B4.2 Functional Analysis II

Melanie Rupflin
HT 2024

April 17, 2024

This course is a continuation of B4.1 Functional Analysis I.

These lecture notes build upon and expand Hilary Priestley's, Gregory Seregin's and Luc Nguyen's lecture notes who taught the course in previous years as well as lecture notes of the lecturer on Functional Analysis 1 from previous years.

Warning. *There has been a change of syllabus between the academic years 2022/23 and 2023/24, which means that some of the material that was previously covered in B4.1 has now been moved to B4.2 (namely aspects of spectral theory), or removed (like the proof of Hahn-Banach) while some material from B4.2 has moved into B4.1 (namely the Riesz-representation theorem and the projection theorem and some aspects of adjoint operators). Additionally, there is a new part on compact operators that was not covered in some of the previous courses. Please keep this in mind when using resources for exam preparation.*

The following literature was also used (either for this set of notes, or for my predecessors'):

B.P. Rynne and M.A. Youngson, *Linear Functional Analysis*, Springer SUMS, 2nd edition, 2008.

E. Kreyszig, *Introductory Functional Analysis with Applications*, Wiley, revised edition, 1989.

N. Young, *An Introduction to Hilbert Space*, Cambridge University Press, 1988.

E.M. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration & Hilbert Spaces*, Princeton Lectures in Analysis III, 2005.

M. Reed and B. Simon, *Methods of Modern Mathematical Physics Vol. I. Functional Analysis*, Academic Press, 1980.

H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Universitext, Springer, 2011.

P.D. Lax, *Functional Analysis*, Wiley, 2002.

R.L. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis*, Dekker, 1977.

Contents

Chapter 1

Banach space, Hilbert spaces and bounded linear operators

1.1 Recap of some relevant material of B4.1 and notation

Normed and inner product spaces:

Throughout this course we will consider normed spaces $(X, \|\cdot\|_X)$ over either $\mathbb{R}$ or $\mathbb{C}$ (both allowed unless specified otherwise) and we will often simply write X for $(X, \|\cdot\|_X)$ if we work with a fixed norm.

We recall that X is a Banach space if it is a complete normed space, that X is an inner product space if $\|\cdot\|$ is induced by an inner product via $\|x\| = (\langle x, x\rangle)^{\frac{1}{2}}$ and that X is a Hilbert space if it is a complete inner product space.

Orthogonal complement and projection theorem:

We recall that if X is an inner product space then x and y are said to be orthogonal if $\langle x, y\rangle = 0$ and given a set $A \subset X$ we can define

$$A^\perp := \{x \in X : \langle a, x\rangle = 0 \text{ for all } a \in A\}$$

which is always a closed subspace of X.

We recall that if Y itself is a closed subspace of X and X is a Hilbert space then the projection theorem ensures that $X = Y \oplus Y^\perp$ and that the corresponding orthogonal projections $P^Y : X \to Y$ and $P^{Y^\perp} : X \to Y^\perp$ are bounded linear operators with $\mathrm{I} = P^Y + P^{Y^\perp}$ and

$$\|P^Y(x)\|^2 + \|P^{Y^\perp}(x)\|^2 = \|x\|^2.$$

We also recall that the closed linear span of a set S in a normed space X is the smallest closed subspace of X containing S, i.e. the intersection of all such subspaces and that this set is always given by the closure of the linear span $\mathrm{Span}\, S := \{\sum_{j=1}^n a_j s_j : s_j \in S, a_j \in \mathbb{F}, n \in \mathbb{N}\}$.

A useful consequence of the projection theorem is that the closed linear span of any subset S of a Hilbert space is $S^{\perp\perp}$.

We recall that the Cauchy-Schwarz inequality holds in inner product spaces, i.e. that $|\langle x, y\rangle| \le \|x\|\|y\|$ with equality if and only if x and y are linearly dependent. In particular,

$$\|x\| = \sup_{\|y\|=1} |\langle x, y\rangle|. \tag{1.1}$$

Bounded linear operators:

We denote the space of bounded linear operators between two normed spaces by

$$\mathscr{B}(X,Y) := \{T : X \to Y \text{ linear s.t. } \exists M \text{ with } \|Tx\|_Y \leq M\|x\|_X \text{ for all } x \in X\}$$

and will always equip this space with the operator norm

$$\|T\| := \inf\{ \text{ such } M\} = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|}.$$

Here and in the following we often drop the subscripts of the norms if it is clear from the context which norms need to be used and will always assume that X is not trivial, i.e. $X \neq \{0\}$.

In the special cases of $Y = X$ (equipped with the same norm) we write $\mathscr{B}(X) = \mathscr{B}(X,X)$ and for $Y = \mathbb{F}$, for $\mathbb{F} = \mathbb{R}$ respectively $\mathbb{C}$ the field over which X is a vector space, we get the dual space $X^* = \mathscr{B}(X,\mathbb{F})$.

We recall that if Y is Banach then so is $\mathscr{B}(X,Y)$, in particular X^* is always Banach.

We also recall that $T \in \mathscr{B}(X,Y)$ is called invertible if there exists $S \in B(Y,X)$ so that

$$S \circ T = \mathrm{I}_X \text{ and } T \circ S = \mathrm{I}_Y$$

and that this is equivalent to

$$T \text{ is bijective and } \exists \delta > 0 \text{ so that } \|Tx\| \geq \delta\|x\| \text{ for all } x \in X. \tag{1.2}$$

Useful Corollaries of Hahn-Banach:

We recall two useful consequences of the theorem of Hahn-Banach to which we will refer to later on in the course, namely that

$$\forall x \in X\, \exists f \in X^* \text{ so that } \|f\| = 1 \text{ and } f(x) = \|x\| \tag{1.3}$$

and that for any closed proper subspace Y of X and any $x \in X \setminus Y$ there exists $f \in X^*$ so that

$$f|_Y = 0 \text{ but } f(x) \neq 0. \tag{1.4}$$

Riesz-Representation theorem and adjoints:

We also recall the following fundamental result for Hilbert spaces:

Theorem 1.1 (Riesz representation theorem)**.** *Let X be a Hilbert space. Then for any $T \in X^*$ there exists a unique element $x_T \in X$ so that*

$$T(y) = \langle y, x_T\rangle \text{ for all } y \in X.$$

Furthermore, this x_T satisfies $\|x_T\| = \|T\|$.

If X is a real vector space then the map $T \mapsto x_T$ is linear, and hence provides an isometric isomorphism between X and X^*. If $\mathbb{F} = \mathbb{C}$ then $T \mapsto x_T$ is still additive, but not linear since $x_{\lambda T} = \bar{\lambda} x_T$.

If X and Y are two Hilbert spaces then we can associate to each $T : X \to Y$ its adjoint operator $T^* : Y \to X$ where $T^* y$ is defined as the unique element x_ℓ (given by Riesz) which represents the bounded linear functional $\ell : x \mapsto \langle Tx, y \rangle$. This adjoint operator $T^* : Y \to X$ is hence the unique operator for which

$$\langle Tx, y \rangle_Y = \langle x, T^* y \rangle_X \text{ for all } x \in X, y \in Y.$$

You have seen in B4.1 that the adjoint operator has some useful properties, including the fact that

- $T^* \in \mathscr{B}(Y, X)$ with $\|T^*\| = \|T\|$ and $T^{**} = T$
- $(ST)^* = T^* S^*$
- T is invertible if and only if T^* is invertible
- the kernel and image of an operator and its adjoint are related by

$$\ker(T) = (\mathrm{Im} T^*)^\perp \text{ and } (\ker T)^\perp = \overline{\mathrm{Im}(T^*)} \tag{1.5}$$

Compactness of sets:

We recall that for metric spaces, and hence in particular for normed spaces, compactness and sequential compactness are equivalent. We also recall that any compact set of a metric space is always closed and bounded. On the other hand, you have seen in B4.1

Theorem 1.2 (Heine-Borel). *Let X be a normed space.*

(i) If $\dim(X) < \infty$ then every bounded closed subset of X is compact.

(ii) If the closed unit ball $\bar{B}^X$ is compact then $\dim(X) < \infty$.

Here and in the following we use the following notation for balls: We write $B_r(x) := \{y : \|x - y\| < r\}$ for the open ball with radius r and centre x, $\bar{B}_r(x) := \{y : \|x - y\| \le r\}$ for the corresponding closed ball which of course agrees with the closure of $B_r(x)$ and will write for short B and $\bar{B}$ for the open/closed unit ball around $x = 0$.

At times, we will be working with balls in different spaces at the same time, and will hence include a superscript to clarify in which space we consider these balls, i.e. write $B_r^X(x)$, B^X,... for the corresponding balls in the space X.

1.2 Normal, Unitary and Selfadjoint operators

In many applications one works with operators that have additional properties and in this section we will discuss the basic properties of special types of operators between inner product spaces, namely of

Definition 1.3. *Let X and Y be inner product spaces.*

We say that an operator $T \in \mathscr{B}(X,Y)$ is ***isometric*** *if $\|Tx\| = \|x\|$ for every $x \in X$ and we say that $T \in \mathscr{B}(X,Y)$ is* ***unitary*** *if it is both isometric and surjective.*

On the other hand, we say that an operator $S \in \mathscr{B}(X)$ from a Hilbert space to itself is ***normal*** *if $S^*S = SS^*$ and S is called* ***selfadjoint*** *if $S^* = S$.*

Every selfadjoint operator is hence trivially normal while the converse is not true. We also observe that for any $S \in \mathscr{B}(X)$ the operator SS^* is selfadjoint as $(ST)^* = T^*S^*$ and $S^{**} = S$ so $(SS^*)^* = (S^*)^*S^* = SS^*$.

We have the following characterization of isometric and unitary operators on Hilbert-spaces.

Proposition 1.4. *Let $T, U : X \to Y$ be bounded linear operators between Hilbert spaces.*

(i) The following are equivalent:

  (a) T is isometric.

  (b) $\langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in X$.

  (c) $T^*T = I_X$.

(ii) The following are equivalent:

  (a) U is unitary.

  (b) $U^*U = I_X$ and $UU^* = I_Y$.

  (c) Both U and U^* are isometric.

Proof. Exercise. □

We hence conclude that an operator $S \in \mathscr{B}(X)$ is unitary if and only if it is isometric and normal.

Examples. *(i) The right-shift operator on ℓ^2 is isometric but not unitary. The left-shift operator on ℓ^2 is not isometric.*

(ii) A multiplication operator $M_h : f \mapsto fh$ is unitary on $L^2(\mathbb{R})$ if and only if $|h| = 1$ a.e.

(iii) If g is a non-negative and measurable function on $\mathbb{R}$, then the map $f \mapsto g^{1/2}f$ is isometric from $L^2(\mathbb{R}, g\,dt)$ to $L^2(\mathbb{R})$. It is unitary if and only if $g > 0$ a.e.

We furthermore have that

Lemma 1.5. *An operator $T \in \mathscr{B}(X)$ on a Hilbert space X is normal if and only if $\|Tx\| = \|T^*x\|$ for every $x \in X$.*

Proof of Lemma 1.5. If T is normal then we have for every x

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle = \langle x, TT^*x\rangle = \langle T^*x, T^*x\rangle = \|T^*x\|^2.$$

Conversely if $\langle Tx, Tx\rangle = \langle T^*x, T^*x\rangle$ for all x then we get by polarisation that $\langle Tx, Ty\rangle = \langle T^*x, T^*y\rangle$ for all $x, y \in X$. Hence for every $x \in X$ we have

$$\langle T^*Tx, y\rangle = \langle Tx, Ty\rangle = \langle T^*x, T^*y\rangle = \langle T^{**}T^*x, y\rangle = \langle TT^*x, y\rangle \text{ for all } y \in X$$

which implies that $T^*Tx = TT^*x$. □

We will furthermore use

Lemma 1.6. *Let X be a Hilbert space.*

(i) If $T \in \mathscr{B}(X)$, then

$$\|T\|_{\mathscr{B}(X)} = \sup\{|\langle Tx, y\rangle| : \|x\| = \|y\| = 1\}.$$

(ii) If $T \in \mathscr{B}(X)$ and T is self-adjoint, then

$$\|T\|_{\mathscr{B}(X)} = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\}.$$

Proof of Lemma 1.6. The first assertion follows from the definition of the operator norm and (1.1).

To prove (ii) we set $K = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\}$ and note that by Cauchy-Schwarz we always have $K \leq \|T\|$. We hence only need to show that if T is selfadjoint then $K \geq \|T\| - \varepsilon$ for all $\varepsilon > 0$ and hence also $K \geq \|T\|$.

Given $\varepsilon > 0$ we can use (i) to find x, y with $\|x\| = \|y\| = 1$ so that $|\langle Tx, y\rangle| > \|T\| - \varepsilon$. Replacing y by ay for some scalar a with $|a| = 1$, we may assume that $|\langle Tx, y\rangle| = \langle Tx, y\rangle$. This implies that

$$\begin{aligned}\|T\| - \varepsilon \leq \langle Tx, y\rangle &= \operatorname{Re}\langle Tx, y\rangle = \tfrac{1}{4}\big[\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle\big] \\ &\leq \tfrac{1}{4}K(\|x+y\|^2 + \|x-y\|^2) = \tfrac{1}{4}K(2\|x\|^2 + 2\|y\|^2) = K,\end{aligned}$$

as desired, where the penultimate step uses the parallelogram identity. □

As A^*A is self-adjoint for any $A \in \mathscr{B}(X)$, and $\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle = \|Ax\|^2$ for every x we deduce following useful result.

Proposition 1.7. *Let X be a Hilbert space and $A \in \mathscr{B}(X)$. Then*

$$\|A^*A\|_{\mathscr{B}(X)} = \|A\|^2_{\mathscr{B}(X)}.$$

In particular, if A is self-adjoint, then $\|A^2\|_{\mathscr{B}(X)} = \|A\|^2_{\mathscr{B}(X)}$.

1.3 Orthonormal sets and bases of Hilbert spaces

In our later discussion of spectral properties of operators on Hilbert spaces and of Fourier series we will use that for Hilbert spaces there is a straightforward extension of the concept of orthonormal bases that you have seen in prelims Linear Algebra and that a Hilbert space always has such an orthonormal basis, see below. Namely we say

Definition 1.8. *A subset S of a Hilbert space X is called an orthonormal set if $\|x\| = 1$ for all $x \in S$ and $\langle x, y\rangle = 0$ for all $x, y \in S$ with $x \neq y$.*

S is called an orthonormal basis (or on-basis for short) for X if S is an orthonormal set and its closed linear span is X.[1]

To analyse orthonormal sets we will often use

[1] In the literature an orthonormal basis is sometimes also called a complete orthonormal set

Theorem 1.9 (Pythagorean theorem)**.** *Let X be a Hilbert space and let $S = \{x_1, x_2, \ldots, x_m\}$ be a finite orthonormal set in X. Then*

$$\|x\|^2 = \sum_{n=1}^{m} |\langle x, x_n\rangle|^2 + \Big\|x - \sum_{n=1}^{m} \langle x, x_n\rangle x_n\Big\|^2 \text{ for every } x \in X.$$

The proof of this is a direct computation and is omitted. An simple consequence is:

Lemma 1.10 (Bessel's inequality)**.** *Let X be a Hilbert space and let $S = \{x_1, x_2, \ldots\}$ be an orthonormal sequence in X. Then*

$$\sum_{n=1}^{\infty} |\langle x, x_n\rangle|^2 \leq \|x\|^2 \text{ for every } x \in X.$$

We furthermore have

Theorem 1.11. *Let X be a Hilbert space and let $S = \{x_1, x_2, \ldots\}$ be an orthonormal sequence in X. Then the closed linear span of S consists of all vectors of the form*

$$x = \sum_{n=1}^{\infty} a_n\, x_n \tag{1.6}$$

for sequences of scalar $(a_1, a_2, \ldots) \in \ell^2(\mathbb{F})$.

Here the sum in (1.6) *converges in the sense of the Hilbert space norm, the element $x \in X$ and the sequence $(a_j) \in \ell^2$ are related by*

$$a_n = \langle x, x_n\rangle. \tag{1.7}$$

and Parseval's Identity holds, i.e.

$$\|x\|^2 = \sum_{n=1}^{\infty} |a_n|^2. \tag{1.8}$$

Proof of Theorem 1.11. We first show that the series in (1.6) converges in X for any $(a_k) \in \ell^2$. To see this we note that the partial sums $S_k := \sum_{n\leq k} a_n x_n$ form a Cauchy sequence since the orthogonality of the x_i means that for $m \geq k \geq N$

$$\|S_k - S_m\|^2 = \sum_{k+1\leq n\leq m} |a_n|^2 \|x_n\|^2 \leq \sum_{n\geq N+1} |a_n|^2 \to 0$$

as $N \to \infty$. As X is Banach the series hence converges.

As the closed linear span Y of S is by definition closed and as it contains span(S) and hence each of the partial sums S_k, we additionally know that $x := \sum_{n=1}^{\infty} a_n x_n = \lim_{k\to\infty} S_k$ is in Y. Also as S is on, we know that $a_n = \langle S_k, x_n\rangle$ for every $k \geq n$ and thus obtain the claimed relation (1.7) in the limit $k \to \infty$. Similarly, as $\|x - S_k\| \to 0$ we obtain Parseval's identity from the Pythagorean theorem.

Finally, we need to show that element x of the closed linear span Y can indeed be written in the form (1.6). To see this we set $a_n := \langle x, x_n\rangle$ use that Bessel's inequality implies that $(a_n) \in \ell^2$. By the previous part of the proof the series $\sum_{n=1}^{\infty} a_n\, x_n$ thus converges and $\tilde{x} = \sum_{n=1}^{\infty} a_n\, x_n$ is an element of Y. We now observe that $x - \tilde{x}$ is orthogonal to all x_n and deduce that $x - \tilde{x} \in (\text{span}S)^\perp = Y^\perp$ since the closed linear span agrees with the closure of the span and since we always have $A^\perp = (\bar{A})^\perp$. But this means that $x - \tilde{x} \in Y \cap Y^\perp = \{0\}$, so $x = \tilde{x}$ has indeed the desired form. $\square$

We can use in particular

Corollary 1.12. *If $\{e_1, e_2, \ldots, ..\}$ is a countable on-basis of a Hilbert space X then*

$$x = \sum \langle x, e_j \rangle e_j \text{ for every } x \in X.$$

We note that X having a countable on-basis implies that X is separable (and is indeed equivalent to this for Hilbert spaces) and that we more generally have

Theorem 1.13. *Every Hilbert space contains an orthonormal basis.*

Proof of Theorem 1.13. We will only give a proof in the case when X is separable, i.e. contains a countable dense subset S. The proof in the more general case draws on more sophisticated arguments such as Zorn's lemma.

We label the elements of S as $y_1, y_2, \ldots$. Applying a slight modification of the Gram-Schmidt process[2], where we simply drop elements whenever they are linearly dependent to the previous elements and continue the process for all n, we obtain an orthonormal set $B = \{x_1, x_2, \ldots\}$ such that, for every n, the span of $\{x_1, \ldots, x_n\}$ contains $y_1, \ldots, y_n$. As $\bar{S} = X$, this implies that $X = \overline{\operatorname{span} B}$. $\square$

There are some well known orthonormal bases of Hilbert spaces such as

(a) The trigonometric functions $\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \sin nx, \frac{1}{\sqrt{\pi}} \cos nx, n = 1, 2, \ldots\}$ and $\{\frac{1}{\sqrt{2\pi}} e^{inx}, n \in \mathbb{Z}\}$ in $L^2([-\pi, \pi], \mathbb{F})$ for $\mathbb{F} = \mathbb{R}$ respectively $\mathbb{F} = \mathbb{C}$.

(b) The Legendre polynomials $p_n(t)$, indexed by their degrees, in $L^2(-1, 1)$.

(c) The Laguerre polynomials $L_n(t)$ in $L^2((0, \infty); e^{-t} dt)$.

(d) The Hermite polynomials $H_n(t)$ in $L^2(\mathbb{R}; e^{-t^2}\, dt)$.

In the next chapter we will prove that the trigonometric system is indeed an on-basis of $L^2([-\pi, \pi])$ and will discuss closely related properties of Fourier series, which are nothing else than the expansion of elements of $L^2(-\pi, \pi)$ with respect to this on-basis as described in Corollary 1.12 above.

1.4 Compact operators

In the last part of this introductory chapter we finally introduce another important class of bounded linear operators, namely compact operators between normed spaces.

Definition 1.14. *Let X, Y be normed spaces, $T \in \mathscr{B}(X, Y)$. Then T is called **compact** if the image of the closed unit ball is precompact, i.e. if $\overline{T(\bar{B}^X)}$ is compact.*

On problem sheet 1 you will establish the following basic properties of compact operators:

[2]The Gram-Schmidt process is usually applied to a set of finitely many linearly independent vectors yielding an orthogonal basis of the same cardinality. In our setting, we will lose the latter property as the vectors y_i's are not necessarily linearly independent.

Lemma 1.15. *Let $T \in \mathscr{B}(X,Y)$ and $S \in \mathscr{B}(Y,Z)$ for X,Y,Z normed spaces. Then*

(i) T is compact if and only if every bounded sequence (x_n) in X has a subsequence x_{n_j} which is so that Tx_{n_j} converges (in Y).

(ii) If $\dim(TX) < \infty$ then T is compact.

(iii) If either S or T is compact then also $S \circ T$ is compact.

(iv) If Y is Banach and $(T_n) \subset \mathscr{B}(X,Y)$ is a sequence of compact operators which converges, $T_n \to T$, in the sense of $\mathscr{B}(X,Y)$ then also T is compact.

Example. *$T : \ell^1 \to \ell^1$ defined by $T(x_1, x_2, ...) := (x_1, \frac{x_2}{2}, \frac{x_3}{3}, ...)$ is a compact operator since $T_j \to T$ for the operators T_j defined by $(T_j(x))_k = \frac{x_k}{k}$ if $k \leq j$ and $(T_j(x))_k = 0$ for $k > j$ which have finite dimensional image.*

Typical examples of compact operators between function spaces are integral operators and to show that such operators are compact we need characterisations of (pre)compact subsets of the corresponding function spaces. The most famous such result is

Theorem 1.16. *(Theorem of Arzela-Ascoli) Let $\Omega \subset \mathbb{R}^n$ be compact, let $X = C(\Omega)$ equipped with the supremum norm and let $\mathcal{F} \subset X$ be any subset of X. Then $\overline{\mathcal{F}}$ is compact if and only if the functions in $\mathcal{F}$ are uniformly bounded and equicontinuous, i.e. if and only if the following two conditions hold*

(i) There exists M so that $\|f\|_{sup} \leq M$ for every $f \in \mathcal{F}$

(ii) For every $\varepsilon > 0$ there exists $\delta > 0$ so that for all $f \in \mathcal{F}$ and all $x, y \in \Omega$ with $|x - y| \leq \delta$ we have $|f(x) - f(y)| \leq \varepsilon$

We will use this result without proof.[3]

We also remark that an analogue result, called the Theorem of Riesz-Kolmogorov, holds for L^p spaces and in discussed in C4.3 Functional Analytic methods for PDEs.

Example. *Integration $I : f \mapsto I(f)$, $I(f)(x) := \int_a^x f(t)dt$ is a compact operator from $X = C([a,b])$ to itself.*

To see this we note that $I(f)$ is continuous if f is continuous and that for every $f \in X$ with $\|f\|_{sup} \leq 1$ the resulting $I(f)$ is so that $\|I(f)\| \leq (b-a)$ and $|I(f)(x) - I(f)(y)| \leq |x - y|$. Hence $I \in \mathscr{B}(X)$ with $\|I\| \leq (b-a)$ (indeed $=$), and $\mathcal{F} := I(\bar{B}^X)$ is uniformly bounded (with $M := (b-a)$) and equicontinuous (we can always choose $\delta = \varepsilon$) so by Arzela Ascoli precompact.

Compact operators have a number of important properties, in particular

[3] The proof of the Theorem of Arzela-Ascoli is non-examinable, but also not difficult. We can e.g. choose a dense countable subset S of Ω, then use a standard diagonal sequence argument to pass to a subsequence of a given sequence (f_n) for which $f_{n_j}(s)$ converges for every $s \in S$ (this part uses uniform boundedness and Bolzano Weierstrass) and finally to use the equicontinuity and density of S to show that this subsequence is indeed a Cauchy sequence in X.

Proposition 1.17. *Let X be a Banach space and let $T \in \mathscr{B}(X)$ be a compact operator. Then*

(i) $\dim(\ker(I-T)) < \infty$

(ii) $(I-T)(X)$ *is closed*

(iii) If X is a Hilbert space and T is additionally selfadjoint then

$$\ker(I-T)^{\perp} = (I-T)(X).$$

We note that the last statement in particular implies that

$$X = (\mathrm{I}-T)(X) \oplus V$$

for a finite dimensional vectorspace V (here $V = \ker(\mathrm{I}-T)$) for which

$$\dim(\ker(\mathrm{I}-T)) = \dim(V)$$

This statement holds true more generally also for compact operators $T \in \mathscr{B}(X)$ on Banach spaces and you can see this as a generalisation of the rank-nullity theorem from prelims Linear Algebra to infinite dimensions as it tells you that the codimension of the image of $I-T$ (which for X finite dimensional is $\dim(X) - \mathrm{rk}(\mathrm{I}-T)$) is the same as the nullity of this operator.

For the proof of (ii) we will use the following lemma, which we will also use later on to analyse the spectrum of more general bounded linear operators on Banach spaces.

Lemma 1.18. *Let X be a Banach space and suppose that $S \in \mathscr{B}(X)$ is so that*

$$\exists \delta > 0 \text{ so that for all } x \in X \text{ we have } \|Sx\| \geq \delta\|x\|. \tag{1.9}$$

Then S is injective and $SX \subset X$ is closed. In particular, if SX is additionally dense in X then S is invertible.

Warning. *This result is wrong if X is not assumed to be complete and we also remark that the image of general bounded linear operators from Banach spaces is not closed. As an example consider the inclusion map $i : (C[0,1], \|\cdot\|_{sup}) \to (L^1[0,1], \|\cdot\|_{L^1})$ which is a bounded linear operator whose image is the subspace of L^1 given by all continuous functions which cannot be closed in L^1 as it is a dense proper subspace of L^1.*

Proof of Lemma 1.18. The only statement whose proof is not trivial is that the image SX is closed and we can prove this as follows: Given any sequence y_n in SX which converges $y_n \to y$ to some $y \in X$, we let $x_n \in X$ be so that $Sx_n = y_n$. We then note that as (y_n) is a Cauchy-sequence, the assumption (1.9) implies that

$$\|x_n - x_m\| \leq \delta^{-1}\|S(x_n - x_m)\| = \delta^{-1}\|y_n - y_m\| \underset{n,m\to\infty}{\longrightarrow} 0,$$

i.e. that also (x_n) is Cauchy and thus, as X is complete, that $x_n \to x$ for some $x \in X$. As S is continuous we thus get that $y = \lim y_n = \lim Sx_n = Sx \in SX$. $\square$

Proof of Proposition 1.17. We write for short $Y := \ker(\mathrm{I} - T)$.

To prove (i) we note that $y = Ty$ for $y \in Y$ so we can view the unit ball $\bar{B}^Y = \{y \in Y : \|y\| \leq 1\}$ in Y as a subset of the image of the unit ball $B_X := \{x \in X : \|x\| \leq 1\}$, i.e. use that $\bar{B}^Y = T(\bar{B}^Y) \subset T(\bar{B}^X) \subset \overline{T(\bar{B}^X)}$. As Y is closed (since it is a kernel) we hence know that $\bar{B}^Y = \bar{B}^X \cap Y$ is a closed subset of the compact set $\overline{T(\bar{B}^X)}$ and hence itself compact. Having thus shown that the closed unit ball in the normed space Y is compact we thus deduce that Y is finite dimensional by the theorem of Heine Borel.

To prove (ii) we recall that a consequence of the Theorem of Hahn-Banach is that every finite dimensional subspace of a Banach space has a closed complement. As Y is finite dimensional there is hence a closed subspace V of X so that $X = Y \oplus V$. As V is a closed subspace of a Banach space it is also Banach and we can consider the map $S := (\mathrm{I} - T)|_V : V \to X$ whose image is $S(V) = (\mathrm{I} - T)(V) = (\mathrm{I} - T)(X)$ as $(\mathrm{I} - T)(Y) = \{0\}$. It hence suffices to show that $S(V)$ is closed and this follows from Lemma 1.18 provided we show that S satisfies (1.9), which we will do by contradiction.

So suppose that there is no $\delta > 0$ so that $\|Sv\| \geq \delta\|v\|$ for all $v \in V$. Then we can choose $v_n \in V$ so that $\|Sv_n\| < \frac{1}{n}\|v_n\|$ and replacing v_n by $\frac{v_n}{\|v_n\|}$ we can assume that $\|v_n\| = 1$ and hence $\|Tv_n - v_n\| = \|Sv_n\| \to 0$. As T is compact we can pass to a subsequence so that Tv_{n_j} converges, say $Tv_{n_j} \to z$. As $T(v_{n_j}) - v_{n_j} \to 0$ this also implies that $v_{n_j} \to z$. As V is closed we must hence have $z \in V$ and as $\|v_{n_j}\| = 1$ we must have $\|z\| = 1$ so $z \neq 0$. However as T is continuous and hence $z = \lim Tv_{n_j} = T(\lim v_{n_j}) = Tz$ we also have $z \in Y$ so $z \in Y \cap V = \{0\}$ which gives a contradiction.

The final claim of the proposition immediately follows since T is assumed to be selfadjoint and since the image of $\mathrm{I} - T$ is closed and hence, by (1.5),

$$(\ker(\mathrm{I} - T))^{\perp} = \overline{(\mathrm{I} - T^*)(X)} = \overline{(\mathrm{I} - T)(X)} = (\mathrm{I} - T)(X).$$

□

Chapter 2

Introduction to convergence of Fourier series

Recall that the Fourier series of a function $f \in L^1(-\pi,\pi) = L^1((-\pi,\pi),\mathbb{C})$ is given by

$$\mathscr{F}(f) = \sum_{n=-\infty}^{\infty} a_n e^{inx}, \qquad a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\, e^{-inx}\, dx.$$

Setting $e_n = \frac{1}{\sqrt{2\pi}} e^{inx}$ we hence have

$$\mathscr{F}(f) = \sum_{n=-\infty}^{\infty} \langle f, e_n\rangle e_n$$

and as the e_n are orthonormal in the complex Hilbert space $L^2(-\pi,\pi)$ know that the infinite sum converges in L^2. We will see soon that $\{e_n\}_{n\in\mathbb{Z}}$ is in fact an orthonormal basis of $L^2(-\pi,\pi)$ and so $f = \mathscr{F}(f)$ as L^2 functions.

The question then arises whether the Fourier series of f converges to f in any better sense. This is a difficult question and to have a satisfactory answer requires knowledge which goes beyond what this course can cover. We are content instead with some brief discussion on the subject.

In the following we can always use that we can extend any function $f \in L^1(-\pi,\pi)$ to a 2π periodic function on $\mathbb{R}$ and that if f is continuous and $f(-\pi) = f(\pi)$ then this extension is also continuous.

We note

Proposition 2.1 (Termwise differentiation of Fourier series)**.** *Suppose that $f \in L^1(-\pi,\pi)$ is so that $a_0(f) = 0$ and let*

$$F(x) = \int_0^x f(t)\, dt.$$

Then $F \in C^0([-\pi,\pi])$ is so that $F(\pi) = F(-\pi)$ so can be extended to a continuous 2π-periodic function on $\mathbb{R}$ and the Fourier coefficients of F and f are related by

$$a_n(f) = in a_n(F) \text{ for } n \neq 0,$$

i.e. if $\mathscr{F}(F) = \sum c_n e^{inx}$, then $\mathscr{F}(f) = \sum in\, c_n\, e^{inx}$.

Note that if $a_0(f) \neq 0$ then we can apply this to $g(x) = f(x) - a_0(f)$, which has $a_0(g) = 0$ since $a_0(1) = 1$, and the corresponding $G(x) = F(x) - a_0(f)x$ to see that

$$\mathscr{F}(F(x) - a_0(f)x) = c_0 + \sum_{n\neq 0} \frac{a_n(f)}{in} e^{inx}$$

for some $c_0 \in \mathbb{C}$.

Proof of Proposition 2.1. As $a_0(f) = 0$ we have

$$F(\pi) - F(-\pi) = \int_{-\pi}^{\pi} f(x)dx = 2\pi a_0(f) = 0$$

and F is continuous by standard properties of integration. To get the claimed relation on the coefficients with $n \neq 0$, we integrate by parts

$$a_n(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-inx}\, dx = \frac{in}{2\pi} \int_{-\pi}^{\pi} F(x)\, e^{-inx}\, dx = in a_n(F)$$

where this integration by parts that involves an L^1 function f can be justified using a density argument and where we use that $F(\pi) = F(-\pi)$. This concludes the proof. $\square$

Theorem 2.2 (Completeness of the trigonometric system)**.** *For every $f \in L^2(-\pi,\pi)$ the partial Fourier sum $S_N f(x) := \sum_{|n|\leq N} a_n(f) e^{inx}$ converges*

$$S_N f \to f \text{ in } L^2(-\pi,\pi).$$

In other words, $\left\{\frac{1}{\sqrt{2\pi}} e^{inx}\right\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2(-\pi,\pi)$.

Proof of Theorem 2.2. Note that if we let $\tilde{f}$ be the limit of $S_N f$, then the Fourier coefficients of $f - \tilde{f}$ are all zero. Thus, to prove the result, it suffices to show that if the Fourier coefficients of a function $f \in L^2(-\pi,\pi)$ are all zero, then $f = 0$ a.e.

We will only consider the case when f is real-valued. The complex-valued case is left as an exercise.

Suppose first that f is a continuous function with $f(-\pi) = f(\pi)$ and that we can hence extend f to a continuous periodic function on $\mathbb{R}$. If $f \neq 0$, then $|f|$ attains it maximum value $M > 0$ at some point $x_0 \in [-\pi,\pi]$ and replacing f by $-f$ if necessary, we may assume that $f(x_0) = M > 0$. As f is continuous we can choose $\delta > 0$ so that $f(x) \geq \frac{1}{2}M$ on $[x_0 - \delta, x_0 + \delta]$ and now consider the function

$$g(x) = 1 + \cos(x - x_0) - \cos\delta.$$

As g^n is a trigonometric polynomial, the fact that the Fourier coefficients of f are zero implies that $\int_{-\pi}^{\pi} f(x)\, g^n(x)\, dx = 0$ for all n.

We can however also use that $|g| \leq 1$ for $\delta \leq |x - x_0| \leq \pi$, while $g \geq 1 > 0$ for $\frac{1}{2}\delta < |x - x_0| \leq \delta$ and $g > 1 + \cos(\delta/2) - \cos(\delta) > 1$ for $|x - x_0| \leq \frac{1}{2}\delta$ to see that

$$\begin{aligned}
\int_{-\pi}^{\pi} f(x)\, g^n(x)\, dx &= \int_{x_0-\pi}^{x_0+\pi} f(x)\, g^n(x)\, dx \\
&\geq \int_{x_0-\delta/2}^{x_0+\delta/2} f(x)\, g^n(x)\, dx - \int_{\delta\leq|x-x_0|\leq\pi} |g^n(x) f(x)|\, dx \\
&\geq \frac{1}{2}M\,(1 + \cos\frac{\delta}{2} - \cos\delta)^n\,\delta - 2\pi M\, 1^n \overset{n\to\infty}{\longrightarrow} \infty
\end{aligned}$$

which contradicts the fact that the right hand side is equal to zero for all n.

Having shown the claim for continuous 2π periodic functions we now consider the general case where $f \in L^2(-\pi,\pi)$. As $a_n(f) = 0$ for all n, so in particular for $n = 0$ we know from Proposition 2.1 that $F(x) = \int_0^x f(t)\,dt$ is 2π-periodic and, by standard properties of integration, also continuous. At the same time, Proposition 2.1 implies that $a_n(F) = 0$ for all $n \neq 0$ so we can apply the first case to $F - a_0(F)$ to see that F is constant and as $F(0) = 0$ must hence be identically zero. This implies that $f = 0$ a.e. as desired. □

Remark. *The proof above actually shows a stronger statement: if f is an integrable function and if all Fourier coefficients of f are zero, then $f = 0$ a.e.*

An immediate consequence of Theorem 1.11 and the fact that the functions e_n form an on-basis of L^2 that we have just established is Parseval's identity for Fourier-series, i.e.

Corollary 2.3. *For every $f \in L^2(-\pi,\pi)$ we have*

$$\sum_{-\infty}^{\infty} |a_n(f)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f|^2\,dx.$$

In particular, the Fourier coefficients of any L^2 function tend to zero as $n \to \pm\infty$ and indeed we have more generally that

Proposition 2.4 (Riemann-Lebesgue Lemma)**.** *Let $f \in L^1(-\pi,\pi)$. Then*

$$\int_{-\pi}^{\pi} f(t)\,e^{ikt}\,dt \to 0 \ \text{ as } k \to \pm\infty$$

Proof of the Riemann-Lebesgue Lemma. If $f \in L^2(-\pi,\pi)$, this follows from Parseval's identity.

For the general case $f \in L^1(-\pi,\pi)$ we fix any $\varepsilon > 0$, use that the continuous functions are dense in L^1 to determine $g \in C[-\pi,\pi]$ so that $h = f - g$ satisfies $\|h\|_{L^1(-\pi,\pi)} \leq \varepsilon$ and use that

$$\int_{-\pi}^{\pi} g(t)\,e^{ikt}\,dt \to 0 \text{ as } k \to \pm\infty$$

since $g \in C[-\pi,\pi] \subset L^2$, while

$$\Big|\int_{-\pi}^{\pi} h(t)\,e^{ikt}\,dt\Big| \leq \int_{-\pi}^{\pi} |h(t)|\,dt \leq \varepsilon.$$

Combined we get that $\limsup_{k\to\pm\infty} |\int_{-\pi}^{\pi} f(t)\,e^{ikt}\,dt| \leq \varepsilon$ which implies the conclusion as $\varepsilon > 0$ was arbitrary. □

We can now ask whether imposing stronger conditions on f allows us to deduce further convergence results on the Fourier series. To this end we first note that asking that f is continuous is *not sufficient* to ensure that the Fourier series converges pointwise everywhere. Indeed, as a consequence of the Uniform boundedness principle that we will prove in the next chapter, we will be able to show

Theorem 2.5. *There exists a 2π-periodic continuous function $f_0 : \mathbb{R} \to \mathbb{R}$ whose Fourier series diverges at $x = 0$.*

Remark. *One can use the above to build a continuous function whose Fourier series diverges at any n arbitrarily given points (cf problem sheets).*

If we want to obtain pointwise convergence of Fourier series then we hence need to impose stronger conditions on the regularity of f. The right condition turns out to be Hölder continuity:

Definition 2.6. *Let $\alpha \in (0, 1]$. We say that a function f is α-Hölder continuous at a point x if there is some $A > 0$ and $\delta > 0$ such that*

$$|f(x+h) - f(x)| \leq A|h|^\alpha \text{ for } |h| \leq \delta.$$

When $\alpha = 1$, we say f is Lipschitz continuous at x.

Given $a < b$ we can also consider the Hölder space $C^{0,\alpha}([a, b])$ which is the set of functions $f : I := [a, b] \to \mathbb{R}$ for which

$$[f]_\alpha := \sup\{\frac{|f(x) - f(y)|}{|x - y|^\alpha} : x, y \in I, x \neq y\} < \infty$$

We remark that this space equipped with $\|\cdot\|_{0,\alpha} := \|f\|_{sup} + [f]_\alpha$ is a Banach space.

Theorem 2.7. *Assume that $f \in L^1_{loc}(-\pi, \pi)$ is 2π-periodic and α-Hölder continuous at a point $x_0 \in [-\pi, \pi]$ for some $\alpha \in (0, 1]$. Then*

$$\lim_{N\to\infty} S_N f(x_0) = f(x_0).$$

To prove this result we note that we can write the N-th partial Fourier sum of a function f as

$$S_N f(x) = \sum_{n=-N}^{N} a_n e^{inx} = \int_{-\pi}^{\pi} f(t)\, k_N(x - t)\, dt$$

where a simple computation gives

$$k_N(x) = \frac{1}{2\pi} \sum_{n=-N}^{N} e^{inx} = \frac{1}{2\pi} \frac{\sin((N + \frac{1}{2})x)}{\sin \frac{x}{2}} = \frac{1}{2\pi} \cos(Nx) + \frac{1}{2\pi} \frac{\cos \frac{x}{2}}{\sin \frac{x}{2}} \sin(Nx). \quad (2.1)$$

Proof of Theorem 2.7. It suffices to consider real valued functions as we can split any complex valued f into its real and imaginary part. Also, as the theorem trivially holds for constant functions f we can assume that $f(x_0) = 0$, as we can otherwise replace f by $f - f(x_0)$ and we can also assume without loss of generality that $x_0 = 0$ as we can otherwise replace f by $f(\cdot - x_0)$ and use that translations of the variable commute with $\mathcal{F}$.

Given any $\delta > 0$ we then note that we can apply the Riemann-Lebesgue lemma to the functions $\frac{1}{2\pi} f \mathbb{1}_{[-\pi,\pi]\setminus[-\delta,\delta]}$ and $\frac{1}{2\pi} \frac{\cos\frac{x}{2}}{\sin\frac{x}{2}} f \mathbb{1}_{[-\pi,\pi]\setminus[-\delta,\delta]}$ to see that

$$\lim_{N\to\infty} \int_{[-\pi,\pi]\setminus[-\delta,\delta]} f(t)\, k_N(t)\, dt \to 0.$$

On the other hand, using the Hölder continuity of f at $x_0 = 0$, we have $|f(h)| \leq A|h|^\alpha$ for small $|h|$, so if δ is small in particular for all $|h| \leq \delta$ and we thus get

$$\Big|\int_{-\delta}^{\delta} f(t)\, k_N(t)\, dt\Big| \leq A \int_{-\delta}^{\delta} |t|^\alpha\, |k_N(t)|\, dt \leq \frac{A}{\pi} \int_0^\delta \frac{t^\alpha}{\sin \frac{t}{2}}\, dt.$$

Using the inequality $\sin \frac{t}{2} \geq \frac{t}{\pi}$ for $0 \leq t \leq \pi$, we see that the right hand side is bounded from above by $\frac{A}{\alpha}\delta^\alpha$. Putting everything together we obtain

$$\limsup_{N\to\infty} |S_N f(0)| \leq \frac{A}{\alpha}\, \delta^\alpha.$$

Since $\delta > 0$ was arbitrary, this implies $S_N f(0) \to 0 = f(0)$, as desired. □

Chapter 3

The Baire category theorem and fundamental results for bounded linear operators

In B4.1 you have already seen one of the fundamental results of functional analysis, namely the Theorem of Hahn-Banach. In this section we discuss other corner stones on which many results in Functional Analysis are built, and which, together with Hahn-Banach, are hence sometimes referred to as the principles of Functional Analysis. These are the uniform boundedness principle and the Open mapping theorem (and the closely related closed graph theorem). The proofs of these theorems are all based on a basic but key result for complete metric spaces, namely the Baire category theorem.

3.1 The Baire category theorem

Let (M, d) be a metric space. We recall that a subset of a metric space M is called dense if $\bar{S} = M$. We furthermore define

Definition 3.1. *A subset S of a metric space M is*

(i) nowhere dense in M if $\bar{S}$ has empty interior.

(ii) has Baire category 1 if it can be written as a countable union $S = \bigcup_{i\in\mathbb{N}} A_i$ of nowhere dense sets A_i.

(iii) has Baire category 2 if it has not category 1.

(iv) residual if S^c has category 1.

Example. *While $\mathbb{Q}$ is dense in $\mathbb{R}$ it is a category 1 set since it is given by the countable union of singletons which have non-empty interior in $\mathbb{R}$. Hence the set of irrationals in $\mathbb{R}$ is a residual set (and as we will see below hence has category 2).*

We will use $\mathrm{cat}(S) \in \{1, 2\}$ as shorthand for the Baire category of a set.

From the definition it is immediately clear that the countable union of category 1 sets is again category 1 and that any subset of a category 1 set is again category 1. This might look similar to the concept of null-sets in integration, and there are further parallels in that both of these concepts can be thought of giving a notion of "sets being small", but we point out not every category 1 set in $\mathbb{R}$ is null and not every null set is category 1.

We note that for closed respectively open sets, the notions of dense and nowhere dense are related by

Lemma 3.2. *Let A be a subset of a metric space. Then the following are equivalent*

(i) A is closed and nowhere dense

(ii) A^c is open and dense.

Proof of Lemma 3.2. If A^c is open and dense then A is of course closed. If it was not nowhere dense then $\bar{A} = A$ would have non-empty interior so would contain some ball $B_r(x)$, $r > 0$, $x \in M$. But then there would be no $y \in A^c$ so that $d(x, y) < r$ which contradicts the assumption that A^c is dense.

If we instead know that A is closed and nowhere dense then we of course know that A^c is open. Given $x \in M$ and $\varepsilon > 0$ we know that $B_\varepsilon(x)$ cannot be fully contained in A (as this would contradict A nowhere dense) so there must be $y \in A^c$ with $d(x, y) < \varepsilon$. As x and $\varepsilon > 0$ were arbitrary we hence get that A^c is dense. □

Theorem 3.3 (The Baire category theorem)**.** *Let (M, d) be a non-empty complete metric space. Then*

(i) cat$(M) = 2$, so M cannot be written as the countable union of nowhere dense sets.

(ii) Every residual set is dense, i.e. cat$(A^c) = 1$ implies that $\overline{A} = M$.

Remark. *(ii) immediately implies (i): If M was category 1 then (ii) would imply that $\emptyset = M^c$ was dense in M which is obviously wrong (as $M \neq \emptyset$) This second statement is hence sometimes called the "strong form" of the Baire category theorem (and (i) the weak form), and while also (i) would suffice to derive the key results on bounded linear operators mentioned above, it turns out that the proof of (ii) is equally simple/difficult, so we prove the full result.*

Proof of the Baire Category Theorem. One can either work directly with nowhere dense sets, or use Lemma 3.2 to instead translate everything into claims about dense sets. Here we choose the later route and as main step of the proof show:

Claim 1: Let U_j, $j \in \mathbb{N}$, be open and dense. Then $\bigcap_{j\in\mathbb{N}} U_j$ is dense.

Proof of Claim 1: Let $x \in M$, $\varepsilon > 0$. We need to show that there exists $y \in \bigcap U_j$ so that $d(x, y) \leq \varepsilon$. To this end we inductively construct a sequence $(y_j)_{j\geq 1}$ by first setting $\varepsilon_0 = \frac{\varepsilon}{2}$ and $y_0 = x$ and then for $j \geq 0$

- use that U_{j+1} is dense to get $y_{j+1} \in U_{j+1}$ with $d(y_j, y_{j+1}) \leq \frac{\varepsilon_j}{2}$
- use that U_{j+1} is open to get $\varepsilon_{j+1} \in (0, \frac{\varepsilon_j}{2}]$ so that $\bar{B}_{\varepsilon_{j+1}}(y_{j+1}) \subset U_{j+1}$.

This construction ensures that

- $\varepsilon_{j+1} \leq \frac{1}{2}\varepsilon_j \leq ... \leq 2^{-j}\varepsilon_1$ so $\varepsilon_j \to 0$ as $j \to \infty$

- As $\varepsilon_{j+1} + d(y_j, y_{j+1}) \leq \varepsilon_j$ we have $\bar{B}_{\varepsilon_{j+1}}(y_{j+1}) \subset \bar{B}_{\varepsilon_j}(y_j)$ for every j and hence

$$\bar{B}_{\varepsilon_j}(y_j) \subset \bigcap_{n=1}^{j} U_n.$$

- (y_j) is a Cauchy sequence since $\varepsilon_N \to 0$ as $N \to 0$ and since for $j, k \geq N$ we have $y_j, y_k \in \bar{B}_{\varepsilon_N}(y_N)$ so $d(y_j, y_k) \leq 2\varepsilon_N$.

As (M, d) is complete we hence find that $y_j \to y$ for some y which needs to be in $\bar{B}_{\varepsilon_j}(y_j)$ for every j as the tail of the sequence is contained in these closed balls. As $\bar{B}_{\varepsilon_j}(y_j) \subset \bigcap_{n=1}^{j} U_n$ we hence get $y \in \bigcap_{n=1}^{j} U_n$ for every j so $y \in \bigcap_{n=1}^{\infty} U_n$. Finally, as $y \in \bar{B}_{\varepsilon_0}(y_0)$ and $y_0 = x$, $\varepsilon_0 = \frac{1}{2}\varepsilon$ we have $d(y, x) \leq \varepsilon$ as desired. This completes the proof of claim 1.

Proof that Claim 1 implies (ii): If $\mathrm{cat}(A) = 1$ then we can write $A = \bigcup_{j \in \mathbb{N}} A_j$ for nowhere dense sets A_j. As $\overline{A_j}$ is closed and nowhere dense we know that the sets $(\overline{A_j})^c$ are open and dense and that their intersection is hence also dense. As $A \subset \bigcup_{j \in \mathbb{N}} \overline{A_j}$ this means that $A^c \supset \big(\bigcup_{j \in \mathbb{N}} \overline{A_j}\big)^c = \bigcap_{j \in \mathbb{N}} (\overline{A_j})^c$ contains a set which is dense in M, so is itself dense in M. □

3.2 Principle of uniform boundedness

Theorem 3.4 (Principle of uniform boundedness; Theorem of Banach-Steinhaus)**.** *Let X be a Banach space and let Y be a normed space. Let $\mathscr{F} \subset \mathscr{B}(X, Y)$ be a set of bounded linear operators from X into Y which is pointwise bounded, i.e. so that for every $x \in X$ we have $\sup_{T \in \mathscr{F}} \|Tx\| < \infty$. Then $\mathcal{F}$ is uniformly bounded, i.e. so that*

$$\sup_{T \in \mathscr{F}} \|T\| < \infty.$$

Here we use the conventions familiar from part A integration that we extend the definition of the supremum also to (non-empty) subsets $A \subset \mathbb{R}$ which are not bounded from above and in this case write $\sup A = \infty$.

The above statement hence means that if for each x the set $\{\|Tx\|, T \in \mathscr{F}\}$ is bounded from above (so as it's a subset of $[0, \infty)$ equivalently bounded) then also $\{\|T\|, T \in \mathscr{F}\}$ is a bounded set.

Warning. *This theorem does not hold if X is not complete. Consider e.g. the space $X = C([0,1])$ equipped with the L^1 norm rather than the usual sup norm. Then it is easy to see that the operators $T_n(f) = n \int_0^{\frac{1}{n}} f(t)dt$ are bounded pointwise (since each element of X is a bounded function) but that their operator norms are not bounded.*

Proof of the Uniform Boundedness Principle. For each $n \in \mathbb{N}$ we set $A_n = \{x \in X : \|Tx\|_Y \leq n \text{ for all } T \in \mathscr{F}\}$ and note that this set is closed as all T are continuous.

By hypothesis, each $x \in X$ belongs to some A_n and so $X = \cup_{n=1}^{\infty} A_n$. By the Baire category theorem, we hence cannot have that all A_n are nowhere dense, so there must be some n_0 such that $A_{n_0} = \bar{A}_{n_0}$ has non-empty interior i.e. contains some ball $B_{r_0}(x_0)$, $x_0 \in X$, $r_0 > 0$.

We now want to get a uniform bound on $\|Tx\|_Y$ that is valid for all $T \in \mathscr{F}$ and all $\|x\|_X < r_0$.

For this we note that for each such x, we have $x_0 + x \in B_{r_0}(x_0)$ and so, by the definition of A_{n_0},

$$\|T(x_0 + x)\|_Y \leq n_0 \text{ for all } T \in \mathscr{F}.$$

As also $x_0 \in B_{r_0}(x_0)$ we also have $\|T(x_0)\|_Y \leq n_0$ for all $T \in \mathscr{F}$ so by the triangle inequality get

$$\|Tx\|_Y \leq \|T(x_0 + x)\|_Y + \|Tx_0\|_Y \leq 2n_0 \text{ for all } T \in \mathscr{F} \text{ and all } \|x\|_X < r_0.$$

Since all T's are linear this implies that $\|Tx\| \leq 2n_0 r_0^{-1}$ for all x with $\|x\| < 1$, so by continuity also for all x with $\|x\| \leq 1$. This shows that $\sup_{T \in \mathscr{F}} \|T\|_{\mathscr{B}(X,Y)} \leq 2n_0\, r_0^{-1} < \infty$. $\square$

The principle of uniform boundedness has far reaching consequences and we will see many results that are based on it throughout the course. One very useful consequences is

Theorem 3.5. *Let X and Y be Banach spaces and consider a sequence $T_n \in \mathscr{B}(X,Y)$. Then the following statements are equivalent.*

(i) There exists $T \in \mathscr{B}(X,Y)$ such that, for every $x \in X$, $T_n x \to Tx$ as $n \to \infty$.

(ii) For each $x \in X$, the sequence $(T_n x)$ is convergent.

(iii) There is a constant M and a dense subset Z of X such that $\|T_n\| \leq M$ and so that the sequence $(T_n z)$ is convergent for each $z \in Z$.

Warning. *In the above theorem, the convergence of T_n to T is in the pointwise sense, and the above statements do not imply convergence in the sense of $\mathscr{B}(X,Y)$. To see this consider for example $X = \ell^2$, $Y = \mathbb{R}$ and $T_n((a_1, a_2, \ldots)) = a_n$. Then, for every $x \in \ell^2$, $T_n x \to 0$, but $\|T_n\| = 1 \not\to 0$.*

Proof. It is clear that (i) $\Rightarrow$ (ii). That (ii) $\Rightarrow$ (iii) is a direct application of the principle of uniform boundedness as pointwise convergence yields pointwise boundedness.

It remains to prove (iii) $\Rightarrow$ (i), so suppose that the assumptions of (iii) hold:

We first claim that, for every $x \in X$, $(T_n x)$ is Cauchy, and hence convergent. To see this, fix any $x \in X$, $\epsilon > 0$, and note that, for every $z \in Z$,

$$\begin{aligned}
\|T_n x - T_m x\| &\leq \|T_n z - T_m z\| + \|T_n(x - z)\| + \|T_m(x - z)\| \\
&\leq \|T_n z - T_m z\| + 2M\|x - z\|.
\end{aligned}$$

In particular, if we choose $z \in Z$ such that $\|x - z\| \leq \frac{\epsilon}{4M}$ and choose N such that $\|T_n z - T_m z\| \leq \frac{\epsilon}{2}$ for $n, m \geq N$, we obtain $\|T_n x - T_m x\| \leq \epsilon$ for all $n, m \geq N$. This proves the claim

For each $x \in X$, we can hence define Tx as the limit of $T_n x$. As each T_n is linear also T is linear. Finally, we have

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leq \limsup_{n\to\infty} \|T_n\|\|x\| \leq M\|x\| \text{ for all } x \in X$$

Thus T is a bounded linear operator on X and we have established (i). $\square$

A useful consequence of the Riesz representation theorem and the uniform boundedness principle is

Theorem 3.6. *Let X be a Hilbert space and suppose that $\mathscr{F}$ is a subset of $\mathscr{B}(X)$ such that*

$$\sup_{T\in\mathscr{F}} |\langle Tx, y\rangle| < \infty \text{ for each } x, y \in X.$$

Then $\sup_{T\in\mathscr{F}} \|T\| < \infty$.

Proof of Theorem 3.6. By the principle of uniform boundedness, it suffices to show that, for each fixed $x \in X$, $\{\|Tx\| : T \in \mathscr{F}\}$ is bounded.

Fix an $x \in X$. Define $K_{T,x} \in X^*$ by $K_{T,x}(y) = \langle y, Tx\rangle$. Then, for each $y \in X$, $\{|K_{T,x}(y)| : T \in \mathscr{F}\}$ is bounded. The principle of uniform boundedness implies then $\{\|K_{T,x}\|_* : T \in \mathscr{F}\}$ is bounded. As $\|K_{T,x}\|_* = \|Tx\|$, we conclude the proof. □

As a consequence of the Uniform boundedness principle we can now also complete the proof of Theorem 2.5 which asserted that there exists a continuous 2π periodic function whose Fourier series at $x_0 = 0$ diverges:

Proof of Theorem 2.5. The convergence of the Fourier series of a function f at $x_0 = 0$ means that

$$\lim_{N\to\infty} S_N f(0) = \lim_{N\to\infty} \int_{-\pi}^{\pi} f(x) k_N(x)\,dx \text{ exists.}$$

Let $X = \{f \in C[-\pi, \pi] : f(\pi) = f(-\pi)\}$ and define $A_N \in X^*$ by

$$A_N(f) = \int_{-\pi}^{\pi} f(x)\,k_N(x)\,dx.$$

Assume by contradiction that the Fourier series of every continuous function converges at $x_0 = 0$. Then $A_N(f)$ is bounded for every f. By the principle of uniform boundedness, this means that $\|A_N\|_*$ is bounded.

However, we can easily check that

$$\|A_N\|_* = \int_{-\pi}^{\pi} |k_N(x)|\,dx.$$

so using the formula for k_N given in (2.1) and the inequality $\sin x \le x$ for $x > 0$, we get

$$\|A_N\|_* \ge \frac{1}{\pi}\int_{-\pi}^{\pi} \Big|\sin((N + \tfrac{1}{2})x)\Big|\frac{dx}{|x|} = \frac{2}{\pi}\int_0^{(N+\frac{1}{2})\pi} |\sin x|\frac{dx}{|x|} \ge C\,\log N$$

for some positive constant C independent of N. This gives a contradiction and concludes the proof. □

3.3 The open mapping theorem

Definition 3.7. *A map f between two metric spaces X and Y (or more generally two topological spaces) is called open if $f(U)$ is open for every open subset U of X.*

For linear operators we have the following simple way of testing whether a map is open:

Lemma 3.8. *Let X, Y be normed spaces, $T : X \to Y$ linear. Then the following are equivalent*

(i) *T is open*

(ii) *There exists $\delta > 0$ so that $B_\delta^Y \subset T(B_1^X)$*

(iii) *There exists $\varepsilon > 0$, $y_0 \in Y$ and $R > 0$ so that $B_\varepsilon^Y(y_0) \subset T(B_R^X)$*

To avoid confusion here we indicate in which space which ball lies by writing $B_r^X(x)$ for the open ball with radius r around x and to shorten notation we abbreviate balls around the origin with B_r^X.

Proof of Lemma 3.8. $(i) \Rightarrow (ii)$ follows as $0 \in T(B_1^X)$ and $(ii) \Rightarrow (iii)$ is trivially true.

Proof that (ii) implies (i):
Let U be open, $y \in T(U)$. Then there exists $x \in U$ so that $y = Tx$ and since U is open there exists $r > 0$ so that $B_r^X(x) \subset U$. By linearity and (ii) hence $T(B_r^X(x)) = \{Tx\} + rT(B_1^X) \supset \{Tx\} + rB_\delta^Y = B_{\delta r}^Y(Tx)$.

Proof that (iii) implies (ii): Note that $\Omega := T(B_R^X)$ is convex (as T is linear and as B_R^X is convex) and symmetric in the sense that if $y \in \Omega$ then also $-y \in \Omega$ (again as the ball has this property and as T is linear).

Hence also $B_\varepsilon^Y(-y_0) \subset \Omega$ and so for each $y \in B_\varepsilon^Y$ we know that both $\pm y_0 + y$ are in Ω and hence, by convexity, so is $y = \frac{1}{2}(y_0 + y) + \frac{1}{2}(-y_0 + y)$.

Thus $B_\varepsilon^Y \subset T(B_R^X)$ so by linearity $B_{\varepsilon/R}^Y \subset T(B_1^X)$. □

Based on the Baire-category theorem we can now establish the following fundamental result about bounded linear operators between Banach spaces:

Theorem 3.9 (Open mapping theorem)**.** *Let X, Y be Banach spaces and let $T \in \mathscr{B}(X, Y)$ be surjective. Then T is an open map.*

Note that for linear maps the converse is trivially true (in particular also if X, Y are not complete).

Proof of the open mapping theorem. We first show

Claim 1: There exists $\varepsilon > 0$ so that $B_{2\varepsilon}^Y \subset \overline{T(B_1^X)}$.
Proof of Claim 1: As $Y = T(X) = T(\bigcup_n (B_n^X)) = \bigcup T(B_n^X)$ and as Y is Banach we know by the Baire category theorem that at least one of the $T(B_n^X)$ is not nowhere dense, so $\overline{T(B_n^X)}$ has nonempty interior, so contains some ball $B_\delta^Y(y_0)$. As in the proof of the above lemma we can easily check that this implies that $B_\delta^Y \subset \overline{T(B_n^X)}$ and we get the claim for $\varepsilon = \frac{\delta}{2n}$.
To complete the proof we now show that

Claim 2: For this ε we have $B_\varepsilon^Y \subset T(B_1^X)$.
Proof of Claim 2: Let $y \in B_\varepsilon^Y$. The idea is to construct a sequence (x_k) so that $\|x_k\| < 2^{-k}$ and so that $\sum_{k=1}^m Tx_k$ converges to y in Y.

To get such a sequence we can use that $B^Y_{2^{-j+1}\varepsilon} \subset \overline{T(B_{2^{-j}})}$ by the first part and linearity. As $y_0 = y \in B^Y_\varepsilon$ we can hence choose x_1 with $\|x_1\| \leq 2^{-1}$ so that $y_1 := y_0 - Tx_1$ is so that $\|y_1\| < \varepsilon 2^{-1}$.

We then proceed inductively, so suppose we have already found $x_1, .., x_j$ and $y_1, \dots, y_j$ with $y_k = y_{k-1} - Tx_k$ and $\|x_k\| < 2^{-k}$ and $\|y_k\| < \varepsilon 2^{-k}$ for all $k = 1, \dots, j$.

Then we use that $y_j \in B^Y_{2^{-j}\varepsilon} \subset \overline{T(B_{2^{-j-1}})}$ to choose the next $x_{j+1} \in B^X_{2^{-(j+1)}}$ so that $\|y_j - Tx_{j+1}\| < \varepsilon 2^{-j}$ and continue the process with $y_{j+1} = y_j - Tx_{j+1}$.

Since $\sum \|x_k\| < 1$ we know that $\sum x_k$ converges in X to some $x \in B^X_1$, as absolute convergence of series in Banach spaces implies convergence. As $y - \sum_{k=1}^n Tx_k = y_n \to 0$ in Y we hence get that

$$y = \sum_{k=1}^\infty Tx_k = \lim_{n\to\infty} \sum_{k=1}^n Tx_k = \lim_{n\to\infty} T(\sum_{k=1}^n x_k) = T(\sum_{k=1}^\infty x_k) = Tx$$

where the third step follows by linearity and the forth by continuity of T. As y was an arbitrary element of B^Y_ε this shows that $B^Y_\varepsilon \subset T(B^X_1)$ and hence that T is open. □

An important consequence is:

Theorem 3.10 (Inverse mapping theorem). *Let X, Y be Banach spaces and let $T \in \mathscr{B}(X, Y)$ be bijective. Then T is invertible.*

Proof. Exercise. □

Application. *Let X be a Banach space with respect to two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ and suppose that there is a constant $C > 0$ such that $\|x\|_1 \leq C\|x\|_2$ for all $x \in X$. Then the two norms are equivalent, i.e. there is a constant C' such that $\|x\|_2 \leq C'\|x\|_1$ for all $x \in X$.*

As a consequence of the inverse mapping theorem we can also show

Theorem 3.11. *Let $T \in \mathscr{B}(X, Y)$ be a bounded linear operator between Hilbert spaces. Then TX is closed if and only if T^*Y is closed.*

In the proof we will use the inverse function theorem and the fact that the adjoint of an invertible operator between Hilbert spaces is again invertible that you have seen in B4.1.

Proof of Theorem 3.11. It suffices to show only one direction, as $T^{**} = T$ so we prove that if $W := T^*Y$ is closed in X then also TX is closed in Y, i.e. $TX = \overline{TX}$.

For this we set $Z = \overline{TX}$ and need to show that $Z \subset TX$ as the reverse inclusion is trivially satisfied. We will prove this by showing that the identity map I_Z can be obtained as a composition of the map T with a suitable map from Z to $W \subset X$ which we construct as follows:

As T maps X into Z we can view it instead as a map into the Banach space Z and we call the resulting map S, i.e. $S \in \mathscr{B}(X, Z)$ is simply given by $Sx = Tx$ for all $x \in X$. The adjoint S^* of S is an operator from Z to X. By (1.5), $Z = \overline{\operatorname{Im} S} = (\operatorname{Ker} S^*)^\perp$, so $\operatorname{Ker} S^* = \{0\}$, i.e. S^* is injective.

We claim that $\operatorname{Im} S^* = W$. To this end we let P be the orthogonal projection from Y onto Z and compute, for $x \in X$ and $y \in Y$,

$$\langle Tx, y\rangle_Y = \langle Sx, Py\rangle_Y = \langle x, S^*Py\rangle_X.$$

This shows that $T^* = S^* \circ P$, and so $\operatorname{Im} S^* = W$, as claimed.

So, S^* can be regarded as a bounded bijective linear operator between Z and W. To make the notation clearer, we rename it as $V \in \mathscr{B}(Z, W)$, $Vz = S^*z$ for all $z \in Z$. As both Z and W are Banach we can apply the inverse mapping theorem to deduce that V is invertible, which in turn implies that also V^* is invertible.

We finally claim that

$$I_Z = T \circ (V^*)^{-1},$$

which then immediately implies that $Z \subset TX$ and hence that indeed $TX = Z = \overline{TX}$.

To show this, i.e. that $z = T((V^*)^{-1}z)$ for any given $z \in Z$, we write for short $w = (V^*)^{-1}z$, note that both z and Tw are elements of Z and that for any $y \in Z$:

$$\langle Tw, y\rangle_Y = \langle Sw, y\rangle_Y = \langle w, S^*y\rangle_X = \langle w, Vy\rangle_X = \langle V^*w, y\rangle_Y = \langle z, y\rangle_Y.$$

Since this holds for all $y \in Z$, so in particular for $y = Tw - z$, we deduce that $Tw = z$ and so $T \circ (V^*)^{-1} = I_Z$ as desired. $\square$

3.4 The closed graph theorem

We recall that if X, Y are normed spaces then we can turn $X \times Y$ into a normed space by defining $\|(x, y)\| = \sqrt{\|x\|_X^2 + \|y\|_Y^2}$ and that this choice of norm ensures that $X \times Y$ inherits all key properties that the spaces X and Y might have, i.e. if X and Y are both Banach or both inner product spaces or Hilbert spaces or separable then also $X \times Y$ has this property.

A very useful result to test whether a linear map is bounded is

Theorem 3.12 (Closed graph theorem)**.** *Let X and Y be Banach spaces and let T be a linear operator from X into Y. Then T is bounded if and only if its graph*

$$\Gamma(T) = \{(x, y) \in X \times Y : y = Tx\}$$

is closed in $X \times Y$.

Proof of the Closed Graph Theorem. If T is bounded, then it is continuous so if a sequence (x_n, Tx_n) of elements of $\Gamma(T)$ converges to some (x, y) we must have $y = \lim Tx_n = T \lim x_n = Tx$, i.e. $(x, y) \in \Gamma(T)$. Thus $\Gamma(T)$ is closed.

Conversely, assume that $\Gamma(T)$ is closed. We note that the graph of any linear operator is a subspace, so as $\Gamma(T)$ is closed and as $X \times Y$ is Banach we find that $\Gamma(T)$ itself is also a Banach space with the norm induced by the norm on $X \times Y$. Consider now the projections onto the components of elements of $\Gamma(T)$, i.e. the continuous maps $P_1 : \Gamma(T) \to X$ and $P_2 : \Gamma(T) \to Y$ defined by

$$P_1(x, Tx) = x \text{ and } P_2(x, Tx) = Tx.$$

It is clear that P_1 is a bijection. By the inverse mapping theorem, P_1 has a continuous inverse P_1^{-1}. The conclusion follows from the fact that $T = P_2 \circ P_1^{-1}$. $\square$

Remark. *Usually, to show that a map A from a normed space X to another normed space Y is continuous, one needs to show that if $x_n \to x$, then $A(x_n) \to A(x)$. In many situations, one struggles to prove some kind of convergence for $A(x_n)$, let alone the* *convergence to $A(x)$. However, if X and Y are Banach spaces and if A is linear, then by virtue of the closed graph theorem, one may assume from the beginning that $A(x_n)$ is convergent in the sense of norm!*

Example. *Let X be a Banach space, and let Y and Z be closed subspaces of X such that $X = Y \oplus Z$. Then the direct sum projection $P : X \to Y$ from X onto the first summand Y is bounded.*

Proof. By the closed graph theorem, it suffices to show that the graph of P is closed, i.e. that if a sequence (x_n, y_n) of elements of the graph converge to some (x, y) then $y = Px$, i.e. $y \in Y$ and $x - y \in Z$.

This immediately follows since the spaces Y and Z are closed and since for element of the graph we have $y_n = Px_n$: namely as $y_n \to y$ we have that $y \in Y$ and as $x_n - y_n = x_n - Px_n$ are by definition of P elements of Z and as they converge to $x - y$ we have that $x - y \in Z$, so indeed $Px = y$, as desired. $\square$

Furthermore, a useful consequence of the closed graph theorem is

Proposition 3.13. *Let X be a Hilbert space and let $T : X \to X$ be linear. If $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in X$, then T is bounded and so self-adjoint.*

Proof of Proposition 3.13. As before, we only need to show that if $x_n \to x$ and $Tx_n \to z$, then $z = Tx$ as we can then apply the closed graph theorem. Indeed, for any $y \in X$, we have

$$\langle Tx, y\rangle = \langle x, Ty\rangle = \lim_{n\to\infty} \langle x_n, Ty\rangle = \lim_{n\to\infty} \langle Tx_n, y\rangle = \langle z, y\rangle,$$

which implies $z = Tx$. $\square$

Example. *It is clear that if $h \in L^\infty(\mathbb{R})$, then the multiplication operator $f \mapsto hf =: M_h f$ maps $L^1(\mathbb{R})$ into itself. The converse of this is also true: If h is some measurable function such that $M_h f \in L^1(\mathbb{R})$ for all $f \in L^1(\mathbb{R})$, then $h \in L^\infty(\mathbb{R})$.*

Proof. By hypothesis M_h maps $L^1(\mathbb{R})$ into itself. We claim that M_h is bounded. To this end, we need show that if $f_n \to f$ and $M_h f_n \to g$, then $g = M_h f$ in the sense of L^1, i.e. a.e. Once we have shown this we deduce that M_h is bounded from the closed graph theorem as L^1 is a Banach space.

For this we use that any sequence that converges in a space L^p has a subsequence that converges pointwise almost everywhere to the same limit.

We first apply this to f_n and use that as $f_n \to f$ in L^1, there is a subsequence, say f_{n_j}, which converges to f a.e. It follows that $M_h f_{n_j} \to M_h f$ a.e. Since we also know that $M_h f_{n_j} \to g$ in L^1, we can pass to a further subsequence so that $M_h f_{n_{j_k}} \to g$ pointwise almost everywhere and hence conclude that $M_h f = g$ a.e. as desired.

Having thus shown that M_h is bounded we now claim that $h \in L^\infty$ and that indeed

$$|h| \le \|M_h\| \text{ a.e.}$$

Suppose that this was not the case i.e. that $Z := \{x : |h(x)| > \|M_h\|\}$ was not a null set. Since we can write $Z = \bigcup_{n\in\mathbb{N}} Z_n$ for $Z_n := \{x \in [-n, n] : |h(x)| > \|M_h\| + \frac{1}{n}\}$ and as the countable union of null sets would be null, there must be some n so that $\mathcal{L}(Z_n) > 0$ and hence so that the corresponding characteristic function $f := \mathbb{1}_{Z_n}$ is a non-zero element of L^1.

However, as $|h| > \|M_h\| + \frac{1}{n}$ on Z_n and as $\|f\| = \mathcal{L}(Z_n) > 0$ we have for this f that

$$\|M_h f\|_{L^1} \geq \int_{Z_n} \|M_h\| + \frac{1}{n} dt = \mathcal{L}(Z_n)(\|M_h\| + \frac{1}{n}) = (\|M_h\| + \frac{1}{n})\|f\| > \|M_h\|\|f\|,$$

contradicting the definition of the operator norm.

$\square$

Chapter 4

Weak convergence

The fact that bounded sequences in infinite dimensional spaces cannot be expected to have a convergent subsequence means that many arguments that you are familiar with from the analysis of functions from $\mathbb{R}$ (or $\mathbb{R}^n$) no longer work in infinite dimensional spaces. This is particularly problematic as most interesting objects (like functions, sequences, curves, surfaces,.....) are not part of a finite dimensional space, but rather form subsets of infinite dimensional spaces. Rather than accepting that we can no longer pass to subsequences which have additional convergence properties, a key idea in the analysis of problems on infinite dimensional spaces is to introduce a weaker notation of convergence. While we cannot expect that such a weaker notion of convergence will still have all of the nice properties of the "standard convergence" (often also called strong convergence or convergence in norm to distinguish clearly from weak convergence), it turns out that several key properties still hold and, most importantly, that in a large class of infinite dimensional spaces (namely all reflexive Banach spaces so in particular all Hilbert spaces) every bounded sequence will have a subsequence that converges in this weak sense.

4.1 Definition and examples

Definition 4.1. *A sequence (x_n) in a normed space X is said to converge weakly to $x \in X$ if*

$$\lim_{n\to\infty} \ell(x_n) = \ell(x) \text{ for all } \ell \in X^*.$$

This relation is indicated by a half arrow

$$x_n \rightharpoonup x,$$

or to make it clearer (and recommended for handwritten notes to avoid confusion between weak and strong convergence) by $x_n \overset{w}{\rightharpoonup} x$.

As elements of X^* are continuous we know that any sequence (x_n) that converges (strongly) to x, i.e. for which $\|x_n - x\| \to 0$, also converges weakly to the same limit. Conversely, in many infinite dimensional spaces, weak convergence does not imply strong convergence.

There are some exceptional spaces where weak and strong convergence are equivalent, most notably finite dimensional spaces and the sequence space ℓ^1.

You will see the following typical example of sequences that converge weakly but not strongly on problem sheet 3

Example. *Let* $X = L^p(\mathbb{R})$, $p \in [1,\infty)$, *let* $f_n = n^{1/p}\mathbb{1}_{[0,\frac{1}{n}]}$ *Then*

- f_n *does not converge strongly*
- *For* $p \in (1,\infty)$ *we have* $f_n \rightharpoonup 0$.
- *For* $p = 1$ *the sequence does not converge weakly in* L^1.

4.2 Basic properties and Mazur's theorem

Lemma 4.2 (Uniqueness of limit)**.** *The weak limit, if it exists, is unique.*

Proof of Lemma 4.2. Suppose that $x_n \rightharpoonup x$ and $x_n \rightharpoonup \tilde{x}$ for $x \neq \tilde{x}$. Then by the Corollary of Hahn-Banach that we recalled in (1.3) there would be $f \in X^*$ so that $f(x) \neq f(\tilde{x})$. This contradicts the uniqueness of limits for the sequence of numbers $f(x_n)$.

□

Proposition 4.3. *If* $x_n \rightharpoonup x$ *then* (x_n) *is bounded and*

$$\|x\| \leq \liminf \|x_n\|.$$

Proof of Proposition 4.3. We recall that the canonical injection $\iota : X \to X^{**}$ is isometric and hence that each x_n defines an element $T_n = \iota_{x_n} \in X^{**}$ with $\|T_n\|_{**} = \|x_n\|$ by

$$T_n(\ell) = \ell(x_n) \text{ for all } \ell \in X^*.$$

Now for each $\ell \in X^*$, $T_n(\ell) = \ell(x_n)$ is convergent, and hence bounded. The principle of uniform boundedness, which is applicable as X^* is Banach, thus implies that $\|T_n\|_{**}$ is bounded and hence that (x_n) is bounded.

To get the second part, we apply the consequence (1.3) of Hahn-Banach to choose $\ell \in X^*$ such that

$$\|x\| = \ell(x) \text{ and } \|\ell\|_* = 1.$$

The conclusion then follows from the inequality

$$|\ell(x_n)| \leq \|\ell\|_* \|x_n\| = \|x_n\|$$

and the fact that $\ell(x_n) \to \ell(x) = \|x\|$. □

The second part of this proposition, is a special case of the following stronger statement:

Theorem 4.4 (Mazur)**.** *Let* K *be a closed convex subset of a normed space* X *and let* (x_n) *be a sequence of elements of* K *which converges weakly to some* $x \in X$. *Then* $x \in K$.

We assume for granted the following result, which can be obtained as a consequence of the Hahn-Banach theorem.

Theorem 4.5 (Hyperplane separation theorem)**.** *Let X be a normed space, let A and B be disjoint convex subsets of X so that $\overset{\circ}{B} \neq \emptyset$. Then A and B can be separated by a hyperplane, i.e. there is $\ell \in X^* \setminus \{0\}$ and a number c such that*

$$Re\,\ell(x) \leq c \leq Re\,\ell(y) \text{ for all } x \in A, y \in B. \tag{4.1}$$

We note that (4.1) indeed also holds for $x \in \bar{A}$ and $y \in \bar{B}$ as ℓ is continuous.

This, and related results, will be discussed in C4.1 Further Functional Analysis.

We will use the following consequence of the above theorem:

Corollary 4.6. *Let K be a closed convex subset of a normed space X and let $x_0 \in X \setminus K$. Then there exists $\ell \in X^*$ and $\delta > 0$ so that*

$$Re(\ell(y)) \leq Re(\ell(x_0)) - \delta \text{ for all } y \in K.$$

Proof of Corollary 4.6. Since $x_0 \in K^c$ and K^c is open, there is some $r > 0$ such that $B_r(x_0) \cap K = \emptyset$. By the hyperplane separation theorem applied for $A = K$ and $B = B_r(x_0)$, there is $\ell \in X^* \setminus \{0\}$ and a number $c \in \mathbb{R}$ such that

$$\mathrm{Re}(\ell(y)) \leq c \leq \mathrm{Re}(\ell(z)) \text{ for all } y \in K \text{ and } z \in B_r(x_0). \tag{4.2}$$

As $\ell \neq 0$ there exists $\bar{x} \in X$ with $\|\bar{x}\| = 1$ and $\ell(\bar{x}) > 0$. Setting $\delta := r\ell(\bar{x}) > 0$ and noting that $x_0 - r\bar{x} \in \bar{B}_r(x_0)$ we then get that

$$\mathrm{Re}(\ell(y)) \leq \mathrm{Re}(\ell(x_0 - r\bar{x})) = \mathrm{Re}(\ell(x_0)) - \delta \text{ for all } y \in K$$

as claimed. □

Proof of Mazur's theorem. Assume by contradiction that $x \notin K$. Then Corollary gives $\ell \in X^*$ and $\delta > 0$ so that

$$\mathrm{Re}(\ell(y)) \leq \mathrm{Re}(\ell(x)) - \delta \text{ for all } y \in K.$$

But this means that $\mathrm{Re}(\ell(x_n)) \leq \mathrm{Re}(\ell(x)) - \delta$ which contradicts $\ell(x_n) \to \ell(x)$. □

A more geometric way of viewing the above results is to use Corollary 4.6 to show

Theorem 4.7 (Geometric version of Mazur)**.** *Every closed convex subset K of a normed space can be written as an intersection of half-spaces, i.e. there exist families of bounded linear operators ℓ_ι, $\iota \in I$, and numbers c_ι, $\iota \in I$, so that the corresponding half-spaces $H_{\ell_\iota, c_\iota} := \{x : Re(\ell_\iota(x)) \leq c_\iota\}$ are so that*

$$K = \bigcap_{\iota \in I} H_{\ell_\iota, c_\iota}.$$

We note that these families are not assumed to be countable and that indeed the easiest way of deriving this result from corollary 4.6 is choose all pairs $(\ell_\iota, c_\iota) \in X^* \times \mathbb{R}$ for which $K \subset H_{\ell_\iota, c_\iota}$. (This proof is an optional exercise on sheet 3.)

4.3 Weak convergence in Hilbert spaces

For sequences in Hilbert spaces we additionally have

Lemma 4.8. *Let X be a Hilbert space, (x_n) a sequence in X and $x \in X$. Then the following are equivalent:*

(i) $x_n \to x$

(ii) $x_n \rightharpoonup x$ *and* $\|x_n\| \to \|x\|$

Proof. Exercise $\square$

Proposition 4.9. *Let X be a Hilbert space, let $\{e_\iota\}_{\iota \in I}$ be an orthonormal basis of X and let (x_n) be a bounded sequence in X and let $x \in X$. Then the following are equivalent:*

(i) (x_n) *converges weakly to* x

(ii) $\langle x_n, e_\iota \rangle \to \langle x, e_\iota \rangle$ *for all* $\iota \in I$.

As we can extend any orthonormal sequence (z_n) to a orthonormal basis, by using an orthonormal basis of $(\overline{\mathrm{span}\{z_n\}})^\perp$ we obtain in particular

Corollary 4.10. *Let (x_n) be an orthonormal sequence. Then x_n tends weakly, but not strongly, to zero.*

Proof of Proposition 4.9. $(i) \Rightarrow (ii)$ is trivially true as $y \mapsto \langle y, e_\iota \rangle$ is an element of X^*.

So suppose that (ii) holds and let ℓ be any element of X^*. By the Riesz representation theorem, there exists $y \in X$ such that

$$\ell(x) = \langle x, y \rangle \text{ for all } x \in X.$$

We thus need to show that

$$\lim_{n \to \infty} \langle x_n - x, y \rangle = 0$$

or equivalently that $\limsup |\langle x_n - x, y \rangle| \leq \varepsilon$ for every $\varepsilon > 0$.

Given $\varepsilon > 0$ we can use that $X = \overline{\mathrm{span}\{e_\iota\}}$ since $\{e_\iota\}$ is an on-basis and hence that there exists $y_\varepsilon \in \mathrm{span}\{e_\iota\}$ with $\|y - y_\varepsilon\| \leq \frac{\varepsilon}{M}$ where we set $M := \sup \|x_n\| + \|x\|$.

As y_ε is a *finite* linear combination of e_ι's the assumption (ii) ensures that $\langle x_n - x, y_\varepsilon \rangle \to 0$ and we hence get that

$$\limsup_{n \to \infty} |\langle x - x_n, y \rangle| \leq \limsup_{n \to \infty} \|x - x_n\| \|y - y_\varepsilon\| \leq \varepsilon.$$

$\square$

4.4 Weak sequential compactness

Definition 4.11. *A subset A of a Banach space X is called weakly sequentially compact if every sequence of A has a subsequence that converges weakly to an element x of A.*

A cruical feature of weak convergence is the following compactness result which has far reaching consequences, most of which go beyond the remit of this course.

Theorem 4.12 (Weak sequential compactness in reflexive Banach spaces)**.** *Let X be a reflexive Banach space. Then the closed unit ball $\{x : \|x\| \leq 1\}$ is weakly sequentially compact.*

This statement is equivalent to

Corollary 4.13. *Every bounded sequence in a reflexive Banach space has a weakly convergent subsequence.*

Proof of Theorem 4.12. We will only prove the theorem in the case of Hilbert spaces, which are reflexive thanks to the Riesz representation theorem.

So let (x_n) be a sequence in the unit ball of a Hilbert space X. We first show:

Claim 1: There exists a subsequence (x_{n_j}) of (x_n), such that $\langle x_{n_j}, x_m\rangle$ converges for every m.

Proof of Claim 1:

We use a diagonal sequence argument: As (x_n) is bounded we know from Cauchy-Schwarz that the sequence $\langle x_n, x_1\rangle$ is a bounded sequence in $\mathbb{F}$. By Bolzano-Weierstrass, we can hence extract a subsequence $n_j^{(1)}$ such that $\langle x_{n_j^{(1)}}, x_1\rangle$ is convergent.

We then consider $\langle x_{n_j^{(1)}}, x_2\rangle$, which is again a bounded sequence in $\mathbb{F}$ and select a convergent subsequence $\langle x_{n_j^{(2)}}, x_2\rangle$. Clearly, $\langle x_{n_j^{(2)}}, x_1\rangle$ is also convergent.

Proceeding in this way, we construct nested subsequence $(n_j^{(k)})$ such that $\langle x_{n_j^{(k)}}, x_m\rangle$ is convergent (with respect to j) for every $m \leq k$.

Finally we set $x_{n_j} = x_{n_j^{(j)}}$ and use that for every fixed m, $(n_j)_{j\geq m}$ is a subsequence of $(n_j^{(m)})_{j\geq m}$. It follows that $\langle x_{n_j}, x_m\rangle$ is convergent for every m.

For this subsequence, we then show

Claim 2: $\langle x_{n_j}, x\rangle$ converges for every $x \in X$.

Proof of Claim 2:

By linearity and claim 1 it is clear that for every $y \in Y := \text{span}\{x_1, x_2, ...\}$ as elements of this space are *finite* linear combinations of elements x_m.

Next we claim that if y is in the the closed linear span $\bar{Y}$ of $\{x_1, \ldots\}$ then $\langle x_{n_j}, y\rangle$ is a Cauchy sequence so also converges.

To see this it suffices to consider $\varepsilon > 0$ with , use that Y is dense in $\bar{Y}$ and that we can hence choose $\tilde{y} \in Y$ with $\|y - \tilde{y}\| \leq \frac{1}{4}\varepsilon$. As we know that $\langle x_{n_j}, \tilde{y}\rangle$ is convergent, and hence Cauchy, we can then choose N so that $|\langle x_{n_j} - x_{n_k}, \tilde{y}\rangle| < \frac{1}{2}\varepsilon$ for all $j, k \geq N$. Applying first the triangle inequality and then Cauchy-Schwarz thus gives

$$|\langle x_{n_j} - x_{n'_j}, y\rangle| \leq |\langle x_{n_j} - x_{n'_j}, \tilde{y}\rangle| + |\langle x_{n_j} - x_{n'_j}, \tilde{y} - y\rangle| \leq \tfrac{1}{2}\varepsilon + 2\|y - \tilde{y}\| \leq \varepsilon.$$

On the other hand, it is clear that $\langle x_{n_j}, z\rangle = 0$ for all $z \in Y^\perp$. Hence, as $X = \bar{Y} \oplus Y^\perp$ by the projection theorem, we have that $\langle x_{n_j}, x\rangle$ is convergent for all $x \in X$.

We can now finally show

Claim 3 x_{n_j} converges weakly to some x_*.

Proof of Claim 3: As $\langle x_{n_j}, x\rangle$ is convergent for all $x \in X$ we can define the linear map

$$\ell(x) = \lim_{j\to\infty} \langle x, x_{n_j}\rangle, \qquad x \in X$$

which is an element of X^* with $\|\ell\| \leq 1$ since Cauchy-Schwarz ensures that $\|\ell(x)\| \leq \sup \|x_{n_j}\|\|x\| \leq \|x\|$ for every x. By the Riesz representation theorem, there is hence some $x_* \in X$ with $\|x_*\| \leq 1$ such that $\ell(x) = \langle x, x_*\rangle$ for all $x \in X$. This implies that $\langle x_{n_j} - x_*, x\rangle = \overline{\langle x, x_{n_j} - x_*\rangle} \to 0$ for every x and hence that $T(x_{n_j} - x_*) \to 0$ for every $T \in X^*$ as the Riesz representation theorem allows us to write any such T as map $z \mapsto \langle z, x\rangle$ for some x. This establishes that $x_{n_j} - x_* \rightharpoonup 0$ i.e. that $x_{n_j} \rightharpoonup x_*$ and hence completes the proof. □

We note that the converse of Theorem 4.12 is true, a result which we will not prove.

Theorem 4.14 (Eberlein)**.** *The closed unit ball in a Banach space X is weakly sequentially compact* only if *X is reflexive.*

As an application of Theorem 4.12, we obtain the following generalization to reflexive Banach spaces of the fact the distance $\mathrm{dist}(x, C)$ between a point x and a convex closed subset C of a Hilbert-space is always achieved.

Theorem 4.15 (Closest point in a closed convex subset)**.** *Let K be a non-empty closed convex subset of a reflexive Banach space X. Then, for every $x \in X$, there is a point $y \in K$ such that*

$$\|x - y\| = dist(x, K) := \inf_{z\in K} \|x - z\|.$$

Note that unlike in the Hilbert space case we do not claim uniqueness.

The proof of this result is a particular instance of a far more general idea, called the "direct method of calculus of variations", which can be used in many instances to establish the existence of a minimiser of a functional.

The basic idea of such proofs is the following: Given a (usually non-linear) function $E : \Omega \to \mathbb{R}$ that is defined on a subset of a reflexive Banach space we

- choose a minimising sequence u_n, i.e. a sequence in Ω so that $E(u_n) \to \inf E(u)$ (which by the definition of the infimum will always exist)
- Try to prove that any such minimising sequence is bounded (whether this works depends on the properties of E)
- Use sequential weak compactness to get a subsequence u_n which converges weakly to some $u \in X$
- Try to prove that u is still in Ω (this depends on the properties of Ω)
- Try to prove that in this process the energy cannot increase in the limit, i.e. that $E(u) \leq \liminf_{n\to\infty} E(u_n)$.

If this all works then we get that $E(u) \leq \inf_\Omega E$, but at the same time also $u \in \Omega$ so $E(u) \geq \inf_\Omega E$ so u is the desired minimiser.

Following this basic principle (also called the direct method of calculus of variations), we can prove Theorem 4.15 as follows:

Proof of Theorem 4.15. Let y_n be a minimising sequence, i.e. a sequence in K so that $\|y_n - x\| \to d(x, K)$. Then (y_n) must be bounded since for n large we have $\|y_n\| \leq \|y_n - x\| + \|x\| \leq d(x, K) + 1 + \|x\|$. By weak compactness we hence get a subsequence y_{n_j} that converges weakly to some $y \in X$. As K is closed and convex Mazur's theorem ensures that indeed $y \in K$ and as $y_{n_j} - x \rightharpoonup y - x$ we get from Proposition 4.3 that $\|y - x\| \leq \liminf \|y_n - x\| = d(x, K)$ so indeed $\|y - x\| = d(x, K)$. $\square$

Chapter 5

Spectral theory

5.1 Definitions and basic properties

The spectral theory for operators on infinite dimensional space is far richer than the one on finite dimensional spaces and of fundamental importance for understanding the operators themselves. Unlike for linear operators from a finite dimensional spaces where the rank nullity theorem ensures that an operator is invertible if and only if it is injective, for linear operator between Banach spaces there are several reasons why an operator might not be invertible.

In the following we will always consider a complex Banach space X and a bounded linear operator $T : X \to X$ and we define

Definition 5.1. *Let X be a complex Banach space and $T \in \mathscr{B}(X)$.*

(i) The resolvent set $\rho(T)$ of T is defined as

$$\rho(T) := \{\lambda \in \mathbb{C} : T - \lambda I \text{ is invertible}\}$$

and for $\lambda \in \rho(T)$, we define the resolvent operator of T at λ as $R_\lambda(T) = (T-\lambda I)^{-1}$.

(ii) The spectrum $\sigma(T)$ of T is the complement of the resolvent set, i.e.

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

(iii) The point spectrum $\sigma_p(T)$ of T is the set of complex numbers λ such that $T - \lambda I$ is not injective, i.e. so that $Ker(T - \lambda I)$ is non-trivial. The elements of $\sigma_p(T)$ are called the eigenvalues of T, and, if $\lambda \in \sigma_p(T)$, the non-trivial elements of $Ker(T - \lambda I)$ are called the eigenvectors of T.

(iv) The residual spectrum $\sigma_r(T)$ of T is the set of complex numbers λ such that $T-\lambda I$ is injective and $Im(T - \lambda I))$ is not dense in X.

(v) The continuous spectrum $\sigma_c(T)$ of T is the set of complex numbers λ for which $T - \lambda I$ is injective and $(T - \lambda I)(X)$ is a proper dense subset of X.

(vi) The approximate point spectrum $\sigma_{ap}(T)$ of T is the set of complex numbers λ such that there is a sequence $x_n \in X$ such that $\|x_n\| = 1$ and $\|Tx_n - \lambda x_n\| \to 0$.

As X is a Banach space the inverse mapping theorem ensures that $T-\lambda I$ is invertible if and only if it is bijective so $\sigma(T)$ is the disjoint union of $\sigma_p(T)$, $\sigma_r(T)$ and $\sigma_c(T)$.

We note that λ is in the approximate point spectrum if and only if there does not exist any $\delta>0$ so that $\|(T-\lambda\mathrm{I})(x)\|\geq\delta\|x\|$ for all $x\in X$, i.e. if and only if the assumption (1.9) of Lemma 1.18 is violated for $S=T-\lambda\mathrm{I}$. As this lemma ensures that the image of operators satisfying (1.9) is always closed we hence immediately deduce

Corollary 5.2.

$$\sigma_c(T)\subset\sigma_{ap}(T).$$

When X is a Hilbert space, we can also give an alternative proof of this Corollary which is not based on Lemma 1.18, but instead uses weak sequential compactness:

Alternative proof of Corollary 5.2 for Hilbert spaces. Let $\lambda\in\sigma_c(T)$ so that $Y:=\mathrm{Im}(\lambda I-T)$ is a dense proper subspace of X. Pick $p\in X\setminus Y$. Then there is some sequence x_n such that $p_n:=(\lambda I-T)x_n\to p$.

If (x_n) is bounded, then, by the weak sequential compactness property of the unit ball, we may assume without loss of generality that x_n converges weakly to some x. This implies, for $z\in X$, that

$$\langle p_n,z\rangle=\langle(\lambda I-T)x_n,z\rangle=\langle x_n,(\bar{\lambda}I-T^*)z\rangle\to\langle x,(\bar{\lambda}I-T^*)z\rangle=\langle(\lambda I-T)x,z\rangle.$$

In other words, p_n converges weakly to $(\lambda I-T)x$. But we also know that p_n converges strongly to p, so obtain $p=(\lambda I-T)x$, which contradicts the choice of p. We thus deduce that (x_n) is unbounded. Replacing (x_n) by a subsequence if necessary, we may assume that $\|x_n\|\to\infty$.

Let $z_n=\|x_n\|^{-1}x_n$. We then have $\|z_n\|=1$ and $\|(\lambda I-T)z_n\|=\|x_n\|^{-1}\|p_n\|\to 0$. Hence $\lambda\in\sigma_{ap}(T)$. □

We also recall from B4.1:

Lemma 5.3. *(Neumann-series) Let X be a Banach space (over $\mathbb{R}$ or $\mathbb{C}$) and let $S\in\mathscr{B}(X)$ be so that $\|S\|<1$. Then $I-S$ is invertible with $(I-S)^{-1}=\sum_{n=0}^{\infty}S^n$.*

From this we can immediately deduce

Corollary 5.4.

$$\sigma(T)\subset\bar{\mathbb{D}}_{\|T\|}=\{\mu\in\mathbb{C}:|\mu|\leq\|T\|\}$$

Proof of Corollary 5.4. If $|\lambda|>\|T\|$ then $\|\lambda^{-1}T\|<1$ so $T-\lambda\mathrm{I}=-\lambda(\mathrm{I}-\lambda^{-1}T)$ is invertible by the above lemma on Neumann-series. □

5.2 Examples

Example 1. (Operators on finite dimensional space) If X is finite dimensional then you know from Linear Algebra (and the rank-nullity theorem) that a linear map $L:X\to X$ is injective if and only if it is surjective. As all linear operators on finite dimensional spaces are continuous we thus have that in this case $\sigma(T)=\sigma_P(T)$.

Example 2. (An operator on ℓ^∞ for which $\sigma_P \neq \sigma_{AP}$) Consider the operator $T \in \mathscr{B}(\ell^\infty)$ defined by $T(x) = (\frac{x_j}{j})_{j\in\mathbb{N}}$. Then each $\lambda = \frac{1}{j}$ is an eigenvalue as we have $T(e^{(j)}) = \frac{1}{j}e^{(j)}$ for $e^{(j)} = (\delta_{jk})_{k\in\mathbb{N}}$. While $\lambda = 0$ is not an eigenvalue it is in the approximate point spectrum as e.g. $e^{(k)}$ gives a sequence in ℓ^∞ with $\|e^{(k)}\|_\infty = 1$ and $\|T(e^{(k)})\| \to 0$.

Example 3. (An Integral operator) Let $X = C([0,1])$, as always equipped with the sup norm, and consider $T \in \mathscr{B}(X)$ defined by $Tx(t) := \int_0^t x(s)ds$.

Claim: $\sigma(T) = \{0\}$ while $\sigma_P(T) = \emptyset$

Proof: We first show that $\lambda = 0 \in \sigma(T) \setminus \sigma_P(T)$. Indeed differentiating the equation $Tx = 0$ (which is allowed as $Tx \in C^1$ for $x \in X$) we immediately get that $x(t) = 0$ for every t and hence that $\lambda = 0$ is not an eigenvalue. On the other hand for any $x \in X$ we have that $Tx(t = 0) = 0$ so $TX \subset \{x \in X : x(0) = 0\} \neq X$. So T is not surjective and thus $0 \in \sigma(T)$. Indeed $0 \in \sigma_r(T)$ as $\{x \in X : x(0) = 0\}$ is a closed proper subspace of X and hence TX cannot be dense in X.

Let now $\lambda \neq 0$. Then we can use that the proof of Picard's Theorem from DE1 shows that for any $y \in C([0,1])$ the integral equation $Tx - \lambda x = y$ has a unique solution $x = (T - \lambda \mathrm{I})^{-1}(y)$; here we note that for $y \in C^1$ the equation is equivalent to the initial value problem $x'(t) - \lambda^{-1}x(t) = -\lambda^{-1}y'(t)$ on $[0,1]$ with $x(0) = 0$, but that the proof of Picard from DE1 actually applies to give the existence of a unique solution of the integral equation also just for y continuous.

Furthermore the fact that this solution depends continuously on y can e.g. be obtained from Gronvall's lemma. Hence λ is not in the spectrum.

An alternative proof that $\sigma(T) = \{0\}$, based on the general properties of the spectrum of compact operators that we prove in a later section can also be given.

Example 4.(Left shift operator on ℓ^1) Consider $T : \ell^1 \to \ell^1$ defined by $T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots)$. We first determine the eigenvalues, i.e. the point spectrum. So suppose that $\lambda \in \mathbb{C}$ is so that for some $x \in \ell^1 \setminus \{0\}$ we have $Tx = \lambda x$. Then $T^j x = \lambda^j x$ so $x_{j+1} = (T^j(x))_1 = \lambda^j x_1$. Hence $x_1 \neq 0$ (as $x \neq 0$) and $x = x_1(1, \lambda, \lambda^2, \dots)$ which satisfies $Tx = \lambda x$ for all values of λ, but is only an element of ℓ^1 if $|\lambda| < 1$. We hence conclude that $\sigma_P(T)$ is the open unit disc $B_1(0) \subset \mathbb{C}$.

We may now check that every point in the closed unit disc $\bar{\mathbb{D}}_1 := \{\lambda : |\lambda| \leq 1\}$ is an approximate eigenvalue and hence that indeed

$$\sigma(T) = \sigma_{AP}(T) = \bar{\mathbb{D}}_1$$

as

$$\bar{\mathbb{D}}_1 \subset \sigma_{AP}(T) \subset \sigma(T) \subset \bar{\mathbb{D}}_{\|T\|} = \bar{\mathbb{D}}_1$$

ensures that all these sets agree.

5.3 The spectrum of bounded linear operators on Banach spaces

Our first main result about the spectrum of bounded linear operators on general complex Banach spaces is

Theorem 5.5 (Properties of the spectrum of bounded linear operators on Banach spaces)**.** *Let $(X, \|\cdot\|)$ be a complex Banach space. Then for any $T \in \mathscr{B}(X)$ we have*

(i) The resolvent set $\rho(T)$ is open and the map

$$\rho(T) \ni \lambda \mapsto R_\lambda(T)$$

is analytic, i.e. for any $\lambda_0 \in \rho(T)$ there exists a neighbourhood U of λ_0 and 'coefficients' $A_j(\lambda_0, T) \in \mathscr{B}(X)$ so that for every $\lambda \in U$ the resolvent operator is given by the convergent power series

$$R_\lambda(T) = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j A_j(\lambda_0, T).$$

(ii) The spectrum $\sigma(T)$ is non-empty, compact and contained in $\bar{\mathbb{D}}_{\|T\|}$.

One of the most important aspects of the above theorem is the last part, i.e. that every bounded operator has non-empty spectrum. Here we crucially use that the vector space is over $\mathbb{C}$. The claim is not true if we were to only consider the real spectrum as you already know from Linear Algebra.

Proof of Theorem 5.5. Let λ_0 be any element of the resolvent set, i.e. so that $(T - \lambda_0 \mathrm{I})$ is invertible, and denote by $R_{\lambda_0}(T)$ its continuous inverse. Then for any $S \in \mathscr{B}(X)$ with $\|S\| < \delta := \|R_{\lambda_0}(T)\|^{-1}$ we can write $T - \lambda_0 \mathrm{I} - S = (T - \lambda_0 \mathrm{I}) \cdot (\mathrm{I} - R_{\lambda_0}(T) S)$ where we note that $\|R_{\lambda_0}(T) S\| < \|R_{\lambda_0}(T)\| \delta = 1$ and hence that $(\mathrm{I} - R_{\lambda_0}(T) S)$ is invertible and its inverse is given by the corresponding Neumann-series. As $\lambda_0 \in \rho(T)$ also $(T - \lambda_0 \mathrm{I})$ is invertible so we find that $T - \lambda_0 \mathrm{I} - S$ is invertible for any $\|S\| < \delta$ with

$$\begin{aligned}(T - \lambda_0 \mathrm{I} + S)^{-1} &= \big((T - \lambda_0 \mathrm{I}) \cdot (\mathrm{I} - R_{\lambda_0}(T) S)\big)^{-1} = (\mathrm{I} - R_{\lambda_0}(T) S)^{-1} R_{\lambda_0}(T) \\ &= \sum_{j=0}^{\infty} (R_{\lambda_0}(T) S)^j R_{\lambda_0}(T).\end{aligned}$$

Given any $\lambda \in \mathbb{C}$ with $|\lambda - \lambda_0| < \delta$, we may apply this argument to $S = (\lambda - \lambda_0)\mathrm{I}$, which has $\|S\| = |\lambda - \lambda_0|$ to obtain that $T - \lambda \mathrm{I} = T - \lambda_0 \mathrm{I} - S$ is invertible with inverse

$$R_\lambda(T) = (T - \lambda \mathrm{I})^{-1} = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j R_{\lambda_0}(T)^{j+1}. \tag{5.1}$$

Hence any such $\lambda \in B_\delta(\lambda_0) \subset \mathbb{C}$ is in the resolvent set, so the resolvent set is open and the resolvent operator is analytic in λ.

To prove (ii) we first note that (i) implies that the spectrum $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is closed. We have already seen that it is a subset of $\bar{\mathbb{D}}_{\|T\|}$ so it is also bounded and as bounded closed subsets of finite dimensional spaces are compact hence compact.

We furthermore note that for $\lambda > \|T\|$

$$\|R_\lambda(T)\| \leq |\lambda|^{-1} \|(\mathrm{I} - \tfrac{1}{\lambda} T)^{-1}\| \leq |\lambda|^{-1} \sum_{j=0}^{\infty} \|\tfrac{T}{\lambda}\|^j \leq |\lambda|^{-1} \sum_{j=0}^{\infty} (\tfrac{\|T\|}{\lambda})^j = \frac{1}{|\lambda| - \|T\|}. \tag{5.2}$$

We now want to combine this with the Theorem of Hahn-Banach (applied to functionals on $\mathscr{B}(X)$, i.e elements of $(\mathscr{B}(X))^*$, instead of X^*) from B4.1 and Liouville's Theorem from A.1 Complex Analysis which we recall says that the only holomorphic maps $g : \mathbb{C} \to \mathbb{C}$ which are bounded are the constant maps.

So suppose that $\sigma(T)$ is empty. Then the resolvent operator R_λ is defined on all of $\mathbb{C}$ so given any $f \in (\mathscr{B}(X))^*$ we can define a function $g_f : \mathbb{C} \to \mathbb{C}$ by

$$g_f(\lambda) := f(R_\lambda(T)).$$

We note that this function is not only well defined, but furthermore that for any $\lambda_0 \in \mathbb{C}$ the function g_f is analytic in a neighbourhood of λ_0, namely

$$g_f(\lambda) = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j f(R_{\lambda_0}(T)^{j+1}) \tag{5.3}$$

in the neighbourhood of λ_0 where the expansion (5.1) converges. In particular, g_f is holomorphic.

We now claim that g_f is also bounded. To see this we first note that as g_f is continuous, it is bounded on any compact set, in particular on the closed disc $\bar{D}_{2\|T\|}$. On the other hand, for any $\lambda \in \mathbb{C}$ with $|\lambda| \geq 2\|T\|$ we know from (5.2) that $\|R_\lambda(T)\| \leq \frac{1}{|\lambda| - \|T\|} \leq \frac{1}{\|T\|}$ and hence

$$|g_f(\lambda)| \leq \|f\|_{(\mathscr{B}(X))^*} \|R_\lambda(T)\| \leq \|f\|_{(\mathscr{B}(X))^*} \|T\|^{-1}$$

so that g_f is also bounded on $(\bar{D}_{2\|T\|})^c$.

From the Theorem of Liouville we thus obtain that g_f must be constant, $g_f(\lambda) = C_f$ for a constant that depends only on the element $f \in (\mathscr{B}(X))^*$ used in the definition of g_f. Returning to the expansion (5.3) we thus conclude that all terms with $j \geq 1$ must be zero, i.e. that for any number $\lambda_0 \in \mathbb{C}$ and any $k \geq 2$ we have that

$$f(R_{\lambda_0}(T)^k) = 0 \text{ for every } f \in (\mathscr{B}(X))^*.$$

But by the Corollary of the theorem of Hahn-Banach that we recalled in (1.3), this implies that all the operators $R_{\lambda_0}(T)^k$, $k \geq 2$, must be zero, which is of course wrong since all of these operators are powers of invertible operators and thus invertible. □

We note that it is not only true that if S is invertible then also S^2 is invertible as we used in the above proof, but that also the converse of this holds and that we indeed have the following useful lemma

Lemma 5.6. *Let $(X, \|\cdot\|)$ be a normed space, $S, T \in \mathscr{B}(X)$. Suppose that $ST = TS$ and that ST is invertible. Then also S and T are invertible.*

Proof of Lemma 5.6. By symmetry it suffices to prove the claim for T and we shall prove this by an argument by contradiction. So suppose that the claim is false. Then we either have that T is not surjective, which is impossible as in this case we would have that $ST(X) = TS(X) = T(SX) \subset TX \subsetneq X$ so ST would not be surjective, or there exists no $\delta > 0$ so that (1.9) holds. In this case we can choose $x_n \in X \setminus \{0\}$ so that $\frac{\|Tx_n\|}{\|x_n\|} \to 0$ and thus conclude that also

$$\frac{\|STx_n\|}{\|x_n\|} \leq \|S\|_{\mathscr{B}(X)} \frac{\|Tx_n\|}{\|x_n\|} \to 0,$$

which means that (1.9) does not hold true for ST, and hence that ST does not have a bounded inverse. $\square$

We note that the above proof works for all normed spaces and could be shortened significantly for Banach spaces by using the inverse mapping theorem.

Based on this lemma we can now prove the following useful result.

Theorem 5.7. *Let X be a complex Banach space, $T \in \mathscr{B}(X)$ and let p be a complex polynomial. Then*

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

Here we set $p(T) := \sum_{j=0}^{n} a_j T^j$ if the polynomial p is given by $p(z) = \sum_{j=0}^{n} a_j z^j$, with the usual convention that $T^0 = \mathrm{I}$.

Proof of Theorem 5.7. We first remark that if p is constant, say $p = c \in \mathbb{R}$, then the spectrum of $p(T) = c\mathrm{I}$ is simply $\{c\}$ while the fact that $\sigma(T)$ is non-empty implies that also $p(\sigma(T)) = \{c\}$. So suppose that p has degree $n \geq 1$ and let $\mu \in \mathbb{C}$ be any given number. As we are working in $\mathbb{C}$ we can factorise $p(\cdot) - \mu$ and write it as $p(z) - \mu = \alpha(z - \beta_1(\mu)) \dots (z - \beta_n(\mu))$ for some $\alpha \neq 0$ and equally factorise

$$p(T) - \mu\mathrm{I} = \alpha(T - \beta_1(\mu)\mathrm{I}) \dots (T - \beta_n(\mu)\mathrm{I}) \tag{5.4}$$

where we note that all operators on the right hand side commute which will allow us to apply Lemma 5.6.

We now note that since the zeros $\beta_j(\mu)$ of $p(\cdot) - \mu$ can be equivalently characterised as the solutions $t = \beta_j(\mu)$ of the equation $p(t) = \mu$ we have that

$$\mu \in p(\sigma(T)) \Leftrightarrow \exists j \text{ so that } \beta_j(\mu) \in \sigma(T).$$

We then note that, applying Lemma 5.6 to (5.4) yields that if $\beta_j(\mu) \in \sigma(T)$ then $p(T) - \mu\mathrm{I}$ cannot be invertible, i.e. μ must be an element of $\sigma(p(T))$. Hence $p(\sigma(T)) \subset \sigma(p(T))$. We now prove that also $p(\sigma(T))^c \subset \sigma(p(T))^c$ and hence that $p(\sigma(T)) \subset \sigma(p(T))$. To see this we note that if $\mu \notin p(\sigma(T))$ then $\beta_j(\mu) \notin \sigma(T)$ so $T - \beta_j(\mu)\mathrm{I}$ is invertible for all $j = 1, \dots, n$. But then (5.4) shows that $p(T) - \mu\mathrm{I}$ is the composition of invertible operators so invertible and thus $\mu \in \rho(T) = \sigma(p(T))^c$. $\square$

We also remark that a similar argument shows that for T invertible λ is an eigenvalue of T if and only if λ^{-1} is an eigenvalue of T^{-1}.

As an immediate consequence of the above lemma we also obtain

Corollary 5.8. *Let $(X, \|\cdot\|)$ be a Banach space, $T \in \mathscr{B}(X)$. Then for any $k \in \mathbb{N}$ and any $\lambda \in \sigma(T)$ we have $\lambda^k \in \sigma(T^k)$. In particular*

$$|\lambda| \leq \inf_j \|T^j\|^{1/j} \text{ for all } \lambda \in \sigma(T), \quad j \in \mathbb{N}.$$

Definition 5.9. *The spectral radius of an operator $T \in \mathscr{B}(X)$ is defined as*

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

From the above we know that $r_\sigma(T) \leq \inf_j \|T^j\|^{1/j}$. Indeed, one can show that $\|T^j\|^{1/j}$ converges as $j \to \infty$ with $\lim_{j\to\infty} \|T^j\|^{1/j} = \inf_{j\in\mathbb{N}} \|T^j\|^{1/j}$ and that

Theorem 5.10 (Gelfand's formula). *Let $(X, \|\cdot\|)$ be a complex Banach space. Then for any $T \in \mathscr{B}(X)$ we have*

$$r(T) = \lim_{j\to\infty} \|T^j\|^{1/j} = \inf_{j\in\mathbb{N}} \|T^j\|^{1/j}.$$

We will use this result without proof.

We recall that in the special case where X is a Hilbert space and where T is selfadjoint we have $\|T^2\| = \|T\|^2$ and hence by iteration $\|T^{2^j}\| = \|T\|^{2^j}$. In this case we hence obtain

Corollary 5.11. *If X is a $\mathbb{C}$-Hilbert space and $T \in \mathscr{B}(X)$ is selfadjoint then*

$$r_\sigma(T) = \|T\|.$$

Furthermore, we have the following close connection between the spectrum of an operator and the spectrum of its dual operator $T' \in \mathscr{B}(X^*)$:

Theorem 5.12. *Let $(X, \|\cdot\|)$ be a Banach space, let $T \in \mathscr{B}(X)$ and let $T' \in \mathscr{B}(X^*)$ be the corresponding dual operator defined by $(T'f)(x) = f(Tx)$. Then*

$$\sigma(T) = \sigma_{AP}(T) \cup \sigma_P(T').$$

Proof of Theorem 5.12. By definition $\sigma_{AP}(T) \subset \sigma(T)$, so it is enough to prove
Claim 1: $\sigma_P(T') \subset \sigma(T)$
and
Claim 2: $\sigma(T) \setminus \sigma_{AP}(T) \subset \sigma_P(T')$

Proof of Claim 1: Let $\lambda \in \sigma_P(T')$. Then there exists $f \in X^*$ with $f \neq 0$ so that $T'f = \lambda f$, i.e. so that for every $x \in X$

$$0 = (T'f - \lambda f)(x) = f(Tx) - \lambda f(x) = f(Tx - \lambda x).$$

Hence the restriction $f|_Y$ of f to the image $Y = (T - \lambda \mathrm{I})X$ of $T - \lambda\mathrm{I}$ is zero, so as f is not the zero element, we must have that $\bar{Y} \neq X$. Thus the image of $T - \lambda\mathrm{I}$ is not dense in X so λ is either in the point spectrum (if it's also not injective) or else in the residual spectrum, and in any case $\lambda \in \sigma(T)$.

Proof of Claim 2: Let $\lambda \in \sigma(T) \setminus \sigma_{AP}(T)$. Then as λ is not an approximate eigenvalue of T we know that there exists some $\delta > 0$ so that $\|Tx - \lambda x\| \geq \delta\|x\|$ for all $x \in X$ which, thanks to Lemma 1.18, implies that the image $Y = (T - \lambda\mathrm{I})(X)$ is closed. At the same time Y cannot be all of X as otherwise $T - \lambda\mathrm{I}$ would have a bounded inverse, so Y is a proper closed subspace of X. We can thus apply the consequence (1.3) of Hahn-Banach, to conclude that there exists some $f \in X^*$ with $f \neq 0$, so that $f|_Y = 0$. This implies that $T'(f) = \lambda f$ and thus that $\lambda \in \sigma_P(T')$ since for every $x \in X$ we have $Tx - \lambda x \in Y$ and thus $(T'(f) - \lambda f)(x) = f(Tx - \lambda x) = 0$.

$\square$

5.4 Spectral theory on Hilbert spaces

In the rest of the chapter, we will specialize to the case where X is a Hilbert space (over $\mathbb{C}$). Note that in this case, the notions of dual operator and adjoint operator can be linked via the Riesz representation theorem and we can hence immediately see that

$$\lambda \in \sigma(T') \Leftrightarrow \bar{\lambda} \in \sigma(T^*).$$

From Theorem 5.12 we hence deduce

Theorem 5.13. *Let X be a complex Hilbert space and let $T \in \mathscr{B}(X)$. Then*

$$\sigma(T) = \sigma_{ap}(T) \cup \sigma_p'(T^*)$$

where $\sigma_p'(T^*) = \{\lambda : \bar{\lambda} \in \sigma_p(T^*)\}$.

This can also be proven directly based on the basic properties of the adjoint that we recalled in the first chapter which, when applied to $(\lambda I - T)^* = \bar{\lambda} I - T^*$, yield that

- $\lambda I - T$ is invertible if and only if $\bar{\lambda} I - T^*$ is invertible, so $\lambda \in \sigma(T)$ if and only if $\bar{\lambda} \in \sigma(T^*)$.
- $\mathrm{Ker}(\lambda I - T) = \mathrm{Im}(\bar{\lambda} I - T^*)^\perp$ and $\mathrm{Ker}(\lambda I - T)^\perp = \overline{\mathrm{Im}(\bar{\lambda} I - T^*)}$.

For selfadjoint operators we already know that $r_\sigma(T) = \|T\|$ and we can indeed prove far more than that:

Theorem 5.14. *Let X be a complex Hilbert space and let $T \in \mathscr{B}(X)$ be self-adjoint. Then*

(i) $\sigma(T) = \sigma_{ap}(T) \subset [a, b] \subset \mathbb{R}$, for

$$a = \inf_{\|x\|=1} \langle x, Tx \rangle \text{ and } b = \sup_{\|x\|=1} \langle x, Tx \rangle$$

and both a and b are in the spectrum.

(ii) T has no residual spectrum, i.e. $\sigma(T) = \sigma_p(T) \cup \sigma_c(T)$.

(iii) Eigenvectors corresponding to different eigenvalues of T are orthogonal.

We note that as T is selfadjoint we have

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

so $\langle Tx, x \rangle$ is real for any $x \in X$ and recall that $\|T\| = \sup_{\|x\|=1} |\langle Tx, x \rangle|$, compare Lemma 1.6.

Proof of Theorem 5.14. We first prove that $\sigma_{ap}(T) \subset [a, b]$.

So let $\lambda \in \sigma_{ap}(T)$. Then there exist x_n with $\|x_n\| = 1$ so that $Tx_n - \lambda x_n \to 0$. Thanks to Cauchy-Schwarz this implies that

$$|\langle Tx_n, x_n \rangle - \lambda| \leq \|Tx_n - \lambda x_n\| \to 0.$$

Hence $\lambda = \lim\langle Tx_n, x_n\rangle \in [a,b] \subset \mathbb{R}$ since we know that $\langle Tx, x\rangle \in [a,b] \subset \mathbb{R}$.

This in particular implies that all eigenvalues of $T = T^*$ are real and hence that $\sigma_P'(T^*) = \sigma_P'(T) = \sigma_P(T)$ so Theorem 5.13 implies that indeed $\sigma(T) = \sigma_{ap}(T) \cup \sigma_P'(T^*) = \sigma_{ap}(T)$ and hence also that the every element of the spectrum is a real number that is contained in $[a,b]$.

We next show that the endpoints of this interval are indeed both in $\sigma(T)$. By definition of a, b, we have $|a| \leq \|T\|$ and $|b| \leq \|T\|$. But as $\sigma(T) \subset [a,b]$ and as $r_\sigma(T) = \|T\|$ since T is selfadjoint, we have also $\|T\| = r_\sigma(T) \leq \max(|a|,|b|)$. Hence at least one of a and b belongs to $\sigma(T)$. Now note that, if c is a real constant, then the spectrum of $cI + T$ is shifted by c and the "a" and "b" of $cI + T$ are also shifted by c. Applying what we just established to $cI + T$ for suitable c, we conclude that both a and b belong to $\sigma(T)$.

This establishes the first statement (i) of the Theorem.

To see that the residual spectrum of T is empty we use that

$$\overline{(T - \lambda \mathrm{I})(X)} = (\ker(T^* - \bar{\lambda}\mathrm{I}))^\perp = (\ker(T - \lambda\mathrm{I}))^\perp$$

where the first equality follows from (1.5) while the second follows as $T^* = T$ and as the spectrum is real.

Hence, if λ is so that the image of $(T - \lambda\mathrm{I})$ is not dense in X then we must have that $\ker(T - \lambda\mathrm{I}) \neq \{0\}$, i.e. that λ will be in the point spectrum rather than the residual spectrum.

The final part of the claim follows by direct computation, exactly as in the case of finite dimensional symmetric matrices discussed in prelims Linear Algebra.

□

Alternatively, we can show that $\sigma(T) \subset [a,b]$ as follows: It suffices to show that if λ is a real number such that $\lambda > b$ then $\lambda I - T$ is invertible. (A similar argument apply to $\lambda < a$.) We have

$$\langle x, (\lambda I - T)x\rangle = \lambda\|x\|^2 - \langle x, Tx\rangle \geq (\lambda - b)\|x\|^2.$$

It thus follows that $\langle x, y\rangle_\lambda := \langle x, (\lambda I - T)y\rangle$ defines a scalar product on X and its associated norm $\|x\|_\lambda := \langle x, (\lambda I - T)x\rangle^{1/2}$ is equivalent to $\|\cdot\|$.

For every $z \in X$, consider the linear functional

$$\ell_z(x) = \langle x, z\rangle.$$

By the Riesz representation theorem, there is some y depending on z such that

$$\ell_z(x) = \langle x, y\rangle_\lambda \text{ i.e. } \langle x, z\rangle = \langle x, (\lambda I - T)y\rangle \text{ for every } x.$$

It thus follows that $\lambda I - T$ is surjective. Since $\lambda I - T$ is self-adjoint, this implies that $\lambda I - T$ is also injective, and hence invertible.

We conclude the section with a result on spectra of unitary operators.

Proposition 5.15. *Let X be a complex Hilbert space and let $U \in \mathscr{B}(X)$ be unitary. Then $|\lambda| = 1$ for all $\lambda \in \sigma(U)$.*

Proof. By Proposition 1.4, U is a surjective isometry and $U^{-1} = U^*$.

As $\|U\| = 1$ it follows that $|\lambda| \leq \|U\| = 1$ for all $\lambda \in \sigma(U)$.

Assume by contradiction that there is some λ with $|\lambda| < 1$ such that $\lambda I - U$ is not invertible. It follows that $\bar{\lambda} I - U^*$ is also not invertible. Consequently, $\bar{\lambda} U - I = (\bar{\lambda} I - U^*)U$ is also not invertible (since U is invertible), and so $\bar{\lambda}^{-1} \in \sigma(U)$. This amounts to a contradiction as $|\bar{\lambda}^{-1}| > 1$. □

5.5 Spectrum of compact operators

We recall that if $T \in \mathscr{B}(X)$ is compact then $\ker(\mathrm{I} - T)$ is finite dimensional and the image of $\mathrm{I} - T$ is closed, compare Proposition 1.17. Additionally we have:

Lemma 5.16. *Let X be a Banach space, $T \in \mathscr{B}(X)$ compact and suppose that $S := I - T$ is injective. Then S is also surjective.*

Proof of Lemma 5.16. We argue by contradiction so assume that $X_1 := S(X)$ is a proper subspace of X. We set $X_k := S^k(X)$, note that these spaces are nested, $X_{k+1} \subset X_k \subset ... \subset X_1 \subsetneq X$ and we claim that these inclusions must all be strict.

Indeed, as $X_1 \neq X$ we can fix some $x \in X \setminus X_1$ and get that $S^n x$ is obviously in X_n, but cannot be in X_{n+1} as otherwise there would be $y \in X$ with $S^n x = S^{n+1} y$, i.e. with $S^n(x - Sy) = 0$ which would be a contradiction to $x \notin SX$ since S^n is injective (since S is injective).

As sums and powers of compact operators are compact we can write $S^n = (\mathrm{I} - T)^n$ also in the form $\mathrm{I} - T_n$ for a compact operator T_n so know from Proposition 1.17 that the spaces X_n are all closed.

By the Riesz-lemma from B4.1 we can hence pick $x_k \in X_k$ with $\|x_k\| = 1$ and $\mathrm{dist}(x_k, X_{k+1}) \geq \frac{1}{2}$ and we want to argue that this bounded sequence does not have a subsequence for which Tx_{n_j} converges.

To see this we note that since $(T - \mathrm{I})(x_k) \in X_{k+1} = (\mathrm{I} - T)(X_k)$ we can write $Tx_k = x_k + (T - \mathrm{I})(x_k)$ to see that

$$\mathrm{dist}(Tx_k, X_{k+1}) = \mathrm{dist}(x_k, X_{k+1}) \geq \frac{1}{2}.$$

On the other hand, we know that for every l we have $Tx_l = x_l + (T - \mathrm{I})(x_l) \in X_l + X_{l+1} = X_l$.

Thus, as the sets are nested we have $Tx_l \in X_{k+1}$ whenever $l \geq k+1$ and thus $\|Tx_l - Tx_k\| \geq \mathrm{dist}(Tx_k, X_{k+1}) \geq \frac{1}{2}$. This means that (Tx_k) cannot have any subsequence that is Cauchy and hence cannot have any subsequence that converges. □

Based on this we can now prove

Theorem 5.17. *(Spectral theorem for compact (selfadjoint) operators) Let X be an infinite dimensional Banach space and let T be a compact selfadjoint operator. Then*

(i) $\sigma(T) \setminus \{0\} = \sigma_P(T) \setminus \{0\}$, *the eigenspace* $\ker(T - \lambda I)$ *is finite dimensional for every $\lambda \in \sigma_p(T) \setminus \{0\}$ and $\lambda = 0$ is in the spectrum.*

(ii) For every $r > 0$ there are at most finitely many eigenvalues λ with $|\lambda| \geq r$. Hence $\lambda = 0$ is the only potential accumulation point of $\sigma(T)$.

(iii) *If X is a Hilbert space and T is additionally selfadjoint then there exists a countable (either finite or countably infinite) on basis $\{e_k\}$ of* $\ker(T)^\perp$ *that consists of eigenvectors e_k to eigenvalues $\lambda_k \neq 0$, with $|\lambda_k| \to 0$ if the basis is infinite, and for each $x \in X$*

$$Tx = \sum \lambda_k \langle x, e_k \rangle e_k$$

In (iii) we can always extend the $\{e_k\}$ to an eigenbasis of the whole space by including additionally an on-basis of $\ker(T)$, though we note that this basis will only be countable if X is separable.

We also remark that an important special case of the above result is

Corollary 5.18. *Suppose that T is a selfadjoint compact operator on an infinite dimensional Hilbert space which is positive definite, i.e. so that $\langle Tx, x \rangle > 0$ for all $x \in X$. Then X must be separable and there exists a (countable) on-basis $e_1, e_2, ...$ of eigenvectors corresponding to eigenvalues $\lambda_1 \geq \lambda_2 \geq ....$ with $\lambda_k \searrow 0$ as $k \to \infty$.*

Proof of Theorem 5.17. The fact that each $\lambda \in \sigma(T) \setminus \{0\}$ must be an eigenvalue with finite dimensional eigenspace follows from Proposition 1.17 and Lemma 5.16 as we can write $(T - \lambda \mathrm{I}) = -\lambda(\mathrm{I} - \lambda^{-1}T)$ for $\lambda \neq 0$ and use that $\lambda^{-1}T$ is compact.

Furthermore, as T is compact, we cannot have that T is invertible as that would imply that the identity map $I = T^{-1} \circ T$ is also compact which would contradict the theorem of Heine Borel as our space is infinite dimensional. Hence $0 \in \sigma(T)$.

To prove (ii) we want to argue by contradiction and suppose that there is some $r > 0$ so that there infinitely many distinct $\lambda_k \in \sigma(T)$ with $|\lambda| \geq r$. By (i) all these λ_k are eigenvalues.

We can use a very similar argument as carried out in the proof of Lemma 5.16 to get a contradiction, now using spaces $X_k = E_{\lambda_1} \oplus ... \oplus E_{\lambda_k}$, E_{λ_i} the (finite dimensional!) eigenspaces to λ_i.

We note that these spaces are all closed (as finite dimensional), so that $TX_k \subset X_k$ (as X_k is built out of eigenspaces), and so that X_k is a proper subspace of X_{k+1}. Additionally we note that $(T - \lambda_k I)(X_k) \subset X_{k-1}$ as E_{λ_k} is the kernel of this map.

Similar to above we can now use the Riesz-lemma to get a sequence of unit vectors $x_k \in X_k$ with $\mathrm{dist}(x_k, X_{k-1}) \geq \frac{1}{2}$.

If $n < m$ then X_{m-1} contains both X_n, and thus $Tx_n \in T(X_n) \subset X_n$, and also $(T - \lambda_m I)(X_m)$, and thus $Tx_m - \lambda x_m$, so we get that

$$\begin{aligned} \|Tx_m - Tx_n\| &= \|\lambda_m x_m + (T - \lambda_m)x_m - Tx_n\| \geq \mathrm{dist}(\lambda_m x_m, X_{m-1}) \\ &\geq \tfrac{1}{2}|\lambda_m| \geq \tfrac{1}{2}r > 0 \end{aligned}$$

The sequence (Tx_n) hence cannot have any subsequence which is Cauchy, which leads to a contradiction as (x_n) is bounded and T is compact.

We note that if X is Hilbert and T is selfadjoint then this argument can be simplified significantly: We first pass to a subsequence to ensure that the λ_j converge, then choose for each λ_j a unit eigenvector e_j and use that as T is selfadjoint these e_j are orthonormal and hence so that $\|e_j - e_k\| = \sqrt{2}$.

At the same time, as T is compact there is a subsequence so that $Te_{j_n} = \lambda_j e_{j_n}$ converges and combined with the convergence of the $\lambda_j \to \lambda \neq 0$ this gives convergence of e_{j_n}, contradiction.

It hence remains to show (iii) and for this we inductively construct the e_k using that Corollary 5.11 (which was a consequence of Gelfand's formula) ensures that the spectral radius of any selfadjoint operator is equal to the norm of the operator.

We first set $S_1 = T$ and consider this as an operator on $X_1 = X$, apply Corollary 5.11 to find $\lambda_1 \in \sigma(S_1)$ so that $|\lambda_1| = \|S_1\|$ and let $e_1, \ldots, e_{k_1}$ be an on-basis of the (finite dimensional as S_1 compact!) space $Z_1 := \ker(S_1 - \lambda_1 \mathrm{I}_{X_1})$. We then let X_2 be the orthogonal complement of Z_1 in X_1. This is a closed subspace of a Hilbert space so again a Hilbert space. We claim that T maps X_2 into itself. Indeed, since Z_1 is an eigenspace we have $T(Z_1) \subset Z_1$ and hence get that if $z \in X_2$ then $\langle Tz, y\rangle = \langle z, Ty\rangle = 0$ for all $y \in Z_1$ so $Tz \in (Z_1)^\perp = X_2$.

We then set $S_2 := T|_{X_2} \in \mathscr{B}(X_2)$ and repeat this argument, obtaining the next eigenvalue λ_2 with an on-eigenbasis $e_{k_1+1}, \ldots, e_{k_1+k_2}$ of $Z_2 = \ker(S_2 - \lambda_2 \mathrm{I})$. We note that Z_2 is the eigenspace of T for λ_2 as $Z_2 = \ker(S_2 - \lambda_2\mathrm{I}) = \ker(T - \lambda_2 \mathrm{I}) \cap X_2 = \ker(T - \lambda_2\mathrm{I})$ since the eigenspaces are orthogonal and hence $\ker(T - \lambda_2\mathrm{I}) \subset (\ker(T - \lambda_1\mathrm{I}))^\perp = X_2$.

We continue to apply this argument to the restriction of T to the next space X_{k+1} obtained as orthogonal complement of Z_k in the Hilbert space X_k.

This process either ends with us obtaining an operator S_k which is identically zero (and hence a finite set of eigenvalues), or continues for all $k \in \mathbb{N}$ to give a countably infinite set of distinct eigenvalues with $|\lambda_k| \geq |\lambda_{k+1}| > 0$ which by (ii) must converge to zero.

To obtain the final claim we first note that the right hand side converges thanks to Bessel's inequality. We then set $K_j := k_1 + \ldots + k_j$ note that $e_1, \ldots, e_{K_j}$ is an on-basis of $Z_1 \oplus \ldots \oplus Z_j = (X_{j+1})^\perp$. Hence $P_j x := \sum_{k \leq K_j} \langle x, e_k\rangle e_k$ is simply the orthogonal projection onto $Z_1 \oplus \ldots \oplus Z_j$ while $I - P_j$ is the orthogonal projection onto X_{j+1}.

Thus

$$\begin{aligned}
\|Tx - \sum_{k \leq K_j} \lambda_k \langle x, e_k\rangle e_k\| &= \|T(x - P_j x)\| = \|T|_{X_{j+1}} \circ (I - P_j)(x)\| \\
&\leq \|T|_{X_{j+1}}\| \|I - P_j\| \|x\| \leq |\lambda_{j+1}| \cdot 1 \cdot \|x\| \to 0
\end{aligned}$$

□