuous if $\mathcal{F}$ is a bounded subset of Y^*.

Proof. Let $K > 0$ be such that $\|f\| \leq K$ for $f \in \mathcal{F}$. Then for $\epsilon > 0$, if $x, y \in X$ have $\|x - y\| < \epsilon/K$, then $|f(x) - f(y)| < \epsilon$ for all $f \in \mathcal{F}$. Thus $B_X^\circ(x, \epsilon/K)$ provide neighbourhoods witnessing equicontinuity. $\square$

Theorem 10.9 (Arzelà-Ascoli). *Let Ω be a compact topological space and suppose that $\mathcal{F}$ is a subset of $C(\Omega)$. Then $\mathcal{F}$ is relatively compact if and only if it is bounded and equicontinuous.*

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

$$|f(s) - f(t)| \leq |f(s) - g(s)| + |g(s) - g(t)| + |g(t) - f(t)| < \varepsilon.$$

Thus $\mathcal{F}$ is equicontinuous.

Conversely, suppose that $\mathcal{F}$ is bounded and equicontinuous, and let $\varepsilon > 0$. Since $\mathcal{F}$ is equicontinuous there exists, for each $t \in \Omega$, an open neighbourhood U_t of t in Ω such that $|f(s) - f(t)| < \varepsilon/3$ for all $f \in \mathcal{F}$ 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

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

and let $S = T(\mathcal{F})$.[53] Then by boundedness of $\mathcal{F}$ the set S is a bounded, and therefore totally bounded, subset of $(\mathbb{F}^n, \|\cdot\|_\infty)$. Let $F \subseteq \mathcal{F}$ be a finite set such

[52]Some of you may have seen this for metric spaces right at the end of the metric spaces course; we will do it here in the slightly more general setting of compact topological spaces,

[53]Notice the technique in this proof of approximating $\mathcal{F}$ by $T(\mathcal{F})$, in a finite dimensional space. While part of a '3ϵ-argument here, finite dimensional approximations of this nature are ubiquitous in functional analysis.

that $T(F)$ is an $\frac{\varepsilon}{3}$-net for S and suppose that $f \in \mathcal{F}$. Then there exists $g \in F$ such that $|f(t_k) - g(t_k)| < \varepsilon/3$ for $1 \le k \le n$. Given $t \in \Omega$ we have $t \in U_{t_k}$ for some $k \in \{1, \dots, n\}$ and hence

$$|f(t) - g(t)| \le |f(t) - g(t_k)| + |f(t_k) - g(t_k)| + |f(t_k) - g(t)| < \varepsilon.$$

Thus $\mathcal{F}$ is totally bounded. □

Example 10.10. Given $k \in L^1(0,1)$, consider the the set

$$\mathcal{F} = \{\Phi(f) : f \in C([0,1]),\ \|f\|_\infty \le 1\},$$

where

$$\Phi(f)(t) = \int_0^t k(s) f(s)\,\mathrm{d}s.$$

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.[54] Hence $\mathcal{F}$ is a bounded equicontinuous subset of $C([0,1])$. By the Arzelà-Ascoli Theorem $\mathcal{F}$ is totally bounded in $C([0,1])$.

In $L^p(\mathbb{R}^n)$, totally bounded sets are characterised by the Kolmogorov-Riez-Fréchet theorem, which is proved in C4.3 (Functional Analytic Methods for PDEs).

Theorem 10.11 (Kolmogorov-Riesz-Fréchet). *Let $1 \le p < \infty$ and $n \ge 1$, and suppose that $M \subseteq L^p(\mathbb{R}^n)$. Then M is relatively compact if and only if M is bounded,*

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

uniformly over $x \in M$.

11 Compact operators

Definition 11.1. 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. We write $\mathcal{K}(X,Y)$ for the set of compact linear operators $T \colon X \to Y$ and we let $\mathcal{K}(X) = \mathcal{K}(X,X)$.

Translating this into sequences, 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.Notice that if T is compact then the closure of $T(B_X)$ is in particular bounded, so T is bounded and hence $\mathcal{K}(X,Y) \subseteq \mathcal{B}(X,Y)$.

[54] This is the absolute continuity of the measure $\mu(A) = \int_A |k(s)|\,\mathrm{d}s$ with respect to Lebesgue measure. One way to see this is as follows. Suppose that there exists $\epsilon > 0$ and Lebesgue measurable sets $A_n \subset [0,1]$ with $|A_n| < 2^{-n}$ (where $|A_n|$ denotes the Lebesgue measure of A_n), yet $\int_{A_n} |k(s)|\,\mathrm{d}s \ge \epsilon$. Let $B_n = \cup_{m=n}^\infty A_n$ so $|B_n| \le 2^{-n+1}$ and so $B = \cap_{n=1}^\infty B_n$ has $|B| = 0$. On the other hand, since $\mu(B_1) \int_{B_1} |k(s)|\,\mathrm{d}s < \infty$ and $B_1 \supseteq B_2 \supseteq B_3 \supseteq \dots$ it follows that $\mu(B) = \lim_n \mu(B_n) \ge \epsilon$, which gives a contradiction.

Example 11.2. (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 10.6 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 10.10 the integral operator

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

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

Proposition 11.3. *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 $\mathcal{B}(X,Y)$.[55]

It remains to show that it is closed if Y is complete. Suppose that $(T_n)_{n=1}^\infty$ is a sequence in $\mathcal{K}(X,Y)$ converging to $T \in \mathcal{B}(X,Y)$. Let $(x_m)_{m=1}^\infty$ be a bounded sequence in X. By a diagonal sequence argument, we can find a subsequence $(y_m)_{m=1}^\infty$ of $(x_n)_{n=1}^\infty$ such that $(T_n(y_m))_{m=1}^\infty$ converges for all $n \in \mathbb{N}$.[56] We claim $(T(y_m))_{m=1}^\infty$ is Cauchy. Indeed let $K = \sup_m \|x_n\|$, and given $\epsilon > 0$ find n such that $\|T - T_n\| \le \epsilon/(3K)$. Then find m_0 such that for $m_1, m_2 \ge m_0$, $\|T_n(y_{m_1}) - T_n(y_{m_2})\| \le \epsilon/3$. Then for $m_1, m_2 \ge m_0$, $\|T(y_{m_1}) - T(y_{m_2})\| \le \epsilon$, as claimed. As Y is complete, T is compact.
(b) Let (w_n) be a bounded sequence in W. Then (Sw_n) is a bounded sequence in X, so has a subsequence (Sw_{n_k}) such that (TSw_{n_k}) converges. As R is continuous $(RTSw_{n_k})$ converges, so RTS is compact. □

Corollary 11.4. *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 11.5. 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 11.4 does hold on many spaces.[57]

We now look the other main operation on operators: taking the dual. Schuader's theorem shows that this preserves compactness.

---

[55] Alternatively one can prove this using a subsequence argument, as in lectures,

[56] In more detail, use compactness to find a subsequence $(x_m^{(1)})_m$ of (x_m) such that $(T_1(x_m^{(1)}))_m$ converges, then use compactness to find a subsequence $(x_m^{(2)})_m$ of $(x_m^{(1)})$ such that $(T_2(x_m^{(2)}))_m$ converges. Carry on in this way and set $y_m = x_m^{(m)}$; the m-th element of the m-th subsequence.

[57] In particular it holds for all spaces with a Schauder basis, as in Section 14, so all classical sequence spaces.

Theorem 11.6 (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. Let $T \in \mathcal{K}(X,Y)$. To show $T^* : Y^* \to X^*$ is compact, fix a bounded sequence $(f_n)_{n=1}^\infty$ in Y^*. As T is compact $\overline{T(B_X)}$ is compact in Y. Since (f_n) is a bounded sequence, this is an equicontinuous family of continuous functions on Y (as discussed in Example 10.8). Therefore by the Arzelà-Ascoli theorem, the restrictions $(f_n|_{\overline{T(B_X)}})$ of the f_n to $\overline{T(B_X)}$ are totally bounded in $C(\overline{T(B_X)})$. Thus there is a subsequence $(f_{n_k})_k$ such that $f_{n_k}|_{\overline{T(B_X)}}$ converges uniformly to some $g \in C(\overline{T(B_X)})$. In particular

$$\sup_{x \in B_X} |f_{n_k}(Tx) - f_{n_l}(Tx)| \to 0,$$

as $k, l \to \infty$. That is

$$\|T^*(f_{n_k} - T^*(f_{n_l})\| \to 0,$$

as $k, l \to \infty$, so $(T^* f_{n_k})_k$ is Cauchy so converges in the complete space X^*. Therefore T^* is compact.

Suppose now that T^* is compact and Y is complete. Then by the first part the operator $T^{**} : X^{**} \to Y^{**}$ 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 10.2, $T(B_X)$ 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 $y \in \operatorname{Ran} T$, and we know from Remark 5.17 that the closure of $\operatorname{Ran} T$ coincides with $(\operatorname{Ker} T^*)_\circ$. In particular, if $\operatorname{Ran} T$ is closed then we have a criterion for our problem to have a solution, namely that $f(y) = 0$ for all $f \in \operatorname{Ker} T^*$.

Definition 12.1. 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. If T is Fredholm we define the *index* of T as

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

We know from Problem Sheet 1 (Question C.3) that any Fredholm operator must have closed range.

Example 12.2. (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 12.3. *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 in this case* $\operatorname{ind} T + \operatorname{ind} T^* = 0$.

Proof. Exercise; see Problem Sheet 4. □

Theorem 12.4. *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*

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

Before giving the proof, recall that problem sheet 2, exercise B.1, shows that if X is a normed space, and Y is a finite dimensional subspace of X, then if Z is any closed subspace of X such that $X = Y \oplus Z$ algebraically, then $X = Y \oplus Z$ topologically.

Proof. Suppose that T is Fredholm. As $\operatorname{Ker} T$ is finite dimensional, there is a closed subspace V of X such that $X = \operatorname{Ker} T \oplus V$ topological by Problem 2.B.3(a) (the first part of this problem). Moreover, T has closed range (by Problem 1.C.3) so we can fix a finite dimensional subspace W which is an algebraic complement of $\operatorname{Ran} T$ so that $Y = \operatorname{Ran} T \oplus W$ topologically by Problem 2.B.3(b). 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. □

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 5.18 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 5.18 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 $g_k(x) = 0$, $1 \le k \le m$, where $\{g_k : 1 \le k \le m\}$ is a basis for V°.

Recall that the collection of isomorphisms between Banach spaces X and Y is open in $\mathcal{B}(X,Y)$. Indeed, 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

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

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 12.5. *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 12.4 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 $Q_S \colon V \times W \to Y$ be given by $Q_S(x,y) = Tx + Sx + y$ for $x \in V$, $y \in W$. Then Q_0 is an isomorphism and $\|Q_0 - Q_S\| = \|S|_V\| \le \|S\|$, $S \in \mathcal{B}(X,Y)$. 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 12.4 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. □

Remark 12.6. Note that by the above proof we have $\operatorname{ind}(T+S) = \operatorname{ind} T$ and $\dim Y/\operatorname{Ran}(T+S) \le \dim Y/\operatorname{Ran} T$ for S of sufficiently small norm. It follows that for such operators S we also have $\dim \operatorname{Ker}(T+S) \le \dim \operatorname{Ker} T$.

We now reach the key perturbation theorem: compact perturbations of Fredholm operators are Fredholm with the same index.

Theorem 12.7. *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 12.4 there exists a closed finite-codimensional subspace V of X such that $T|_V$ is an isomorphism from V onto $\operatorname{Ran}(T)$. 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 12.3, so the above argument shows that $\operatorname{Ker}(T^*+S^*)$ is also finite-dimensional. We aim to show that $\operatorname{Ran}(T+S)$ is closed, so that it will be equal to $\operatorname{Ker}(T^*+S^*)_\circ$.

As V is finite co-dimensional in X, it suffices to show that $(T+S)(V)$ is closed.[58] As $\operatorname{Ker}(T+S) \cap V$ is finite dimensional, we can write $V = (\operatorname{Ker}(T+S) \cap V) \oplus V_1$ for some closed subspace V_1 in V, and note that $T+S$ is injective on V_1. We will show that $(T+S)$ is bounded below on V_1. Suppose not, then there would exist a sequence $(x_n)_{n=1}^{\infty}$ of unit vectors in V_1 such that $(T+S)(x_n) \to 0$. By compactness of S, we may assume (after passing to a subsequence) that $S(x_n) \to y \in Y$ say. Then $Tx_n = (T+S)(x_n) - S(x_n) \to -y$. Since T is an isomorphism from V to $T(V)$ (which is closed in Y) it follows that $x_n \to -T^{-1}(y) = z \in V_1$ (as V_1 is closed) and as $\|x_n\| = 1$, we have $\|z\| = 1$. Then $(T+S)(z) = \lim (T+S)(x_n) = 0$, a contradiction as $T+S$ is injective on V_1. It then follows that $(T+S)(V_1) = (T+S)(V)$ is closed.[59]

For $0 \leq t \leq 1$ let $Q_t = T + tS$. Then each Q_t is Fredholm and Theorem 12.5 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. □

Corollary 12.8 (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.*
- $\operatorname{Ran}(\lambda - T) = \operatorname{Ker}(\lambda - T^*)_\circ$.
- $\dim \operatorname{Ker}(\lambda - T^*) = \dim \operatorname{Ker}(\lambda - T)$ *and both numbers are finite.*

Proof. For $\lambda \neq 0$ the operator λI is Fredholm with index zero, so by Theorem 12.7 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 5.18 we have

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

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

Note that this theorem can be applied, for example, to the compact integral operators we gave in Example 11.2.

13 Spectral Theory for compact operators

We now turn to look at some spectral theory for compact operators. The spectrum is the appropriate generalisation of eigenvalues to operators on infinite dimensional spaces, and an instrumental tool for working with operators. First, let's recall the definition of the spectrum of an operator.

[58] As $(T+S)(V)$ is finite codimensional closed subspace of $(T+S)(X)$, it is complemented in $(T+S)(X)$, so we can write $(T+S)(X) = (T+S)(V) \oplus W_0$ as a topological direct sum for some finite dimensional W_0. Since $(T+S)(V)$ is closed in Y it is complete, and W is complete, so $(T+S)(X)$ is complete, so closed in Y.

[59] This is a standard Banach space argument: suppose that $(T+S)(x_n) \to y$ say for $x_n \in V_1$. Then $((T+S)(x_n))_{n=1}^{\infty}$ is Cauchy, and as $(T+S)$ is bounded below on V_1, so $(x_n)_{n=1}^{\infty}$ is Cauchy, so converges to $x \in V_1$ say. Then $(T+S)(x_n) \to (T+S)(x)$, so $y = (T+S)(x) \in (T+S)(V_1)$.

Definition 13.1. Let X ne a complex Banach space and $T \in \mathcal{B}(X)$, then the *spectrum* of T is

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

Recall that the spectrum is a non-empty compact subset of $\mathbb{C}$.[60]. Recall too that the *spectral radius* $r(T) = \sup_{\lambda\in\sigma(T)} |\lambda|$ satisfies $r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$. [61]

Definition 13.2. 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.[62] However, for compact operators we at least have $\sigma(T) \setminus \{0\} \subseteq \sigma_p(T)$, as we now show.

Theorem 13.3. *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. Suppose $\lambda \neq 0$ is not an eigenvalue of T. We claim that there exists $r > 0$ such that $\|(T - \lambda)x\| \geq r\|x\|$ for all $x \in X$. If this fails, find a sequence $(x_n)_{n=1}^\infty$ of unit vectors in X with $(T - \lambda)x_n \to 0$. Passing to a subsequence, we can assume $Tx_n \to y$, so $x_n \to y/\lambda$. Since each x_n is a unit vector, $y \neq 0$, so y is an eigenvector with eigenvalue λ, giving the required contradiction.

Then $\mathrm{Ran}(T - \lambda)$ is complete, so closed, and hence $T - \lambda : X \to \mathrm{Ran}(T - \lambda)$ is an isomorphism by the open mapping theorem. If $T - \lambda$ is not surjective then $X_1 = \mathrm{Ran}(T - \lambda)$ is a proper closed subspace of X, and in general $X_{n+1} = (T - \lambda)(X_n)$ is a proper closed subspace of X_n. Then by Reisz's lemma, choose a sequence of unit vectors $(x_n)_{n=1}^\infty$ in X_n such that $d(x_n, X_{n+1}) \geq 1/2$. Note that for $n > m$,

$$\|Tx_m - Tx_n\| = \|\lambda x_m + (T - \lambda x_m) - (T - \lambda)x_n + \lambda x_n\| \geq |\lambda| d(x_m, X_{m+1}) \geq |\lambda|/2$$

as $(T-\lambda)x_m, x_n, (T-\lambda)x_n \in X_{m+1}$. Therefore $(Tx_n)_{n=1}^\infty$ does not have a convergent subsequence. Accordingly $T - \lambda$ is surjective, so $\lambda \notin \sigma(T)$.

[60] It is cruicial that we work with complex scalars here, as this result relies on complex analysis, normally through Louivilles' Theorem. Indeed, over $\mathbb{R}$ there are 2×2 matrices with no eigenvalues.

[61] These results are sketched in Example Sheet 0. Depending on your previous courses some of you may have only seen these results when X is a Hilbert Space; but the Banach space case works essentially identically. The tiny difference between the two cases is that in a Hilbert space setting, we have plenty of functionals on $\mathcal{B}(X)$ (given by maps of the form $T \mapsto (Tx, y)$ for $x, y \in X$). In the Banach space setting it is necessary to use the Hahn-Banach theorem to produce enough functionals to learn that every bounded analytic function $\mathbb{C} \to \mathcal{B}(X)$ is constant.

In fact the general framework in which the spectrum can be defined is that of a *Banach algebra*. A Banach algebra $\mathcal{A}$ is a Banach space, equipped with an associative multiplication which distributes over addition (so the addition and multiplication make $\mathcal{A}$ a ring) which satisfies $\|ab\| \leq \|a\|\|b\|$ for all $a, b \in \mathcal{A}$. We call $\mathcal{A}$ unital if there is an element $I \in \mathcal{A}$ with $Ia = aI = a$ for all $a \in \mathcal{A}$. The bounded operators $\mathcal{B}(X)$ on a Banach space form a key example of a Banach algebra. All the work on the spectrum of an operator in $\mathcal{B}(X)$ goes through mutais mutandis in the setting of a unital Banach alegebra (in the non-unital case we add an identity).

[62] Indeed there are many bounded operators with no eigenvalues.

A very similar subsequence argument then shows that the subspace of E spanned by all eigenvectors of T corresponding to eigenvalues in $\{\lambda \in \mathbb{C} : |\lambda| \geq r\}$ is finite-dimensional for each $r > 0$.[63] This shows that the non-zero eigenvalues have finite dimensional eigenspaces, and that all non-zero eigenvalues are isolated. Moreover, for each $n \in \mathbb{N}$, there are at most finitely many elements in $\sigma(T)$ with $|\lambda| > 1/n$. Hence $\sigma(T)$ is countable.

If $\sigma(T)$ is infinite, then, as $\sigma(T)$ is a compact metric space, it must have a limit point, which must be 0. So 0 is the unique limit point of $\sigma(T)$.

Finally note 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 $B_X(r)$ is a closed subset of the compact set $\overline{T(B_X)}$, so compact. Then Theorem 10.6 forces X to be finite-dimensional. Thus $0 \in \sigma(T)$ whenever X is infinite-dimensional. $\square$

We now turn to the Hilbert space setting, aiming for the spectral theorem for compact self-adjoint operators. This is a generalisation of the diagonalisation of hermitian complex matrices. For the rest of this section, let X be a complex Hilbert space with inner product $(\cdot\,,\cdot)$.[64]

Definition 13.4. For $T \in \mathcal{B}(X)$, recall that the *adjoint operator* $T^\star \in \mathcal{B}(X)$ of T is defined by the identity $(Tx, y) = (x, T^\star y)$, $x, y \in X$.[65] Recall too that T is said to be *self-adjoint* if $T^\star = T$.

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:

$$\begin{tikzcd} X^* \arrow[r, "T^*"] \arrow[d, shift left, "\Phi_X^{-1}"] & X^* \arrow[d, shift left, "\Phi_X^{-1}"] \\ X \arrow[r, "T^\star"] \arrow[u, shift left, "\Phi_X"] & X \arrow[u, shift left, "\Phi_X"] \end{tikzcd},$$

and in practise we typically identify the dual operator T^* with the adjoint $T^\star$.

---

[63] If the space E was infinite dimensional, fix an infinite LI sequence $(x_n)_{n=1}^\infty$ of eigenvectors in E with corresponding eigenvalues (λ_n) and write $E_n = \mathrm{Span}(x_1, \ldots, x_n)$ so that $T(E_n) = E_n$ for all n. Choose a sequence of unit vectors $(y_n)_{n=1}^\infty$ with $y_n \in E_n$ and $d(y_n, E_{n-1}) \geq 1/2$. (by Riesz's lemma, though we're only using it in finite dimensions). Note that $(T - \lambda_n y_n) \in E_{n-1}$ as y_n is of the form $z + \alpha x_n$ for some $z \in E_{n-1}$ and $\alpha \in \mathbb{F}$. For $n > m$,

$$\|Ty_n - Ty_m\| = \|\lambda_n y_n + (T - \lambda_n)y_n - Ty_m\| \geq |\lambda_n| d(y_n, E_{n-1}) \geq r/2.$$

Therefore $(Ty_n)_{n=1}^\infty$ does not have a convergent subsequence, contradicting compactness of T.

[64] My personal preferance is for the inner product to be linear in the first variable, and conjugate linear in the second, but of course it doesn't really matter.

[65] Note the cunning bit of typesetting which avoids abusing notation. The star on the operator $T^\star$ on $\mathcal{B}(X)$ has 5-points as opposed to the dual operator T^* where the star has 6-points! Of course we should think of these operators as being the same after we identify X with X^* using the Riesz Representation Theorem, so you don't need to spend time counting the points on the stars.

To give the spectral theorem, recall the appropriate notion of a basis in the setting of Hilbert spaces.

Definition 13.5. 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 $\operatorname{span}\{x_\alpha : \alpha \in A\}$ is dense in X.

Theorem 13.6 (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 13.3. 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 $\operatorname{Ker}(\lambda_n - T)$ and if we use Zorn's Lemma to obtain an orthonormal basis B_0 for $\operatorname{Ker} T$, 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 13.3 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 $Z \subseteq \operatorname{Ker} T \subseteq Y = Z^\perp$. Hence $Z = \{0\}$, so $X = Y$. □

Remark 13.7. The Spectral Theorem can be extended to the case of compact *normal* operators, that is to say compact operators T such that $T^\star T = TT^\star$.[66]

14 Schauder bases

We now return to looking at Banach spaces, with the aim of further examining the classical sequence spaces ℓ^p. We start by looking at the apprporiate notion of a basis for a Banach space.

[66] There's a more general spectral theorem for all normal (not necessarily compact) operators - leading us in the direction of spectral theory. These need not have eigenvalues but we can suitably decompose such an operator in the form $\int_\Omega \lambda d\mu(\lambda)$, where μ is a *projection*-valued measure on Ω, i.e. a measure taking values in the orthogonal projections in $\mathcal{B}(X)$. Of course, we need to think about in what sense such an integral converges. In the case of compact normal operators, the measure space is countable, and the integral becomes the sum $\sum_i \lambda_i E_i$, where E_i is the projection onto the eigenspace corresponding to λ_i (and in the self adjoint case we end up with exactly the situation of Theorem 13.6).

Definition 14.1. 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

$$x = \sum_{n=1}^{\infty} \lambda_n x_n \tag{14.1}$$

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 (14.1). 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 (14.1).

Definition 14.2. The maps $(P_n)_{n=1}^{\infty}$ above are called the *basis projections* associated with the Schauder basis $\{x_n : n \geq 1\}$ and the functionals $(f_n)_{n=1}^{\infty}$, are the associated *basis functionals*, sometimes also referred to as *coordinate functionals*.

Theorem 14.3. *Let X be a Banach space and suppose that $\{x_n : n \geq 1\}$ is a Schauder basis for X. Then the basis projections $(P_n)_{n=1}^{\infty}$, and the basis functionals $(f_n)_{n=1}^{\infty}$, 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. $\square$

One consequence of the existence of a Schauder basis is that it enables us to approximate the compact operators by finite rank operators.

Theorem 14.4. *Let X be a Banach space with a Schauder basis. Then every compact operator on X is a norm limit of finite rank operators.*

Proof. See example sheet 4. □

By Remark 11.5 it follows that 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, and so the previous theorem applies to all these spaces, and in particular to separable Hilbert spaces.[67]

Example 14.5. (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
$$y_n(t) = 2^{n-1} \int_0^t x_{n-1}(s)\, \mathrm{d}s, \quad 0 \leq t \leq 1,\ n \geq 1.$$
Then $\{y_n : n \geq 0\}$ is a Schauder basis for X. It is known as *Schauder's basis.*

15 Subspaces of classical sequence spaces

Definition 15.1. Given a Banach space X, a sequence $(y_n)_{n=1}^\infty$ in X is said to be *basic* if it is a Schauder basis for its closed linear span $Y = \overline{\operatorname{span}}\, S$

When $(y_n)_{n=1}^\infty$ is a basic sequence, then there exist, by Theorem 14.3, 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 $M \geq 1$.

The following 'principle of small perturbations'

Theorem 15.2. *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.

[67]It also applies to non-separable Hilbert spaces: every compact operator is still a norm limit of finite rank operators.

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. $\square$

One consequence of the previous result is that we can perturb Schauder's basis for $C([0,1])$ and obtain a basis consisting of polynomial functions.

Proposition 15.3. *Let $X = \ell^p$ for $1 \leq p < \infty$ or $X = c_0$.*

(a) *If $S = \{y_n : n \geq 1\} \subseteq S_X$ 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 $\{y_n : n \geq 1\} \subseteq S_X$ and vectors $z_n \in Z$, $n \geq 1$, 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. $\square$

Theorem 15.4. *Let $X = \ell^p$ for $1 \leq p < \infty$ 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 15.2 and Proposition 15.3. $\square$

Corollary 15.5. (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 15.6 (Pełczyński). *Let $X = \ell^p$ for $1 \leq p < \infty$ 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 15.4 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. □

A Baire's Category Theorem

In this appendix to the notes I briefly review 4 fundamental theorems for Banach spaces which are normally deduced from the Baire category theorem: the open mapping therorem, the inverse mapping theorem, the closed graph theorem, and the principle of uniform boundedness. These are discussed in B4.2. First the statement of Baire's category[68] theorem:

Theorem A.1 (Baire's Category Theorem). *Let (X, d) be a complete metric space 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.*

If you've not seen a proof of this, I encourage you to have a go at proving it: the proof can be found in B4.2 if you want a hint. If you have seen a proof of this before, you could have a go at one of the extensions: for example, the Baire category theorem also holds for locally compact Hausdorff spaces; [7, 2.2].

Remark A.2. Before discussing the applications of Baire's category theorem to Banach spaces, I want to take a brief aside. One of the reasons Baire's theorem such an important result is it allows us to think of the notion of a typical element in any space where Baire's category theorem holds. For example, consider the space $C[0, 1]$. It's not at all easy to put a reasonable probability measure on this space, so it doesn't make sense to talk of a 'random continuous function'. However we can use Baire's category theorem to talk about the properties which are satisfied by a *generic continuous function*. In a topological space X where the Baire category theorem holds (i.e. the countable intersection of open dense sets is dense), we will say a subset A of X is *generic* if $A \subset \cap_{n=1}^{\infty} U_n$ for open dense subsets U_n of X. The point is that countable intersections of generic sets remain generic — this way of thinking heads towards descriptive set theory, and gives us a framework for discussing, for example, typical properties of representations. Moreover Baire's category theorem gives a non constructive method of proving the existance of various objects — by showing that they are generic (without ever constructing a single example), and this has been widely used across analysis.[69]

Now back to the big consequences of Baire's category theorem for Banach spaces.

[68]This has absolutely nothing to do with category theory: Baire's work is in the very late 1890's — category theory started to emerge in around 50 years later

[69]Indeed, in prelims analysis you may remember a brief discussion of the existence of a continuous function which is nowhere differentiable. This can be established through Baire's category theorem: the nowhere differentiable functions are genenric in $C[0, 1]$.

Theorem A.3 (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.*

In addition to the Baire category theorem, the other ingredient in the proof of the open mapping theorem is the successive approximation lemma.

Proof of Theorem A.3. Suppose that X, Y are Banach spaces and that $T \in \mathcal{B}(X,Y)$ is surjective. Then $Y = \bigcup_{n=1}^{\infty} F_n$, where F_n denotes the closure of $T(B_X(n))$, $n \geq 1$. Let $U_n = Y \setminus F_n$, $n \geq 1$. Then each U_n, $n \geq 1$, is open and $\bigcap_{n\geq 1} U_n = \emptyset$, so by the Baire Category Theorem there exists $k \geq 1$ such that U_k fails to be dense in Y. Hence F_k has non-empty interior, so $F_k \supseteq B_Y^\circ(y,\varepsilon)$ for some $y \in Y$ and $\varepsilon > 0$. By symmetry and convexity $B_Y^\circ(\varepsilon) \subseteq F_k$ (see for example the proof of Theorem 4.13(d)⇒(e)). Then the sucessive approximation lemma (Lemma 4.14) gives $B_Y^\circ \subseteq T(B_X^\circ(M))$ for $M = \varepsilon^{-1}k$. It follows using linearity that T is an open map. □

The inverse mapping theorem is then an immediate corollary of the open mapping theorem. So once we have a linear bijection between Banach spaces, we only have to show continuity to get an isomorphism.

Theorem A.4 (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.*

Next up is the closed graph theorem. This is a very effective way of proving that various maps are bounded in concrete examples.

Theorem A.5 (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 $X \times Y$.*

Proof. It is easy to see that if $T \in \mathcal{B}(X,Y)$, then G_T is closed in $X \times Y$. Conversely, if G_T is closed in $X \times Y$, then since $X \times Y$ is a Banach space (with say the ℓ^1-norm, see Section 2), so too is G_T. But the map $\theta : G_T \to X$ given by $\theta(x, Tx) = x$ is a continuous linear bijection from G_T to X, so an isomorphism of Banach spaces. Since the map $S : G_T \to Y$, $S(x, Tx) = Tx$ is continuous, $T = S\theta^{-1}$ is continuous. □

The other important consequence of Baire's category theorem that it's useful to know about for the course is the principle of uniform boundedness. It doesn't appear so explicitly in the notes, but might well be useful for one or two exercises (and is most often applied to note that if a sequence $(T_n)_{n=1}^{\infty}$ of bounded operators converges pointwise, i.e. $T_n(x) \to T(x)$ say for each $x \in X$, then the limit operator T is bounded.[70]

Theorem A.6 (Principle of uniform boundedness). *Let X be a Banach space and Y a normed space. Given a family $\mathcal{F}$ of continuous linear operators from X to Y such that for each $x \in X$, $\sup_{T\in\mathcal{F}} \|T(x)\| < \infty$. Then $\sup_{T\in\mathcal{F}} \|T\| < \infty$.*

Proof. Here's the standard Baire category proof. For each n, let $X_n = \{x \in X : \sup_{T\in\mathcal{F}} \|Tx\| \leq n\}$. These are closed sets with $\bigcup_{n=1}^{\infty} X_n = X$, and so by Baire's category theorem there is some $n \in \mathbb{N}$ such that X_n has non-empty interior (just as

[70] Note though that there is no reason to expect T_n to converge to T in uniform norm.

in the proof of the open mapping theorem). Let $\epsilon > 0$ and $x_0 \in X$ be such that $B^\circ_X(x_0, 2\epsilon) \subseteq X_n$. Then for $\|x\| \leq 1$ write $x = \epsilon^{-1}((x_0 + \epsilon x) - x_0)$, so that both $x_0 + \epsilon x$ and x_0 lie in X_n. Then for $T \in \mathcal{F}$,

$$T(x) \leq \epsilon^{-1}(\|T(x_0 + \epsilon x)\| + \|T(x_0)||) \leq 2\epsilon^{-1}n.$$

so $\sup_{T\in\mathcal{F}} \|T\| \leq 2\epsilon^{-1}n$. $\square$

It turns out that all these 4 consequences of Baire's category theorem are equivalent; I'll ask you to deduce the principle of uniform boundedness from the closed graph theorem, and go back from the principle of uniform boundedness to the open mapping theorem on the first example sheet. 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.