MATH411

Operator Theory

Hyperlinked edition

Contents

Contents i

Preface iii

Acknowledgements . . . iv

Conventions and notation . . . iv

1 Basic theory of Hilbert and Banach spaces 1

1.1 Inner products . . . 1

1.2 Norms . . . 2

1.3 Polarisation . . . 3

1.4 Convergence and continuity . . . 4

1.5 Completeness . . . 7

1.6 Orthogonal decomposition . . . 8

1.7 The Riesz–Fréchet theorem . . . 9

2 Linear operators on Hilbert spaces 11

2.1 Linear transformations . . . 11

2.2 Boundedness . . . 12

2.3 Adjoints . . . 15

2.4 Kernel and range . . . 17

2.5 Isometries, co-isometries and unitaries . . . 19

2.6 Orthogonal projections . . . 19

3 Some results involving completeness 21

3.1 Extension of densely defined bounded operators . . . 21

3.2 Banach's completeness criterion . . . 22

3.3 Completeness of $B(\mathsf{H};\mathsf{K})$ . . . 23

4 Invertibility and the spectrum **25**

4.1 The group of units 25

4.2 Spectrum and resolvent 27

4.3 The Neumann series 29

4.4 Compactness of the spectrum 30

4.5 Non-emptiness of the spectrum 31

4.6 The numerical range 32

5 Functional calculus **35**

5.1 The polynomial spectral mapping theorem 35

5.2 The spectral radius formula 36

5.3 The continuous functional calculus 38

6 Operator order and applications **43**

6.1 Operator order 43

6.2 Square roots and absolute values 45

6.3 Polar decomposition 48

7 Spectral theory for compact self-adjoint operators **49**

7.1 Dirac dyads 49

7.2 Finite-rank operators 50

7.3 Compact operators 51

7.4 The spectral theorem 54

Bibliography **59**

Index **61**

Preface

These notes are an introduction to the theory of bounded linear operators on Hilbert space. They take the single-operator viewpoint, rather than operator-algebraic one. Their intended audience is the cohort of fourth-year undergraduate students at Lancaster University taking the *MATH411 Operator Theory* module. These students will have taken courses on real and complex analysis, and have seen the basic theory of Hilbert and Banach spaces.

As this theory is so important, and underpins all subsequent developments, the first chapter contains a review of it, together with some key ideas from analysis. For the setting of these notes, it is more efficient to use sequences to characterise continuity and closure, rather than the more abstract and general topological theory *via* open sets.

Bounded linear operators is introduced are Chapter 2. The existence and key properties of adjoint operators are proved, and certain classes of operator are defined, including isometries, unitary operators and orthogonal projections.

Chapter 3 contains three results which involve completeness: the extension of bounded linear operators which are only densely defined, the characterisation of completeness *via* absolute convergence, as used by Banach, and the fact that the set of bounded operators between Hilbert spaces is complete for the operator norm.

Basic properties of the spectrum are deduced in Chapter 4; the fact that the spectrum is non-empty is derived with the help of the Hahn–Banach theorem. This result is stated in the form required, but its proof would be too much of a digression and is therefore omitted.

The spectral theory is then put to use in Chapter 5, to obtain the continuous functional calculus for self-adjoint operators. This is employed in Chapter 6 in the development of the theory of operator order.

The final chapter, Chapter 7, explores spectral theory for compact self-adjoint operators. This may be viewed as a generalisation of the fact that any Hermitian matrix is unitarily conjugate to a diagonal one, which should be familiar from the second-year Lancaster University module on linear algebra. As well as providing this echo from an earlier course, the spectral theory serves to show the application of some of the ideas from previous chapters.

These notes attempt to sacrifice generality for clarity; many of the results have extensions and ramifications which are not developed here. The reader who want to go further may enjoy the books [5], [4] and [3], which are ordered in increasing level of sophistication (in the opinion of this author). Two more excellent texts which cover topics related to these notes are [2] and [6]. The lecture notes [1] may also hold some interest.

Acknowledgements

The choice of topics for these notes was determined by the MATH411 syllabus, and the approach to many of them was influenced by the notes of Professor J. Martin Lindsay. The idea of using a Taylor series, rather than a Laurent series, to prove the Beurling–Gelfand spectral radius formula, Theorem 5.6, is due to Dr Daniel Elton. Thanks are given to Professor Garth Dales, for comments on an earlier version.

This document was typeset using LaTeX 2ε with Peter Wilson's memoir class and the AMS-LaTeX packages. The index was produced with the aid of the *MakeIndex* program.

Alexander C. R. Belton
Lancaster, 5th January 2018

Conventions and notation

The notation "$P := Q$" means that the quantity P is defined to equal Q. For example,

$$\mathbb{R}_+ := \{x \in \mathbb{R} : x \geqslant 0\}.$$

If f is a function from the set A, its *domain*, to the set B, its *co-domain*, which is written as $f : A \to B$, and C is a subset of A, written as $C \subseteq A$, then the *restriction* of f to C is the function

$$f|_C : C \to B;\ x \mapsto f(x).$$

Further notation will be introduced as required.

One Basic theory of Hilbert and Banach spaces

1.1 Inner products

Definition 1.1. Let V be a complex vector space. An *inner product* is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C} : \ (x, y) \mapsto \langle x, y \rangle$$

such that

(i) the map $V \to \mathbb{C}$; $y \mapsto \langle x, y \rangle$ is linear for every $x \in V$,

(ii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for every x, $y \in V$, and

(iii) $\langle x, x \rangle \geqslant 0$ for every $x \in V$, with equality if and only if $x = 0$.

Remark 1.2. These properties are (i) *linearity*, (ii) *Hermitian* (or *conjugate*) *symmetry* and (iii) *positive definiteness*.

Remark 1.3. Linearity in (i) means that

$$\langle x, y + \lambda z \rangle = \langle x, y \rangle + \lambda \langle x, z \rangle \qquad \text{for every } x, y, z \in V \text{ and } \lambda \in \mathbb{C}.$$

Together with (ii), this implies conjugate linearity in the first argument:

$$\langle x + \lambda y, z \rangle = \langle x, z \rangle + \overline{\lambda} \langle y, z \rangle \qquad \text{for every } x, y, z \in V \text{ and } \lambda \in \mathbb{C}.$$

Some authors prefer to have linearity in the first argument, and conjugate linearity in the second.

Example 1.4. Let $n \geqslant 1$. The finite-dimensional space $\mathbb{C}^n$ has the standard inner product

$$\big\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \big\rangle := \overline{z_1} w_1 + \cdots + \overline{z_n} w_n = \sum_{j=1}^{n} \overline{z_j} w_j$$

for every $(z_1, \ldots, z_n)$, $(w_1, \ldots, w_n) \in \mathbb{C}^n$.

Theorem 1.5. (The Cauchy–Schwarz inequality) If $\langle \cdot, \cdot \rangle$ is an inner product on the complex vector space V then

$$|\langle x, y \rangle|^2 \leqslant \langle x, x \rangle \langle y, y \rangle \qquad \text{for every } x, y \in V,$$

with equality if and only if x and y are linearly dependent.

Proof. Fix x, $y \in V$, let $\lambda \in \mathbb{C}$ be such that $|\lambda| = 1$ and $\lambda \langle x, y \rangle = |\langle x, y \rangle| \geqslant 0$, and let $t \in \mathbb{R}$. Then

$$\begin{aligned} 0 \leqslant \langle x + t\lambda y, x + t\lambda y \rangle &= \langle x, x \rangle + t\lambda \langle x, y \rangle + t\overline{\lambda} \langle y, x \rangle + t^2 |\lambda|^2 \langle y, y \rangle \\ &= \langle x, x \rangle + 2t|\langle x, y \rangle| + t^2 \langle y, y \rangle. \end{aligned}$$

The last expression is a quadratic polynomial in t, with real coefficients and at most one real root. Hence its discriminant is non-positive, so

$$4|\langle x, y \rangle|^2 - 4\langle x, x \rangle \langle y, y \rangle \leqslant 0 \qquad \Longleftrightarrow \qquad |\langle x, y \rangle|^2 \leqslant \langle x, x \rangle \langle y, y \rangle.$$

Finally, if equality holds then the discriminant is equal to zero, so the polynomial has a real root, so $x + t\lambda y = 0$ for some $t \in \mathbb{R}$, by positive definiteness; this implies that when equality holds, the vectors x and y are linearly dependent. Conversely, if x and y are linearly dependent then, without loss of generality, $x = \lambda y$ for some $\lambda \in \mathbb{C}$, so

$$|\langle x, y \rangle|^2 = |\langle \lambda y, y \rangle|^2 = |\lambda| \langle y, y \rangle^2 = \langle x, x \rangle \langle y, y \rangle. \qquad \square$$

Example 1.6. The space of square-summable sequences

$$\ell^2 := \left\{ z = (z_0, z_1, z_2, \ldots) : \sum_{n=0}^{\infty} |z_n|^2 < \infty \right\}$$

has the inner product

$$\langle z, w \rangle := \sum_{n=0}^{\infty} \overline{z_n} w_n \qquad \text{for every } z, w \in \ell^2.$$

The Cauchy–Schwarz inequality may be used to show this series is absolutely convergent, so convergent, and thus the inner product is well defined.

1.2 Norms

Definition 1.7. A *norm* on a complex vector space X is a map

$$\| \cdot \| : X \to \mathbb{R}_+; \ x \mapsto \|x\|$$

such that

(i) $\|x + y\| \leqslant \|x\| + \|y\|$ for every x, $y \in X$,

(ii) $\|\lambda x\| = |\lambda| \, \|x\|$ for every $\lambda \in \mathbb{C}$ and $x \in X$, and

(iii) $\|x\| = 0$ if and only if $x = 0$.

Remark 1.8. These properties are known as (i) *subadditivity* or the *triangle inequality*, (ii) *homogeneity* and (iii) *faithfulness*.

Proposition 1.9. If $\langle \cdot, \cdot \rangle$ is an inner product on the complex vector space V then setting

$$\| \cdot \| : V \to \mathbb{R}_+; \; x \mapsto \langle x, x \rangle^{1/2}$$

defines a norm on V.

Proof. Positive definiteness implies that this function is well defined, with co-domain as claimed, and that $\| \cdot \|$ is faithful.

For homogeneity, note that linearity and conjugate linearity imply that

$$\|\lambda x\|^2 = \langle \lambda x, \lambda x \rangle = \lambda \overline{\lambda} \langle x, x \rangle = |\lambda|^2 \, \|x\|^2$$

for every $\lambda \in \mathbb{C}$ and $x \in \mathbb{C}$; taking square roots gives the identity required.

Finally, if x, $y \in V$ then

$$\begin{aligned} \|x + y\|^2 = \langle x + y, x + y \rangle &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \\ &\leqslant \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \\ &\leqslant \|x\|^2 + 2\|x\| \, \|y\| + \|y\|^2 = \big(\|x\| + \|y\|\big)^2, \end{aligned}$$

using the fact that $\operatorname{Re} \lambda \leqslant |\lambda|$ for any $\lambda \in \mathbb{C}$ together with Theorem 1.5, the Cauchy–Schwarz inequality. Taking square roots gives subadditivity. □

Remark 1.10. As seen in the proof above, with the introduction of the norm, the Cauchy–Schwarz inequality takes the following, more compact form:

$$|\langle x, y \rangle| \leqslant \|x\| \, \|y\| \qquad \text{for every } x, y \in V.$$

1.3 Polarisation

The following trick shows that it is possible to recover an inner product from its norm.

Proposition 1.11. Let $\langle \cdot, \cdot \rangle$ be an inner product on the complex vector space V, with corresponding norm $\| \cdot \|$. Then

$$\langle x, y \rangle = \frac{1}{4} \sum_{n=0}^{3} \mathrm{i}^{-n} \|x + \mathrm{i}^n y\|^2 \qquad \text{for every } x, y \in V.$$

Proof. The sum on the right-hand side equals

$$\begin{aligned}
&\langle x+y, x+y\rangle - \mathrm{i}\langle x+\mathrm{i}y, x+\mathrm{i}y\rangle - \langle x-y, x-y\rangle + \mathrm{i}\langle x-\mathrm{i}y, x-\mathrm{i}y\rangle \\
&= \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 - \mathrm{i}\|x\|^2 + \langle x,y\rangle - \langle y,x\rangle - \mathrm{i}\|y\|^2 \\
&\qquad - \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle - \|y\|^2 + \mathrm{i}\|x\|^2 + \langle x,y\rangle - \langle y,x\rangle + \mathrm{i}\|y\|^2 \\
&= 4\langle x,y\rangle.
\end{aligned}$$

□

1.4 Convergence and continuity

As well as results from linear algebra, we need some fundamental ideas from analysis.

Throughout this section, unless otherwise stated, we let X denote a complex vector space equipped with a norm $\|\cdot\|$, and we fix a subset $X_0 \subseteq X$; later, this choice will be clear from the context. Once we have a norm we can use topological ideas, such as convergence and continuity. As this is not a topology course, we will introduce a minimal amount of these concepts.

Definition 1.12. (Open and closed sets) A set $U \subseteq X_0$ is *open* if, for every $u \in U$, there exists $\varepsilon > 0$ such that the *open ball*

$$\{x \in X_0 : \|x - u\| < \varepsilon\} \subseteq U.$$

A set $F \subseteq X_0$ is *closed* if $X_0 \setminus F$ is open.

Example 1.13. Every open ball is itself an open set. The whole set X_0 and the empty set $\emptyset$ are always open subsets of X_0.

Remark 1.14. Note that the union of an arbitrary collection of open sets is itself open; consequently, the intersection of an arbitrary collection of closed sets is closed. Similarly, a finite intersection of open sets is open, and a finite union of closed sets is closed.

Definition 1.15. (Convergence) A sequence $(x_n)_{n\geqslant 1} \subseteq X_0$ is *convergent* if there exists $x \in X_0$ such that the real sequence $\|x_n - x\| \to 0$ as $n \to \infty$, *i.e.,* for every $\varepsilon > 0$, there exists $N \geqslant 1$ such that $\|x_n - x\| < \varepsilon$ for every $n \geqslant N$. In this case, we write $x_n \to x$, or $x_n \to x$ as $n \to \infty$ if the index needs to be made clear.

Remark 1.16. It follows from the triangle inequality and faithfulness that limits are unique: if $x_n \to x$ and $x_n \to x'$ then

$$0 \leqslant \|x - x'\| \leqslant \|x - x_n\| + \|x_n - x'\| \to 0,$$

so $\|x - x'\| = 0$ and $x = x'$. Thus we may write $\lim_{n\to\infty} x_n = x$ as well as $x_n \to x$.

The following proposition gives a characterisation of closed sets which is extremely useful in practice.

Proposition 1.17. A set $F \subseteq X_0$ is closed if and only if it contains its limit points: whenever $(x_n)_{n\geqslant 1} \subseteq F$ is convergent to some point $x \in X_0$, then in fact $x \in F$.

Proof. Suppose F is closed, and let $x_n \to x$, where $x_n \in F$ for every $n \geqslant 1$. Suppose for contradiction that $x \in X_0 \setminus F$; as this set is open, there exists $\varepsilon > 0$ such that, whenever $\|y - x\| < \varepsilon$, it holds that $y \in X_0 \setminus F$. However, this gives the desired contradiction, since, by the definition of convergence, there exists some $x_n \in F$ with $\|x_n - x\| < \varepsilon$.

Conversely, suppose F contains its limit points. If $X_0 \setminus F$ is not open, then there exists $x \in X_0 \setminus F$ such that, for every $\varepsilon > 0$, there exists $y = y(\varepsilon) \in F$ such that $\|x - y\| < \varepsilon$. Taking $\varepsilon = 1/n$ and $x_n = y(1/n)$ for every $n \geqslant 1$ gives a sequence $(x_n)_{n\geqslant 1} \subseteq F$ such that $x_n \to x$, but then $x \in F$, a contradiction. Hence F is closed. ☐

Definition 1.18. If $D \subseteq X_0$ then the *closure* $\overline{D}$ of D is the intersection of all the closed sets which contain D.

The set $\overline{D}$ is closed and $D \subseteq \overline{D}$, with equality if and only if D is closed.

Proposition 1.19. If $D \subseteq X_0$ then

$$\overline{D} = \{x \in X_0 : d(x, D) = 0\},$$

where $d(x, D) := \inf\big\{\|x - y\| : y \in D\big\}$ is the *distance* from x to D.

Proof. Let $x \in X_0$, and set $\delta := d(x, D)$.

If $\delta = 0$ then x is a limit point of D, so is contained in any closed set in X_0 containing D, by Proposition 1.17. Thus $x \in \overline{D}$.

If $\delta > 0$ then the open ball $B := \{y \in X_0 : \|x-y\| < \delta\}$ contains x but does not meet D. Thus $X_0 \setminus B$ is a closed set in X_0 containing D, so $x \notin \overline{D}$. ☐

Definition 1.20. Let X and Y be complex vector spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and let $X_0 \subseteq X$. A function $f : X_0 \to Y$ is *continuous* if the *pre-image*

$$f^{-1}(U) := \{x \in X_0 : f(x) \in U\}$$

is open for every open set $U \subseteq Y$.

Theorem 1.21. Let X_0, X, Y and f be as in Definition 1.20. The following are equivalent.

(i) The function f is continuous.

(ii) For every $x \in X_0$ and $\varepsilon > 0$ there exists $\delta > 0$ such that $\|f(x') - f(x)\|_Y < \varepsilon$ whenever $x' \in X_0$ is such that $\|x' - x\|_X < \delta$.

(iii) If $x_n \to x$ in X_0 then $f(x_n) \to f(x)$ in Y.

Proof. (i) $\implies$ (ii). Suppose f is continuous, and let $x \in X_0$ and $\varepsilon > 0$. Since the open ball $U := \{y \in Y : \|y - f(x)\|_Y < \varepsilon\}$ is open, so is its pre-image $f^{-1}(U)$. Thus, as $x \in f^{-1}(U)$, there exists $\delta > 0$ such that $\|x' - x\|_X < \delta$ implies that $x' \in f^{-1}(U)$. However, this means that $\|f(x') - f(x)\|_Y < \varepsilon$, so (ii) holds.

(ii) $\implies$ (iii). Suppose (ii) holds, let $x_n \to x$ and $\varepsilon > 0$, and suppose $\delta > 0$ is as in the statement of (ii). Choose $N > 0$ such that $\|x_n - x\|_X < \delta$ for every $n \geqslant N$, and note this implies that $\|f(x_n) - f(x)\|_Y < \varepsilon$ for every such n. Hence $f(x_n) \to f(x)$, as required.

(iii) $\implies$ (i). Suppose (iii) holds and let $U \subseteq Y$ be open; it suffices to prove that $F := X_0 \setminus f^{-1}(U) = f^{-1}(Y \setminus U)$ is closed. Let $(x_n)_{n\geqslant 1} \subseteq F$ be convergent, with limit $x' \in X_0$. Then $f(x_n) \to f(x')$ and $f(x_n) \in Y \setminus U$ for all $n \geqslant 1$, which is closed. Hence $f(x') \in Y \setminus U$ also, so $x' \in f^{-1}(Y \setminus U) = F$ and F is closed, as required. □

The following proposition shows that vector addition and scalar multiplication are jointly continuous and the norm is continuous.

Proposition 1.22. Let sequences $(x_n)_{n\geqslant 1} \subseteq X$, $(y_n)_{n\geqslant 1} \subseteq X$ and $(\lambda_n)_{n\geqslant 1} \subseteq \mathbb{C}$ be such that $x_n \to x$, $y_n \to y$ and $\lambda_n \to \lambda$. Then $x_n + \lambda_n y_n \to x + \lambda y$ and $\|x_n\| \to \|x\|$.

Proof. The first is an immediate consequence of subadditivity and homogeneity of the norm:

$$\begin{aligned}\big\|(x_n + \lambda_n y_n) - (x + \lambda y)\big\| &= \|x_n - x + \lambda_n y_n - \lambda_n y + \lambda_n y - \lambda y\| \\ &\leqslant \|x_n - x\| + |\lambda_n|\,\|y_n - y\| + |\lambda_n - \lambda|\,\|y\| \to 0,\end{aligned}$$

since $(\lambda_n)_{n\geqslant 1}$ is bounded, being convergent.

That the second holds is also a consequence of subadditivity: if z, $w \in X$ then

$$\|z\| = \|z - w + w\| \leqslant \|z - w\| + \|w\| \quad \implies \quad \|z\| - \|w\| \leqslant \|z - w\|$$

so, swapping z and w if necessary,

$$\big|\|z\| - \|w\|\big| \leqslant \|z - w\| \qquad \text{for every } z, w \in X.$$

Thus $\big|\|x_n\| - \|x\|\big| \leqslant \|x_n - x\| \to 0$. □

Definition 1.23. (Compactness) A set $K \subseteq X_0$ is *compact* if every sequence in K has a convergent subsequence with limit in K.

Theorem 1.24. (Heine–Borel) If $X_0 = \mathbb{C}^n$ for some $n \geqslant 1$, with the norm coming from the usual inner product, then a set $K \subseteq X_0$ is compact if and only if it is closed and bounded, where the latter means there exists $R > 0$ such that $K \subseteq \{x \in X_0 : \|x\| \leqslant R\}$.

Proof. At its heart, this result is a consequence of the Bolzano–Weierstrass theorem from MATH114: every bounded sequence in $\mathbb{R}$ has a convergent subsequence. It is left as an exercise to see why. □

The following theorem may be summarised by stating that the continuous image of a compact set is compact. It is a generalisation of the result from real analysis which states that a continuous function on a closed and bounded interval is bounded and attains its bounds.

Theorem 1.25. Let X and Y be normed vector spaces, and let $X_0 \subseteq X$. If $f : X_0 \to Y$ is continuous and $K \subseteq X_0$ is compact then $f(K)$ is compact as well.

Proof. Let $(y_n)_{n \geqslant 1}$ be a sequence in $f(K)$. Then, for each $n \geqslant 1$, there exists $x_n \in K$ such that $f(x_n) = y_n$. Since K is compact, the sequence $(x_n)_{n \geqslant 1}$ has a convergent subsequence, say $(x_{n_k})_{k \geqslant 1}$, with limit $x' \in K$. As f is continuous, Theorem 1.21 implies that $y_{n_k} = f(x_{n_k}) \to f(x') =: y'$. Hence $(y_n)_{n \geqslant 1}$ has the convergent subsequence $(y_{n_k})_{k \geqslant 1}$ with limit $y' \in f(K)$, and therefore $f(K)$ is compact. $\square$

Example 1.26. Given a compact set $K \subseteq \mathbb{C}$, let

$$C(K) := \{f : K \to \mathbb{C} \mid f \text{ is continuous}\}.$$

Then $C(K)$ is a complex vector space, with vector-space operations defined pointwise: if $f,\, g \in C(K)$ and $\lambda \in \mathbb{C}$ then $f + \lambda g \in C(K)$, where

$$(f + \lambda g)(z) := f(z) + \lambda g(z) \qquad \text{for every } z \in K.$$

Setting

$$\|f\|_K := \sup\big\{|f(z)| : z \in K\big\} \qquad \text{for every } f \in C(K)$$

defines a norm on $C(K)$, the *uniform norm*, and convergence with respect to this norm corresponds to *uniform convergence* on K. The norm is well defined by Theorem 1.25.

The space $C(K)$ is also closed under multiplication and conjugation: if $f,\, g \in C(K)$ then $fg \in C(K)$ and $\overline{f} \in C(K)$, where

$$(fg)(z) := f(z)g(z) \text{ and } \overline{f}(z) := \overline{f(z)} \qquad \text{for every } z \in K.$$

Furthermore, the norm is submultiplicative and conjugation is isometric: if $f,\, g \in C(K)$ then

$$\|fg\|_K \leqslant \|f\|_K \, \|g\|_K \qquad \text{and} \qquad \big\|\overline{f}\big\|_K = \|f\|_K.$$

Remark 1.27. The overline notation $\overline{\,\cdot\,}$ is used to denote both closure of a set and the complex conjugate; it should be clear from the context which is meant.

1.5 Completeness

Throughout this section, let X be a complex vector space with a norm $\|\cdot\|$.

Definition 1.28. A *Cauchy sequence* in X is a sequence $(x_n)_{n \geqslant 1} \subseteq X$ such that, for all $\varepsilon > 0$, there exists $N \geqslant 1$ with $\|x_m - x_n\| < \varepsilon$ whenever $m,\, n \geqslant N$, in which case we write that $\|x_m - x_n\| \to 0$ as $m,\, n \to \infty$.

Informally, Cauchy sequences are those where the terms get arbitrarily close together once one goes far enough along the sequence.

Every convergent sequence is Cauchy, but the converse is false in general. However, as for convergent sequences, Cauchy sequences are necessarily bounded.

Definition 1.29. The space X is *complete* if every Cauchy sequence is convergent.

A complete normed vector space is also known as a *Banach space*, or a *Hilbert space* if the norm comes from an inner product.

Proposition 1.30. Let Y be a subspace of the Banach space X. Then Y is closed if and only if Y is complete.

Proof. Suppose that Y is complete and $(x_n)_{n\geqslant 1} \subseteq Y$ is convergent, with limit $x' \in X$. Then $(x_n)_{n\geqslant 1}$ is Cauchy in Y, so convergent there, and hence $x' \in Y$, by the uniqueness of limits. Thus Y is closed.

Conversely, if Y is closed and $(x_n)_{n\geqslant 1} \subseteq Y$ is Cauchy then $x_n \to x'$ for some $x' \in X$, since X is complete. However, it follows that $x' \in Y$, as Y is closed, and therefore Y is complete. □

Example 1.31. The space $C(K)$ of Example 1.26 is complete for the uniform norm. If $K \subseteq \mathbb{C}$ is compact and $(f_n)_{n\geqslant 1} \subseteq C(K)$ is Cauchy then

$$|f_m(z) - f_n(z)| \leqslant \|f_m - f_n\|_K \to 0 \qquad \text{as } m, n \to \infty,$$

for every $z \in K$. Hence $f(z) := \lim_{n\to\infty} f_n(z)$ exists for every $z \in K$, since $\mathbb{C}$ is complete. Furthermore, if $\varepsilon > 0$, and $N \geqslant 1$ is such that $\|f_m - f_n\|_K < \varepsilon$ whenever m, $n > N$, then

$$|f(z) - f_n(z)| = \lim_{m\to\infty} |f_m(z) - f_n(z)| \leqslant \varepsilon \qquad \text{for every } z \in K,$$

so $\|f - f_n\|_K < \varepsilon$ whenever $n \geqslant N$. Thus $\|f_n - f\|_K \to 0$.

Finally, to see that f is continuous, let $(z_n)_{n\geqslant 1} \subseteq K$ be such that $z_n \to z$, let $m \geqslant 1$, and note that

$$\begin{aligned}
|f(z) - f(z_n)| &\leqslant |f(z) - f_m(z)| + |f_m(z) - f_m(z_n)| + |f_m(z_n) - f(z_n)| \\
&\leqslant 2\|f - f_m\|_K + |f_m(z) - f_m(z_n)| \\
&\to 2\|f - f_m\|_K
\end{aligned}$$

as $n \to \infty$, since f_m is continuous. As $\|f - f_m\|_K$ may be made arbitrarily small, it follows that $f(z_n) \to f(z)$ and f is continuous, as claimed.

1.6 Orthogonal decomposition

Definition 1.32. Let V be a complex vector space, and let L and M be subspaces of V. Then V is the *direct sum* of L and M, denoted $V = L \oplus M$, if and only if

(i) $L + M = V$, so that every $v \in V$ may be written as $v = l + m$, where $l \in L$ and $m \in M$, and

(ii) $L \cap M = \{0\}$.

It follows from (ii) that the decomposition of v given by (i) is unique: if l, $l' \in L$ and m, $m' \in M$ are such that $l + m = l' + m'$ then $l - l' = m' - m \in L \cap M$, so $l = l'$ and $m = m'$.

This is sometimes called the *internal direct sum*, to distinguish it from the external direct sum obtained by combining vector spaces: see Exercise Sheet 1.

It follows from Proposition 1.30 that, for a Hilbert space, the best-behaved subspaces are the closed ones: they are themselves Hilbert spaces for the restriction of the inner product.

Theorem 1.33. If L is a closed subspace of the Hilbert space H then $\mathsf{H} = \mathsf{L} \oplus \mathsf{L}^\perp$, where the closed subspace

$$\mathsf{L}^\perp := \{x \in \mathsf{H} : \langle x, y\rangle = 0 \text{ for every } y \in \mathsf{L}\}.$$

Proof. This result is proved in MATH317, as a consequence of the closest-point property. □

Remark 1.34. (Pythagoras) Let L be a closed subspace of the Hilbert space H. If $x \in \mathsf{L}$ and $y \in \mathsf{L}^\perp$ then

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2.$$

1.7 The Riesz–Fréchet theorem

Throughout this section, let H be a Hilbert space.

Theorem 1.35. Let $\phi : \mathsf{H} \to \mathbb{C}$ be continuous and linear. There exists a unique vector $x \in \mathsf{H}$ such that $\phi(y) = \langle x, y\rangle$ for every $y \in \mathsf{H}$.

Proof. To see that such a vector x must be unique, suppose x, $x' \in \mathsf{H}$ are such that

$$\langle x, y\rangle = \langle x', y\rangle \qquad \text{for every } y \in \mathsf{H}.$$

Then, taking $y = x - x'$ and re-arranging, it follows that $\langle x - x', x - x'\rangle = 0$, so $x - x' = 0$, by faithfulness, and $x = x'$.

For existence, note first that if $\phi = 0$ then $x = 0$ suffices, so we may assume that

$$K = \ker\phi := \{x \in \mathsf{H} : \phi(x) = 0\} \neq \mathsf{H}.$$

Since ϕ is linear, the set K is a subspace of H, and since ϕ is continuous, the set K is closed:

$$\text{if } (x_n)_{n\geqslant 1} \subseteq K \text{ and } x_n \to x \in X \qquad \text{then } 0 = \phi(x_n) \to \phi(x), \text{ so } x \in K.$$

Thus, by Theorem 1.33, we may write $\mathsf{H} = K \oplus K^\perp$, and $K^\perp \neq \{0\}$.

Choose $z \in K^\perp$ such that $\|z\| = 1$, and let $x = \overline{\phi(z)}\, z$. Then $\phi(z) \neq 0$, since $z \notin K$, and

$$\phi\big(\phi(y)\phi(z)^{-1}z - y\big) = \phi(y) - \phi(y) = 0 \qquad \text{for every } y \in \mathsf{H},$$

so $\phi(y)\phi(z)^{-1}z - y \in K$. As $x \in K^\perp$, it follows that

$$0 = \langle x, \phi(y)\phi(z)^{-1}z - y\rangle = \phi(y)\langle z, z\rangle - \langle x, y\rangle \quad \Longrightarrow \quad \phi(y) = \langle x, y\rangle. \qquad \square$$

Remark 1.36. The converse of Theorem 1.35 is an immediate consequence of the Cauchy–Schwarz inequality: if $x \in \mathsf{H}$ and

$$\phi : \mathsf{H} \to \mathbb{C};\ y \mapsto \langle x, y\rangle$$

then ϕ is linear, by the definition of an inner product, and continuous, since if $y_n \to y$ in H then

$$|\phi(y_n) - \phi(y)| = |\langle x, y_n - y\rangle| \leqslant \|x\|\, \|y_n - y\| \to 0.$$

Thus we have identified all the *continuous linear functionals* on a Hilbert space: they are exactly the maps given by taking the inner product with some vector. The subject of functional analysis uses continuous linear functionals on Banach spaces and other, more exotic spaces, where the situation may be a lot more complicated.

Two — Linear operators on Hilbert spaces

Throughout this chapter, let H and K be Hilbert spaces.

2.1 Linear transformations

Throughout this section, let V and W be complex vector spaces.

Definition 2.1. A *linear transformation* is a map $T : V \to W$ which is linear:

$$T(x + \lambda y) = Tx + \lambda Ty \qquad \text{for every } x, y \in V \text{ and } \lambda \in \mathbb{C}.$$

It follows from linearity that $T(-x) = -Tx$ for all $x \in V$, and that $T0 = 0$.

Example 2.2. Let $M_{n\times m}(\mathbb{C})$ denote the complex vector space of $n \times m$ matrices with complex entries. If V and W have bases $\{e_1, \ldots, e_m\}$ and $\{f_1, \ldots, f_n\}$, respectively, then any $n \times m$ complex matrix $A = (a_{jk})$ gives rise to a linear transformation $T : V \to W$, by setting

$$T \sum_{k=1}^{m} \lambda_k e_k = \sum_{j=1}^{n} \sum_{k=1}^{m} a_{jk} \lambda_k f_j \qquad \text{for every } \lambda_1, \ldots, \lambda_m \in \mathbb{C}.$$

Conversely, if $T : V \to W$ then there exists some $A \in M_{n\times m}(\mathbb{C})$ such that this identity holds.

If W has an inner product $\langle \cdot, \cdot \rangle_W$ for which the basis is *orthonormal*, so that $\langle f_j, f_j \rangle_W = 1$ for $j = 1, \ldots, n$ and $\langle f_j, f_k \rangle_W = 0$ whenever $j \neq k$, then

$$\langle f_j, Te_k \rangle_W = \sum_{l=1}^{n} a_{lk} \langle f_j, f_l \rangle_W = a_{jk} \qquad \text{for } j = 1, \ldots, n \text{ and } k = 1, \ldots, m.$$

Example 2.3. Two simple but important linear transformations are the zero mapping

$$0 : V \to W;\ x \mapsto 0$$

and the identity mapping

$$I : V \to V;\ x \mapsto x.$$

Remark 2.4. The set $L(V;W)$ of linear transformations from V to W is itself a complex vector space, with the algebraic operators defined by setting

$$(S+\lambda T)x := Sx + \lambda Tx \qquad \text{for every } x \in V$$

whenever S, $T \in L(V;W)$ and $\lambda \in \mathbb{C}$. The map

$$M_{n\times m}(\mathbb{C}) \to L(V;W);\ A \mapsto T$$

of Example 2.2 is an invertible linear transformation which depends on the choice of bases for V and W.

2.2 Boundedness

On Hilbert spaces we have a norm, so we can consider those linear transformations which are continuous. Linearity reduces the question of continuity to one of *boundedness*.

Proposition 2.5. A linear transformation $T \in L(\mathsf{H};\mathsf{K})$ is continuous if and only if there exists $M > 0$ such that $\|Tx\| \leqslant M\|x\|$ for every $x \in \mathsf{H}$.

Proof. First, let T be continuous. By Theorem 1.21 and the fact that $T0 = 0$, there exists $\delta > 0$ such that $\|x\| < \delta$ implies that $\|Tx\| < 1$. Thus if $y \in \mathsf{H} \setminus \{0\}$ then $\big\|(\delta/2\|y\|)y\big\| = \delta/2 < \delta$, and so

$$\big\|T(\delta/2\|y\|)y\big\| < 1 \iff \|Ty\| < (2/\delta)\|y\|.$$

Hence the desired inequality holds with $M = 2/\delta$.

Conversely, suppose $M > 0$ is as in the statement of the proposition. If $x_n \to x$ then

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \leqslant M\|x_n - x\| \to 0,$$

so $Tx_n \to Tx$ and T is continuous, by another application of Theorem 1.21. $\square$

Definition 2.6. A linear transformation $T \in L(\mathsf{H};\mathsf{K})$ is *bounded* if there exists $M > 0$ such that $\|Tx\| \leqslant M\|x\|$ for every $x \in \mathsf{H}$; by Proposition 2.5, boundedness and continuity are equivalent for linear transformations between Hilbert spaces. For such a *bounded operator*, its *operator norm* is

$$\|T\| := \inf\{M > 0 : \|Tx\| \leqslant M\|x\| \text{ for every } x \in \mathsf{H}\}$$

and the *operator-norm inequality* holds:

$$\|Tx\| \leqslant \|T\|\,\|x\| \qquad \text{for every } x \in \mathsf{H}.$$

The set of such bounded operators is denoted $B(\mathsf{H};\mathsf{K})$, or $B(\mathsf{H})$ if $\mathsf{H} = \mathsf{K}$.

Proposition 2.7. The set $B(\mathsf{H};\mathsf{K})$ is a subspace of $L(\mathsf{H};\mathsf{K})$ and the operator norm $\|\cdot\|$ is a norm on $B(\mathsf{H};\mathsf{K})$.

Proof. Let S, $T \in L(\mathsf{H};\mathsf{K})$ and suppose there exist $M > 0$ and $N > 0$ such that $\|Sx\| \leqslant M\|x\|$ and $\|Tx\| \leqslant N\|x\|$ for every $x \in \mathsf{H}$. If $\lambda \in \mathbb{C}$ then

$$\|(S+\lambda T)x\| = \|Sx + \lambda Tx\| \leqslant \|Sx\| + |\lambda|\,\|Tx\| \leqslant M\|x\| + |\lambda|N\|x\| = (M + |\lambda|N)\|x\|$$

for every $x \in \mathsf{H}$, so $S + \lambda T$ is bounded. Thus $B(\mathsf{H};\mathsf{K})$ is a subspace of $L(\mathsf{H};\mathsf{K})$, as claimed.

To prove that $\|\cdot\|$ is a norm, three things must be established: subadditivity, homogeneity and faithfulness. The first and last are immediate consequences of the operator-norm inequality: if S, $T \in B(\mathsf{H};\mathsf{K})$ then

$$\|(S+T)x\| = \|Sx + Tx\| \leqslant \|Sx\| + \|Tx\| \leqslant \|S\|\,\|x\| + \|T\|\,\|x\| = \big(\|S\| + \|T\|\big)\|x\|$$

for every $x \in \mathsf{H}$, so $\|S + T\| \leqslant \|S\| + \|T\|$. Furthermore, if $\|T\| = 0$ then $\|Tx\| \leqslant 0$ for every $x \in \mathsf{H}$, so $Tx = 0$, by faithfulness of the Hilbert-space norm, and therefore $T = 0$.

For homogeneity, note that $\|\lambda T\| = 0 = |\lambda|\,\|T\|$ if $\lambda = 0$, so suppose $|\lambda| > 0$. If $M > 0$ then $\|Tx\| \leqslant M\|x\|$ if and only if $\|\lambda Tx\| \leqslant |\lambda|M\|x\|$, by homogeneity of the Hilbert-space norm, so $\|\lambda T\| = |\lambda|\,\|T\|$. ☐

Remark 2.8. Let $T \in L(\mathsf{H};\mathsf{K})$. It is sometimes useful to let $\|T\| = \infty$ when T is not bounded, *i.e.,* if there does not exist any $M > 0$ such that $\|Tx\| \leqslant M\|x\|$ for every $x \in \mathsf{H}$.

Proposition 2.9. Let $T \in L(\mathsf{H};\mathsf{K})$, where $\mathsf{H} \neq \{0\}$ and $\mathsf{K} \neq \{0\}$. Then

$$\|T\| = \sup\big\{\|Tx\| : \|x\| = 1\big\} = \sup\big\{|\langle y, Tx\rangle| : \|x\| = 1,\ \|y\| = 1\big\},$$

in the sense that either all of these are infinite or all are finite, and in the latter case they are equal.

Proof. Given $x \in \mathsf{H}$, there exists $y \in \mathsf{K}$ such that $\|y\| = 1$ and $|\langle y, Tx\rangle| = \|Tx\|$: if $Tx = 0$ then take y to be any unit vector, and if $Tx \neq 0$ then take $y = \|Tx\|^{-1}Tx$. Hence

$$\|Tx\| = \sup\big\{|\langle y, Tx\rangle| : \|y\| = 1\big\},$$

since the Cauchy–Schwarz inequality implies $|\langle y, Tx\rangle| \leqslant \|Tx\|$ if $\|y\| = 1$. It follows that

$$\sup\big\{\|Tx\| : \|x\| = 1\big\} = \sup\big\{|\langle y, Tx\rangle| : \|x\| = 1,\ \|y\| = 1\big\},$$

whether these quantities are both finite or both infinite.

Now suppose T is bounded: there exists $M > 0$ such that $\|Tx\| \leqslant M\|x\|$ for every $x \in \mathsf{H}$. Then $M \geqslant \sup\big\{\|Tx\| : \|x\| = 1\big\}$, so

$$\|T\| = \inf\{M > 0 : \|Tx\| \leqslant M\|x\| \text{ for every } x \in \mathsf{H}\} \geqslant \sup\big\{\|Tx\| : \|x\| = 1\big\} =: S_T.$$

For the converse, note that if $x \neq 0$ then $x' = \|x\|^{-1}x$ has norm 1 and

$$\|Tx\| = \|Tx'\|\,\|x\| \leqslant S_T\|x\| \qquad \Longleftrightarrow \qquad \|T\| \leqslant S_T.$$

Finally, if T is unbounded then, for all $n \geqslant 1$, there exists $x_n \in \mathsf{H}$ with $\|Tx_n\| > n\|x_n\|$, so $x_n \neq 0$ and $\sup\bigl\{\|Tx\| : \|x\| = 1\bigr\} > n$. Conversely, if $\bigl\{\|Tx\| : \|x\| = 1\bigr\}$ is unbounded then, for all $n \geqslant 1$, there exists $x_n \in \mathsf{H}$ such that $\|x_n\| = 1$ and $\|Tx_n\| > n$. Therefore it is not the case that $\|Tx\| \leqslant n\|x\|$ for every $x \in \mathsf{H}$; since this inequality holds for no n, the operator T is unbounded. □

The following result will be helpful in establishing certain identities: one need only check they hold on "diagonal elements". For an example of its use, see the proof of Proposition 2.28.

Proposition 2.10. If $T \in B(\mathsf{H})$ is such that $\langle x, Tx\rangle = 0$ for all $x \in \mathsf{H}$ then $T = 0$.

Proof. This is another example of polarisation, as in Proposition 1.11. Let x, $y \in \mathsf{H}$ and note that

$$\begin{aligned}
0 &= \langle x+y, T(x+y)\rangle - \mathrm{i}\langle x+\mathrm{i}y, T(x+\mathrm{i}y)\rangle - \langle x-y, T(x-y)\rangle + \mathrm{i}\langle x-\mathrm{i}y, T(x-\mathrm{i})y\rangle \\
&= \langle x, Tx\rangle + \langle x, Ty\rangle + \langle y, Tx\rangle + \langle y, Ty\rangle - \mathrm{i}\langle x, Tx\rangle + \langle x, Ty\rangle - \langle y, Tx\rangle - \mathrm{i}\langle y, Ty\rangle \\
&\qquad - \langle x, Tx\rangle + \langle x, Ty\rangle + \langle y, Tx\rangle - \langle y, Ty\rangle + \mathrm{i}\langle x, Tx\rangle + \langle x, Ty\rangle - \langle y, Tx\rangle + \mathrm{i}\langle y, Ty\rangle \\
&= 4\langle x, Ty\rangle.
\end{aligned}$$

Hence $Ty = 0$ for all $y \in \mathsf{H}$, so $T = 0$. □

Remark 2.11. Proposition 2.10 is false for inner-product spaces with real scalar field: if

$$T : \mathbb{R}^2 \to \mathbb{R}^2;\ (x, y) \mapsto (-y, x)$$

and $\mathbb{R}^2$ has the standard inner product then

$$\bigl\langle (x, y), T(x, y)\bigr\rangle = \bigl\langle (x, y), (-y, x)\bigr\rangle = x(-y) + yx = 0 \qquad \text{for every } (x, y) \in \mathbb{R}^2,$$

but $T(1, 0) = (0, 1)$, so $T \neq 0$.

As well as having a linear structure, so that linear transformations may be added together and multiplied by scalars, there is a multiplication given by composition.

Definition 2.12. Let U, V and W be complex vector spaces. If $S \in L(U; V)$ and $T \in L(V; W)$ then $TS \in L(U; W)$, where

$$TS : U \to W;\ x \mapsto T(Sx) \qquad \text{for every } x \in U.$$

Example 2.13. Suppose U, V and W are finite-dimensional complex vector spaces. If $S \in L(U; V)$ and $T \in L(V; W)$ correspond to matrices A and B, as in Example 2.2, then $TS \in L(U; W)$ corresponds to the usual matrix product BA.

This composition preserves boundedness.

Proposition 2.14. Let H, K and L be Hilbert spaces. If $S \in B(\mathsf{H};\mathsf{K})$ and $T \in B(\mathsf{K};\mathsf{L})$ then $TS \in B(\mathsf{H};\mathsf{L})$, with $\|TS\| \leqslant \|T\|\,\|S\|$: the operator norm is *submultiplicative*. Consequently, the multiplication of operators is continuous: if $(S_n)_{n\geqslant 1} \subseteq B(\mathsf{H};\mathsf{K})$ and $(T_n)_{n\geqslant 1} \subseteq B(\mathsf{K};\mathsf{L})$ are such that $S_n \to S$ and $T_n \to T$ then $T_nS_n \to TS$.

Proof. The first claim is an immediate consequence of the operator-norm inequality:

$$\|TSx\| \leqslant \|T\|\,\|Sx\| \leqslant \|T\|\,\|S\|\,\|x\| \qquad \text{for every } x \in \mathsf{H},$$

so TS is bounded, with $\|TS\| \leqslant \|T\|\,\|S\|$.

For the second, recall that $\|T_n\| \to \|T\|$, so the sequence $\big(\|T_n\|\big)_{n\geqslant 1}$ is bounded, and therefore

$$\|T_nS_n - TS\| \leqslant \|T_n(S_n - S)\| + \|(T_n - T)S\| \leqslant \|T_n\|\,\|S_n - S\| + \|T_n - T\|\,\|S\| \to 0. \qquad \square$$

Remark 2.15. It follows from results in the final chapter that $L(\mathsf{H};\mathsf{K})$ and $B(\mathsf{H};\mathsf{K})$ are equal if either H or K is finite dimensional.

2.3 Adjoints

Theorem 2.16. Let $T \in B(\mathsf{H};\mathsf{K})$. There exists a unique operator $T^* \in B(\mathsf{K};\mathsf{H})$ such that

$$\langle y, Tx\rangle = \langle T^*y, x\rangle \qquad \text{for every } x \in \mathsf{H} \text{ and } y \in \mathsf{K}.$$

[Note that there are two different inner products in this expression: the one of the left is taken in K, whereas the one on the right is taken in H.]

Furthermore, it holds that $\|T^*\| = \|T\|$.

Proof. To see uniqueness, suppose that T^* is as above and $T' \in B(\mathsf{K};\mathsf{H})$ is such that

$$\langle y, Tx\rangle = \langle T'y, x\rangle \qquad \text{for every } x \in \mathsf{H} \text{ and } y \in \mathsf{K}.$$

Then

$$\langle T^*y, x\rangle = \langle T'y, x\rangle \iff \langle (T^* - T')y, x\rangle = 0 \qquad \text{for every } x \in \mathsf{H} \text{ and } y \in \mathsf{K};$$

taking $x = (T^* - T')y$ gives that $(T^* - T')y = 0$ for every $y \in \mathsf{K}$, so $T^* = T'$.

For existence, fix $y \in \mathsf{K}$ and note that

$$\mathsf{H} \to \mathbb{C};\ x \mapsto \langle y, Tx\rangle$$

is a continuous linear functional, since

$$|\langle y, Tx\rangle| \leqslant \|y\|\,\|Tx\| \leqslant \|y\|\,\|T\|\,\|x\| \qquad \text{for every } x \in \mathsf{H},$$

by the Cauchy–Schwarz and operator-norm inequalities. Hence, by the Riesz–Fréchet theorem, Theorem 1.35, there exists $z = z(y) \in \mathsf{H}$ such that

$$\langle z, x\rangle = \langle y, Tx\rangle \qquad \text{for every } x \in \mathsf{H}.$$

Let $T^*y := z$ and note that the required inner-product identity holds. It remains to show that this gives an operator $T^* \in B(\mathsf{K};\mathsf{H})$, *i.e.*, that the map $y \mapsto T^*y$ is linear and bounded.

For linearity, let y, $z \in \mathsf{K}$ and $\lambda \in \mathbb{C}$, and note that

$$\begin{aligned}\langle T^*(y+\lambda z), x\rangle = \langle y+\lambda z, Tx\rangle &= \langle y, Tx\rangle + \overline{\lambda}\langle z, Tx\rangle \\ &= \langle T^*y, x\rangle + \overline{\lambda}\langle T^*z, x\rangle = \langle T^*y + \lambda T^*z, x\rangle\end{aligned}$$

for every $x \in \mathsf{H}$. Since x is arbitrary, it follows that $T^*(y+\lambda z) = T^*y + \lambda T^*z$, so T^* is linear.

For boundedness and the final claim, note that, by Proposition 2.9,

$$\begin{aligned}\|T\| &= \sup\big\{|\langle y, Tx\rangle| : \|x\| = 1,\ \|y\| = 1\big\} \\ &= \sup\big\{|\langle T^*y, x\rangle| : \|x\| = 1,\ \|y\| = 1\big\} \\ &= \sup\big\{|\langle x, T^*y\rangle| : \|x\| = 1,\ \|y\| = 1\big\} \\ &= \|T^*\|.\end{aligned}$$

□

Remark 2.17. Note that $0^* = 0$ and $I^* = I$.

Example 2.18. Let $n \geqslant 1$ The *standard basis* for $\mathbb{C}^n$ is the set $\{e_1, \ldots, e_n\}$, where e_j is the column vector with 1 in the jth row and 0 elsewhere.

If $T \in B(\mathbb{C}^m;\mathbb{C}^n)$ corresponds to the matrix $A = (a_{jk}) \in M_{n\times m}(\mathbb{C})$ as in Example 2.2, where $\mathbb{C}^m$ and $\mathbb{C}^n$ are equipped with the standard bases, then its adjoint $T^* \in B(\mathbb{C}^n;\mathbb{C}^m)$ corresponds to the matrix $A^* = (\overline{a_{kj}}) \in M_{m\times n}(\mathbb{C})$, the *conjugate transpose* of A.

Proposition 2.19. The map

$$B(\mathsf{H};\mathsf{K}) \to B(\mathsf{K};\mathsf{H});\ T \mapsto T^*$$

is conjugate linear, self inverse and such that the C^* *identity* holds:

$$(S+\lambda T)^* = S^* + \overline{\lambda}T^*, \qquad (T^*)^* = T \qquad \text{and} \qquad \|T^*T\| = \|T\|^2$$

for every S, $T \in B(\mathsf{H};\mathsf{K})$ and $\lambda \in \mathbb{C}$.

Furthermore, if $S \in B(\mathsf{H};\mathsf{K})$ and $T \in B(\mathsf{K};\mathsf{L})$ then $(TS)^* = S^*T^*$.

Proof. Since

$$\begin{aligned}\big\langle (S^* + \overline{\lambda} T^*)y, x\big\rangle &= \langle S^*y + \overline{\lambda} T^*y, x\rangle \\ &= \langle S^*y, x\rangle + \lambda\langle T^*y, x\rangle \\ &= \langle y, Sx\rangle + \lambda\langle y, Tx\rangle \\ &= \langle y, (S+\lambda T)x\rangle \qquad \text{for every } x \in \mathsf{H} \text{ and } y \in \mathsf{K},\end{aligned}$$

the first claim follows from the uniqueness part of Theorem 2.16. Uniqueness of the adjoint also gives the second and last claims, since

$$\langle x, T^*y\rangle = \overline{\langle T^*y, x\rangle} = \overline{\langle y, Tx\rangle} = \langle Tx, y\rangle \qquad \text{for every } x \in \mathsf{H} \text{ and } y \in \mathsf{K},$$

and

$$\langle z, TSx\rangle = \langle T^*z, Sx\rangle = \langle S^*T^*z, x\rangle \qquad \text{for every } x \in \mathsf{H} \text{ and } z \in \mathsf{L}.$$

For the C^* identity, note that Proposition 2.9 gives that

$$\begin{aligned}\|T^*T\| &= \sup\big\{|\langle y, T^*Tx\rangle| : \|x\| = 1,\ \|y\| = 1\big\} \\ &\geqslant \sup\big\{|\langle x, T^*Tx\rangle| : \|x\| = 1\big\} \\ &= \sup\big\{\|Tx\|^2 : \|x\| = 1\} \\ &= \|T\|^2,\end{aligned}$$

and the operator-norm inequality gives the reverse, that $\|T^*T\| \leqslant \|T^*\|\,\|T\| = \|T\|^2$. □

2.4 Kernel and range

Throughout this section, let U and V be complex vector spaces, and let H and K be Hilbert spaces.

Definition 2.20. Let $T \in L(U; V)$. Then the *kernel*

$$\ker T := \{x \in U : Tx = 0\}$$

is a subspace of U, and the *image* or *range*

$$\operatorname{im} T := \{Tx : x \in U\}$$

is a subspace of V.

Proposition 2.21. If $T \in B(\mathsf{H}; \mathsf{K})$ then $\ker T$ is a closed subspace of H.

Proof. Let $(x_n)_{n\geqslant 1} \subseteq \ker T$ be convergent, to $x \in \mathsf{H}$. Then $0 = Tx_n \to Tx$, since T is continuous, so $Tx = 0$ and $x \in \ker T$. Thus $\ker T$ is closed, as claimed. □

Definition 2.22. Let $D \subseteq \mathsf{H}$. Recall that the *orthogonal complement* of D is

$$D^\perp := \{y \in \mathsf{H} : \langle x, y\rangle = 0 \text{ for every } x \in D\}.$$

Proposition 2.23. If $D \subseteq \mathsf{H}$ then $D^\perp$ is a closed subspace of H. If $\mathsf{L} \subseteq \mathsf{H}$ is a closed subspace of H then $\mathsf{L} = (\mathsf{L}^\perp)^\perp$.

Proof. Let $x \in D$, y, $z \in D^\perp$ and $\lambda \in \mathbb{C}$. Then

$$\langle x, y + \lambda z\rangle = \langle x, y\rangle + \lambda\langle x, z\rangle = 0,$$

so $y+\lambda z \in D^\perp$ and $D^\perp$ is a subspace of H. To see that it is closed, suppose $(w_n)_{n\geqslant 1} \subseteq D^\perp$ is convergent, with limit $w \in \mathsf{H}$. Then $0 = \langle x, w_n\rangle \to \langle x, w\rangle$ and so $w \in D^\perp$, as required.

For the second claim, suppose $x \in \mathsf{L}$ and $y \in \mathsf{L}^\perp$. Then $\langle y, x\rangle = 0$, so $x \in (\mathsf{L}^\perp)^\perp$ and therefore $\mathsf{L} \subseteq (\mathsf{L}^\perp)^\perp$. Furthermore, if $x \in (\mathsf{L}^\perp)^\perp$ then, by Theorem 1.33, $x = y + z$, where $y \in \mathsf{L}$ and $z \in \mathsf{L}^\perp$, but this implies that $z = x-y \in \mathsf{L}^\perp \cap (\mathsf{L}^\perp)^\perp = \{0\}$, so $x = y \in \mathsf{L}$. Hence $(\mathsf{L}^\perp)^\perp \subseteq \mathsf{L}$ and equality follows. □

In some circumstances, the following result is a useful replacement for the rank-nullity theorem, which doesn't make sense in infinite dimensions.

Theorem 2.24. (Kernel-adjoint-range relation) Let $T \in B(\mathsf{H};\mathsf{K})$. Then

$$\ker T^* = (\operatorname{im} T)^\perp.$$

Proof. Note that

$$\begin{aligned} x \in (\operatorname{im} T)^\perp &\iff \langle Ty, x\rangle = 0 \qquad \text{for every } y \in \mathsf{H} \\ &\iff \langle y, T^*x\rangle = 0 \qquad \text{for every } y \in \mathsf{H} \\ &\iff T^*x = 0 \iff x \in \ker T^*. \end{aligned}$$ □

Although the kernel of a bounded operator must be closed, the range need not be. Here is one situation when a closed range must occur.

Lemma 2.25. Let $T \in B(\mathsf{H};\mathsf{K})$ and $c > 0$ be such that

$$\|Tx\| \geqslant c\|x\| \qquad \text{for every } x \in \mathsf{H}.$$

Then T is injective and its range $\operatorname{im} T$ is a closed subspace of K.

Proof. Injectivity is immediate, by the faithfulness of the norm. If $(y_n)_{n\geqslant 1} \subseteq \operatorname{im} T$ is convergent, to $y \in \mathsf{K}$, then, for every $n \geqslant 1$ there exists $x_n \in \mathsf{H}$ such that $y_n = Tx_n$, and the sequence $(x_n)_{n\geqslant 1}$ is Cauchy, because

$$\|x_m - x_n\| \leqslant c^{-1}\,\|T(x_m - x_n)\| = c^{-1}\,\|y_m - y_n\| \qquad \text{for every } m, n \geqslant 1.$$

Thus $x_n \to x$ for some $x \in \mathsf{H}$, and then $y_n = Tx_n \to Tx$, so $y = Tx \in \operatorname{im} T$. This shows that $\operatorname{im} T$ is closed. □

2.5 Isometries, co-isometries and unitaries

Definition 2.26. Let $T \in L(\mathsf{H};\mathsf{K})$. Then T is an *isometry* if and only if $\|Tx\| = \|x\|$ for every $x \in \mathsf{H}$. Clearly, isometries are bounded and have norm 1 (unless $\mathsf{H} = \{0\}$).

Remark 2.27. If we interpret $\|x-y\|$ as the distance between the vectors x and y then isometries are the linear transformations that preserve distance, since

$$\|Tx - Ty\| = \|T(x-y)\| = \|x-y\| \qquad \text{for every } x, y \in \mathsf{H}$$

if and only if $T \in L(\mathsf{H};\mathsf{K})$ is an isometry.

Proposition 2.28. Let $T \in B(\mathsf{H};\mathsf{K})$. Then T is an isometry if and only if $T^*T = I$.

Proof. If $T^*T = I$ then

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle = \langle x, x\rangle = \|x\|^2 \qquad \text{for every } x \in \mathsf{H},$$

so T is an isometry.

Conversely, if T is an isometry then the previous working shows that $\langle x, (T^*T - I)x\rangle = 0$ for every $x \in \mathsf{H}$. Proposition 2.10 implies that $T^*T - I = 0$, so $T^*T = I$, as claimed. ☐

Remark 2.29. It follows from Proposition 2.28 that isometries preserve the inner product: if $T \in B(\mathsf{H};\mathsf{K})$ is an isometry then $\langle Tx, Ty\rangle = \langle x, y\rangle$ for every x, $y \in \mathsf{H}$. (This also follows by using polarisation, Proposition 1.11, directly.)

Definition 2.30. An operator $T \in B(\mathsf{H};\mathsf{K})$ such that T^* is an isometry is called a *co-isometry*. Note that, by Proposition 2.28, the operator T is a co-isometry if and only if $TT^* = I$.

An operator $U \in B(\mathsf{H};\mathsf{K})$ is a *unitary* if U is both an isometry and a co-isometry; equivalently, the operator U is a unitary if and only if $U^*U = I$ and $UU^* = I$.

2.6 Orthogonal projections

Throughout this section, let L be a closed subspace of the Hilbert space H. Recall that Theorem 1.33 gives the orthogonal decomposition $\mathsf{H} = \mathsf{L} \oplus \mathsf{L}^\perp$, and that this means every vector in H can be written uniquely in the form $x + y$, where $x \in \mathsf{L}$ and $y \in \mathsf{L}^\perp$.

Definition 2.31. The map

$$P_{\mathsf{L}} : \mathsf{H} \to \mathsf{H};\ x + y \mapsto x \qquad (x \in \mathsf{L},\ y \in \mathsf{L}^\perp)$$

is the *orthogonal projection* onto L.

Theorem 2.32. The orthogonal projection $P_{\mathsf{L}} \in B(\mathsf{H})$ and is such that $P_{\mathsf{L}}^2 = P_{\mathsf{L}} = P_{\mathsf{L}}^*$. Furthermore, the range $\operatorname{im} P_{\mathsf{L}} = \mathsf{L}$, the kernel $\ker P_{\mathsf{L}} = \mathsf{L}^\perp$ and $I - P_{\mathsf{L}} = P_{\mathsf{L}^\perp}$.

Proof. Let x, $x' \in \mathsf{L}$, y, $y' \in \mathsf{L}^\perp$ and $\lambda \in \mathbb{C}$.

Note first that

$$P_\mathsf{L}\big(x + y + \lambda(x' + y')\big) = P_\mathsf{L}(x + \lambda x' + y + \lambda y') = x + \lambda x' = P_\mathsf{L}(x + y) + \lambda P_\mathsf{L}(x' + y'),$$

so P_L is linear. Furthermore,

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 \geqslant \|x\|^2 = \|P_\mathsf{L}(x + y)\|^2,$$

so $\|P_\mathsf{L}(x + y)\| \leqslant \|x + y\|$ and therefore $P_\mathsf{L} \in B(\mathsf{H})$, with $\|P_\mathsf{L}\| \leqslant 1$.

Since

$$P_\mathsf{L}^2(x + y) = P_\mathsf{L}\big(P_\mathsf{L}(x + y)\big) = P_\mathsf{L}(x) = x = P_\mathsf{L}(x + y),$$

it follows that $P_\mathsf{L}^2 = P_\mathsf{L}$. Furthermore,

$$\langle x + y, P_\mathsf{L}(x' + y')\rangle = \langle x + y, x'\rangle = \langle x, x'\rangle + \langle y, x'\rangle = \langle x, x'\rangle$$

and

$$\langle P_\mathsf{L}(x + y), x' + y'\rangle = \langle x, x' + y'\rangle = \langle x, x'\rangle + \langle x, y'\rangle = \langle x, x'\rangle,$$

so $P_\mathsf{L}^* = P_\mathsf{L}$.

The fact that $\operatorname{im} P_\mathsf{L} = \mathsf{L}$ is immediate from the definition, as is the fact that $\ker P_\mathsf{L} = \mathsf{L}^\perp$.

The final claim is straightforward: since $(\mathsf{L}^\perp)^\perp = \mathsf{L}$, it follows that

$$P_{\mathsf{L}^\perp}(x + y) = y = x + y - x = (I - P_\mathsf{L})(x + y). \qquad \square$$

The following result shows that orthogonal projections may be characterised in purely algebraic terms.

Proposition 2.33. If $P \in B(\mathsf{H})$ is such that $P^2 = P = P^*$ then $\mathsf{L} := \operatorname{im} P$ is a closed subspace of H such that $P = P_\mathsf{L}$.

Proof. To see that L is closed, let $(x_n)_{n\geqslant 1} \subseteq \mathsf{H}$ be such that $(Px_n)_{n\geqslant 1}$ is convergent, with limit $y \in \mathsf{H}$. Then

$$P(Px_n) = P^2 x_n = Px_n \to y \qquad \text{and} \qquad P(Px_n) \to Py,$$

so $y = Py$ by the uniqueness of limits and $y \in \mathsf{L}$. Thus L is closed.

Now, if $x \in \mathsf{L} = \operatorname{im} P$ then $x = Pz$ for some $z \in \mathsf{H}$, so $Px = P^2 z = Pz = x$. Moreover, if $y \in \mathsf{L}^\perp$ then $\langle y, Pz\rangle = 0$ for every $z \in \mathsf{H}$, so $\langle Py, z\rangle = 0$ for such z and $Py = 0$. Thus

$$P(x + y) = Px + Py = x = P_\mathsf{L}(x + y) \qquad \text{for every } x \in \mathsf{L} \text{ and } y \in \mathsf{L}^\perp,$$

so $P = P_\mathsf{L}$, as claimed. $\square$

Notation 2.34. If $P \in B(\mathsf{H})$ is an orthogonal projection, then so is $P^\perp := I - P$. It is a straightforward consequence of Theorem 2.32 and Proposition 2.33 that $\operatorname{im} P^\perp = \ker P$ and $\ker P^\perp = \operatorname{im} P$.

Three Some results involving completeness

3.1 Extension of densely defined bounded operators

Let X and Y be complex vector spaces equipped with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. The definitions of boundedness and the operator norm make sense for every $T \in L(X;Y)$.

Definition 3.1. The operator $T \in B(X;Y)$ is *bounded* if there exists $M > 0$ such that $\|Tx\|_Y \leqslant M\|x\|_X$ for every $x \in X$, in which case

$$\begin{aligned}\|T\| &:= \inf\{M > 0 : \|Tx\|_Y \leqslant M\|x\|_X \text{ for every } x \in X\}\\ &= \sup\big\{\|Tx\| : x \in X,\ \|x\| = 1\big\}.\end{aligned}$$

The operator-norm inequality still holds: $\|Tx\| \leqslant \|T\|\,\|x\|$ for every $x \in X$.

We write $T \in B(X;Y)$ to denote this situation; it is proved as above that $B(X;Y)$ is a subspace of $L(X;Y)$ and the operator norm is a norm on this space.

Definition 3.2. Let H be a Hilbert space. A set $D \subseteq \mathsf{H}$ is *dense* in H if, given any vector $x \in \mathsf{H}$, there exists a sequence $(x_n)_{n\geqslant 1} \subseteq D$ such that $x_n \to x$.

Theorem 3.3. Let $T_0 \in B(\mathsf{H}_0;\mathsf{K})$, where H_0 is a dense subspace of the Hilbert space H and K is a Hilbert space. There exists a unique operator $T \in B(\mathsf{H};\mathsf{K})$ which extends T_0, so that $Tx_0 = T_0x_0$ for every $x_0 \in \mathsf{H}_0$. Furthermore, it holds that $\|T_0\| = \|T\|$.

Proof. Given $x \in \mathsf{H}$, let $(x_n)_{n\geqslant 1} \subseteq \mathsf{H}_0$ be such that $x_n \to x$; such a sequence exists by the density of H_0 in H. Note that $(x_n)_{n\geqslant 1}$ is Cauchy, because it is convergent, and thus $(T_0x_n)_{n\geqslant 1}$ is also Cauchy, as $\|T_0x_m - T_0x_n\| \leqslant \|T_0\|\,\|x_m - x_n\|$ for every m, $n \geqslant 1$. Hence, as K is complete, we may define

$$T : \mathsf{H} \to \mathsf{K};\ Tx := \lim_{n\to\infty} T_0x_n.$$

This is a good definition, independent of the choice of sequence, since if $(x'_n)_{n\geqslant 1} \subseteq \mathsf{H}_0$ is such that $x'_n \to x$ then $\|T_0x_n - T_0x'_n\| \leqslant \|T_0\|\,\|x_n - x'_n\|$ and $(T_0x'_n)_{n\geqslant 1}$ has the same limit as $(T_0x_n)_{n\geqslant 1}$. Moreover, if $x_0 \in \mathsf{H}_0$ and $x_n = x$ for all $n \geqslant 1$ then $Tx_0 = \lim T_0x_n = T_0x_0$, so the mapping T extends T_0.

If $(y_n)_{n\geqslant 1} \subseteq \mathsf{H}_0$ converges to $y \in \mathsf{H}$ and $\lambda \in \mathbb{C}$ then $x_n + \lambda y_n \to x + \lambda y$ and therefore, by Proposition 1.22,

$$T(x+\lambda y) = \lim_{n\to\infty} T_0(x_n + \lambda y_n) = \lim_{n\to\infty} T_0 x_n + \lambda \lim_{n\to\infty} T_0 y_n = Tx + \lambda Ty,$$

so T is linear. Furthermore,

$$\|Tx\| = \lim_{n\to\infty} \|T_0 x_n\| \leqslant \lim_{n\to\infty} \|T_0\|\,\|x_n\| = \|T_0\|\,\|x\|,$$

and thus $T \in B(\mathsf{H};\mathsf{K})$ with $\|T\| \leqslant \|T_0\|$. The reverse inequality holds because

$$\|T\| = \sup\big\{\|Tx\| : x \in \mathsf{H},\ \|x\| = 1\big\} \geqslant \sup\big\{\|T_0 x\| : x \in \mathsf{H}_0,\ \|x\| = 1\big\} = \|T_0\|. \qquad \square$$

3.2 Banach's completeness criterion

Definition 3.4. Let X be a complex vector space with norm $\|\cdot\|$, and let $(x_n)_{n\geqslant 1} \subseteq X$. The series $\sum_{n=1}^{\infty} x_n$ is convergent if and only if the sequence of partial sums $(\sum_{n=1}^{N} x_n)_{N\geqslant 1}$ is convergent, in which case the *sum* of the series is

$$\sum_{n=1}^{\infty} x_n := \lim_{N\to\infty} \sum_{n=1}^{N} x_n.$$

The series $\sum_{n=1}^{\infty} x_n$ is *absolutely convergent* if and only if the real series $\sum_{n=1}^{\infty} \|x_n\|$ is convergent.

Theorem 3.5. Let X be a complex vector space with norm $\|\cdot\|$. Then X is a Banach space if and only if every absolutely convergent series in X is convergent.

Proof. Suppose first that X is a Banach space, and let $(x_n)_{n\geqslant 1} \subseteq X$ be such that the series $\sum_{n=1}^{\infty} x_n$ is absolutely convergent. If $N > M \geqslant 1$ then the subadditivity of the norm gives that

$$\Big\|\sum_{n=1}^{M} x_n - \sum_{n=1}^{N} x_n\Big\| \leqslant \sum_{n=M+1}^{N} \|x_n\| \leqslant \sum_{n=M+1}^{\infty} \|x_n\| \to 0$$

as $M \to \infty$. Hence the sequence of partial sums is Cauchy, so convergent.

For the converse, suppose every absolutely convergent series in X is convergent, and let $(y_n)_{n\geqslant 1} \subseteq X$ be Cauchy. For all $k \geqslant 1$ there exists $M_k \geqslant 1$ such that $\|y_m - y_n\| < 2^{-k}$ whenever $m, n \geqslant M_k$; without loss of generality, we may assume that $(M_k)_{k\geqslant 1}$ is strictly increasing. Now let

$$x_1 = y_{M_1} \qquad \text{and} \qquad x_k = y_{M_k} - y_{M_{k-1}} \qquad \text{for every } k \geqslant 2,$$

so that $\|x_k\| \leqslant 2^{-k+1}$ for all $k \geqslant 2$. Then $\sum_{k=1}^{\infty} x_k$ is absolutely convergent, so convergent, with limit $y \in X$; moreover,

$$\sum_{k=1}^{K} x_k = y_{M_1} + (y_{M_2} - y_{M_1}) + \cdots + (y_{M_K} - y_{M_{K-1}}) = y_{M_K},$$

so $y_{M_K} \to y$ as $K \to \infty$. Thus if $\varepsilon > 0$, and K is sufficiently large so that $\|y_{M_K} - y\| < \varepsilon/2$ and $2^{-K} < \varepsilon/2$, then

$$\|y_n - y\| \leqslant \|y_n - y_{M_K}\| + \|y_{M_K} - y\| < \varepsilon \qquad \text{for every } n \geqslant M_K,$$

so $y_n \to y$. Hence X is complete. □

3.3 Completeness of $B(\mathsf{H};\mathsf{K})$

Theorem 3.6. Let H and K be Hilbert spaces. Then $B(\mathsf{H};\mathsf{K})$ is a Banach space for the operator norm.

Proof. Let $(T_n)_{n\geqslant 1} \subseteq B(\mathsf{H};\mathsf{K})$ be a Cauchy sequence, and let $x \in \mathsf{H}$. Then $(T_n x)_{n\geqslant 1}$ is a Cauchy sequence in K, since $\|T_m x - T_n x\| \leqslant \|T_n - T_m\| \, \|x\|$, and so we may define

$$T : \mathsf{H} \to \mathsf{K};\ x \mapsto \lim_{n\to\infty} T_n x.$$

To see that T is linear, note that if $x, y \in \mathsf{H}$ and $\lambda \in \mathbb{C}$ then, by Proposition 1.22,

$$T(x + \lambda y) = \lim_{n\to\infty} T_n(x + \lambda y) = \lim_{n\to\infty} T_n x + \lambda \lim_{n\to\infty} T_n y = Tx + \lambda T y.$$

Furthermore, since $(T_n)_{n\geqslant 1}$ is Cauchy, it is bounded, by $M > 0$ say. Then

$$\|T_n x\| \leqslant \|T_n\| \, \|x\| \leqslant M\|x\| \qquad \text{for every } n \geqslant 1 \text{ and } x \in \mathsf{H},$$

so $\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leqslant M\|x\|$. Hence $T \in B(\mathsf{H};\mathsf{K})$, with $\|T\| \leqslant M$.

It remains to prove that $T_n \to T$. Fix $\varepsilon > 0$ and suppose $N \geqslant 1$ is such that $\|T_m - T_n\| < \varepsilon$ whenever $m, n \geqslant N$. If $n \geqslant N$ then

$$\|(T_n - T)x\| = \lim_{m\to\infty} \|(T_n - T_m)x\| \leqslant \varepsilon\|x\| \qquad \text{for every } x \in \mathsf{H},$$

by the operator-norm inequality, so $\|T_n - T\| < \varepsilon$ whenever $n \geqslant N$. Thus $T_n \to T$, as required. □

Remark 3.7. The combination of Theorems 3.6 and 3.5 allows us to work with series of operators in a straightforward manner.

Four Invertibility and the spectrum

Throughout this chapter, let H, K and L be Hilbert spaces.

4.1 The group of units

Proposition 4.1. Let $T \in B(\mathsf{H};\mathsf{K})$. There exists at most one operator $S \in B(\mathsf{K};\mathsf{H})$ such that $ST = I$ and $TS = I$.

Proof. Suppose R, $S \in B(\mathsf{K};\mathsf{H})$ are such that $RT = I$, $TR = I$, $ST = I$ and $TS = I$. Then
$$R = RI = R(TS) = (RT)S = IS = S. \qquad \square$$

Remark 4.2. In fact, the proof of Proposition 4.1 shows that if T has a left inverse R and a right inverse S, then it has a two-sided inverse $R = S$, which is necessarily unique.

We will see below examples of an operator with a right inverse but no left inverse, and an operator with a left inverse but no right inverse.

Definition 4.3. An operator $T \in B(\mathsf{H};\mathsf{K})$ is *invertible* if there exists $S \in B(\mathsf{K};\mathsf{H})$ such that $ST = I$ and $TS = I$.

Remark 4.4. If such an operator, the *inverse* of $T \in B(\mathsf{H};\mathsf{K})$, exists then it is unique, by Proposition 4.1, and is denoted T^{-1}. If T is invertible then so is T^{-1}, with $(T^{-1})^{-1} = T$.

Notation 4.5. Let
$$B(\mathsf{H};\mathsf{K})^\times := \{T \in B(\mathsf{H};\mathsf{K}) : T \text{ is invertible}\}$$
be the set of invertible elements of $B(\mathsf{H};\mathsf{K})$, with $B(\mathsf{H})^\times := B(\mathsf{H};\mathsf{H})^\times$.

Proposition 4.6. Let $S \in B(\mathsf{H};\mathsf{K})^\times$ and $T \in B(\mathsf{K};\mathsf{L})^\times$. The operator $TS \in B(\mathsf{H};\mathsf{L})^\times$, with $(TS)^{-1} = S^{-1}T^{-1}$.

Proof. Note that
$$(TS)(S^{-1}T^{-1}) = T(SS^{-1})T^{-1} = TIT^{-1} = TT^{-1} = I$$

and

$$(S^{-1}T^{-1})(TS) = S^{-1}(T^{-1}T)S = S^{-1}IS = S^{-1}S = I. \qquad \square$$

Remark 4.7. It follows from Proposition 4.6 that the set $B(\mathsf{H})^\times$ forms a group when equipped with the usual operator multiplication, with identity element I, known as the *group of units*.

Remark 4.8. Note that unitary operators are invertible: by definition, if $U \in B(\mathsf{H}; \mathsf{K})$ is unitary then $U^*U = I$ and $UU^* = I$, so U is invertible with $U^{-1} = U^*$. Another way of thinking of unitaries is as invertible isometries, and the inverse is automatically isometric itself. This explains why another name for unitary operator is *isometric isomorphism*.

The set of unitary operators from H to itself forms a subgroup of $B(\mathsf{H})^\times$.

Remark 4.9. If $T \in B(\mathsf{H}; \mathsf{K})^\times$ then its adjoint $T^* \in B(\mathsf{K}; \mathsf{H})^\times$, with $(T^*)^{-1} = (T^{-1})^*$, since

$$T^*(T^{-1})^* = (T^{-1}T)^* = I^* = I$$

and

$$(T^{-1})^*T^* = (TT^{-1})^* = I^* = I.$$

Since $(T^*)^* = T$, it follows that $T \in B(\mathsf{H}; \mathsf{K})^\times$ if and only if $T^* \in B(\mathsf{K}; \mathsf{H})^\times$.

Lemma 4.10. Suppose $T_1, \ldots, T_n \in B(\mathsf{H})$ commute with each other. Then

$$T_1, \ldots, T_n \in B(\mathsf{H})^\times \qquad \iff \qquad T_1 \cdots T_n \in B(\mathsf{H})^\times.$$

Proof. Since $B(\mathsf{H})^\times$ is a group, the forward implication is immediate. For the reverse, suppose $T := T_1 \cdots T_n \in B(\mathsf{H})^\times$ and note that if $k \in \{1, \ldots, n\}$ then

$$S_k := T^{-1} \prod_{j \neq k} T_j$$

is an inverse for T_k, since $S_k T_k = T^{-1}T = I$ and

$$T^{-1}T_k = T^{-1}T_k T T^{-1} = T^{-1}T T_k T^{-1} = T_k T^{-1} \qquad \Longrightarrow \qquad T_k S_k = T^{-1}T = I. \qquad \square$$

The following result extends Lemma 2.25.

Lemma 4.11. Let $T \in B(\mathsf{H}; \mathsf{K})$ and $c > 0$ be such that

$$\|Tx\| \geqslant c\|x\| \qquad \text{for every } x \in \mathsf{H}.$$

If $(\operatorname{im} T)^\perp = \{0\}$ then $T \in B(\mathsf{H}; \mathsf{K})^\times$ and $\|T^{-1}\| \leqslant c^{-1}$.

Proof. Lemma 2.25 implies that T is injective and $\operatorname{im} T$ is closed. If $(\operatorname{im} T)^\perp = \{0\}$ then $\operatorname{im} T = \mathsf{K}$, by Theorem 1.33, so T is a bijection and T^{-1} is a linear transformation. To see that T^{-1} is bounded with norm as claimed, let $y \in \mathsf{K}$. Then $y = Tx$ for some $x \in \mathsf{H}$, and

$$\|T^{-1}y\| = \|x\| \leqslant c^{-1}\|Tx\| = c^{-1}\|y\|. \qquad \square$$

4.2 Spectrum and resolvent

Definition 4.12. Let $T \in B(\mathsf{H})$. The *spectrum* of T is the set

$$\sigma(T) := \{\lambda \in \mathbb{C} : \lambda I - T \text{ is } \textbf{not} \text{ invertible}\}.$$

Remark 4.13. It follows from Remark 4.9 that if $T \in B(\mathsf{H})$ then

$$\sigma(T^*) = \overline{\sigma(T)} := \{\overline{\lambda} : \lambda \in \sigma(T)\},$$

since

$$\lambda \notin \sigma(T) \iff \lambda I - T \in B(\mathsf{H})^\times \iff \overline{\lambda} I - T^* \in B(\mathsf{H})^\times \iff \overline{\lambda} \notin \sigma(T^*).$$

Remark 4.14. Let H be finite dimensional, so that $\mathsf{H} = \mathbb{C}^n$ for some $n \geqslant 0$, but perhaps with a non-standard innner product. For every $S \in L(\mathsf{H})$ we have the *rank-nullity theorem*:

$$\dim \ker S + \dim \operatorname{im} S = n.$$

Thus

$$S \text{ is not invertible} \iff S \text{ is not injective} \iff S \text{ is not surjective}.$$

Furthermore, $\lambda I - T$ is not injective if and only if there exist distinct vectors x, $y \in \mathbb{C}^n$ such that

$$(\lambda I - T)x = (\lambda I - T)y \iff (\lambda I - T)(x - y) = 0 \iff T(x - y) = \lambda(x - y).$$

Hence the non-injectivity of $\lambda I - T$ is equivalent to λ being an eigenvalue of T.

To summarise: if H is finite dimensional, the spectrum of an operator T is its set of eigenvalues.

Remark 4.15. If H is infinite dimensional, then the rank-nullity theorem doesn't apply, and it is possible that an operator may be injective but not surjective, or surjective but not injective.

Example 4.16. To see that the first possibility in Remark 4.15 may occur, consider the *right shift* $R : \ell^2 \to \ell^2$, defined by setting

$$R(z_0, z_1, z_2, \ldots) := (0, z_0, z_1, z_2, \ldots) \qquad \text{for every } z = (z_0, z_1, z_2, \ldots) \in \ell^2.$$

It is immediate that R is linear, and since

$$\|Rz\|^2 = 0^2 + |z_0|^2 + |z_1|^2 + \cdots = \|z\|^2 \qquad \text{for every } z \in \ell^2,$$

the operator R is an isometry; in particular, it is bounded and injective. However, this operator is not surjective, since

$$(1, 0, 0, \ldots) \in \ell^2 \setminus \operatorname{im} R.$$

Thus R is injective but not surjective.

Example 4.17. For an example of the second possibility in Remark 4.15, an operator which is surjective but not injective, consider the *left shift* $L : \ell^2 \to \ell^2$, where

$$L(z_0, z_1, z_2, \dots) := (z_1, z_2, z_3, \dots) \qquad \text{for every } z = (z_0, z_1, z_2, \dots) \in \ell^2.$$

Again, it is straightforward to check that L is linear, and

$$\|Lz\|^2 = \sum_{n=1}^{\infty} |z_n|^2 \leqslant \sum_{n=0}^{\infty} |z_n|^2 = \|z\|^2 \qquad \text{for every } z \in \ell^2,$$

so L is bounded, with $\|L\| \leqslant 1$. It is straightforward to prove that, in fact, this operator has norm 1.

To see that L is surjective, note that $LRz = z$, for every $z \in \ell^2$, where R is the right shift of the previous example. To see that L is not injective, note that

$$z = (1, 0, 0, \dots) \in \ell^2 \setminus \{0\} \qquad \text{and} \qquad Lz = 0.$$

Remark 4.18. The previous two examples show that a bounded operator may fail to be invertible by being non-injective, or by begin non-surjective. In principle, there could be a third possibility for failure.

Recall that $T \in B(\mathsf{H})$ is invertible if there exists $S \in B(\mathsf{H})$ such that

$$ST = I \qquad \text{and} \qquad TS = I.$$

The first equation gives that T is injective: if $x \in \mathsf{H}$ is such that $Tx = 0$ then

$$x = Ix = STx = S0 = 0.$$

The second equation that T is surjective: if $x \in \mathsf{H}$ then

$$T(Sx) = (TS)x = Ix = x.$$

Conversely, if $T \in B(\mathsf{H})$ is bijective then we can define a map

$$S : \mathsf{H} \to \mathsf{H};\ Ty \mapsto y.$$

It is easily verified that S is a well-defined linear transformation, such that $ST = I$ and $TS = I$. However, it is far from clear whether or not S is bounded. The following theorem decides this.

Theorem 4.19. (Banach) If $T \in B(\mathsf{H})$ is bijective, then T is invertible. In other words, if T is invertible as a function, then it is invertible as a bounded operator.

Proof. The proof of this result relies on one of the three big theorems of functional analysis, the open-mapping theorem, which is beyond the scope of this module. $\square$

The concrete Lemma 4.11 is often more useful than the abstract Theorem 4.19, as it gives an explicit estimate for the norm of the inverse operator.

Definition 4.20. Let $T \in B(\mathsf{H})$. The function

$$\mathbb{C} \setminus \sigma(T) \to B(\mathsf{H});\ \lambda \mapsto (\lambda I - T)^{-1}$$

is the *resolvent* of T.

Thus the spectrum is the set of points at which the resolvent is undefined. Consequently, we can investigate the spectrum by looking at properties of the resolvent, and this is done in the following sections.

4.3 The Neumann series

The geometric series should be very familiar: if $z \in \mathbb{C}$ is such that $|z| < 1$ then

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n = 1 + z + z^2 + \cdots .$$

If you remember how to prove that the sum of this series is as claimed, then you can obtain a similar result for operators.

Theorem 4.21. (C. Neumann) If $T \in B(\mathsf{H})$ is such that $\|T\| < 1$ then $I - T$ is invertible and the series

$$\sum_{n=0}^{\infty} T^n = I + T + T^2 + \cdots$$

is absolutely convergent, its sum equals $(I-T)^{-1}$ and $\|(I-T)^{-1}\| \leqslant \big(1 - \|T\|\big)^{-1}$.

Proof. Note first that $\|T^n\| \leqslant \|T\|^n$ for every $n \geqslant 0$, by the submultiplicativity of the operator norm. Thus the series

$$\sum_{n=0}^{\infty} T^n$$

is absolutely convergent, by comparison with the geometric series with ratio $\|T\|$. By Banach's convergence criterion, Theorem 3.5, this series is convergent; let S denote its sum. Then

$$(I-T)S = (I-T) \lim_{N\to\infty} \sum_{n=0}^{N} T^n = \lim_{N\to\infty} \sum_{n=0}^{N} (I-T)T^n = \lim_{N\to\infty} I - T^{N+1} = I,$$

since $\|T^{N+1}\| \leqslant \|T\|^{N+1} \to 0$ as $N \to \infty$.

Similarly,

$$S(I-T) = \Big(\lim_{N\to\infty} \sum_{n=0}^{N} T^n\Big)(I-T) = \lim_{N\to\infty} \sum_{n=0}^{N} T^n(I-T) = \lim_{N\to\infty} I - T^{N+1} = I,$$

and therefore, by the uniqueness of inverses, $S = (I - T)^{-1}$.

The final claim holds because

$$\Big\|\sum_{n=0}^{N} T^n\Big\| \leqslant \sum_{n=0}^{N} \|T\|^n \leqslant \sum_{n=0}^{\infty} \|T\|^n = \big(1 - \|T\|\big)^{-1} \qquad \text{for any } N \geqslant 0. \qquad \square$$

Corollary 4.22. Let $T \in B(\mathsf{H})$. Then the spectrum

$$\sigma(T) \subseteq \big\{\lambda \in \mathbb{C} : |\lambda| \leqslant \|T\|\big\}.$$

Proof. If $\lambda \in \mathbb{C}$ is such that $|\lambda| > \|T\|$ then $\lambda \neq 0$ and $\|\lambda^{-1}T\| < 1$, so $I - \lambda^{-1}T$ is invertible, by Theorem 4.21. Hence $\lambda I - T = \lambda(I - \lambda^{-1}T)$ is also invertible, so $\lambda \notin \sigma(T)$. Thus the spectrum is as claimed. $\square$

Theorem 4.23. The group of units $B(\mathsf{H})^\times$ is open: if $T \in B(\mathsf{H})^\times$ and $\|S-T\| < \|T^{-1}\|^{-1}$ then $S \in B(\mathsf{H})^\times$ also. Furthermore, the inversion map

$$B(\mathsf{H})^\times \to B(\mathsf{H})^\times;\ T \mapsto T^{-1}$$

is continuous, so a *homeomorphism*, *i.e.*, a continuous function with continuous inverse.

Proof. Suppose S and T are as in the statement of the theorem. Then

$$\|T^{-1}(T-S)\| \leqslant \|T^{-1}\|\,\|T-S\| = \|T^{-1}\|\,\|S-T\| < 1,$$

so $I - T^{-1}(T-S)$ is invertible, by Theorem 4.21. As $B(\mathsf{H})^\times$ is a group, it follows that

$$T\big(I - T^{-1}(T-S)\big) = T - T + S = S$$

is invertible too.

Furthermore, this working shows that

$$S^{-1} - T^{-1} = \big(I - T^{-1}(T-S)\big)^{-1}T^{-1} - T^{-1} = \sum_{n=1}^{\infty}\big(T^{-1}(T-S)\big)^n T^{-1},$$

so

$$\|S^{-1} - T^{-1}\| \leqslant \frac{\|T^{-1}(T-S)\|\,\|T^{-1}\|}{1 - \|T^{-1}(T-S)\|} \leqslant \frac{\|T^{-1}\|^2\|T-S\|}{1 - \|T^{-1}(T-S)\|} \to 0$$

as $S \to T$. Hence $T \mapsto T^{-1}$ is continuous, as claimed. $\square$

4.4 Compactness of the spectrum

Corollary 4.22 shows that the spectrum is a bounded set. In fact, it is also closed, and therefore compact. (Recall that, by the Heine–Borel theorem, Theorem 1.24, a subset of the complex plane is compact if and only if it is closed and bounded.)

Proposition 4.24. Let $T \in B(\mathsf{H})$. Then $\sigma(T)$ is closed.

Proof. Let $(\lambda_n)_{n\geqslant 1} \subseteq \sigma(T)$ be convergent, to $\lambda \in \mathbb{C}$. Then $\lambda_n I - T \in B(\mathsf{H}) \setminus B(\mathsf{H})^\times$ for all $n \geqslant 1$, and this set is closed, by Theorem 4.23. Since

$$\|(\lambda_n I - T) - (\lambda I - T)\| = |\lambda_n - \lambda| \to 0 \qquad \text{as } n \to \infty,$$

so $\lambda I - T \in B(\mathsf{H}) \setminus B(\mathsf{H})^\times$ also. Thus $\lambda \in \sigma(T)$, as required. □

4.5 Non-emptiness of the spectrum

We have already seen that the spectrum is a generalisation of the set of eigenvalues of a linear transformation on a finite-dimensional vector space. Recall that eigenvalues are calculated as the roots of the characteristic equation: if a linear transformation on $\mathbb{C}^n$ has matrix A with respect to some basis, then the *characteristic polynomial* for A is

$$\chi_A(\lambda) := \det(\lambda I - A).$$

The eigenvalues are the roots of the equation $\chi_A(\lambda) = 0$. In order to know that this equation has any roots, we invoke the Fundamental Theorem of Algebra: every complex polynomial has a root.

This is a deep theorem, and any proof that the spectrum is non-empty might be expected to use something at least as deep. The MATH215 proof of the Fundamental Theorem uses Liouville's theorem, and we will use a slight extension of the same idea.

Theorem 4.25. (Liouville) If $f : \mathbb{C} \to \mathbb{C}$ is entire, *i.e.*, holomorphic (also known as complex differentiable) on the whole of $\mathbb{C}$, and bounded then f is constant.

Definition 4.26. Consider a function $f : U \to X$, where $U \subseteq \mathbb{C}$ is an open set and X is a Banach space. Then f is *strongly holomorphic* if

$$f'(w) := \lim_{z\to w} \frac{f(z) - f(w)}{z - w} \qquad \text{exists in } X \text{ for every } w \in U.$$

(The limit is defined here exactly as in the complex case. Equivalently, the limit must exist for every sequence $(z_n)_{n\geqslant 1} \subseteq U \setminus \{w\}$ which converges to w.)

Proposition 4.27. (The resolvent identity) Let $T \in B(\mathsf{H})$. If $\lambda, \mu \in \mathbb{C} \setminus \sigma(T)$ then

$$(\lambda I - T)^{-1} - (\mu I - T)^{-1} = (\mu - \lambda)(\lambda I - T)^{-1}(\mu I - T)^{-1}.$$

Consequently, the resolvent function $\lambda \mapsto (\lambda I - T)^{-1}$ is strongly holomorphic on $\mathbb{C} \setminus \sigma(T)$.

Proof. Note that

$$(\lambda I - T)\big((\lambda I - T)^{-1} - (\mu I - T)^{-1}\big)(\mu I - T) = (\mu I - T) - (\lambda I - T) = (\mu - \lambda)I.$$

Multiplying this on the left by $(\lambda I - T)^{-1}$ and on the right by $(\mu I - T)^{-1}$ gives the identity claimed.

For the second part, note that if $\lambda \neq \mu$ then

$$\frac{(\lambda I - T)^{-1} - (\mu I - T)^{-1}}{\lambda - \mu} = -(\lambda I - T)^{-1}(\mu I - T)^{-1} \to -(\mu I - T)^{-2}$$

as $\lambda \to \mu$, since $\lambda \mapsto \lambda I - T$ and the inversion map are both continuous. □

The following theorem is a fundamental result from the subject called functional analysis. It states that every Banach space X has enough continuous linear functionals to separate points.

Theorem 4.28. (Hahn–Banach) Let X be a Banach space. If $x \in X \setminus \{0\}$ then there exists $\phi \in X^* := B(X; \mathbb{C})$ such that $\|\phi\| = 1$ and $\phi(x) = \|x\|$.

Proof. This is another result with a proof which is beyond the scope of this module. □

Theorem 4.29. Let $T \in B(\mathsf{H})$. Then $\sigma(T)$ is non empty.

Proof. Suppose otherwise, for contradiction. Then T is invertible, since $0 \notin \sigma(T)$, so Theorem 4.28 gives the existence of $\phi \in B(\mathsf{H})^*$ such that $\|\phi\| = 1$ and $\phi(T^{-1}) \neq 0$. The function

$$f : \mathbb{C} \to \mathbb{C};\ \lambda \mapsto \phi\big((\lambda I - T)^{-1}\big)$$

is entire and tends to zero as $|\lambda| \to \infty$, since if $|\lambda| > \|T\|$ then Theorem 4.21 implies that

$$\begin{aligned} |f(\lambda)| &\leqslant \|\phi\| \, \|(\lambda I - T)^{-1}\| \\ &= |\lambda|^{-1} \, \|(I - \lambda^{-1}T)^{-1}\| \\ &\leqslant |\lambda|^{-1} \left(1 - \|\lambda^{-1}T\|\right)^{-1} \\ &= \big(|\lambda| - \|T\|\big)^{-1}. \end{aligned}$$

It follows that f is bounded: there exists $R > 0$ such that $|f(\lambda)| < 1$ if $|\lambda| > R$, by the above, and f is bounded on the compact set $\{\lambda \in \mathbb{C} : |\lambda| \leqslant R\}$, since f is continuous there.

Thus f is entire and bounded, so constant, by Liouville's theorem. Furthermore, the only possible value for the constant is 0, since $f(\lambda) \to 0$ as $|\lambda| \to \infty$, but $f(0) = -\phi(T^{-1}) \neq 0$. This contradiction gives the result. □

4.6 The numerical range

Definition 4.30. Let $T \in B(\mathsf{H})$. The *numerical range* of T is

$$\nu(T) := \{\langle x, Tx \rangle : x \in \mathsf{H},\ \|x\| = 1\}.$$

Proposition 4.31. Let $T \in B(\mathsf{H})$. The numerical range $\nu(T)$ is a bounded subset of $\mathbb{C}$ which contains all the eigenvalues of T. Furthermore,

$$\nu(\lambda T + \mu) = \lambda \nu(T) + \mu := \{\lambda z + \mu : z \in \nu(T)\} \qquad \text{for all } \lambda, \mu \in \mathbb{C}.$$

Proof. This is left as an exercise. □

Theorem 4.32. Let $T \in B(\mathsf{H})$. Then $\sigma(T) \subseteq \overline{\nu(T)}$.

Proof. Suppose that $\lambda \in \mathbb{C} \setminus \overline{\nu(T)}$, so that $\delta := d(\lambda, \nu(T)) > 0$, by Proposition 1.19. If $x \in \mathsf{H}$ with $\|x\| = 1$ then

$$\delta \leqslant |\lambda - \langle x, Tx \rangle| = |\langle x, (\lambda I - T)x \rangle| \leqslant \|(\lambda I - T)x\|.$$

Thus $\lambda I - T$ is injective and has closed range $\mathsf{K} := \operatorname{im}(\lambda I - T)$, by Lemma 2.25.

If $y \in \mathsf{K}^\perp$ is such that $\|y\| = 1$ then

$$0 = \langle y, (\lambda I - T)y \rangle = \lambda - \langle y, Ty \rangle,$$

contradicting the fact that $\lambda \notin \nu(T)$. Thus $\mathsf{K}^\perp = \{0\}$ and therefore $\lambda \notin \sigma(T)$, by Lemma 4.11. □

It is straightforward to see that if T is self adjoint then the numerical range $\nu(T)$ is a subset of $\mathbb{R}$. Together with the previous result, this shows that the spectrum of a self-adjoint operator contains only real numbers. More precisely, we have the following.

Corollary 4.33. Let $T \in B(\mathsf{H})$ be self adjoint. If

$$m_T := \inf\{\langle x, Tx \rangle : x \in \mathsf{H},\ \|x\| = 1\} \quad \text{and} \quad M_T := \sup\{\langle x, Tx \rangle : x \in \mathsf{H},\ \|x\| = 1\}$$

then $\sigma(T) \subseteq [m_T, M_T] \subseteq [-\|T\|, \|T\|]$.

Five — Functional calculus

Throughout this chapter, let H be a Hilbert space.

We consider here how we to define functions of a given bounded operator. For motivation, consider the following example.

Example 5.1. Let $A \in M_n(\mathbb{C})$ be an $n \times n$ Hermitian matrix, so that $A^* = A$. There exists a unitary matrix $U \in M_n(\mathbb{C})$ and a diagonal matrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ such that $A = UDU^*$, where $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ are the eigenvalues of A, repeated according to multiplicity. Hence

$$A^2 = UDU^*UDU^* = UD^2U^* \qquad \text{and} \qquad D^2 = \operatorname{diag}(\lambda_1^2, \ldots, \lambda_n^2).$$

More generally, this holds with 2 replaced by any non-negative integer n, and by linearity,

$$p(A) = Up(D)^*U \qquad \text{and} \qquad p(D) = \operatorname{diag}\big(p(\lambda_1), \ldots, p(\lambda_n)\big)$$

for any complex polynomial $p(z) \in \mathbb{C}[z]$. In particular, if A has eigenvalues $\lambda_1, \ldots, \lambda_n$ then $p(A)$ has eigenvalues $p(\lambda_1), \ldots, p(\lambda_n)$. This observation extends in great generality, as Theorem 5.3 shows.

This example provides some guidance to help move beyond matrices. The importance of eigenvalues here is a clue that further investigation of the spectrum may be worthwhile.

5.1 The polynomial spectral mapping theorem

Definition 5.2. If $p(z) \in \mathbb{C}[z]$ is a complex polynomial, say $p(z) = a_n z^n + \cdots + a_1 z + a_0$, then

$$p(T) := a_n T^n + \cdots + a_1 T + a_0 I \in B(\mathsf{H}) \qquad \text{for every } T \in B(\mathsf{H}).$$

Note that $\mathbb{C}[z]$ is a complex vector space, with the usual polynomial addition and scalar multiplication, and the mapping

$$\epsilon_T : \mathbb{C}[z] \to B(\mathsf{H});\ p(z) \mapsto p(T)$$

is a linear transformation for every $T \in B(\mathsf{H})$. This mapping is also *multiplicative*:

$$\epsilon_T\big(p(z)q(z)\big) = p(T)q(T) = \epsilon_T\big(p(z)\big)\epsilon_T\big(q(z)\big) \qquad \text{for every } p(z), q(z) \in \mathbb{C}[z],$$

where polynomial multiplication is defined in the usual manner. Furthermore, it respects the adjoint:

$$\epsilon_T\big(p(z)\big)^* = \epsilon_{T^*}\big(p^*(z)\big) \qquad \text{for every } p(z) \in \mathbb{C}[z],$$

where $p^*(z) := \overline{a_n} z^n + \cdots + \overline{a_1} z + \overline{a_0}$.

Theorem 5.3. (The polynomial spectral mapping theorem) If $T \in B(\mathsf{H})$ then

$$\sigma\big(p(T)\big) = p\big(\sigma(T)\big) := \{p(\lambda) : \lambda \in \sigma(T)\} \qquad \text{for all } p(z) \in \mathbb{C}[z].$$

Proof. If $p(z)$ is constant, with value $\lambda \in \mathbb{C}$, then $p(T) = \lambda I$ and

$$\sigma\big(p(T)\big) = \{\lambda\} = p\big(\sigma(T)\big).$$

Now suppose $p(z)$ is non-constant. By the Fundamental Theorem of Algebra, if $\mu \in \mathbb{C}$ then

$$\mu - p(z) = \beta \prod_{j=1}^{n} (\lambda_j - z)^{\alpha_j},$$

where $\beta \in \mathbb{C} \setminus \{0\}$, $n \geqslant 1$, $\alpha_1, \ldots, \alpha_n \geqslant 1$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$; note that $p(z) = \mu$ if and only if each side of this identity is zero, so $p^{-1}(\mu) = \{\lambda_1, \ldots, \lambda_n\}$. Hence

$$\mu I - p(T) = \beta \prod_{j=1}^{n} (\lambda_j I - T)^{\alpha_j}. \tag{5.1}$$

Since all the factors in the right-hand side (5.1) commute, the whole product is not invertible if and only if one of its term is not invertible, by Lemma 4.10. Thus

$$\mu \in \sigma\big(p(T)\big) \iff \lambda \in \sigma(T) \quad \text{for some } \lambda \in p^{-1}(\mu) \iff \mu \in p\big(\sigma(T)\big). \qquad \square$$

5.2 The spectral radius formula

Definition 5.4. Let $T \in B(\mathsf{H})$. The *spectral radius*

$$\rho(T) := \sup\big\{|\lambda| : \lambda \in \sigma(T)\big\}.$$

This is the radius of the smallest disc centred at the origin which contains the spectrum of T. Note that $\rho(T) \leqslant \|T\|$, by Corollary 4.22.

To prove the key result in this section, we need some more ideas from complex analysis.

Lemma 5.5. Let $T \in B(\mathsf{H})$ and let

$$D := \{\mu \in \mathbb{C} : |\mu| < \rho(T)^{-1}\}, \qquad \text{where} \qquad D := \mathbb{C} \quad \text{if } \rho(T) = 0.$$

Given $\phi \in B(\mathsf{H})^*$, the function

$$f : D \to \mathbb{C};\ \mu \mapsto \mu\, \phi\big((I - \mu T)^{-1}\big)$$

is holomorphic on D and has the Taylor series

$$f(\mu) = \sum_{n=1}^{\infty} \phi(T^{n-1})\, \mu^n \qquad \text{for every } \mu \in D.$$

Furthermore, if $0 < r < \rho(T)^{-1}$ then

$$\phi(T^{n-1}) = \frac{1}{2\pi \mathrm{i}} \int_{\{z\in\mathbb{C}:|z|=r\}} \frac{\phi\big((I-zT)^{-1}\big)}{z^n}\,\mathrm{d}z \qquad \text{for every } n \geqslant 1, \tag{5.2}$$

with the contour described once anticlockwise.

Proof. Proposition 4.27 implies that the function

$$\{\lambda \in \mathbb{C} : |\lambda| > \rho(T)\} \to \mathbb{C};\ \lambda \mapsto \phi\big((\lambda I - T)^{-1}\big) = \lambda^{-1}\phi\big((I-\lambda^{-1}T)^{-1}\big)$$

is holomorphic. Since $z \mapsto z^{-1}$ is holomorphic on $\mathbb{C}\setminus\{0\}$, it follows that f is holomorphic on $D \setminus \{0\}$. Furthermore, Theorem 4.23 implies that

$$\frac{f(\mu) - f(0)}{\mu} = \phi\big((I-\mu T)^{-1}\big) \to \phi(I) \qquad \text{as } \mu \to 0,$$

so $f'(0) = \phi(I)$ and f is holomorphic on D.

If $|\mu| < \|T\|^{-1}$ then the Neumann series of Theorem 4.21 and the continuity of ϕ give that

$$f(\mu) = \mu \sum_{n=0}^{\infty} \mu^n \phi(T^n) = \sum_{n=1}^{\infty} \phi(T^{n-1})\, \mu^n;$$

by the uniqueness of Taylor series, this identity is valid everywhere on D. The final claim is an immediate consequence of Cauchy's formula for Taylor coefficients. □

Theorem 5.6. (Beurling–Gelfand) If $T \in B(\mathsf{H})$ then

$$\rho(T) = \lim_{n\to\infty} \|T^n\|^{1/n} = \inf\big\{\|T^n\|^{1/n} : n \geqslant 1\big\}.$$

Proof. Note first that, by Corollary 4.22, the spectral radius $\rho(T) \leqslant \|T\|$. By the polynomial spectral mapping theorem, Theorem 5.3, if $n \geqslant 1$ then $\sigma(T^n) = \sigma(T)^n$ and therefore

$$\rho(T) = \rho(T^n)^{1/n} \leqslant \|T^n\|^{1/n}.$$

To find an upper bound, we use Lemma 5.5. Fix r such that $0 < r < \rho(T)^{-1}$ and recall that, by Proposition 4.27, the resolvent function $\lambda \mapsto (\lambda I - T)^{-1}$ is strongly holomorphic, so continuous, on $\mathbb{C} \setminus \sigma(T)$; thus

$$M_r := \sup\big\{\|(I-zT)^{-1}\| : |z| = r\big\} < \infty.$$

Now let $\phi \in B(\mathsf{H})^*$ and note that, by (5.2) from Lemma 5.5,

$$|\phi(T^{n-1})| \leqslant \|\phi\|\, M_r\, r^{-n+1} \qquad \text{for every } n \geqslant 1.$$

Hence the Hahn–Banach theorem, Theorem 4.28, implies that

$$\|T^n\|^{1/n} \leqslant M_r^{1/n} r^{-1} \to r^{-1} \qquad \text{as } n \to \infty,$$

and therefore $\lim_{n\to\infty} \|T^n\|^{1/n} = \rho(T)$, as claimed.

For the final claim, note that if $(a_n)_{n\geqslant 1} \subseteq \mathbb{R}$ is such that $a_n \to a$ and $a_n \geqslant a$ for every $n \geqslant 1$, then $a = \inf\{a_n : n \geqslant 1\}$. ☐

Corollary 5.7. If $T \in B(\mathsf{H})$ is *self adjoint*, *i.e.*, $T = T^*$, then $\rho(T) = \|T\|$.

Proof. The C^* identity of Proposition 2.19 gives that $\|T^2\| = \|T^*T\| = \|T\|^2$, and if $\|T^{2^n}\| = \|T\|^{2^n}$ for some $n \geqslant 1$ then

$$\|T^{2^{n+1}}\| = \|T^{2^n}T^{2^n}\| = \|T^{2^n}\|^2 = \|T\|^{2^{n+1}}.$$

It follows by induction that $\|T^{2n}\| = \|T\|^{2n}$ for every $n \geqslant 1$, and therefore, since every subsequence of a convergent sequence is convergent to the same limit,

$$\rho(T) = \lim_{n\to\infty} \|T^{2^n}\|^{1/2^n} = \|T\|. \qquad ☐$$

5.3 The continuous functional calculus

The following observation is the key to extending beyond polynomial functions. It uses both the results developed so far in this chapter.

Proposition 5.8. Let $T \in B(\mathsf{H})$ be self adjoint. If $p(z) \in \mathbb{C}[z]$ then

$$\|p(T)\| = \sup\big\{|p(\lambda)| : \lambda \in \sigma(T)\big\} =: \|p\|_{\sigma(T)}.$$

Proof. Note first that, by the C^* identity and the spectral radius formula for self-adjoint operators,

$$\|p(T)\|^2 = \|p(T)^*p(T)\| = \big\||p|^2(T)\big\| = \rho\big(|p|^2(T)\big),$$

where $|p|^2(z) := \overline{p}(z)p(z) \in \mathbb{C}[z]$. Furthermore, by the polynomial spectral mapping theorem and Corollary 4.33, which implies that $\sigma(T) \subseteq \mathbb{R}$,

$$\rho\big(|p|^2(T)\big) := \sup\big\{|\lambda| : \lambda \in \sigma\big(|p|^2(T)\big)\big\} = \sup\big\{|p(\lambda)|^2 : \lambda \in \sigma(T)\big\},$$

which gives the result upon taking square roots. ☐

Corollary 5.9. Let $T \in B(\mathsf{H})$ be self adjoint. If $p(z),\ q(z) \in \mathbb{C}[z]$ are such that $p(\lambda) = q(\lambda)$ for every $\lambda \in \sigma(T)$ then $p(T) = q(T)$.

Proof. If $r(z) := p(z) - q(z) \in \mathbb{C}[z]$ then $r(\lambda) = p(\lambda) - q(\lambda) = 0$ for every $\lambda \in \sigma(T)$. Thus, by Proposition 5.8,

$$\|p(T) - q(T)\| = \|r(T)\| = \sup\big\{|r(\lambda)| : \lambda \in \sigma(T)\big\} = 0. \qquad ☐$$

The next tool we need is the following, which shows that continuous functions can be well approximated on compact sets by polynomials.

Theorem 5.10. (Stone–Weierstrass) Suppose $K \subseteq \mathbb{R}$ is compact and let $f : K \to \mathbb{C}$ be continuous. There exists a sequence of polynomials $\bigl(p_n(z)\bigr)_{n\geqslant 1} \subseteq \mathbb{C}[z]$ such that

$$\|f - p_n\|_K := \sup\bigl\{|f(x) - p_n(x)| : x \in K\bigr\} \to 0 \qquad \text{as } n \to \infty.$$

Proof. This result is proved in MATH317 for the case where K is a compact interval, as a corollary of Fejér's theorem.

If $K \subseteq [a, b]$, where a, $b \in \mathbb{R}$ are such that $a < b$, then $[a, b] \setminus K$ is open, so is a countable union of disjoint open intervals. On each of these intervals, one may extend f by linearity, and the resulting function g will be continuous. The MATH317 version can now be applied to g, and the resulting sequence will also approximate f uniformly on K, as required. □

If the operator $T \in B(\mathsf{H})$ is self adjoint then $\sigma(T)$ is a compact subset of the real line, by Corollary 4.33. Hence any continuous function on $\sigma(T)$ may be uniformly approximated by polynomials, by the Stone–Weierstrass theorem. We exploit this fact to define a continuous function of T as a limit of polynomial functions of T, as follows.

Theorem 5.11. (Continuous functional calculus) Let $T \in B(\mathsf{H})$ be self adjoint. There exists a linear transformation

$$\epsilon_T : C\bigl(\sigma(T)\bigr) \to B(\mathsf{H});\ f \mapsto f(T)$$

which is multiplicative, $*$-preserving and isometric:

$$(fg)(T) = f(T)g(T), \qquad f(T)^* = \overline{f}(T) \quad \text{and} \quad \|f(T)\| = \|f\|_{\sigma(T)}.$$

Furthermore, if f is a polynomial function, *i.e.*, there exists $p(z) \in \mathbb{C}[z]$ with $f(x) = p(x)$ for all $x \in \sigma(T)$, then $\epsilon_T(f) = \epsilon_T\bigl(p(z)\bigr)$ as in Definition 5.2.

Proof. Given $f \in C\bigl(\sigma(T)\bigr)$, let $\bigl(p_n(z)\bigr)_{n\geqslant 1} \subseteq \mathbb{C}[z]$ be such that $\|p_n - f\|_{\sigma(T)} \to 0$. The sequence $\bigl(p_n(T)\bigr)_{n\geqslant 1}$ is Cauchy in $B(\mathsf{H})$, because

$$\|p_m(T) - p_n(T)\| = \|(p_m - p_n)(T)\| = \|p_m - p_n\|_{\sigma(T)} \leqslant \|p_m - f\|_{\sigma(T)} + \|f - p_n\|_{\sigma(T)} \to 0$$

as m, $n \to \infty$ Thus $\bigl(p_n(T)\bigr)_{n\geqslant 1}$ is convergent, since $B(\mathsf{H})$ is complete; we let

$$f(T) := \lim_{n\to\infty} p_n(T).$$

To see this definition is independent of the choice of approximating sequence $\bigl(p_n(z)\bigr)_{n\geqslant 1}$, suppose $\bigl(p'_n(z)\bigr)_{n\geqslant 1} \subseteq \mathbb{C}[z]$ is such that $\|p'_n - f\|_{\sigma(T)} \to 0$. As above, the limit

$$f(T)' := \lim_{n\to\infty} p'_n(T)$$

exists, and

$$\|f(T)' - f(T)\| = \lim_{n\to\infty} \|p_n'(T) - p_n(T)\| = \lim_{n\to\infty} \|p_n' - p_n\|_{\sigma(T)} = \|f - f\|_{\sigma(T)} = 0,$$

so $f(T)' = f(T)$ and $f(T)$ is well defined. It also follows from this that $f(T) = \epsilon_T\big(p(z)\big)$ if $p(z) \in \mathbb{C}[z]$ is such that $p(x) = f(x)$ for every $x \in \sigma(T)$, since in this case we can take $p_n(z) = p(z)$ for every $n \geqslant 1$.

To see that the mapping $f \mapsto f(T)$ is linear and multiplicative, let $g \in C\big(\sigma(T)\big)$ and suppose $\big(q_n(z)\big)_{n\geqslant 1} \subseteq \mathbb{C}[z]$ is such that $\|q_n - g\|_{\sigma(T)} \to 0$. Then

$$\|(p_n + \lambda q_n) - (f + \lambda g)\|_{\sigma(T)} \leqslant \|p_n - f\|_{\sigma(T)} + |\lambda|\, \|q_n - g\|_{\sigma(T)} \to 0$$

and therefore

$$(f + \lambda g)(T) = \lim_{n\to\infty} (p_n + \lambda q_n)(T) = \lim_{n\to\infty} p_n(T) + \lambda \lim_{n\to\infty} q_n(T) = f(T) + \lambda g(T).$$

Similarly, since $\big(\|p_n\|_{\sigma(T)}\big)_{n\geqslant 1}$ is convergent, it is bounded, and therefore

$$\begin{aligned}\|p_n q_n - fg\|_{\sigma(T)} &\leqslant \|p_n(q_n - g)\|_{\sigma(T)} + \|(p_n - f)g\|_{\sigma(T)} \\ &\leqslant \|p_n\|_{\sigma(T)}\, \|q_n - g\|_{\sigma(T)} + \|p_n - f\|_{\sigma(T)}\, \|g\|_{\sigma(T)} \to 0.\end{aligned}$$

Hence

$$(fg)(T) = \lim_{n\to\infty} (p_n q_n)(T) = \lim_{n\to\infty} p_n(T) q_n(T) = \lim_{n\to\infty} p_n(T) \lim_{n\to\infty} q_n(T) = f(T)g(T).$$

Furthermore, since $\big\|\overline{f}\big\|_{\sigma(T)} = \|f\|_{\sigma(T)}$, it follows that

$$\big\|\overline{p_n} - \overline{f}\big\|_{\sigma(T)} = \big\|\overline{p_n - f}\big\|_{\sigma(T)} = \|p_n - f\|_{\sigma(T)}$$

and therefore, since $\|S^*\| = \|S\|$ for any $S \in B(\mathsf{H})$, so

$$f(T)^* = \lim_{n\to\infty} p_n(T)^* = \lim_{n\to\infty} \overline{p_n}(T) = \overline{f}(T).$$

Finally,

$$\|f(T)\| = \lim_{n\to\infty} \|p_n(T)\| = \lim_{n\to\infty} \|p_n\|_{\sigma(T)} = \|f\|_{\sigma(T)}. \qquad \square$$

Remark 5.12. Let $T \in B(\mathsf{H})$ be self adjoint and let $f \in C\big(\sigma(T)\big)$.

(i) If f is real valued then $\overline{f} = f$, so $f(T)^* = \overline{f}(T) = f(T)$ and thus $f(T)$ is self adjoint.

(ii) If f takes values in the unit circle $\{z \in \mathbb{C} : |z| = 1\}$ then $|f|^2 = \overline{f} f = 1$, regarded as the constant function $\sigma(T) \to \mathbb{C};\ z \mapsto 1$, and

$$f(T)^* f(T) = \overline{f}(T) f(T) = |f|^2(T) = I = f(T) f(T)^*,$$

so $f(T)$ is a unitary operator.

The polynomial spectral mapping theorem extends as we might hope.

Theorem 5.13. (Continuous spectral mapping) Let $T \in B(\mathsf{H})$ be self adjoint, and let $f : \sigma(T) \to \mathbb{C}$ be continuous. Then

$$\sigma\big(f(T)\big) = f\big(\sigma(T)\big).$$

Proof. Let $\lambda \in \sigma(T)$ and let $\big(p_n(z)\big)_{n\geqslant 1} \subseteq \mathbb{C}[z]$ be such that $\|p_n - f\|_{\sigma(T)} \to 0$. The polynomial spectral mapping theorem implies that $p_n(\lambda) \in \sigma\big(p_n(T)\big)$ for every $n \geqslant 1$ and

$$B(\mathsf{H}) \setminus B(\mathsf{H})^\times \ni p_n(\lambda)I - p_n(T) \to f(\lambda)I - f(T);$$

as $B(\mathsf{H})^\times$ is open, by Theorem 4.23, it follows that this last operator is also not invertible and so $f(\lambda) \in \sigma\big(f(T)\big)$. Thus $f\big(\sigma(T)\big) \subseteq \sigma\big(f(T)\big)$.

Now suppose $\mu \in \mathbb{C} \setminus f\big(\sigma(T)\big)$. Then

$$g : \sigma(T) \to \mathbb{C};\ z \mapsto \big(\mu - f(z)\big)^{-1}$$

is continuous and

$$g(T)\big(\mu I - f(T)\big) = I = \big(\mu I - f(T)\big)g(T),$$

so $\mu \in \mathbb{C} \setminus \sigma\big(f(T)\big)$. Thus $\mathbb{C} \setminus f\big(\sigma(T)\big) \subseteq \mathbb{C} \setminus \sigma\big(f(T)\big)$ and the result follows. $\square$

Since spectra behave as one would expect under the continuous functional calculus, we may also consider the composition of continuous functions.

Theorem 5.14. Let $T \in B(\mathsf{H})$ be self adjoint, let $f : \sigma(T) \to \mathbb{R}$ be continuous, and let $S = f(T)$. If $g : \sigma(S) \to \mathbb{C}$ is continuous then $g \circ f : \sigma(T) \to \mathbb{C}$ is well defined, continuous and such that

$$(g \circ f)(T) = g\big(f(T)\big).$$

Proof. Note first that S is self adjoint, by Remark 5.12(i), so $g(S) = g\big(f(T)\big)$ is well defined. Furthermore, Theorem 5.13 implies that $f\big(\sigma(T)\big) = \sigma\big(f(T)\big) = \sigma(S)$, so $g \circ f$ and $(g \circ f)(T)$ are well defined. Since the functional calculus is multiplicative,

$$f(T)^n = (f^n)(T) \qquad \text{for every } n \geqslant 0,$$

where $f^n(z) := f(z)^n$ for every $z \in \sigma(T)$. By linearity of the functional calculus, it follows that

$$p\big(f(T)\big) = (p \circ f)(T) \qquad \text{for every } p(z) \in \mathbb{C}[z].$$

Finally, let $\big(p_n(z)\big)_{n\geqslant 1} \subseteq \mathbb{C}[z]$ be such that $\|p_n - g\|_{\sigma(S)} \to 0$, and note that

$$g\big(f(T)\big) = g(S) = \lim_{n\to\infty} p_n(S) = \lim_{n\to\infty} p_n\big(f(T)\big) = \lim_{n\to\infty} (p_n \circ f)(T) = (g \circ f)(T),$$

since

$$\begin{aligned}\|p_n \circ f - g \circ f\|_{\sigma(T)} &= \sup\big\{\big|p_n\big(f(\lambda)\big) - g\big(f(\lambda)\big)\big| : \lambda \in \sigma(T)\big\}\\ &= \sup\big\{|p_n(\mu) - g(\mu)| : \mu \in \sigma\big(f(T)\big)\big\}\\ &= \|p_n - g\|_{\sigma(S)}.\end{aligned}$$

$\square$

Six Operator order and applications

Throughout this chapter, let H be a Hilbert space.

6.1 Operator order

Remark 6.1. If $T \in B(\mathsf{H})$ is self adjoint and $x \in \mathsf{H}$ then

$$\langle x, Tx\rangle = \langle T^* x, x\rangle = \langle Tx, x\rangle = \overline{\langle x, Tx\rangle},$$

so $\langle x, Tx\rangle$ is a real number.

Definition 6.2. Let $T \in B(\mathsf{H})$ be self adjoint. Then T is *positive* if and only

$$\langle x, Tx\rangle \geqslant 0 \qquad \text{for every } x \in \mathsf{H},$$

in which case we write $T \geqslant 0$ or $0 \leqslant T$.

More generally, if S, $T \in B(\mathsf{H})$ are self adjoint then $S \leqslant T$ (or, equivalently, $T \geqslant S$) if and only if

$$\langle x, Sx\rangle \leqslant \langle x, Tx\rangle \qquad \text{for every } x \in \mathsf{H}.$$

Remark 6.3. The binary relation $\leqslant$ on the set of self-adjoint elements of $B(\mathsf{H})$ is a *partial order*: it is

(i) *reflexive*, *i.e.*, $T \leqslant T$ for every self-adjoint $T \in B(\mathsf{H})$,

(ii) *antisymmetric*, *i.e.*, if $S, T \in B(\mathsf{H})$ are self adjoint and such that $S \leqslant T$ and $T \leqslant S$ then $S = T$, and

(iii) *transitive*, *i.e.*, if R, S, $T \in B(\mathsf{H})$ are self adjoint and such that $R \leqslant S$ and $S \leqslant T$ then $R \leqslant T$.

The first and last are immediate; the second follows from Proposition 2.10.

Remark 6.4. The set of positive operators on H is denoted $B(\mathsf{H})_+$ and is a *cone*: it is closed under addition and multiplication by non-negative scalars.

Furthermore, for every self-adjoint operator $T \in B(\mathsf{H})$, the Cauchy–Schwarz inequality and the fact that $\|-T\| = \|T\|$ imply that

$$-\|T\|\, I \leqslant T \leqslant \|T\|\, I.$$

Proposition 6.5. If $S, T \in B(\mathsf{H})$ are self adjoint and such that $S \leqslant T$ then

(i) $S + R \leqslant T + R$ for every self-adjoint operator $R \in B(\mathsf{H})$,

(ii) $C^*SC \leqslant C^*TC$ for every $C \in B(\mathsf{H})$, and

(iii) given $c \in \mathbb{R}$, it holds that $cS \leqslant cT$ if $c \geqslant 0$ and $cS \geqslant cT$ if $c \leqslant 0$.

Proof. These claims are straightforward consequences of the definition, together with properties of the inner product and the adjoint. □

Remark 6.6. Since $I \geqslant 0$, it follows from Proposition 6.5(ii) that $C^*C \in B(\mathsf{H})_+$ for any $C \in B(\mathsf{H})$.

Proposition 6.7. If $T \in B(\mathsf{H})_+$ is invertible then $T^{-1} \in B(\mathsf{H})_+$.

Proof. If $y \in \mathsf{H}$ then $y = Tx$ for some $x \in \mathsf{H}$, so

$$\langle y, T^{-1}y\rangle = \langle Tx, T^{-1}Tx\rangle = \langle Tx, x\rangle = \langle x, Tx\rangle \geqslant 0. \qquad \square$$

Definition 6.8. If $T \in B(\mathsf{H})$ is self adjoint and such that $T \geqslant cI$ for some $c > 0$ then T is *uniformly positive*.

Theorem 6.9. Let $T \in B(\mathsf{H}; \mathsf{K})$. Then T is invertible if and only if both T^*T and TT^* are uniformly positive.

Proof. Suppose first that T is invertible. Then the operator-norm inequality implies that

$$\|x\| = \|T^{-1}Tx\| \leqslant \|T^{-1}\|\,\|Tx\| \qquad \text{for every } x \in \mathsf{H},$$

so

$$\langle x, T^*Tx\rangle = \|Tx\|^2 \geqslant \|T^{-1}\|^{-2}\|x\|^2 = \langle x, cx\rangle \qquad \text{for every } x \in \mathsf{H},$$

where $c := \|T^{-1}\|^{-2}$, so $T^*T \geqslant cI$. Furthermore, since T^* is also invertible and

$$\|(T^*)^{-1}\| = \|(T^{-1})^*\| = \|T^{-1}\|,$$

the same working, with T replaced by T^*, shows that $TT^* \geqslant cI$ also.

For the converse, suppose $c > 0$ is such that $T^*T \geqslant cI$ and $TT^* \geqslant cI$. Then

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle \geqslant \langle x, cx\rangle = c\|x\|^2 \qquad \text{for every } x \in \mathsf{H},$$

so T is injective and has closed range, by Lemma 2.25. Similar working shows that $\ker T^* = \{0\}$. By the kernel-adjoint-range relation, Theorem 2.24, it follows that

$$\operatorname{im} T = \big((\operatorname{im} T)^\perp\big)^\perp = (\ker T^*)^\perp = \{0\}^\perp = \mathsf{K},$$

so T is invertible with $\|T^{-1}\| \leqslant c^{-1/2}$, since

$$\|T^{-1}y\| \leqslant c^{-1/2}\,\|TT^{-1}y\| = c^{-1/2}\,\|y\| \qquad \text{for every } y \in \mathsf{K}. \qquad \square$$

Proposition 6.10. If $S, T \in B(\mathsf{H})_+$ are such that $S \leqslant T$ then $\|S\| \leqslant \|T\|$.

Proof. Note first that $S \leqslant T \leqslant \|T\|I$, so $S \leqslant \|T\|I$. For convenience, let $c := \|T\|$. As $cI - S \geqslant 0$, Corollary 4.33 implies that $\sigma(cI - S) \subseteq \big[0, \|cI - S\|\big]$. Theorem 5.3, the polynomial spectral mapping theorem, and the positivity of S now give that

$$\sigma(S) \subseteq [c - \|cI - S\|, c] \cap \mathbb{R}_+ \subseteq [0, c].$$

Hence $\|T\| = c \geqslant \rho(S) = \|S\|$, by Corollary 5.7, the spectral radius formula for self-adjoint operators. $\square$

Theorem 6.11. Let $T \in B(\mathsf{H})$ be self adjoint and let $f, g \in C\big(\sigma(T)\big)$ be real valued. Then

$$f(T) \geqslant g(T) \qquad \iff \qquad f(t) \geqslant g(t) \quad \text{for every } t \in \sigma(T).$$

Proof. If $f(T) \geqslant g(T)$ then $f(T) - g(T) = (f - g)(T) \geqslant 0$. Hence Corollary 4.33 and the continuous spectral mapping theorem, Theorem 5.13, imply that

$$\mathbb{R}_+ \supseteq \sigma\big((f-g)(T)\big) = \{f(t) - g(t) : t \in \sigma(T)\},$$

so $f(t) \geqslant g(t)$ for every $t \in \sigma(T)$.

Conversely, if $f(t) \geqslant g(t)$ for every $t \in \sigma(T)$ then

$$h : \sigma(T) \to \mathbb{R}_+;\ t \mapsto \big(f(t) - g(t)\big)^{1/2}$$

is well defined and continuous, and $h(T)$ is self adjoint, by Theorem 5.11. Hence

$$f(T) - g(T) = (f - g)(T) = h(T)^2 = h(T)^* h(T) \geqslant 0. \qquad \square$$

6.2 Square roots and absolute values

Lemma 6.12. Let $T \in B(\mathsf{H})$ be self adjoint and let $f \in C\big(\sigma(T)\big)$. If $S \in B(\mathsf{H})$ commutes with T, *i.e.*, $ST = TS$, then S commutes with $f(T)$.

Proof. As S commutes with T, it commutes with $p(T)$ for every polynomial $p(z) \in \mathbb{C}[z]$. Now let $\big(p_n(z)\big)_{n\geqslant 1}$ be such that $\|p_n - f\|_{\sigma(T)} \to 0$ and note that

$$f(T)S = \lim_{n\to\infty} p_n(T)S = \lim_{n\to\infty} S\,p_n(T) = Sf(T). \qquad \square$$

Theorem 6.13. Let $T \in B(\mathsf{H})_+$. There exists a unique operator $S \in B(\mathsf{H})_+$ such that $T = S^2$.

Proof. Since $\sigma(T) \subseteq \mathbb{R}_+$, by Corollary 4.33, and

$$f : \mathbb{R}_+ \to \mathbb{R}_+;\ t \mapsto t^{1/2}$$

is continuous, the operator $S := f|_{\sigma(T)}(T)$ given by the functional calculus, Theorem 5.11, is well defined and self adjoint. Furthermore, the multiplicativity in Theorem 5.11 implies that

$$S^2 = \big(f|_{\sigma(T)}(T)\big)^2 = T.$$

To see that S is positive, note that $\sigma(S) \subseteq \mathbb{R}_+$, by the continuous spectral mapping theorem, Theorem 5.13, and so $R := f|_{\sigma(S)}(S)$ is also well defined and self adjoint, and $S = R^2 = R^*R \geqslant 0$.

For uniqueness, suppose S is as above and let $R \in B(\mathsf{H})_+$ be such that $R^2 = T$. Then R commutes with T, since $RT = R^3 = TR$, and therefore R commutes with S, by Lemma 6.12. Furthermore, as $S - R$ is self adjoint,

$$(S-R)S(S-R) + (S-R)R(S-R) = (S^2 - R^2)(S-R) = 0.$$

As both terms on the left-hand side of this equation are positive, they must both be zero, by Proposition 2.10. Hence their difference

$$(S-R)S(S-R) - (S-R)R(S-R) = (S-R)^3$$

is zero, and so

$$0 = \|(S-R)^3(S-R)\| = \|(S-R)^4\| = \|(S-R)^2\|^2 = \|S-R\|^4,$$

by the C^* identity applied twice. Thus $R = S$ and the square root is unique. □

Definition 6.14. Let $T \in B(\mathsf{H})_+$. The operator $S \in B(\mathsf{H})_+$ given by Theorem 6.13 is the *square root* of T, and is denoted $T^{1/2}$.

Proposition 6.15. Let $T \in B(\mathsf{H})$. The following are equivalent.

(i) T is positive.

(ii) There exists $C \in B(\mathsf{H})$ such that $T = C^*C$.

(iii) There exists $S \in B(\mathsf{H})_+$ such that $T = S^2$.

(iv) T is self adjoint and $\sigma(T) \subseteq \mathbb{R}_+$.

Proof. The proof of existence in Theorem 6.13 shows that (iv) implies (iii). That (iii) implies (ii) and (ii) implies (i) are immediate, and that (i) implies (iv) follows from Corollary 4.33. □

Proposition 6.16. Let $T \in B(\mathsf{H})_+$ be invertible. Then the square root $T^{1/2}$ is invertible, and $(T^{-1})^{1/2} = (T^{1/2})^{-1}$.

Proof. Recall that T^{-1} is positive, by Proposition 6.7. Invertibility of $T^{1/2}$ follows from its definition and the continuous spectral mapping theorem, Theorem 5.13. Finally,

$$\big((T^{1/2})^{-1}\big)^2 = (T^{1/2})^{-1}(T^{1/2})^{-1} = (T^{1/2}\,T^{1/2})^{-1} = T^{-1},$$

so $(T^{1/2})^{-1} = (T^{-1})^{1/2}$ by the uniqueness part of Theorem 6.13; note that $(T^{1/2})^{-1}$ is positive, because $T^{1/2}$ is. $\square$

Remark 6.17. Since $(T^{1/2})^{-1}$ and $(T^{-1})^{1/2}$ are equal, henceforth we let this operator be denoted by $T^{-1/2}$.

The square root allows us to prove that invertibility and uniform positivity are equivalent for positive operators.

Proposition 6.18. Let $T \in B(\mathsf{H})_+$. Then T is invertible if and only if T is uniformly positive.

Proof. If T is uniformly positive, so that $T \geqslant cI$ for some $c > 0$, then, by the second part of Remark 6.4 and Proposition 6.5,

$$\|T\|I \geqslant T \geqslant cI \qquad \Longrightarrow \qquad 0 \leqslant I - \|T\|^{-1}T \leqslant (1 - \|T\|^{-1}\,c)I.$$

Hence, by Proposition 6.10, it follows that

$$\|I - \|T\|^{-1}\,T\| \leqslant 1 - \|T\|^{-1}\,c < 1,$$

so $\|T\|^{-1}\,T$ is invertible, by Theorem 4.21. Hence T is invertible.

Conversely, if T is invertible then the operator-norm inequality gives that

$$\|x\| = \|T^{-1/2}\,T^{1/2}x\| \leqslant \|T^{-1/2}\|\,\|T^{1/2}x\|$$

and therefore T is uniformly positive, since

$$\langle x, Tx\rangle = \|T^{1/2}x\|^2 \geqslant \|T^{-1/2}\|^{-2}\,\|x\|^2 = \langle x, \|T^{-1/2}\|^2 x\rangle \qquad \text{for every } x \in \mathsf{H}. \qquad \square$$

The previous result allows us to give a more precise description of the spectrum of a self-adjoint operator.

Theorem 6.19. Let $T \in B(\mathsf{H})$ be self adjoint. Then $\{m_T, M_T\} \subseteq \sigma(T) \subseteq [m_T, M_T]$ and $\|T\| = \max\big\{|m_T|, |M_T|\big\}$, where

$$m_T := \inf\{\langle x, Tx\rangle : \|x\| = 1\} \qquad \text{and} \qquad M_T := \sup\{\langle x, Tx\rangle : \|x\| = 1\}.$$

Proof. If $m_T \notin \sigma(T)$ then $T - m_T I$ is positive and invertible, so uniformly positive, but then $\langle x, Tx\rangle$ is bounded away from m_T whenever $\|x\| = 1$, a contradiction. A similar argument shows that $M_T \in \sigma(T)$ also. The other inclusion holds by Corollary 4.33, and the last claim now follows because $\|T\| = \rho(T)$, by Corollary 5.7. $\square$

6.3 Polar decomposition

Definition 6.20. If $T \in B(\mathsf{H};\mathsf{K})$ then the *absolute value* of T is $|T| := (T^*T)^{1/2} \in B(\mathsf{H})$.

Lemma 6.21. If $T \in B(\mathsf{H};\mathsf{K})$ is invertible then so is $|T| \in B(\mathsf{H})^\times$.

Proof. If T is invertible then so is T^*, and therefore T^*T is invertible. Thus, by the continuous spectral mapping theorem, Theorem 5.13,

$$0 \neq \sigma(T^*T)^{1/2} = \sigma\big(|T|\big),$$

so $|T|$ is invertible. □

Every non-zero complex number z may be written in the form $|z|w$, where $|z|$ is non-negative and invertible, and w has modulus one, so $\overline{w}w = w\overline{w} = 1$. Essentially the same result holds for operators.

Theorem 6.22. Let $T \in B(\mathsf{H};\mathsf{K})^\times$. There exists a unique unitary operator $U \in B(\mathsf{H};\mathsf{K})$ such that $T = U|T|$.

Proof. For existence, let $U := T|T|^{-1}$. Then

$$U^*U = |T|^{-1}T^*T|T|^{-1} = |T|^{-1}|T|^2|T|^{-1} = I$$

and

$$UU^* = T|T|^{-1}|T|^{-1}T^* = T|T|^{-2}T^* = TT^{-1}(T^*)^{-1}T^* = I.$$

Uniqueness is immediate, by the invertibility of $|T|^{-1}$. □

Remark 6.23. Polar decomposition can be extended more generally, but the unitary operator U is replaced by a partial isometry V; this is an operator which is isometric on the orthogonal complement of its kernel. Uniqueness of V is a more delicate matter, and these issues are beyond the scope of this module.

Seven Spectral theory for compact self-adjoint operators

Throughout this chapter, let G, H, K and L be Hilbert spaces.

7.1 Dirac dyads

Definition 7.1. Let $x \in \mathsf{H}$ and $y \in \mathsf{K}$. Then the *Dirac dyad*

$$|y\rangle\langle x| : \mathsf{H} \to \mathsf{K};\ z \mapsto \langle x, z\rangle y.$$

Proposition 7.2. If $x \in \mathsf{H}$ and $y \in \mathsf{K}$ then $|y\rangle\langle x| \in B(\mathsf{H};\mathsf{K})$, with $\big\| |y\rangle\langle x| \big\| = \|y\|\,\|x\|$ and $|y\rangle\langle x|^* = |x\rangle\langle y|$.

Furthermore, if $S \in B(\mathsf{G};\mathsf{H})$ and $T \in B(\mathsf{K};\mathsf{L})$ then $T|y\rangle\langle x|S = |Ty\rangle\langle S^*x|$.

Finally, if $x \in \mathsf{H}$, y, $y' \in \mathsf{K}$ and $z \in \mathsf{L}$ then $|z\rangle\langle y'||y\rangle\langle x| = \langle y', y\rangle |z\rangle\langle x|$.

Proof. The dyad $|y\rangle\langle x|$ inherits linearity from the inner product, and the Cauchy–Schwarz inequality implies that

$$\| |y\rangle\langle x|u\| = \|\langle x, u\rangle y\| = |\langle x, u\rangle|\,\|y\| \leqslant \|x\|\,\|y\|\,\|u\| \qquad \text{for every } u \in \mathsf{H}.$$

Thus $|y\rangle\langle x| \in B(\mathsf{H};\mathsf{K})$, with $\big\| |y\rangle\langle x| \big\| \leqslant \|y\|\,\|x\|$. For the reverse inequality, note that

$$\| |y\rangle\langle x|x\| = \|x\|^2\,\|y\| \quad \Longrightarrow \quad \big\| |y\rangle\langle x| \big\| \geqslant \|x\|\,\|y\|.$$

Furthermore, if $u \in \mathsf{H}$ and $v \in \mathsf{K}$ then

$$\langle v, |y\rangle\langle x|u\rangle = \langle v, \langle x, u\rangle y\rangle = \langle v, y\rangle\,\langle x, u\rangle = \langle \langle y, v\rangle x, u\rangle = \langle |x\rangle\langle y|v, u\rangle,$$

so $|y\rangle\langle x|^* = |x\rangle\langle y|$, and

$$\langle v, T|y\rangle\langle x|Su\rangle = \langle T^*v, \langle x, Su\rangle y\rangle = \langle v, \langle S^*x, u\rangle Ty\rangle = \langle v, |Ty\rangle\langle S^*x|u\rangle,$$

so $T|y\rangle\langle x|S = |Ty\rangle\langle S^*x|$. Finally,

$$|z\rangle\langle y'||y\rangle\langle x|u = |z\rangle\langle y'|\big(\langle x, u\rangle y\big) = \langle x, u\rangle\langle y', y\rangle z = \langle y', y\rangle|z\rangle\langle x|u. \qquad \square$$

7.2 Finite-rank operators

Definition 7.3. An linear transformation $T \in L(\mathsf{H};\mathsf{K})$ has *finite rank* if and only if $\operatorname{im} T$ is finite dimensional.

Proposition 7.4. Let $T \in L(\mathsf{H};\mathsf{K})$. Then T has finite rank if and only if there exist vectors $x_1, \ldots, x_n \in \mathsf{H}$ and $y_1, \ldots, y_n \in \mathsf{K}$ such that

$$T = \sum_{j=1}^{n} |y_j\rangle\langle x_j|.$$

Proof. If T is of this form then $\operatorname{im} T$ is spanned by the finite set $\{y_1, \ldots, y_n\}$. Hence $\operatorname{im} T$ is finite dimensional and T has finite rank.

Conversely, if $\operatorname{im} T$ is finite dimensional then it has a finite spanning set and so, by the Gram–Schmidt process, a finite orthonormal basis $\{y_1, \ldots, y_n\}$. Hence if $x \in \mathsf{H}$ then

$$Tx = \sum_{j=1}^{n} \langle y_j, Tx\rangle\, y_j = \sum_{j=1}^{n} \langle T^* y_j, x\rangle\, y_j = \sum_{j=1}^{n} |y_j\rangle\langle T^* y_j| x,$$

so the desired representation holds with $x_j = T^* y_j$ for $j = 1, \ldots, n$. □

Notation 7.5. The set of finite-rank operators from H to K is denoted $B_{00}(\mathsf{H};\mathsf{K})$, or $B_{00}(\mathsf{H})$ if $\mathsf{H} = \mathsf{K}$.

It is immediate from Proposition 7.4 that $B_{00}(\mathsf{H};\mathsf{K})$ is a subspace of $B(\mathsf{H};\mathsf{K})$.

Corollary 7.6. If $S \in B_{00}(\mathsf{H};\mathsf{K})$, $R \in B(\mathsf{G};\mathsf{H})$ and $T \in B(\mathsf{K};\mathsf{L})$ then $TSR \in B_{00}(\mathsf{G};\mathsf{L})$.

Proof. This is immediate from Propositions 7.4 and 7.2. □

Corollary 7.7. Let $T \in B(\mathsf{H};\mathsf{K})$. Then T has finite rank if and only if T^* has finite rank.

Proof. This is also immediate from Propositions 7.4 and 7.2. □

Remark 7.8. In algebraic terms, the above results imply that the set $B_{00}(\mathsf{H})$ of finite-rank operators on H is a two-sided $*$-ideal in the $*$-algebra of bounded operators $B(\mathsf{H})$.

A *complex associative algebra* or *algebra* A is simultaneously a ring and a complex vector space, such that the scalar and ring multiplications are compatible:

$$\lambda(ab) = (\lambda a)b = a(\lambda b) \qquad \text{for every } a, b \in A \text{ and } \lambda \in \mathbb{C}.$$

Key examples are the $n \times n$ complex matrices $M_n(\mathbb{C})$, the linear transformations $L(V)$ on a complex vector space V, the bounded operators $B(\mathsf{H})$ on a Hilbert space H and the complex-valued functions $C(K)$ on a compact set $K \subseteq \mathbb{C}$.

A complex algebra A is *unital* if there is a identity element for the ring multiplication, *i.e.,* there exists $1 \in A$, called the *unit*, such that $1a = a1 = a$ for every $a \in A$.

A $*$-*algebra* is an algebra A equipped with an *involution*, which is a conjugate-linear map $A \to A$; $a \mapsto a^*$ such that $(a^*)^* = a$ and $(ab)^* = b^* a^*$ for every $a, b \in A$. If A is unital then it is required that the unit $1^* = 1$.

A *two-sided ideal* or *ideal* I of a complex algebra A is a subspace such that $cba \in I$ whenever $c, a \in A$ and $b \in I$. The ideal I is a $*$-*ideal* if A is a $*$-algebra and I is closed under the involution, *i.e.,* $a^* \in I$ for every $a \in I$.

7.3 Compact operators

Definition 7.9. A linear transformation $T \in L(\mathsf{H}; \mathsf{K})$ is compact if, for any bounded sequence $(x_n)_{n \geqslant 1} \subseteq \mathsf{H}$, *i.e.,* there exists $M > 0$ such that $\|x_n\| \leqslant M$ for every $n \geqslant 1$, then $(Tx_n)_{n \geqslant 1}$ has a convergent subsequence.

Proposition 7.10. If $S \in L(\mathsf{H}; \mathsf{K})$ is compact then S is bounded. Furthermore, if, also, $R \in B(\mathsf{G}; \mathsf{H})$ and $T \in B(\mathsf{K}; \mathsf{L})$ then $TSR \in B(\mathsf{G}; \mathsf{L})$ is compact.

Proof. Suppose for contradiction that S is not bounded: for every $n \geqslant 1$ there exists $x_n \in \mathsf{H}$ such that $\|Sx_n\| > n\|x_n\|$. In particular, $x_n \neq 0$, so we may assume that $\|x_n\| = 1$, replacing x_n by $\|x_n\|^{-1} x_n$ as necessary. But then $(x_n)_{n \geqslant 1}$ is a bounded sequence in H such that $(Sx_n)_{n \geqslant 1}$ has no convergent subsequence, as every subsequence is unbounded. This contradicts the compactness of S, hence S is bounded as claimed.

Let $(x_n)_{n \geqslant 1} \subseteq \mathsf{G}$ be bounded. The operator-norm inequality implies that $(Rx_n)_{n \geqslant 1}$ is bounded, and so, since S is compact, it follows that $(SRx_n)_{n \geqslant 1}$ has a convergent subsequence, say $(SRx_{n_k})_{k \geqslant 1}$. Since T is continuous, it follows that $(TSRx_n)_{n \geqslant 1}$ has the convergent subsequence $(TSRx_{n_k})_{k \geqslant 1}$, so TSR is compact. $\square$

Notation 7.11. The set of compact operators from H to K is denoted $B_0(\mathsf{H}; \mathsf{K})$, or $B_0(\mathsf{H})$ if $\mathsf{H} = \mathsf{K}$.

Remark 7.12. Note that $B_{00}(\mathsf{H}; \mathsf{K}) \subseteq B_0(\mathsf{H}; \mathsf{K})$, by Theorem 1.24, the Heine–Borel theorem: every finite-rank operator is compact.

More generally, if $T \in L(\mathsf{H}; \mathsf{K})$ is such that $\operatorname{im} T$ is finite dimensional, then T is compact and so $T \in B(\mathsf{H}; \mathsf{K})$. In particular, if H or K is finite dimensional then $L(\mathsf{H}; \mathsf{K}) = B(\mathsf{H}; \mathsf{K})$.

Remark 7.13. Note that $I \in B_0(\mathsf{H})$ if and only if $B_0(\mathsf{H}) = B(\mathsf{H})$, by the second part of Proposition 7.10. Furthermore, if H is finite dimensional then $B_{00}(\mathsf{H}) = B_0(\mathsf{H}) = B(\mathsf{H})$, whereas these three spaces are distinct if H is infinite dimensional.

For example, if H is infinite dimensional then, by the Gram–Schmidt process, there exists an *orthonormal sequence* $(e_n)_{n \geqslant 1} \subseteq \mathsf{H}$, so that $\|e_n\| = 1$ for every $n \geqslant 1$ and $\langle e_m, e_n \rangle = 0$

whenever $m \neq n$. In particular, if m and n are distinct then

$$\|e_m - e_n\|^2 = \langle e_m - e_n, e_m - e_n \rangle = \|e_m\|^2 - \langle e_m, e_n \rangle - \langle e_n, e_m \rangle + \|e_n\|^2 = 2,$$

so $(e_n)_{n\geqslant 1}$ is a bounded sequence with no Cauchy subsequence, and so no convergent subsequence. Thus the identity operator $I \in B(\mathsf{H}) \setminus B_0(\mathsf{H})$.

For a compact operator that does not have finite rank, see Example 7.18.

Proposition 7.14. The set $B_0(\mathsf{H};\mathsf{K})$ is a subspace of $B(\mathsf{H};\mathsf{K})$.

Proof. Let S, $T \in B_0(\mathsf{H};\mathsf{K})$ and $\lambda \in \mathbb{C}$, and suppose $(x_n)_{n\geqslant 1} \subseteq \mathsf{H}$ is bounded. As S is compact, so $(Sx_n)_{n\geqslant 1}$ has a convergent subsequence, say $(Sx_{n_k})_{n\geqslant 1}$. Then $(x_{n_k})_{k\geqslant 1}$ is also bounded and T is compact, so $(Tx_{n_k})_{k\geqslant 1}$ has a convergent subsequence, say $(Tx_{n_{k_l}})_{l\geqslant 1}$. As a subsequence of a convergent sequence is itself convergent, it follows that

$$\lim_{l\to\infty} (S + \lambda T)x_{n_{k_l}} = \lim_{l\to\infty} Sx_{n_{k_l}} + \lambda \lim_{l\to\infty} Tx_{n_{k_l}} \qquad \text{exists,}$$

and therefore $S + \lambda T$ is compact. □

Theorem 7.15. Let $T \in B(\mathsf{H};\mathsf{K})$. Then T is compact if and only if T^* is compact.

Proof. Since $(T^*)^* = T$, it suffices to assume that T is compact and prove that T^* is compact also. Let $(y_n)_{n\geqslant 1} \subseteq \mathsf{K}$ be bounded, and note that TT^* is compact, by Proposition 7.10. Thus $(TT^*y_n)_{n\geqslant 1}$ has a convergent subsequence, $(TT^*y_{n_k})_{k\geqslant 1}$, and therefore

$$\begin{aligned}\|T^*y_{n_k} - T^*y_{n_l}\|^2 &= \langle y_{n_k} - y_{n_l}, TT^*(y_{n_k} - y_{n_l}) \rangle \\ &\leqslant \|y_{n_k} - y_{n_l}\| \, \|TT^*(y_{n_k} - y_{n_l})\| \to 0 \qquad \text{as } k, l \to \infty,\end{aligned}$$

since $(y_{n_k})_{k\geqslant 1}$ is bounded. Hence $(T^*y_{n_k})_{k\geqslant 1}$ is Cauchy, so convergent, as required. □

Theorem 7.16. Let $(T_n)_{n\geqslant 1} \subseteq B_0(\mathsf{H};\mathsf{K})$ be such that $T_n \to T$, where $T \in B(\mathsf{H};\mathsf{K})$. Then T is compact, and so $B_0(\mathsf{H};\mathsf{K})$ is a closed subspace of $B(\mathsf{H};\mathsf{K})$.

Proof. The proof involves passing to successive subsequences, as in Proposition 7.14, but infinitely often rather than twice.

Let $(x_n^{(0)})_{n\geqslant 1} \subseteq \mathsf{H}$ be bounded, and successively choose $(x_n^{(j)})_{n\geqslant 1}$ for every $j \geqslant 1$ such that $(x_n^{(j)})_{n\geqslant 1}$ is a subsequence of $(x_n^{(j-1)})_{n\geqslant 1}$ and $(T_j x_n^{(j)})_{n\geqslant 1}$ is convergent. Then let $x'_n := x_n^{(n)}$ for every $n \geqslant 1$, and note that this is a subsequence of $(x_n^{(0)})_{n\geqslant 1}$. We claim that $(Tx'_n)_{n\geqslant 1}$ is Cauchy, so convergent, and hence T is compact.

Let $\varepsilon > 0$. Since $(x_n^{(0)})_{n\geqslant 1}$ is bounded, so is $(x'_n)_{n\geqslant 1}$, and there exists $N \geqslant 1$ such that $\|T - T_N\| \, \|x'_n\| < \varepsilon/2$ for every $n \geqslant 1$. Then

$$\|Tx'_m - Tx'_n\| \leqslant \|T - T_N\| \, \|x'_m\| + \|T_N(x'_m - x'_n)\| + \|T_N - T\| \, \|x'_n\| < \|T_N(x'_m - x'_n)\| + \varepsilon \to \varepsilon$$

as m, $n \to \infty$, since after some point the sequence $(x'_n)_{n\geqslant 1}$ is a subsequence of $(x_n^{(N)})_{n\geqslant 1}$. As ε is arbitrary, it follows that $(Tx'_n)_{n\geqslant 1}$ is Cauchy, as claimed. □

Remark 7.17. A *normed algebra* A is an algebra equipped with a norm which is submultiplicative. If A is unital then it is required that $\|1\| = 1$, and if A is a $*$-algebra then the norm should be invariant under the involution, *i.e.*, $\|a^*\| = \|a\|$ for every $a \in A$.

If the algebra A is complete for the norm, then A is a *Banach algebra*, and if the norm also satisfies the C^* identity, *i.e.*, $\|a^*a\| = \|a\|^2$ for every $a \in A$, then A is a *C^* algebra*.

The key examples of C^* algebras are $B(\mathsf{H})$ and $C(K)$; both of these are unital. The compact operators $B_0(\mathsf{H})$, with H infinite dimensional, is a good example of a non-unital C^* algebra.

Example 7.18. Let $A = (a_{jk})_{j,k=0}^{\infty} \subseteq \mathbb{C}$ be an infinite matrix of complex numbers, and suppose that

$$M_A := \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |a_{jk}|^2 < \infty.$$

Then

$$A : \ell^2 \to \ell^2; \ (Az)_j := \sum_{k=0}^{\infty} a_{jk} z_k \qquad \text{for every } z = (z_0, z_1, z_2, \ldots) \in \ell^2 \text{ and } j \geqslant 0$$

is a compact operator on ℓ^2 with $\|A\| \leqslant M_A^{1/2}$.

To see this, note first that if $z \in \ell^2$ then, by the Cauchy–Schwarz inequality for ℓ^2,

$$\sum_{j=0}^{\infty} \Big(\sum_{k=0}^{\infty} |a_{jk} z_k| \Big)^2 \leqslant \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |a_{jk}|^2 \sum_{k=0}^{\infty} |z_k|^2 = M_A \|z\|^2.$$

Hence $A \in B(\ell^2)$ with norm as claimed.

Now let $e_j := (0, \ldots, 0, 1, 0, \ldots) \in \ell^2$ have 1 in the jth place and 0 elsewhere, so that $(e_j)_{j \geqslant 0}$ is an orthonormal sequence in ℓ^2, and let

$$P_n := \sum_{k=0}^{n} |e_k\rangle\langle e_k|.$$

Then $P_n^2 = P_n = P_n^*$, so P_n is an orthogonal projection onto

$$\operatorname{im} P_n = \{(z_0, \ldots, z_n, 0, \ldots,) \in \ell^2 : z_0, \ldots, z_n \in \mathbb{C}\}.$$

It follows from Corollary 7.6 that $P_n A$ has finite rank, so is compact, and

$$\big((A - P_n A)z\big)_j = \begin{cases} 0 & \text{if } j = 0, \ldots, n, \\ \displaystyle\sum_{k=0}^{\infty} a_{jk} z_k & \text{if } j \geqslant n+1, \end{cases}$$

so $A - P_n A$ is of the same form as A, but with the entries of the first $n+1$ rows of the matrix A set to 0. Hence the previous working gives that

$$\|A - P_n A\|^2 \leqslant M_{A - P_n A}^2 = \sum_{j=n+1}^{\infty} \sum_{k=0}^{\infty} |a_{jk}|^2 \to 0 \qquad \text{as } n \to \infty.$$

Hence $A \in B_0(\ell^2)$, by Theorem 7.16.

7.4 The spectral theorem

As recalled earlier, if $A \in M_n(\mathbb{C})$ is Hermitian, *i.e.*, $A^* = A$, then there exists a unitary matrix U and a diagonal matrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ such that $A = UDU^*$.

Consider the columns of $U = (u_{jk})_{j,k=1}^n$ as column vectors $v_1, \ldots, v_n$, so that $(v_j)_k = u_{kj}$ whenever $1 \leqslant j, k \leqslant n$. The identity $U^*U = I$ is equivalent to the fact that

$$\langle v_j, v_k \rangle_{\mathbb{C}^n} = \sum_{m=1}^n \overline{u_{mj}} u_{mk} = (U^*U)_{jk} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k, \end{cases}$$

i.e., that $\{v_1, \ldots, v_n\}$ is an orthonormal set. Thus $I = \sum_{j=1}^n |v_j\rangle\langle v_j|$ and so, letting A act as a linear transformation on $\mathbb{C}^n$, as in Example 2.2,

$$A = IAI = \sum_{j,k=1}^n |v_j\rangle\langle v_j|A|v_k\rangle\langle v_k| = \sum_{j,k=1}^n \langle v_j, Av_k\rangle |v_j\rangle\langle v_k| = \sum_{j=1}^n \lambda_j |v_j\rangle\langle v_j|,$$

since

$$\langle v_j, Av_k\rangle = \langle U^*v_j, DU^*v_k\rangle = \langle e_j, De_k\rangle = \langle e_j, \lambda_k e_k\rangle,$$

where $e_i \in \mathbb{C}^n$ has 1 in the ith position and 0 elsewhere. In this section, we will show that a similar representation exists for any compact self-adjoint operator.

Lemma 7.19. Let $T \in B_0(\mathsf{H})$ and $\lambda \in \mathbb{C} \setminus \{0\}$. Then $\ker(\lambda I - T)$ is finite dimensional.

Proof. Suppose for contradiction that λ is such that $\ker(\lambda I - T)$ is not finite dimensional, and let $(e_n)_{n \geqslant 1}$ be an orthonormal sequence in $\ker(\lambda I - T)$. Then

$$\|Te_m - Te_n\| = |\lambda|\, \|e_m - e_n\| = \sqrt{2}\, |\lambda| \qquad \text{whenever } m \neq n,$$

so the bounded sequence $(e_n)_{n \geqslant 1}$ is such that $(Te_n)_{n \geqslant 1}$ has no Cauchy, so no convergent, subsequence. This contradiction gives the result. □

Lemma 7.20. Let $T \in B_0(\mathsf{H})$ be self adjoint, where $\mathsf{H} \neq \{0\}$. There exists $x \in \mathsf{H} \setminus \{0\}$ such that $Tx = \lambda x$, where $\lambda \in \mathbb{R}$ is such that $|\lambda| = \|T\|$.

Proof. The result is immediate if $T = 0$, so suppose otherwise. Since T is self adjoint, there exists $\lambda \in \sigma(T) \subseteq \mathbb{R}$ such that $|\lambda| = \|T\|$, by Theorem 6.19.

Now, for every $c > 0$ there exists $x \in \mathsf{H}$ such that $\|(\lambda I - T)x\| < c\|x\|$, for otherwise there exists some $c > 0$ such that

$$(\lambda I - T)^*(\lambda I - T) = (\lambda I - T)(\lambda I - T)^* \geqslant c^2 I$$

and therefore $\lambda I - T$ is invertible, by Theorem 6.9. Taking $c = 1/n$ gives the existence of a sequence $(x_n)_{n \geqslant 1} \subseteq \mathsf{H}$ such that $\|x_n\| = 1$ for every $n \geqslant 1$ and $(\lambda I - T)x_n \to 0$. As T is compact, there exists a subsequence $(x_{n_k})_{k \geqslant 1}$ such that $Tx_{n_k} \to x$ for some $x \in \mathsf{H}$, and

$$\|x\| = \lim_{k \to \infty} \|Tx_{n_k}\| = |\lambda| = \|T\| \neq 0,$$

since

$$x = \lim_{k\to\infty} Tx_{n_k} = \lim_{k\to\infty} Tx_{n_k} + (\lambda I - T)x_{n_k} = \lim_{k\to\infty} \lambda x_{n_k}.$$

Finally,

$$Tx = \lim_{k\to\infty} T(Tx_{n_k}) = \lim_{k\to\infty} \lambda T x_{n_k} = \lambda x. \qquad \square$$

Lemma 7.21. Let $T \in B(\mathsf{H})$ be self adjoint. If $D \subseteq \mathsf{H}$ is *invariant* under T, so that

$$T(D) := \{Tx : x \in D\} \subseteq D,$$

then $D^\perp$ is invariant under T also: $T(D^\perp) \subseteq D^\perp$.

Proof. Let $x \in D^\perp$ and $y \in D$. Then $Ty \in D$ and therefore

$$\langle Tx, y\rangle = \langle x, Ty\rangle = 0 \qquad \Longrightarrow \qquad Tx \in D^\perp. \qquad \square$$

Lemma 7.22. Let $T \in B(\mathsf{H})$ be self adjoint, and suppose x, $y \in \mathsf{H}$ are eigenvectors corresponding to distinct eigenvalues, so that $Tx = \lambda x$ and $Ty = \mu y$, where λ, $\mu \in \mathbb{R}$ are such that $\lambda \neq \mu$. Then $\langle x, y\rangle = 0$.

Proof. Note that

$$\lambda\langle x, y\rangle = \langle \lambda x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \langle x, \mu y\rangle = \mu\langle x, y\rangle \quad \Longrightarrow \quad (\lambda - \mu)\langle x, y\rangle = 0. \quad \square$$

These four lemmas provide the tools required to prove the following spectral theorem.

Theorem 7.23. Let $T \in B_0(\mathsf{H})$ be self adjoint. Then

$$T = \sum_{j=1}^{N} \lambda_j |e_j\rangle\langle e_j|,$$

where $\{\lambda_j : j = 1, \ldots, N\}$ is a set of real numbers, $\{e_j : j = 1, \ldots, N\} \subseteq \mathsf{H}$ is an orthonormal set and $N \geqslant 0$ or $N = \infty$. In the latter case, the sequence $(\lambda_j)_{j\geqslant 1}$ converges to 0 and the series is convergent in $B(\mathsf{H})$.

Proof. The case $N = 0$ corresponds to $T = 0$, so suppose $T \neq 0$. By Lemma 7.20, there exists $\lambda \in \mathbb{R}$ such that $|\lambda| = \|T\|$ and $\mathsf{L} := \ker(\lambda I - T) \neq \{0\}$; let $P = P_{\mathsf{L}}$ be the orthogonal projection onto this eigenspace, which is finite dimensional, by Lemma 7.19. Then

$$T = T(P + P^\perp) = \lambda P + TP^\perp = \lambda P + P^\perp T P^\perp,$$

where the final equality is a consequence of Lemma 7.21.

If $P^\perp T P^\perp = 0$ then we stop. Otherwise, the operator $P^\perp T P^\perp$ is compact, self adjoint and has norm $\|P^\perp T P^\perp\| \leqslant \|T\|$, so applying the previous working gives $\mu \in \mathbb{R}$ such that $|\mu| = \|P^\perp T P^\perp\| > 0$ and an orthogonal projection Q onto $\ker(\mu I - P^\perp T P^\perp)$. Note that

$$x \in \operatorname{im} Q \quad \Longrightarrow \quad \mu x - P^\perp T P^\perp x = 0 \quad \Longrightarrow \quad x \in \operatorname{im} P^\perp T P^\perp \subseteq \operatorname{im} P^\perp,$$

so $PQy = 0$ for every $y \in \mathsf{H}$. It follows that $P + Q$ is an orthogonal projection, by Proposition 2.33, and

$$T = \lambda P + \mu Q + Q^\perp P^\perp T P^\perp Q^\perp = \lambda P + \mu Q + (P+Q)^\perp T (P+Q)^\perp;$$

note that $P^\perp Q^\perp = (I-P)(I-Q) = I - P - Q + PQ = (P+Q)^\perp$ because $PQ = 0$.

Continuing in this way yields distinct non-zero eigenvalues $\mu_1, \ldots, \mu_n$ and finite-rank orthogonal projections $P_1, \ldots, P_n$ such that $|\mu_1| \geqslant \cdots \geqslant |\mu_n|$, the operator $R_n := P_1 + \cdots + P_n$ is an orthogonal projection such that $R_n P_{n+1} = 0$ and

$$T = \sum_{k=1}^{n} \mu_k P_k + R_n^\perp T R_n^\perp.$$

If, at any stage, $R_n^\perp T R_n^\perp = 0$, we stop. Otherwise we have that

$$\Big\| T - \sum_{k=1}^{n} \mu_k P_k \Big\| = \|R_n^\perp T R_n^\perp\| = |\mu_{n+1}| \qquad \text{for every } n \geqslant 1.$$

Suppose for contradiction that $\mu_n \not\to 0$. Then, since $\big(|\mu_n|\big)_{n\geqslant 1}$ is non-increasing, there exists $\varepsilon > 0$ such that $|\mu_n| \geqslant \varepsilon$ for every $n \geqslant 1$. As μ_n is an eigenvalue of T, for every $n \geqslant 1$ there exists $x_n \in \mathsf{H}$ such that $\|x_n\| = 1$ and $Tx_n = \mu_n x_n$. Then $(x_n)_{n\geqslant 1}$ is a bounded sequence such that $(Tx_n)_{n\geqslant 1}$ has no convergent subsequence, since

$$\|Tx_m - Tx_n\|^2 = \|\mu_m x_m - \mu_n x_n\|^2 = \mu_m^2 + \mu_n^2 \geqslant 2\varepsilon^2 \qquad \text{whenever } m \neq n,$$

by Lemma 7.22. This contradicts the fact that T is compact, and so

$$T = \sum_{k=1}^{M} \mu_k P_k,$$

where $M \geqslant 0$ or $M = \infty$; in the latter case the series converges in $B(\mathsf{H})$ and $\mu_n \to 0$.

To get the representation claimed, note first that the ranges of the P_k are eigenspaces for distinct non-zero eigenvalues; in particular, the finite-dimensional ranges of these projections are mutually orthogonal.

Let $\{e_1, \ldots, e_{n_1}\}$ be an orthonormal basis for P_1 and let $\lambda_j = \mu_1$ for $j = 1, \ldots, n_1$, then continue this for successive P_k: for $k = 1, \ldots, M-1$, let $\{e_{n_k+1}, \ldots, e_{n_{k+1}}\}$ be a basis for P_{k+1} and let $\lambda_j = \mu_{k+1}$ for $j = n_k + 1, \ldots, n_{k+1}$. Then $\{e_j : j = 1, \ldots, N\}$ is an orthonormal set and

$$T = \sum_{k=1}^{M} \mu_k P_k = \sum_{j=1}^{N} \lambda_j |e_j\rangle\langle e_j|$$

where $N = n_M$ if $M < \infty$ and $N = \infty$ otherwise. In this last case, to see that convergence holds, note that if $n \geqslant n_1$ is such that $n_m \leqslant n < n_{m+1}$ then

$$\sum_{j=1}^{n} \lambda_j |e_j\rangle\langle e_j| = \sum_{k=1}^{m} \mu_k P_k + \sum_{j=n_m+1}^{n} \mu_{m+1} |e_j\rangle\langle e_j|$$

and this last term has norm $|\mu_{m+1}|$, so tends to 0 as $n \to \infty$. ☐

The following result may be restated as saying that the closure of the set of finite-rank operators is the set of compact operators: $\overline{B_{00}(\mathsf{H})} = B_0(\mathsf{H})$.

Corollary 7.24. An element of $B(\mathsf{H})$ is compact if and only if it is the limit of a sequence of finite-rank operators.

Proof. Given $T \in B_0(\mathsf{H})$, we can apply Theorem 7.23 to the operators $\operatorname{Re} T := (T+T^*)/2$ and $\operatorname{Im} T := (T - T^*)/(2\mathrm{i})$, noting that $T = \operatorname{Re} T + \mathrm{i} \operatorname{Im} T$.

The converse follows from Remark 7.12 and Theorem 7.16. ☐

Bibliography

[1] A. C. R. Belton, *Functional analysis*, lecture notes, University of Oxford, 2006. http://www.maths.lancs.ac.uk/~belton/www/notes/fa_notes.pdf [Cited on page iv.]

[2] B. Bollobás, *Linear analysis*, corrected printing, Cambridge University Press, Cambridge, 1992. [Cited on page iv.]

[3] G. K. Pedersen, *Analysis now*, revised printing, Graduate Texts in Mathematics 118, Springer-Verlag, New York, 1995. [Cited on page iv.]

[4] W. Rudin, *Real and complex analysis*, third edition, McGraw-Hill, Inc., New York, 1986. [Cited on page iv.]

[5] G. F. Simmons, *Introduction to topology and modern analysis*, McGraw-Hill, Inc., New York, 1963. [Cited on page iv.]

[6] N. Young, *An introduction to Hilbert space*, Cambridge University Press, Cambridge, 1998. [Cited on page iv.]

Index

Numbers set in italic type refer to the page where the term is first defined.

Symbols

$*$-algebra . . . *51*
$*$-ideal . . . *51*

A

absolute value . . . *48*
absolutely convergent . . . *22*
algebra . . . *50*
  Banach . . . *53*
  C^* . . . *53*
  normed . . . *53*
antisymmetric . . . *43*

B

ball
  open . . . *4*
Banach algebra . . . *53*
Banach space . . . *8*
basis
  standard . . . *16*
bounded . . . *12*, *21*
bounded operator . . . *12*

C

C^* algebra . . . *53*
C^* identity . . . *16*
Cauchy sequence . . . *7*
characteristic polynomial . . . *31*
closed . . . *4*
closure . . . *5*
co-domain . . . *iv*
co-isometry . . . *19*
compact operator . . . *51*
compact set . . . *6*
complement
  orthogonal . . . *18*
complete . . . *8*
cone . . . *43*
conjugate transpose . . . *16*
continuous . . . *5*
continuous linear functional . . . *10*
convergence
  uniform . . . *7*
convergent . . . *4*
  absolutely . . . *22*
convergent series . . . *22*

D

dense . . . *21*
Dirac dyad . . . *49*
direct sum . . . *8*
  internal . . . *9*
distance . . . *5*
domain . . . *iv*

F

finite rank . . . *50*
functional
  continuous linear . . . *10*

G

group of units . . . *26*

H

Hilbert space . . . *8*
holomorphic
  strongly . . . *31*
homeomorphism . . . *30*

I

ideal . . . 51
identity
  C^* . . . 16
image . . . 17
inner product . . . 1
internal direct sum . . . 9
invariant . . . 55
inverse . . . 25
invertible . . . 25
involution . . . 51
isometric isomorphism . . . 26
isometry . . . 19
isomorphism
  isometric . . . 26

K

kernel . . . 17

L

left shift . . . 28
linear functional
  continuous . . . 10
linear transformation . . . 11

M

multiplicative . . . 35

N

norm . . . 2
  operator . . . 12
  uniform . . . 7
normed algebra . . . 53
numerical range . . . 32

O

open . . . 4
open ball . . . 4
operator
  bounded . . . 12
  compact . . . 51
operator norm . . . 12
order
  partial . . . 43
orthogonal complement . . . 18
orthogonal projection . . . 19
orthonormal . . . 11
orthonormal sequence . . . 51

P

partial order . . . 43
polynomial
  characteristic . . . 31
positive . . . 43
  uniformly . . . 44
pre-image . . . 5
projection
  orthogonal . . . 19

R

radius
  spectral . . . 36
range . . . 17
  numerical . . . 32
rank-nullity theorem . . . 27
reflexive . . . 43
resolvent . . . 29
restriction . . . iv
right shift . . . 27

S

self adjoint . . . 38
sequence
  Cauchy . . . 7
  orthonormal . . . 51
series
  convergent . . . 22
set
  compact . . . 6
shift
  left . . . 28
  right . . . 27
space
  Banach . . . 8
  Hilbert . . . 8
spectral radius . . . 36
spectrum . . . 27
square root . . . 46
standard basis . . . 16
strongly holomorphic . . . 31
submultiplicative . . . 15
sum . . . 22
  direct . . . 8
  internal direct . . . 9

T

theorem
  rank-nullity . . . 27

transformation
  linear 11
transitive 43

U

uniform convergence 7
uniform norm 7
uniformly positive 44
unit 51
unital 51
unitary 19
units
  group of 26