BERGISCHE UNIVERSITÄT WUPPERTAL

Functional Analysis 1

Advanced Spectral Theory and Functional Calculus

Jochen Glück

Lecture Notes
Winter term 2022/23

Contents

Version

October 7, 2023

Acknowledgements

I am indebted to Dominik Briganti for providing me with a great wealth of corrections for the manuscript.

Part I

Spectral Theory in Banach Algebras

Chapter 1

Banach Algebras

1.1 Banach algebras with and without neutral element

Recall that, for vector spaces X, Y, Z over the same field, a map $b : X \times Y \to Z$ is called *bilinear* if for each $x \in X$ and each $y \in Y$ the maps

$$b(x, \cdot\,) : Y \to Z \qquad \text{and} \qquad b(\,\cdot\,, y) : X \to Z$$

are linear.

Definition 1.1.1 (Algebras and Banach Algebras)**.** (a) An *algebra* is a complex vector space A together with an associative and bilinear mapping $\cdot : A \times A \to A$.[1,2] An algebra A is called *commutative* if $ab = ba$ for all $a, b \in A$.

(b) An element e of an algebra A is called a *neutral* element of A if $ea = ae = a$ for all $a \in A$.[3,4] We will often use the notation 1 (rather than e) for a neutral element in an algebra.

(c) A *Banach algebra* is an algebra A which is also endowed with a norm that is complete[5] and *submultiplicative*, which means that $\|ab\| \le \|a\|\,\|b\|$ for all $a, b \in A$.

(d) A Banach algebra A is called a *unital Banach algebra* if there exists a neutral element $1 \in A$ which satisfies $\|1\| \le 1$.[6]

---

[1] Note that bilinearity of $\cdot$ simply means that the usual distributive laws are satisfied.

[2] We use the common convention to give the operation $\cdot$ higher priority than $+$, and for $a, b \in A$ we often write ab for $a \cdot b$.

[3] Note that there exists at most one neutral element in an algebra A: if e_1 and e_2 are neutral elements, then $e_1 = e_1 e_2 = e_2$.

[4] Also note that a neutral element e is automatically non-zero if if $A \neq \{0\}$. Indeed, let $0 \neq a \in A$. If e were zero, then $a = ea = 0$.

[5] I.e., it turns A into a Banach space – thus the name *Banach algebra*.

[6] If a Banach algebra A satisfies $A \neq \{0\}$, then a neutral element 1 in A always satisfies $\|1\| \ge 1$. This follows from $\|1\| = \|1 \cdot 1\| \le \|1\|^2$ together with $\|1\| \neq 0$. In the exercise you will see that there

One can also define Banach algebras over the real scalar field. But throughout the lecture we will focus in so-called *spectral theoretic* questions, and in order to develop spectral theory in its full strength one needs to work over the complex field. So in order to keep the situation simple, we will focus on the complex case throughout.

Example 1.1.2 (The space of bounded operators)**.** Let X be a complex Banach space. The space $\mathcal{L}(X)$ of bounded linear operators from X to X is a unital Banach algebra when endowed with the operator norm and with the composition of linear operators as multiplication. The neutral element in $\mathcal{L}(X)$ is the identity operator $1 := \mathrm{id}_X : X \to X$.[7]

It is not difficult to check that the Banach algebra $\mathcal{L}(X)$ is not commutative if $\dim X \geq 2$.

Example 1.1.3 (Spaces of continuous functions)**.** (a) Let (K, d) be a compact metric space – or, more generally, let K be a compact topological Hausdorff space[8] – and let $\mathrm{C}(K)$ denote the space of all continuous functions $K \to \mathbb{C}$. This is a unital commutative Banach algebra when endowed with pointwise addition, pointwise scalar multiplication, pointwise multiplication, and the supremum norm $\|\cdot\|_\infty$ that is given by $\|f\|_\infty := \max\{|f(x)| \mid x \in K\}$ for all $f \in \mathrm{C}(K)$. The neutral element in this Banach algebra is the constant functiont $\mathbb{1}$ with value 1.

(b) The space $\mathrm{C}_0(\mathbb{R})$ of all continuous functions $f : \mathbb{R} \to \mathbb{C}$ which satisfy the properties $\lim_{x\to\infty} f(x) = 0$ and $\lim_{x\to-\infty} f(x) = 0$ is a commutative Banach algebra when endowed the the same operations as in the previous example and the supremum norm $\|\cdot\|_\infty$ given by $\|f\|_\infty := \max\{|f(x)| \mid x \in \mathbb{R}\}$ for all $f \in \mathrm{C}_0(\mathbb{R})$.

This Banach algebra does not contain a neutral element and hence, it is not unital. Indeed, assume that $e \in \mathrm{C}_0(\mathbb{R})$ is a unit. Let $x \in \mathbb{R}$ and choose a function $f \in \mathrm{C}_0(\mathbb{R})$ such that $f(x) \neq 0$.[9] Then $f(x) = (ef)(x) = e(x)f(x)$, so $e(x) = 1$. So we showed that $e(x) = 1$ for all $x \in X$; but this contradicts the condition $\lim_{x\to\infty} e(x) = 0$ which must be true since $e \in \mathrm{C}_0(\mathbb{R})$.

Example 1.1.4 (The space $\boldsymbol{\ell^1}(\mathbb{Z})$ with convolution)**.** Consider the space $\ell^1(\mathbb{Z})$ of all complex sequences $f = (f_n)_{n\in\mathbb{Z}}$ that satisfy $\|f\|_1 := \sum_{n\in\mathbb{Z}} |f_n| < \infty$. Recall that $\ell^1(\mathbb{Z})$ is a Banach space with respect to the norm $\|\cdot\|_1$. Now we define the so-called *convolution* $\star : \ell^1(\mathbb{Z}) \times \ell^1(\mathbb{Z}) \to \ell^1(\mathbb{Z})$ by

$$(f \star g)_n := \sum_{k\in\mathbb{Z}} f_{n-k} g_k \qquad \text{for all } n \in \mathbb{Z}$$

exist examples of Banach algebras which have a neutral element 1 such that $\|1\| \neq 1$. Compare, however, Proposition 2.1.1.

So if A is a unital Banach algebra, then either $A = \{0\}$ or $\|1\| = 1$.

[7] Note that $\mathrm{id}_X = 0$ if $X = \{0\}$.

[8] In case that you do not know, yet, what this is, just think about a compact metric space for now; we will give a brief introduction to topological spaces later in this course.

[9] Why does such a function f exist?

for all $f, g \in \ell^1(\mathbb{Z})$.

To show that this is well-defined note that, for all $f, g \in \ell^1(\mathbb{Z})$, one has

$$\sum_{n\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} |f_{n-k}g_k| = \sum_{k\in\mathbb{Z}} |g_k| \sum_{n\in\mathbb{Z}} |f_{n-k}| = \sum_{k\in\mathbb{Z}} |g_k| \sum_{n\in\mathbb{Z}} |f_n| = \|g\|_1 \|f\|_1 < \infty.$$

This implies two things: first, the series $\sum_{k\in\mathbb{Z}} f_{n-k}g_k$ is absolutely convergence for each $n \in \mathbb{Z}$, so $f \star g$ is a well-defined sequence; and second, the sequence $f \star g$ satisfies

$$\|f \star g\|_1 = \sum_{n\in\mathbb{Z}} |(f \star g)_n| \leq \|g\|_1 \|f\|_1 < \infty,$$

so $f \star g$ is indeed an element of $\ell^1(\mathbb{Z})$. Hence, $\star$ is indeed a well-defined mapping from $\ell^1(\mathbb{Z}) \times \ell^1(\mathbb{Z})$ to $\ell^1(\mathbb{Z})$.

One can check[10] that $\star$ is associative, bilinear and commutative, so – as we have shown above that $\|f \star g\|_1 \leq \|f\|_1 \|f\|_2$ for all $f, g \in \ell^1(\mathbb{Z})$ – $\ell^1(\star)$ is a commutative Banach algebra with respect to the multiplication $\star$.

Finally, the 0-th canonical unit vector[11] $\mathrm{e}^{(0)}$ is a neutral element in $\ell^1(\mathbb{Z})$ since

$$(f \star \mathrm{e}^{(0)})_n = \sum_{k\in\mathbb{Z}} f_{n-k}\, \mathrm{e}^{(0)}_k = f_n$$

for all $f \in \ell^1(\mathbb{Z})$ and all $n \in \mathbb{Z}$. As $\left\|\mathrm{e}^{(0)}\right\|_1 = 1$, the Banach algebra $\ell^1(\mathbb{Z})$ is unital.

Recall that, if metric spaces (M_k, d_k) are metric spaces for each $k \in \{1, 2, 3\}$, then a mapping $\varphi : M_1 \times M_2 \to M_2$ is continuous if and only if $\varphi(x_n, y_n) \to \varphi(\lim_n x_n, \lim_n y_n)$ for all convergent sequences (x_n) in M_1 and (y_n) in M_2.

Proposition 1.1.5 (Joint continuity of multiplication)**.** *Let A be a Banach algebra. Then the multiplication $A \times A \to A$ is continuous.*

Proof. Let (a_n) and (b_n) be sequences in A that converge to element A and B, respectively. Then there exists a number $M \geq 0$ such that $\|a_n\| \leq M$ for all $n \in \mathbb{N}$, so one has

$$\|a_nb_n - ab\| \leq \|a_nb_n - a_nb\| + \|a_nb - ab\| \leq M\,\|b_n - b\| + \|a_n - a\|\,\|b\| \to 0$$

as $n \to \infty$, which shows that the sequence (a_nb_n) converges to ab. □

When introducing new structures, it is common to also define appropriate "substructures" – for instance, one consider vector subspaces of vector spaces, subgroups of groups, subfields of fields, and so on. We do the same now for (Banach) algebras.

Definition 1.1.6 (Subalgebras)**.** Let A be an algebra. A *subalgebra* of A is a vector subspace B of A such that $bc \in B$ for all $b, c \in B$.

[10] This is part of Exercise Sheet 1.

[11] I.e., the sequence which is 1 at position 1 and 0 elsewhere.

Obviously, a subalgebra of an algebra A is an algebra in its own right. In a Banach algebra A, every closed subalgebra B of A is itself a Banach algebra. Moreover, the closure of a subalgebra of A is again a subalgebra of A.[12]

The intersection of arbtitrarily many subalgebras is clearly a subalgbera, and the intersection of arbtitrarily many closed subalgebras of a Banach algebra is clearly a closed subalgebra. From this, one can immediately derive the following proposition:

Proposition 1.1.7 (Algebras generated by subsets)**.** (a) *Let A be an algebra and $M \subseteq A$. Then the intersection of all subalgebras of A that contain M is the smallest subalgebra of A that contains M.*

It is called the subalgebra generated by M within A.

(b) *Let A be a Banach algebra and let $M \subseteq A$. Then the intersection of all closed subalgebras of A that contain M is the smallest closed subalgebra of A the contains M.*[13]

It is called the Banach algebra generated by M within A.

In the continuous $\mathbb{C}$-valued functions over a compact space, there is a a very useufal sufficient criterion for a subalgebra to be dense. We recall this criterion in the following theorem. The proof is not particularly difficult, but rather topological in nature; hence, it is set a bit apart from the main thrust of this course, so we refer to the literature, for instance to [Ped89, Theorem 4.3.4], for the proof.

Theorem 1.1.8 (Stone–Weierstrass approximation theorem)**.** *Let (K, d) be a compact metric space (or, more generally, let K be a compact topological Hausdorff space) and let A be a subalgebra of $\mathrm{C}(K)$ with the following two properties:*

(1) *The constant function $\mathbb{1}$ is an element of A.*

(2) *The set A is* invariant under complex conjugation*, i.e., for every $f \in A$ the complex conjugate function*[14] *$\overline{f}$ is also in A.*

(3) *The set A* separates the points of K*, i.e., for any pair of distinct elements $x_1, x_2 \in K$ there exists $f \in A$ such that $f(x_1) \neq f(x_2)$.*

Then A is dense in $\mathrm{C}(K)$.

The following simple example demonstrates how the theorem can be applied:

Example 1.1.9 (Polynomials in $\boldsymbol{z}$ and $\overline{\boldsymbol{z}}$ on the unit circle)**.** Let $\mathbb{T}$ denote the complex unit cirlce, i.e.

$$\mathbb{T} := \{z \in \mathbb{C} \mid \ |z| = 1\}.$$

[12] Why?

[13] And one easily checks it coincides with the closure of the subalgebra generated by M.

[14] The complex conjugate function $\overline{f}$ of f is simply defined by $\overline{f}(x) := \overline{f(x)}$ for all $x \in K$.

Let A denote the set of all polyomials in both z and $\overline{z}$ on $\mathbb{T}$, i.e. the set of all functions $f : \mathbb{T} \to \mathbb{C}$ for which there exists a integer $n \geq 0$ and coefficients $c_{jk} \in \mathbb{C}$ for all integers $0 \leq j \leq k \leq n$ such that

$$f(z) = \sum_{k=0}^{n} \sum_{j=0}^{k} c_{jk} z^j \overline{z}^{k-j} \qquad \text{for all } z \in \mathbb{T}.$$

Then it is easy to see that A is a subalgebra of $\mathrm{C}(\mathbb{T})$ that contains $\mathbb{1}$, and A is invariant under complex conjugation. Moreover, A clearly separates the points in $\mathbb{T}$ since A contains the identity function. Hence, A is dense in $\mathrm{C}(K)$ according to the Stone–Weierstraß approximation theorem 1.1.8.

When thinking of the unital Banach algebra of bounded linear operators $\mathcal{L}(X)$ on a complex Banach space X, it is natural to consider *invertibility* of elements is an essential concept – since it is, in $\mathcal{L}(X)$ related to solving linear equations.

Definition 1.1.10 (Invertible elements)**.** Let A be an algebra that contains a neutral element. An element $a \in A$ is called *invertible* if there exists an element $b \in A$ such that $ab = ba = 1$. In this case, the element b is uniquely determined[15] and we denote it as a^{-1}.

The set of all invertible elements in A is denoted by $\mathrm{Inv}(A)$.

Note that, if a, b are invertible elements of a unital algebra A, then ab is also invertible and $(ab)^{-1} = b^{-1}a^{-1}$; moreover, if a is invertible, then so is a^{-1}, and in this case one has $(a^{-1})^{-1} = a$. If the unital algebra A is non-zero, then all invertible elements are non-zero, too.

Note that, if a unital Banach algebra A is not $\{0\}$, then one has

$$\|a\|^{-1} \leq \left\|a^{-1}\right\|;$$

for every invertible element $a \in A$. This follows from $1 = \|1\| = \left\|a^{-1}a\right\| \leq \left\|a^{-1}\right\| \|a\|$.

In unital Banach algebras, there is a simple but very powerful tool which allows us to check that many element are invertible. It is a Banach algebra version of the geometric series in $\mathbb{C}$, and you might already know it for the special case of matrices:

Proposition 1.1.11 (Neumann series)**.** *Let A be a unital Banach algebra and let $a \in A$ be an element of norm $\|a\| < 1$. Then $1 - a$ is invertible and*[16]

$$(1-a)^{-1} = \sum_{n=0}^{\infty} a^n,$$

[15] Why?

[16] We use the standard convention that $a^n := a \cdots a$ with n factors for each integer $n \geq 1$, and that $a^0 := 1$ in algebras which have a neutral element.

where the series converges absolutely in the Banach space A; moreover, one has

$$\|1-a\|^{-1} \leq \left\|(1-a)^{-1}\right\| \leq (1-\|a\|)^{-1}.$$

where the second inequality is always true and the first inequality is true if $A \neq \{0\}$.

Proof. As the norm as submultiplicative on A, one has $\|a^n\| \leq \|a\|^n$ for each $n \in \mathbb{N}$, and since $\|1\| \leq 1$, this is also true for $n = 0$. Hence,

$$\sum_{n=0}^{\infty} \|a^n\| \leq \sum_{n=0}^{\infty} \|a\|^n \leq \frac{1}{1-\|a\|}$$

since $\|a\| < 1$. Therefore, the series $\sum_{n=0}^{\infty} a^n$ converges absolutely to an element $b \in A$, and this element has norm $\|b\| \leq \sum_{n=0}^{\infty} \|a^n\| \leq (1-\|a\|)^{-1}$. Moreover, one has

$$(1-a)b = (1-a)\sum_{n=0}^{\infty} a^n = \lim_{N\to\infty} \big(\sum_{n=0}^{N} a^n - \sum_{n=1}^{N+1} a^n\big) = \lim_{N\to\infty}(1-a^{N+1}) = 1;$$

the last equality follows from $\left\|a^{N+1}\right\| \leq \|a\|^{N+1} \to 0$ and, again, $\|a\| < 1$. A similar computation shows that $b(1-a) = 1$.

So the element $1-a$ is indeed invertible with inverse b. We already proved the upper norm estimate for $(1-a)^{-1} = b$, and the lower norm estimate is true for inverses of all invertible elements in non-zero unital Banach algebras, as remarked before the proposition. □

Corollary 1.1.12 (The set of invertible elements is open and the inverse map is continuous)**.** *Let A be a unital Banach algebra. Then the set* $\mathrm{Inv}(A)$ *of invertible elements of A is open and the mapping*

$$\mathrm{Inv}(A) \to \mathrm{Inv}(A), \qquad a \mapsto a^{-1}$$

is continuous.

More precisely, let $a \in A$ be invertible and let $b \in A$ such that $\|b-a\| < \left\|a^{-1}\right\|^{-1}$. Then b is also invertible, its inverse b^{-1} satisfies the norm estimate

$$\left\|b^{-1}\right\| \leq \frac{1}{\|a^{-1}\|^{-1} - \|b-a\|},$$

and the distance of b^{-1} to a^{-1} satisfies the norm estimate

$$\left\|b^{-1}-a^{-1}\right\| \leq \frac{\left\|a^{-1}\right\| \|b-a\|}{\|a^{-1}\|^{-1} - \|b-a\|}.$$

Proof. As a is invertible, we have

$$b = a + b - a = a(1 + a^{-1}(b-a)).$$

Since $\left\|-a^{-1}(b-a)\right\| \leq \left\|a^{-1}\right\| \|b-a\| < 1$, the element $1 + a^{-1}(b-a)$ is invertible and its inverse is given as the Neumann series $\sum_{n=0}^{\infty} \big(- a^{-1}(b-a)\big)^n$ and has norm

$$\left\|1 + a^{-1}(b-a)\right\| \leq \frac{1}{1 - \|a^{-1}\| \, \|b-a\|}.$$

Thus, b is also invertible with inverse

$$b^{-1} = (1 + a^{-1}(b-a))^{-1} a^{-1} = \sum_{n=0}^{\infty} \big(- a^{-1}(b-a)\big)^n a^{-1}.$$

The claimed norm estimate for b^{-1} follows readily from this series representation,[17] the claimed norm estimate for $b^{-1} - a^{-1}$ follows from the series represenation

$$b^{-1} - a^{-1} = \sum_{n=1}^{\infty} \big(- a^{-1}(b-a)\big)^n a^{-1}.$$

Note that the claimed continuity of $a \mapsto a^{-1}$ is an immediate consequence of the norm estimate for $b^{-1} - a^{-1}$. □

1.2 Homomorphisms and ideals

Definition 1.2.1 (Algebra homomorphisms)**.** A mapping $\phi : A \to B$ between two algebras A and B is called a *algebra homomorphism* if ϕ is linear and $\varphi(a_1 a_2) = \varphi(a_1)\varphi(a_2)$ for all $a_1, a_2 \in A$.

Definition 1.2.2 (Ideals)**.** A subset I of an algebra A is called an *ideal* if I is a vector subspace of A and for all $a \in A$ and $c \in I$ one has $ac \in I$ and $ca \in I$.[18]

Note that, every ideal in an algebra A is also a subalgebra of A.[19] If A contains a neutral element 1 and an ideal $I \subset A$ satisfies $1 \in I$, then $A = I$.[20] Moreover, if A is a Banach algebra, then the closure of every ideal is an ideal.[21]

Example 1.2.3 (Ideals in spaces of continuous functions)**.** Let (K, d) be a compact metric space (or, more generally, let K be a compact topological Hausdorff space). If $C \subseteq K$ is closed, then the set

$$I_C := \{ f \in \mathrm{C}(K) \mid \; f|_C = 0 \}$$

is a closed ideal in $\mathrm{C}(K)$.

[17] Or alternatively from the upper norm estimate that is given in Proposition 1.1.11; but this is essentially the same argument.
[18] We can formulate this a bit shorter as the conditions $AI \subseteq I$ and $IA \subseteq I$.
[19] Why?
[20] Why?
[21] Why?

One can show that, in fact, all closed ideals in $\mathrm{C}(K)$ are of this form; see for instance [MN91, Proposition 2.1.9 on p. 57].

One way in which closed ideal occur naturally is as kernels of continuous algebra homomorphisms:

Proposition 1.2.4 (Kernels of algebra homomorphisms)**.** *Let A, B be Banach algebras and let $\phi : A \to B$ be a continuous Banach algebra homomorphism. Then $\ker \phi$ is a closed ideal in A.*

Proof. Since ϕ is linear, its kernel is a vector subspace of A, and the kernel is closed since ϕ is continuous. Finally, let $a \in A$ and $c \in \ker \phi$. Then

$$\phi(ac) = \phi(a)\phi(c) = \phi(a) \cdot 0 = 0,$$

and a similar computation shows that $\phi(ca) = 0$. Hence, $ac \in \ker \phi$ and $ca \in \ker \phi$, i.e., $\ker \phi$ is indeed an ideal in A. □

The preceding proposition has, in a sense, a converse: every closed ideal in a Banach algebra is the kernel of a certain continuous algebra homomorphismus. We will prove this in the following proposition by means of a quotient space construction.

Proposition 1.2.5 (Quotient algebras)**.** *Let I be a closed ideal in a Banach algebra A. Then the multiplication*

$$(a + I) \cdot (b + I) := ab + I \qquad \text{for all } a + I, b + I \in A/I$$

is a well-defined mapping $A/I \times A/I \to A/I$, and it turns A/I into a Banach algebra with respect to the quotient norm given by

$$\|a + I\| := \operatorname{dist}(a, I) = \inf\{\|a - c\| \mid c \in I\}.$$

The quotient mapping

$$q : A \to A/I, \qquad a \mapsto a + I$$

is a continuous algebra homomorphism.

If the Banach algebra A is unital with neutral element 1, then A/I is unital with neutral element $1 + I$.

Proof. All claims can be straightforwardly deduced from the definitions. □

The following example of a closed ideal and its quotient space will become very important later on in the course.

Example 1.2.6 (The ideal of compact operators and the Calkin algebra)**.** Let X be a complex Banach space and let $\mathcal{K}(X)$ denote the set of all compact linear operators from X to X. Then $\mathcal{K}(X)$ is a closed vector subspace of the space $\mathcal{L}(X)$ of bounded linear operators on X, and $\mathcal{K}(X)$ is even an ideal in $\mathcal{L}(X)$.[22] If X is infinite-dimensional, then $\mathcal{K}(X) \neq \mathcal{L}(X)$.[23]

The Banach algebra $\mathcal{L}(X)/\mathcal{L}(X)$ is called the *Calkin algebra* over X; it is a unital Banach algebra according to the preceding proposition.

1.3 Intermezzo: Calculus with values in Banach spaces

We are going to need some of the standard results from calculus for vector-valued functions, where the notion *vector-valued* is meant in the sense of *having values in a Banach space*. We start with derivatives, and proceed to Riemann integrals afterwards. To treat both derivatives and integrals in a quite efficient way, it is convenient to first introduce the notion of a *net* – which we will also need later when we treat topological spaces.

Definition 1.3.1 (Directed sets)**.** Let J be a set. A *direction* on J is a relation $\preceq$ on J which satisfied the following axioms:

(I) *Reflexivity:* For all $j \in J$ one has $j \preceq j$.

(II) *Transitivity:* For all $j_1, j_2, j_3 \in J$ which satisfy $j_1 \preceq j_2$ and $j_2 \preceq j_3$ one also has $j_1 \preceq j_3$.

(III) *Directedness:* For all $j_1, j_2 \in J$ there exists $j_3 \in J$ such that $j_1 \preceq j_3$ and $j_2 \preceq j_3$.

A *directed set* is pair $(J, \preceq)$, where J is a set and $\preceq$ is a direction on J.

Note that we do not require a direction to be anti-symmetric. The following notational shortcut is very common: one often just speaks of a directed set J, thus suppressing the symbol for the direction in the notation.

Definition 1.3.2 (Nets)**.** Let X be a set. A *net* in X is a family of elements $(x_j)_{j\in J}$ of X, where J is a non-empty directed set.[24]

You are already familiar with a very important example of nets: every sequence $(x_n)_{n\in\mathbb{N}}$ in X is also a net in X, if we endow $\mathbb{N}$ with its usual order.[25]

Similarly as for sequences, one can define convergence for nets. We will later do this in the very general setting of topological space; but for now, we first discuss this concept in metric spaces.

[22] Since for a compact linear operator $K : X \to X$ and a bounded linear operator $T : X \to X$, the compositions TK and KT are also compact.

[23] Why?

[24] Note that the direction on J is – though typically suppressed in the notation – part of the definition of the net $(x_j)_{j\in J}$. In other words, it only makes sense to speak of a net $(x_j)_{j\in J}$ if one specifies (or if it is clear from the context) which direction one uses on J.

[25] Here we use that the usual order on $\mathbb{N}$ is also a direction.

Definition 1.3.3 (Convergence of nets)**.** Let (M, d) be a metric space, let $(x_j)_{j\in J}$ be a net in M.

(a) Let $x \in M$. We say that the net $(x_j)_{j\in J}$ *converges to* x, or that x is the *limit* of the net $(x_j)_{j\in J}$ if the following holds: for very $\varepsilon > 0$ there exists an index $j_0 \in J$ such that $\mathrm{d}(x_j, x) < \varepsilon$ for all $j \succeq j_0$.[26]

We sometimes write $x = \lim_j x_j$ to express that $(x_j)_{j\in J}$ converes to x.[27]

(b) We say that the net $(x_j)_{j\in J}$ *converges* if there exists $x \in M$ such that $(x_j)_{j\in J}$ converges to x.

(c) We say that the net $(x_j)_{j\in J}$ is a Cauchy net if the following holds: for every $\varepsilon > 0$ there exists $j_0 \in J$ such that $\mathrm{d}(x_{j_1}, x_{j_2}) < \varepsilon$ for all indices $j_1, j_2 \succeq j_0$ in J.

Recall that a metric space is said to be *complete* if every Cauchy sequence in it is convergent. It is very useful that this property already implies the same property for Cauchy nets:

Proposition 1.3.4 (Convergence of Cauchy nets)**.** *Let* (M, d) *be a metric space. The following are equivalent:*

(i) *The space* (M, d) *is complete (i.e., every Cauchy sequence in* M *converges).*

(ii) *Every Cauchy net in* M *converges.*

Proof. This will be part of the second exercise sheet. □

Now we can define derivatives of vector-valued functions in the following way:

Definition 1.3.5 (Derivatives of vector-valued functions on intervals)**.** Let X be a real or complex Banach space, let $I \subseteq \mathbb{R}$ be an interval that is non-empty and not a singleton, and let $f : I \to X$.

(a) Let $t_0 \in I$. The function f is called *differentiable* at x if there exists a vector $v \in E$ with the following propperty: for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $t \in I \setminus \{t_0\}$ with $|t - t_0| \leq \delta$ one has $\left\| \frac{f(t)-f(t_0)}{t-t_0} - v \right\| < \varepsilon$.

In this case, the vector v is called the *derivative* of f at t_0 and is denoted by $v =: f'(t_0)$.

(b) The function f is called *differentiable* if it is differentiable at every point $t \in I$. In this case, the mapping $f' : I \to E$, $t \mapsto f'(t)$ is called the *derivative* of f.

[26] Of course, the notation $j \succeq j_0$ is simply defined to mean $j_0 \preceq j$.

[27] Note that this notation makes sense as it follows from the positive definiteness of the metric d that a net in a metric space has at most one limit. For the same reason, it makes sense to speak of "the" limit of a convergent net, rather than only of "a" limit.

Note that, in the situation if the previous definition, we can rephrase differentiability of f at a point $t_0 \in I$ in terms of nets: endow the set $I \setminus \{t_0\}$ with the direction $\preceq$ given by $t_1 \preceq t_2$ iff $|t_1 - t_0| \geq |t_2 - t_0|$. then f is differentiable at x if and only if the net $\big(\frac{f(t)-f(t_0)}{t-t_0}\big)_{t\in I\setminus\{t_0\}}$ converges to a point in X, and in this case, the limit of this net is the derivative $f'(t_0)$.[28]

Let us list a few properties of the derivative of vector-valued functions:

Proposition 1.3.6 (Properties of the derivative). *Let X be a real or complex Banach space, let $I \subseteq \mathbb{R}$ be an interval that is non-empty and not a singleton, and let $f, g : I \to X$ and $t_0 \in I$.*

(a) *If f and g are differentiable at t_0 and α, β are scalars, then $\alpha f + \beta g$ is differentiable at t_0 and*
$$(\alpha f + \beta g)'(t_0) = \alpha f'(t_0) + \beta g'(t_0).$$

(b) *Let Y be a Banach space over the same field as X and let $T : X \to Y$ be a bounded linear operator. If f is differeniable at t_0, then $Tf := T \circ f : I \to F$ is also differentiable at t_0 and*
$$(Tf)'(t_0) = T\big(f'(t_0)\big).$$

Proof. The proof is straightforward, so we omit it. □

As in the scalar-valued case, the derivative of a function is constantly 0 if and only if the function is constant:

Proposition 1.3.7. *Let X be a real or complex Banach space, let $I \subseteq \mathbb{R}$ be an interval that is non-empty and not a singleton, and let $f : I \to X$ be differentiable.*

If $f'(t) = 0$ for all $t \in I$, then f is constant.

Proof. We can easily reduce this to the scalar valued case: Fix $t_0 \in I$ and let $\varphi \in X'$ be a bounded linear functional on X.[29] According to Proposition 1.3.6(b) the mapping $\varphi \circ f : I \to \mathbb{K}$ (where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ denotes the underlying scalar field of X) is differentiable with derivative $(\varphi \circ f)'(t) = \varphi(f'(t)) = \varphi(0) = 0$ for all $t \in I$. Hence, as you know from scalar-valued calculus one has $\varphi(f(t)) = \varphi(f(t_0))$, or equivalently,
$$\langle \varphi, f(t) - f(t_0)\rangle = 0$$
for all $t \in I$.[30] As X' separates X (recall that this is a consequence of the Hahn–Banach theorem), it follows that $f(t) = f(t_0)$ for all $t \in I$. □

[28]This also implies that the vector v in the definition is uniquely determined, so the notation $f'(t_0) := v$ is indeed justified.

[29]Throughout, we will denote the dual space of a Banach space X by X'.

[30]Here we used the very common notation $\langle \varphi, x\rangle := \varphi(x)$ for all $x \in X$.

After briefly discussing derivatives in Banach spaces, we are now going to discuss integration. There are various ways to define integrals of functions with values in Banach spaces.

For most parts of this course, the *vector-valued Riemann integral* will do a good job. The following example of a directed set will be very useful to define the (vector-valued) Riemann integral.

Definition 1.3.8 (Partitions of intervals)**.** Let $\emptyset \neq I \subseteq \mathbb{R}$ be a compact interval.

(a) A *partition* p of I is a tuple $p = (p_0, \dots, p_n)$, where $n \geq 1$ is an integer and $p_0, \dots, p_n$ are elements of I such that $\min I = p_0 \leq p_1 \leq \dots \leq p_n = \max I$.

In this case, the number n is called the *number of grid points* of p, and we denote it by $\mathrm{n}(p)$.

(b) The *grid width* of a partition p of I with number of grid points $\mathrm{n}(p) := n$ is the number[31]

$$\mathrm{w}(p) := \max\{p_1 - p_0, \dots, p_n - p_{n-1}\}.$$

(c) A *sampled partition*[32] of I is a pair (p, s), where p is a partition of I and $s = (s_1, \dots, s_{\mathrm{n}(p)})$ is a tuple of $\mathrm{n}(p)$ points in I such that $s_k \in [p_{k-1}, p_k]$ for all $k \in \{1, \dots, \mathrm{n}(p)\}$.

(d) For two sampled partitions (p, s) and $(\tilde{p}, \tilde{s})$ of I we define $(p, s) \preceq (\tilde{p}, \tilde{s})$ if $\mathrm{w}(p) \geq \mathrm{w}(\tilde{p})$.

Note that, for a compact interval $\emptyset \neq I \subseteq \mathbb{R}$, the relation $\preceq$ is a direction on the set of all sampled partitions of I; it is, however, not a partial order since it is not anti-symmetric. For vector-valued continuous functions on compact intervals we can now define the Riemann integral by using the following result:

Theorem 1.3.9 (Convergence of Riemann sums for continuous functions)**.** *Let X be a real or complex Banach space, let $\emptyset \neq I \subseteq \mathbb{R}$ be a compact interval, and let $f : I \to X$ be continuous. Then the net*

$$\Big(\sum_{k=0}^{\mathrm{n}(p)-1} f(s_{k+1})(p_{k+1} - p_k) \Big)_{(s,p)}$$

(where (s, p) runs through the set of all sampled partitions of I) converges in X.

Proof. Since the norm on X is complete, it suffices according to Proposition 1.3.4 to show that our net is a Cauchy net. So let $\varepsilon > 0$. As I is compact, the continuous function f is even uniformly continuous, i.e., there exists a number $\delta > 0$ such that $\|f(t) - f(s)\| \leq \varepsilon$ for all $s, t \in I$ that satisfy $|t - s| \leq \delta$.

[31] Note that one always has $\mathrm{w}(p) \geq 0$, and $\mathrm{w}(p) = 0$ if and only if I is a singleton.

[32] In German: *punktierte Zerlegung*.

Next we note the following: consider a sampled partition (p, s) of I with grid width $\mathrm{w}(p) \leq \delta$, and another sampled partition $(\tilde{p}, \tilde{s})$ of I such that every point in p also occurs at least once in $\tilde{p}$.[33] For each $k \in \{0, \ldots, \mathrm{n}(p)\}$ let ℓ_k be the first index in the set $\{0, \ldots, \mathrm{n}(\tilde{p})\}$ such that $\tilde{p}_{\ell_k}$ coincides with p_k. Then one has, for each $k \in \{0, \ldots, \mathrm{n}(p) - 1\}$,

$$\begin{aligned}
&\left\| f(s_{k+1})(p_{k+1} - p_k) - \sum_{j=\ell_k}^{\ell_{k+1}-1} f(\tilde{s}_{j+1})(\tilde{p}_{j+1} - \tilde{p}_j) \right\| \\
&= \left\| \sum_{j=\ell_k}^{\ell_{k+1}-1} f(s_{k+1})(\tilde{p}_{j+1} - \tilde{p}_j) - \sum_{j=\ell_k}^{\ell_{k+1}-1} f(\tilde{s}_{j+1})(\tilde{p}_{j+1} - \tilde{p}_j) \right\| \\
&\leq \sum_{j=\ell_k}^{\ell_{k+1}-1} \underbrace{\left\| f(s_{k+1}) - f(\tilde{s}_{j+1}) \right\|}_{\substack{\leq \varepsilon \\ \text{since } \left| s_{k+1} - \tilde{s}_{j+1} \right| \leq \delta}} (\tilde{p}_{j+1} - \tilde{p}_j) \leq \varepsilon (p_{k+1} - p_k).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\left\| \sum_{k=0}^{\mathrm{n}(p)-1} f(s_{k+1})(p_{k+1} - p_k) - \sum_{j=0}^{\mathrm{n}(\tilde{p})-1} f(\tilde{s}_{j+1})(\tilde{p}_{j+1} - \tilde{p}_j) \right\| \\
&= \left\| \sum_{k=0}^{\mathrm{n}(p)-1} \Big(f(s_{k+1})(p_{k+1} - p_k) - \sum_{j=\ell_k}^{\ell_{k+1}-1} f(\tilde{s}_{j+1})(\tilde{p}_{j+1} - \tilde{p}_j) \Big) \right\| \\
&\leq \sum_{k=0}^{\mathrm{n}(p)-1} \varepsilon (p_{k+1} - p_k) = \varepsilon \left| I \right|,
\end{aligned}$$

where $|I|$ denotes the length of I.[34] Now choose any sampled partition $(\hat{p}, \hat{s})$ of I of gridwidth $\mathrm{w}(\hat{p}) \leq \delta$. Every sampled partiation (p, s) that satisfies $(p, s) \succeq (\hat{p}, \hat{s})$ also has grid width $\mathrm{w}(p) \leq \delta$ (due to the definition of the relation $\succeq$), so for any such (p, s) we can find another sampled partition $(\tilde{p}, \tilde{s})$ such that all points in $\hat{p}$ and p also occur in $\tilde{p}$. By what we proved above it follows that

$$\left\| \sum_{k=0}^{\mathrm{n}(p)-1} f(s_{k+1})(p_{k+1} - p_k) - \sum_{j=0}^{\mathrm{n}(\hat{p})-1} f(\hat{s}_{j+1})(\hat{p}_{j+1} - \hat{p}_j) \right\| \leq 2 \left| I \right| \varepsilon,$$

so our net is indeed a Cauchy net. □

[33] Since we allow a point to occur multiple times in a partition, we note here that we do not require each point of p to occur in $\tilde{p}$ at least the same number of times as it occurs in p; this is simply not necessary for the subsequent argument.

[34] Note that, between the first and the second line of this computation, several summands of the latter sum are missing in case that some of the points in the partition $\tilde{p}$ coincide; this is not a problem since those summands are 0.

Definition 1.3.10 (Riemann integral of continuous functions)**.** In the situation of Theorem 1.3.9 the limit of the net is called the *Riemann integral of* f; it is denoted by $\int_I f(t)\ \mathrm{d}t$ or by $\int_{\min I}^{\max I} f(t)\ \mathrm{d}t$.

In the following proposition we list a number of useful properties of the Riemann integral with values in Banach spaces:

Proposition 1.3.11 (Properties of the vector-valued Riemann integral)**.** *Let X be a real or complex Banach space, let $a \le b \le c$ be real numbers, and let $f, g : [a, c] \to X$ be continuous.*

(a) *If $a = c$, then $\int_a^c f(t)\ \mathrm{d}t = 0$.*

(b) *One has*

$$\int_a^c f(t)\ \mathrm{d}t = \int_a^b f(t)\ \mathrm{d}t + \int_b^c f(t)\ \mathrm{d}t.$$

(c) *For all scalars α, β one has*

$$\int_a^c \alpha f(t) + \beta g(t)\ \mathrm{d}t = \alpha \int_a^c f(t)\mathrm{d}t + \beta \int_a^c g(t)\ \mathrm{d}t.$$

(d) *Let Y be a Banach space over the same field as X and let $T : X \to Y$ be linear and continuous. Then*

$$\int_a^c (Tf)(t)\ \mathrm{d}t = T \int_a^c f(t)\ \mathrm{d}t.$$

(e) *One has the* fundamental estimate

$$\left\| \int_a^c f(s)\ \mathrm{d}s \right\| \le \int_a^c \|f(s)\|\ \mathrm{d}s.$$

The proof of the proposition is quite straightforward, so we omit it. Of course we are going to need the fundamental theorem of calculus:

Theorem 1.3.12 (The fundamental theorem of vector-valued calculus)**.** *Let X be a real or complex Banach space, and let $a \le b$ be real numbers, and let $f : [a, b] \to X$.*

(a) *If f is continuous, then the function $F : [a, b] \to X$, $t \mapsto \int_a^t f(s)\ \mathrm{d}s$, is differentiable and $F' = f$.*

(b) *If f is differentiable and $f' : [a, b] \to E$ is continuous, then*

$$f(b) - f(a) = \int_a^b f'(x)\ \mathrm{d}x.$$

Proof. (a) Let $t_0 \in I$ and let ε. There exists $\delta > 0$ such that for all $t \in I$ with $|t - t_0| \leq \delta$ one has $\|f(t) - f(t_0)\| \leq \varepsilon$. Fix such a t, and first assume that $t > t_0$. Then[35]

$$\begin{aligned}\left\| \frac{F(t) - F(t_0)}{t - t_0} - f(t_0) \right\| &= \frac{\left\| \int_{t_0}^t f(s) - f(t_0)\, \mathrm{d}s \right\|}{t - t_0} \\ &\leq \frac{\int_{t_0}^t \|f(s) - f(t_0)\|\, \mathrm{d}s}{t - t_0} \\ &\leq \frac{\int_{t_0}^t \varepsilon\, \mathrm{d}s}{t - t_0} = \varepsilon.\end{aligned}$$

By a similar computation the same estimate can be shown if $t < t_0$. - This proves that F is differentiable at t_0 with $F'(t_0) = f(t_0)$.

(b) Define a function $g : [a, b] \to X$ by $g(t) = \int_a^t f'(s)\, \mathrm{d}s$ for all $t \in [a, b]$. As shown in ((a)), the function g is differentiable with derivative f'. Hence, $g - f$ is differentiable with derivative 0 and is thus constant according to Proposition 1.3.7. So

$$g(b) - f(b) = g(a) - f(a) = -f(a),$$

and hence, $\int_a^b f'(s)\, \mathrm{d}s = g(b) = f(b) - f(a)$. □

1.4 The exponential function in Banach algebras

In this section we introduce the *exponential function* in unital Banach algebras. This function will sometimes serves as a nice illustration of some of our results, and on some occassions the exponential function also turns out to be useful in proofs.

Definition 1.4.1 (The exponential function)**.** Let A be a unital Banach algebra. We define $\exp : A \to A$ by

$$\exp(a) := \sum_{n=0}^{\infty} \frac{a^n}{n!} \qquad \text{for every } a \in A,$$

where the series converges absolutes in A,[36] and call exp the *exponential function* on A.

Let a be an element of a unital Banach algebra A. Sometimes one uses the notation $e^a := \exp(a)$. In the following we list a few properties of the exponential function:

[35] In this computation we use that, for all $x \in X$ and all real numbers $a \leq b$, one has $\int_a^b x\, \mathrm{d}s = (b - a)s$. How can this be derived from Proposition 1.3.11(d)?

[36] Since $\sum_{n=0}^{\infty} \left\| \frac{a^n}{n!} \right\| \leq \sum_{n=0}^{\infty} \frac{\|a\|^n}{n!} = e^{\|a\|} < \infty$. Note that this also shows that $\|\exp(a)\| \leq \exp(\|a\|)$.

Proposition 1.4.2 (Properties of the exponential function)**.** *Let A be a unital Banach algebra.*

(a) *The map* $\exp : A \to A$ *is continuous.*

(b) *If $a, b \in A$ commute, then* $\exp(a+b) = \exp(a)\exp(b) = \exp(b)\exp(a)$.

(c) *Let $a \in A$. The mapping*

$$f:\ \mathbb{R} \to X, \qquad t \mapsto \exp(ta)$$

is differentiable, and its derivative at any point $t \in \mathbb{R}$ equals $a\exp(ta)$.

(d) *One has* $\exp(0) = 1$. *More generally, for every $\lambda \in \mathbb{C}$ one has*

$$\exp(\lambda \cdot 1) = \exp(\lambda) \cdot 1,$$

where 1 *denotes the neutral element in A on both sides of the equation.*

(e) *For every $a \in A$ the element* $\exp(a)$ *is invertible and* $\big(\exp(a)\big)^{-1} = \exp(-a)$.

Proof. (a) This is part of Exercise Sheet 2.

(b) For commuting a and b one can inductively show the usual binomial formula

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

for all integers $n \geq 0$. From this one can then derive the claimed functional equality for exp, precisely as it is done for scalars or matrices.

(c) We first show the differentiability and the claimed formula for the derivative at the point 0: for all $t \in \mathbb{R} \setminus \{0\}$ one has

$$\begin{aligned} \left\| \frac{f(t) - f(0)}{t - 0} - a \right\| &= \left\| \frac{\exp(ta) - 1}{t} - a \right\| \\ &= \left\| \sum_{n=2}^{\infty} \frac{(ta)^n}{n!} \right\| \leq \sum_{n=2}^{\infty} \frac{|t|^n \, \|a\|^n}{n!}, \end{aligned}$$

which converges to 0 as $t \to 0$.

The claim at general points $t_0 \in \mathbb{R}$ can be derived from the claim at 0 and assertion (b).[37]

(d) This follows readily from the definition of exp.

(e) Since a and $-a$ commute we have, according to (b),

$$\exp(a)\exp(-a) = \exp(a + (-a)) = \exp(0) = 1,$$

where we used (d) for the last equality. Since $\exp(a)$ and $\exp(-a)$ commute, the claim follows. □

[37] If one uses that ta and $t_0 a$ commute for all $t \in \mathbb{R}$.

Examples 1.4.3 (Matrix exponential function)**.** Let $d \in \mathbb{N}$, let $\mathbb{C}^d$ be endowed wtih any norm, and let $\mathbb{C}^{d\times d} = \mathcal{L}(\mathbb{C}^d)$ be endowed with the induced operator norm.

For every $A \in \mathbb{C}^{d\times d}$ the mapping

$$\mathbb{R} \to \mathbb{C}^{d\times d}, \qquad t \mapsto \exp(tA)$$

is the matrix exponential function that is an important tool to solve linear autonomous ordinary differential equations in finite dimensions.

Let us extend the previous example to the infinite-dimensional case:

Example 1.4.4 (Linear autonomous ODEs in Banach spaces)**.** Let X be a complex Banach space, let $A \in \mathcal{L}(X)$ and $x_0 \in X$. Then the mapping

$$u : \mathbb{R} \to X, \qquad t \mapsto u(t) := \exp(tA)x_0$$

is differentiable and solves the initial value problem[38]

$$\begin{cases} u(0) &= x_0, \\ \dot{u}(t) &= Au(t) \qquad \text{for all } t \in \mathbb{R}. \end{cases}$$

Proof. The mapping $T : \mathcal{L}(X) \to X$, $B \mapsto Bx_0$ is linear and bounded.[39] Since the mapping $f : \mathbb{R} \ni t \mapsto \exp(tA) \in \mathcal{L}(X)$ is, according to Proposition 1.4.2(c), differentiable with derivative $\dot{f} : t \mapsto A\exp(tA)$, it follows from Proposition 1.3.6(b) that $u := T \circ f : \mathbb{R} \ni t \mapsto \exp(tA)x_0 \in X$ is differentiable with derivative

$$\dot{u}(t) = T\big(\dot{f}(t)\big) = T\big(A\exp(tA)\big) = A\exp(tA)x_0 = Au(t)$$

at each point $t \in \mathbb{R}$. □

Example 1.4.5 (The exponential function in spaces of continuous functions)**.** Let (K, d) be a compact metric space (or, more generally, a compact topological Hausdorff space), and let $f \in \mathrm{C}(K)$. Then $\exp(f) = \exp \circ f$.

Proof. Let $x \in K$. The mapping $\delta_x : \mathrm{C}(K) \to \mathbb{C}$, $g \mapsto g(x)$ is a continuous linear functional on $\mathrm{C}(K)$, so one has

$$\begin{aligned} \big(\exp(f)\big)(x) = \delta_x\big(\exp(f)\big) &= \sum_{n=0}^{\infty} \frac{\delta_x\big(f^n\big)}{n!} \\ &= \sum_{n=0}^{\infty} \frac{\big(f(x)\big)^n}{n!} = \exp(f(x)) = (\exp \circ f)(x), \end{aligned}$$

so indeed $\exp(f) = \exp \circ f$. □

[38] Here, we use the notation $\dot{u}(t)$ rather than $u'(t)$ to denote the derivative of u at t; this notation is common in the theory of ordinary differential equations.

[39] Why?

Chapter 2

Spectrum and Resolvent

2.1 Prologue: the unitization of a Banach algebra

For every algebra, no matter whether it has a neutral element or not, we can define a larger algebra which has a neutral element:

Proposition 2.1.1 (Adjoining a neutral element)**.** *Let A be an algebra. The vector space $\tilde{A} := A \times \mathbb{C}$ is an algebra with respect to the product $\cdot$ given by*

$$\begin{pmatrix} a \\ \alpha \end{pmatrix} \cdot \begin{pmatrix} b \\ \beta \end{pmatrix} := \begin{pmatrix} ab + \alpha b + \beta a \\ \alpha\beta \end{pmatrix}$$

for all (a, α), $(b, \beta) \in \tilde{A}$, and the element $(0, 1)$ of $\tilde{A}$ is a neutral in $\tilde{A}$. If we identify A with the subset $A \times \{0\}$ of $\tilde{A}$, then A is an ideal (and thus, in particular, a subalgebra) in $\tilde{A}$.

If A is a Banach algebra and we endow $\tilde{A}$ with the norm given by

$$\left\| \begin{pmatrix} a \\ \alpha \end{pmatrix} \right\| := \|a\| + |\alpha|$$

for all $(a, \alpha) \in \tilde{A}$, then $\tilde{A}$ is a unital Banach algebra.

Proof. All the claims can be checked by straightforward computations. $\square$

In the situation of the preceding proposition we will, as indicated in the proposition, typically consider A to be a subalgebra of $\tilde{A}$ by identifying it with $A \times \{0\}$, i.e. each element $a \in A$ will be identified with the element $(a, 0)$ of $\tilde{A}$. If we denote the neutral element $(0, 1)$, as usual, by 1, this means that every element $(a, \alpha) \in \tilde{A}$ can be written as $(a, \alpha) = a + \alpha 1$.

Another nice observation is that a Banach algebra with a neutral element can always be equivalent renormed in order to become a unital Banach algebra.

Proposition 2.1.2 (Renorming a Banach algebra to make it unital)**.** *Let $A \neq \{0\}$ be a Banach algebra and let $1 \in A$ be a neutral element.*[1] *There exists an equivalent norm*[2] *on A which turns A into a unital Banach algebra.*

Proof. If $\|1\| \leq 1$ there is nothing to prove, so let us assume that $\|1\| > 1$. Let us define a new norm $\|\cdot\|_1$ on A by the formula

$$\|a\|_1 := \sup\{\|ab\| \mid b \in A \text{ and } \|b\| \leq 1\}$$

for each $a \in A$. One readily checks that

$$\frac{1}{\|1\|}\|a\| \leq \|a\|_1 \leq \|a\|$$

for each $a \in A$ and that $\|\cdot\|_1$ is indeed a norm on A. Submulitplicativity of $\|\cdot\|_1$ is also easy to check, and one has

$$\|1\|_1 = \sup\{\|b\| \mid b \in A \text{ and } \|b\| \leq 1\} = 1.$$

So A is indeed a unital Banach algebra with respect to $\|\cdot\|_1$. □

Note that if a Banach algebra A is already unital (i.e., if $\|1\| \leq 1$), one can still define the norm $\|\cdot\|_1$ in the proof above, and this norm then coincides with the given norm on A.

What we have done so far in this section suggests the following definition:

Definition 2.1.3 (Unitization of a Banach algebra)**.** Let A be a Banach algebra. A unital Banach algebra $A^\sharp$ is called *a unitization of A* if one of the following conditions is satisfied:

(1) The algebra A has a neutral element, $A^\sharp$ is the same algebra as A, and the norms of A and $A^\sharp$ are equivalent.

(2) The algebra A does not have a neutral element, the algebra $A^\sharp$ is the algebra $\tilde{A}$ from Proposition 2.1.1, and the norm on $A^\sharp$ is equivalent to the norm introduced in this proposition.[3]

According this definition, there is only one unitization of a Banach algebra A from a purely algebraic point of view (since the algebra $A^\sharp$ is determined in the Definition) and also from a purely topological point of view (since all norms that we allow on $A^\sharp$ are equivalent). Only the norm on $A^\sharp$ is not uniquely determined.

In particular, the question whether an element $a \in A$ is invertible in $A^\sharp$ does not depend on the choice of them unitization since it is a purely algebraic question that does not depend on the choice of the norm.

[1]But the point of the proposition is that we do not assume $\|1\| \leq 1$ now.

[2]Recall that two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a (real or complex) vector space X are called *equivalent* if there exist real numbers $c, C > 0$ such that $c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$ for all $x \in X$. This is equivalently to the assertion that both norms induce the same topology on X.

[3]One might wonder why we leave some freedom in the choice of the norm here and doesn't simply require $A^\sharp$ to carry the specific norm introduced in Proposition 2.1.1. The reason is that, when one introduced so-called *C^*-algebras* (as we will do later on in this course), the norm from this proposition is not particularly well-suited; we will come back to this point later on.

2.2 The spectrum

Let A be an algebra with a neutral element which we denote, just for this paragraph, by e. For every $\lambda \in \mathbb{C}$ we use λ as a short cut for λe; for instance, for any element $a \in A$ we write $\lambda - a$ as an abbreviation for $\lambda e - a$. If we adopt againt the usual notation 1 (rather than e) for the neutral element, this short cut is consists with the ambiguity that 1 denotes both the neutral element in A and in $\mathbb{C}$. If course, this simplified notation comes at the price that we always have to infer from the context whether, for a number $\lambda \in \mathbb{C}$, the symbol λ denotes the complex number λ, or the elment of A that is given as λ times the neutral element of e.

Definition 2.2.1 (Spectrum and resolvent)**.** Let A be an algebra.

(a) Assume that A contains a neutral element 1. For every $a \in A$ we define the *spectrum* and the *resolvent set* of A as

$$\begin{aligned} & \sigma(a) := \{\lambda \in \mathbb{C} \mid \ \lambda - a \notin \operatorname{Inv}(A)\} \\ \text{and} \quad & \rho(a) := \{\lambda \in \mathbb{C} \mid \ \lambda - a \in \operatorname{Inv}(A)\}, \end{aligned}$$

respectively.[4]

For every $a \in A$ and every $\lambda \in \rho(a)$ the element $\mathcal{R}(\lambda, a) := (\lambda - a)^{-1}$ is called the *resolvent* of a at λ.

(b) Assume that A does not contain a neutral element, and let $\tilde{A}$ denote the algebra given in Proposition 2.1.1.

For every $a \in A$ we define the spectrum $\sigma(a)$ and the resolvent set $\rho(a)$ as the spectrum and the resolvent set, respectively, of a in the algebra $\tilde{A}$.

Remarks 2.2.2. Let A be an algebra which does not have a neutral element, and let $\tilde{A}$ be the algebra from Proposition 2.1.1.

(a) No element of A is in $\operatorname{Inv}(\tilde{A})$; otherwise, the neutral element of $\tilde{A}$ would be in A since A is an ideal.

So in particular, one has $0 \in \sigma(a)$ for each $a \in A$.

(b) If an algebra A does not have a neutral element, a is an element of A and $\lambda \in \mathbb{C}$ is not in the spectrum of a, then the inverse $(\lambda - a)^{-1} \in \tilde{A}$ is not in A (this follows from the previous point since $(\lambda - a)^{-1}$ is invertible in $\tilde{A}$).

Remarks 2.2.3. (a) If A is a Banach algebra, then Definitions 2.2.1 and 2.1.3 show that the spectrum of an element $a \in A$ is given by

$$\sigma(a) = \{\lambda \in \mathbb{C} \mid \ \lambda - a \notin \operatorname{Inv}(A^\sharp)\}.$$

[4]So $\mathbb{C}$ is the disjoint union of the spectrum $\sigma(a)$ and the resolvent set $\rho(a)$.

(b) In the unital Banach algebra $A = \{0\}$ the element 0 has empty spectrum since $1 = 0$ and thus,

$$\lambda - 0 = 0$$

is invertible in A (with inverse 0) for every $\lambda \in \mathbb{C}$.

Proposition 2.2.4 (Properties of spectrum and resolvent)**.** *Let A be a unital Banach algebra and let $a \in A$.*

(a) *The resolvent set $\rho(a)$ is open and the spectrum $\sigma(a)$ is closed.*

(b) *The resolvent mapping*

$$\mathcal{R}(\,\cdot\,, a) : \rho(a) \to \mathrm{Inv}(A)$$

is continuous.

(c) *Let $\lambda \in \rho(a)$ and $\mu \in \mathbb{C}$ such that $|\mu - \lambda| < \|\mathcal{R}(\lambda, a)\|^{-1}$. Then also $\mu \in \rho(a)$ and*

$$\mathcal{R}(\mu, a) = \sum_{n=0}^{\infty} (\lambda - \mu)^n \mathcal{R}(\lambda, a)^{n+1},$$

where the series converges absolutely in A. Moreover, one has the upper resolvent estimate

$$\|\mathcal{R}(\mu, a)\| \leq \frac{1}{\|\mathcal{R}(\lambda, a)\|^{-1} - |\mu - \lambda|}.$$

(d) *For every $\lambda \in \rho(A)$ one has the lower resolvent estimate*[5]

$$\frac{1}{\mathrm{dist}(\lambda, \sigma(a))} \leq \|\mathcal{R}(\lambda, a)\|.$$

(e) *The spectrum $\sigma(a)$ is bounded (hence, compact).*

More precisely, let $\lambda \in \mathbb{C}$ be of modulus $|\lambda| > \|a\|$. Then λ is contained in the resovlent set of a, and

$$\mathcal{R}(\lambda, a) = \sum_{n=0}^{\infty} \frac{a^n}{\lambda^{n+1}},$$

where the series converges absolutely in A. Moreover, one has the upper resolvent estimate

$$\|\mathcal{R}(\lambda, a)\| \leq \frac{1}{|\lambda| - \|a\|}.$$

[5] Here one defines $\mathrm{dist}(\lambda, \emptyset) := \infty$ and $1/\infty := 0$. But, as we shall see in Theorem 2.4.1, this specific case is only relevant if $A = \{0\}$.

(f) *One has* $\|\mathcal{R}(\lambda, a)\| \to 0$ *as* $|\lambda| \to \infty$.

Proof. (a) This follows immediately from the fact that $\mathrm{Inv}(A)$ is open, which we showed in Corollary 1.1.12.

(b) This follows immediately from the continuity of the mapping

$$\mathrm{Inv}(A) \to \mathrm{Inv}(A), \quad a \mapsto a^{-1}$$

that we proved in Corollary 1.1.12.

(c) For λ and μ as given in the assertion the distance between $\mu - a$ and $\lambda - a$ is strictly less than $\left\|(\lambda - a)^{-1}\right\|^{-1}$, so Corollary 1.1.12 yields that $\mu - a$ is invertible, too. Moreover, similarly as in the proof of the corollary, we can write $\mu - a$ as

$$\mu - a = (\lambda - a)\big(1 - (\lambda - \mu)\mathcal{R}(\lambda, a)\big),$$

so the claimed series expansion follows by applying the Neumann series.

(d) It follows from (c) that points $\lambda \in \mathbb{C}$ which are strictly closer to a than $\|\mathcal{R}(\lambda, a)\|^{-1}$ are not in the spectrum of a, so

$$\mathrm{dist}(\lambda, \sigma(a)) \geq \|\mathcal{R}(\lambda, a)\|^{-1}.$$

This gives the claimed estimate.

(e) As $\lambda \neq 0$, one has $\lambda - a = \lambda(1 - a/\lambda)$, so the claim follows from the Neumann series (Proposition 1.1.11).

(f) This is an immediate consequence of (e). □

Note that the compactness of the spectrum clearly remains true if the Banach algebra A is not unital.

Example 2.2.5 (Matrices and Operators). (a) Let $n \geq 1$ be an integer. In the algebra $\mathbb{C}^{n \times n}$, the spectrum of a matrix $A \in \mathbb{C}^{n \times n}$ is simply the set of all eigenvalues of A. Indeed, it is a standard result in linear algebra that a complex number $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only of $\lambda - A$ is not invertible (i.e., if λ is in the spectrum of A) in the sense of Definition 2.2.1.

(b) More generally, let X be a complex Banach space. Then the spectrum of an element $T \in \mathcal{L}(X)$ coincides with the set of all $\lambda \in \mathbb{C}$ such that $\lambda - T : X \to X$ is not bijective; this follows from the bounded inverse theorem.

Note that, specifically in the case $X = \{0\}$, the only operator on X is the 0 operator. It coincides with the identity operator on $X = \{0\}$ and has empty spectrum since

$$\lambda \cdot 0 - 0 = 0$$

is bijective on $X = \{0\}$ for every $\lambda \in \mathbb{C}$.[6]

[6]Note how this is consistent with Remark 2.2.3(b): for $X = \{0\}$ the unital Banach algebra $\mathcal{L}(X)$ is equal to $\{0\}$, and we have already discussed in this remark that the element 0 of this algebra has empty spectrum.

Example 2.2.6 (The spectrum in spaces of continuous functions)**.** Let K be a non-empty compact Hausdorff space and let f be an element of the unital Banach algebra $\mathrm{C}(K)$. Then

$$\sigma(f) = f(K).$$

Proof. Let $\lambda \in \mathbb{C}$.

„$\subseteq$“ Assume that $\lambda \notin f(K)$. Then the function $g : K \to \mathbb{C}$ that is given by

$$g(\omega) = \frac{1}{\lambda - f(\omega)} \quad \text{for all } \omega \in K$$

is well-defined and continuous, i.e., $g \in \mathrm{C}(K)$. Clearly, $g(\lambda - f) = (\lambda - f)g = 1$, so $\lambda - f$ is invertible.

„$\supseteq$“ Let $\lambda \notin \sigma(f)$. Then $g := \mathcal{R}(\lambda, f) = (\lambda - f)^{-1}$ satisfies $(\lambda - f)g = 1$. Hence, for every $\omega \in K$ one has

$$(\lambda - f(\omega))g(\omega) = 1$$

and thus $\lambda \neq f(\omega)$. So $\lambda \notin f(K)$. $\square$

We conclude this section with two nice properties of resolvents. For the proof of the first one we need the second part of the following nice lemma about invertible elements:

Lemma 2.2.7. *Let A be an algebra with neutral element and let $a, b \in A$.*

(a) *If there exist $c, d \in A$ such that $ac = da = 1$, then a is invertible and $c = d = a^{-1}$.*

(b) *Assume that ab is invertible. If a is invertible, then so is b; and if b is invertible, then so is a.*

(c) *Assume that a and b commute. Then ab is invertible if and only if a and b are invertible.*[7]

Proof. (a) We have

$$d = d \cdot 1 = dac = 1 \cdot c = c,$$

which implies, by definition of invertibility, that a is invertible with inverse $c = d$.

(b) By assumption there exists $c \in \operatorname{Inv}(A)$ such that

$$c\,ab = ab\,c = 1.$$

[7] Can you find an example to show that this is false without the assumption that a and b commute?

If a is invertible, then it follows from $abc = 1$ that $b = a^{-1}c^{-1}$, so b is invertible as a product of invertible elements. And if b is invertible, then it follows from $cab = 1$ that $a = c^{-1}b^{-1}$, so a is invertible as a product of invertible elements.

(c) „$\Rightarrow$“ Let ab be invertible and set $c := (ab)^{-1}$. Then $a(bc) = (ab)c = 1$ and $(cb)a = c(ab) = 1$, so it follows from (a) that a is invertible. Thus, (b) implies that b is also invertible.

„$\Leftarrow$“ This implication is clear. $\square$

Proposition 2.2.8 (Spectral mapping theorem for the resolvent)**.** *Let A be a unital Banach algebra, let $a \in A$ and let $\lambda_0 \in \rho(a)$. Then*

$$\sigma\big(\mathcal{R}(\lambda_0, a)\big) = \left\{ \frac{1}{\lambda_0 - \lambda} \mid \lambda \in \sigma(a) \right\}.$$

Proof. First note that if λ, μ are complex numbers such that $\lambda \neq \lambda_0$ and $\mu = \frac{1}{\lambda_0 - \lambda}$, then one can readily check that

$$(\lambda - a) = (\lambda_0 - \lambda)\,(\lambda_0 - a)\big(\mu - \mathcal{R}(\lambda_0, a)\big) \tag{2.2.1}$$

„$\subseteq$“ Let $\mu \in \sigma\big(\mathcal{R}(\lambda_0, a)\big)$. Then μ is non-zero since $\mathcal{R}(\lambda_0, a)$ is invertible, so we can define $\lambda := \lambda_0 - \frac{1}{\mu}$. Then $\lambda \neq \lambda_0$ and $\mu = \frac{1}{\lambda_0 - \lambda}$, so (2.2.1) holds. Since $\mu - \mathcal{R}(\lambda_0, a)$ is not invertible and $(\lambda_0 - \lambda)(\lambda_0 - a)$ is invertible, it follows from Lemma 2.2.7(b) that $\lambda - a$ is not invertible, i.e., $\lambda \in \sigma(a)$. This proves that μ is in the set on the right hand side.

„$\supseteq$“ Let μ be an element of the set on the right hand side, i.e., let μ be of the form $\mu = \frac{1}{\lambda_0 - \lambda}$ for some $\lambda \in \sigma(a)$. Then equality (2.2.1) holds. Since $\lambda - a$ is not invertible and $(\lambda_0 - a)(\lambda_0 - \lambda)$ is invertible, it follows that $\mu - \mathcal{R}(\lambda_0, a)$ is not invertible, so $\mu \in \sigma\big(\mathcal{R}(\lambda_0, a)\big)$. $\square$

We close this section with the following simple but very useful properties of resolvents.

Proposition 2.2.9 (Resolvent identity)**.** *Let A be a unital Banach algebra and let $a \in A$. For all $\lambda, \mu \in \rho(a)$ the so-called* resolvent identity[8]

$$\mathcal{R}(\lambda, a) - \mathcal{R}(\mu, a) = (\mu - \lambda)\mathcal{R}(\lambda, a)\mathcal{R}(\mu, a)$$

holds.

Proof. Fix $\lambda, \mu \in \rho(a)$. One has

$$(\lambda - a)\big(\mathcal{R}(\lambda, a) - \mathcal{R}(\mu, a)\big) = 1 - (\lambda - \mu \;+\; \mu - a)\mathcal{R}(\mu, a) = -(\lambda - \mu)\mathcal{R}(\mu, a).$$

If we multiply this equality with $\mathcal{R}(\lambda, a)$ we obtain the claimed resolvent identity. $\square$

[8] Some people also call it *resolvent equation.* In German, it is typically called *Resolventengleichung.*

2.3 Intermezzo: A first glance at complex analysis with values in Banach spaces

In this section we discuss a bit of complex analysis for functions with values in Banach spaces. As in the scalar-valued case there is a significant difference between differentiability of functions that are defined on a real interval and of functions that are defined on an (open) subset of the complex plane: in the latter case, differentiability is a much stronge property – called holomorphy – and leads to a very rich and often surprising theory.

Definition 2.3.1 (Vector-valued holomorphic functions)**.** Let $\Omega \subseteq \mathbb{C}$ be non-empty and open, let X be a complex Banach space, and let $f : \Omega \to \mathbb{C}$.

(a) Let $z_0 \in \mathbb{C}$. The function f is called *holomorphic* or *analytic* at z_0 if there exists an element $x \in X$ such that

$$\frac{f(z) - f(z_0)}{z - z_0}$$

converges to x in the following sense: for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left\| \frac{f(z) - f(z_0)}{z - z_0} - x \right\| \le \varepsilon$$

for all $z \in \Omega$ that satisfy $|z - z_0| \le \delta$.[9]

In this case, the vector x is called the *derivative of f at z_0* and is denoted by $f'(z_0)$.[10]

(b) The function f is called *holomorphic* or *analytic* if it is holomorphic at each point $z_0 \in \Omega$.

Example 2.3.2 (Resolvents are holomorphic)**.** Let A be a unital Banach algebra and let $a \in A$. Then the mapping

$$\rho(a) \to A, \qquad \lambda \mapsto \mathcal{R}(\lambda, a)$$

is holomorphic, and one has $\mathcal{R}'(\lambda, a) = -\mathcal{R}(\lambda, a)^2$ for each $\lambda \in \rho(a)$.

Proof. This follows immediately from the resolvent identity (Proposition 2.2.9) and the continuity of the resolvent (Proposition 2.2.4(b)). □

As in the scalar-valued case, the identity theorem for holomorphic functions holds:

[9] How can this be rephrase in terms of nets?
[10] Why is x uniquely determined?

Theorem 2.3.3 (Identity theorem for holomorphic functions). *Let $\Omega \subseteq \mathbb{C}$ be non-empty and open, let X be a complex Banach space, and let $f, g : \Omega \to X$ be holomorphic.*

Assume that Ω is connected and that there exists a set $S \subseteq \Omega$ which has an accumulation point in Ω and such that $f(z) = g(z)$ for all $z \in S$, then $f = g$.

Proof. Let $x' \in X'$. Then $x' \circ f$ and $x' \circ g$ are holomorphic functions from Ω to $\mathbb{C}$, and they coincide on S. Thus, by the identity theorem for scalar-valued holomorphic functions one has $x' \circ f = x' \circ g$, i.e.,

$$\langle x', f(\omega) \rangle = \langle x', g(\omega) \rangle$$

for all $\omega \in \Omega$. Since X' separates X, it follows that $f(\omega) = g(\omega)$ for all $\omega \in \Omega$. $\square$

Arguments as in the previous proof – which use that X' separates X – are often referred to as *Hahn–Banach arguments*. By such a Hahn–Banach argument when can also derive the following result from the scalar-valued case:

Theorem 2.3.4 (Liouville's theorem). *Let X be a complex Banach space. If $f : \mathbb{C} \to X$ be holomorphic and bounded*[11]*. Then f is constant.*

Let X be a complex Banach space and let $\Omega \subseteq \mathbb{C}$ be non-empty and open. It is easy to check that every holomorphic function $f : \Omega \to X$ is continuous, and hence it is *locally bounded* in the sense that for every $z_0 \in \Omega$ there exists a neighbourhood U of z_0 in Ω such that $\sup_{z \in U} \|f(z)\| < \infty$. Conversely, if a function is already known to be locally bounded, then it can be tested for holomorphy by testing it against linear functionals and looking for scalar-valued holomorphy:

Theorem 2.3.5 (Weak vs. strong holomorphy). *Let $\Omega \subseteq \mathbb{C}$ be non-empty and open, let X be a complex Banach space, and let $f : \Omega \to X$ be locally bounded. Then the following are equivalent:*

(i) *The function f is holomorphic.*

(ii) *For every $x' \in X'$ the function $x' \circ f : \Omega \to \mathbb{C}$ is holomorphic.*

We refrain from giving a proof here, and instead refer to the literature, for instance to [ABHN11, Proposition A.3 on p. 462]. As a consequence of the theorem one obtains the following regularity result for holomorphic functions from the same result for scalar-valued functions.

Corollary 2.3.6 (Regularity of holomorphic functions). *Let $\Omega \subseteq \mathbb{C}$ be non-empty and open, let X be a complex Banach space, and let $f : \Omega \to X$ be holomorphic.*

Then $f' : \Omega \to \mathbb{C}$ is holomorphic, too, and consequentally, the n-th iterated derivative $f^{(n)}$ equists for each integer $n \geq 0$.

[11] I.e., $\sup_{z \in \mathbb{C}} \|f(z)\| < \infty$.

Proof. For every $x' \in X$ the function $x' \circ f : \Omega \to \mathbb{C}$ is holomorphic, and it is know from scalar-valued complex analysis that its derivative $(x' \circ f)' : \Omega \to \mathbb{C}$ is holomorphic, too. Moreover, one clearly has $(x' \circ f)' = x' \circ f'$.

Let us now show that f' is locally bounded: fix a point $z_0 \in \Omega$, and let γ by a positively oriented circle in Ω around z_0 with radius $r > 0$. Then Cauchy's integral formula for the derivative of holomorphic functions gives for every $z_1 \in \Omega$ that is enclosed by γ

$$\langle x', f'(z_1)\rangle = \frac{1}{2\pi \mathrm{i}} \int_\gamma \Big\langle x', \frac{f(z)}{(z - z_1)^2} \Big\rangle \,\mathrm{d}z,$$

and thus

$$\big|\langle x', f(z_1)\rangle\big| \le \frac{rM}{\operatorname{dist}(z_1, \gamma)},$$

where M is the supremum of f over γ. By taking the supremum over all $x' \in X'$ of norm ≤ 1 we thus obtain $\|f'(z_1)\| \le rM/\operatorname{dist}(z_1, \gamma)$. Thus, f' is locally bounded, so it follows from Theorem 2.3.5 that f' is holomorphic.

The claim for the n-th derivatives of f now follows readily by induction. □

Finally, we discuss the Taylor series expansion of holomorphic functions. As in the scalar-valued case, the

Theorem 2.3.7 (Taylor series expansion of holomorphic functions)**.** *Let X be a complex Banach space.*

(a) *Let $z_0 \in \mathbb{C}$ and consider a sequence $(a_n)_{n\in\mathbb{N}_0}$ in X which satisfies $r_0 := \liminf_{n\to\infty} \|a_n\|^{-1/n} > 0$. Define a function $f : \mathrm{B}_{<r_0}(z_0) \to X$ by*[12]

$$f(z) := \sum_{n=0}^\infty a_n (z - z_0)^n$$

for all $z \in \mathrm{B}_{<r_0}(z_0)$, where the series converges absolutely.[13]

Then f is holomorphic, and $f^{(n)}(z_0) = n!\, a_n$ for each $n \in \mathbb{N}_0$.

(b) *Let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be open, and let $f : \Omega \to X$ be holomorphic. Let $z_0 \in \Omega$ and $r \in (0, \infty]$ such that $\mathrm{B}_{<r}(z_0) \subseteq \Omega$. Define $a_n := f^{(n)}(z_0)/n!$ for each $n \in \mathbb{N}_0$.*

Then $\liminf_{n\to\infty} \|a_n\|^{-1/n} \ge r$, and one has

$$f(z) = \sum_{n=0}^\infty a_n (z - z_0)^n$$

for all $z \in \mathrm{B}_{<r}(z_0)$, where the series converges absolutely.

[12] Here, $\mathrm{B}_{<r_0}(z_0)$ denotes the open ball in $\mathbb{C}$ with center z_0 and radius r_0.

[13] Why?

(c) *Let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be open, and let $f : \Omega \to X$ be holomorphic. Let $z_0 \in \Omega$ and let $(a_n)_{n\in\mathbb{N}_0}$ be a sequence in X such that, for all z in some open ball with center z_0, the series $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ converges absolutely and is equal to $f(z)$.*

Then $a_n = f^{(n)}(z_0)/n!$ for each $n \in \mathbb{N}_0$.

Sketch of proof. (a) The absolute convergence of the series follows from the root test for series convergence.[14]

The other assertions can be shown similarly as in the scalar-valued case: one first shows a dominated convergence theorem for vector-valued series which allows one to interchange limits and series, and then one can immediately check the claims by using the definition of the complex derivative.[15]

(b) By testing against linear functionals $x' \in X'$ and using Cauchy's integral formula for the n-th derivatrive $f^{(n)}(z_0)$ (where one integrates along a circle with center z_0 and a radius $\tilde{r}$ that is slightly smaller than r) one can, similarly as in the proof of Corollary 2.3.6, estimate the coefficients $a_n = f^{(n)})(z_0)/n!$ to see that, indeed, $\liminf_{n\to\infty} \|a_n\|^{-1/n} \geq r$.

The formula for $f(z)$ follows from the scalar-valued case by testing against linear functionals.

(c) According to (a) the function $\tilde{f} : z \mapsto \sum_{n=0}^{\infty} a_n(z - z_0)^n$, defined on an open ball with center z_0, is holomorphic, and its n-th derivative at z_0 is given by $n!\, a_n$. But one has, by assumption, $f(z) = \tilde{f}(z)$ in a neighbourhood of z_0, so the n-th derivatives of f and $\tilde{f}$ at z_0 coincide. □

Example 2.3.8 (Taylor series expansion of the resolvent)**.** Let A be a unital Banach algebra and let $a \in A$. We know from Example 2.3.2 that $\mathcal{R}(\,\cdot\,, a)$ is a holomorphic mapping from $\rho(a)$ to A.

Now, fix a point $\lambda \in \rho(a)$ and set $r_0 := \|\mathcal{R}(\lambda, a)\|^{-1} \in (0, \infty]$. According to Proposition 2.2.4(c) the open ball with center λ and radius r_0 is contained in $\rho(a)$, and for each μ in this ball one has

$$\mathcal{R}(\mu, a) = \sum_{n=0}^{\infty} (\lambda - \mu)^n \mathcal{R}(\lambda, a)^{n+1},$$

where the series convergences absolutely. According to Theorem 2.3.7(c) this is the Taylor series expansion of the resolvent at λ, so $\mathcal{R}^{(n)}(\lambda, a) = (-1)^n n! \mathcal{R}(\lambda, a)^{n+1}$ for each integer $n \geq 0$.

In Corollary 2.4.7 in the next section we will come back to the previous example and have a closer look at the radius of convergence of the Taylor series of the resolvent.

[14] Note that this is a purely scalar-valued result, since absolute convergence of a series of vectors simply means that the series over the norms convergences.

[15] Alternatively, one can show that f is holomorphic by testing against functionals, using the same result for the scalar-valued case, and then employing Theorem 2.3.5.

2.4 The spectral radius

As you may known, one can derive the fundamental theorem of algebra[16] from Liouville's theorem: if $p : \mathbb{C} \to \mathbb{C}$ is a polynomial function that does not vanish anywhere on $\mathbb{C}$, then $1/p : \mathbb{C} \to \mathbb{C}$ is a holomorphic function which is bounded; hence, it is constant as a consequence of Liouvilles theorem and thus, p is constant. As a consequence of the fundamental theorem of algebra, every matrix in $\mathbb{C}^{n\times n}$ for $n \geq 0$ has an eigenvalue.

We will now use a very similar argument to show that every element in a non-zero unital Banach algebra has non-empty spectrum.

Theorem 2.4.1 (The spectrum is non-empty)**.** *Let A be a unital Banach algebra and $a \in A$. If $A \neq \{0\}$, then $\sigma(a) \neq \emptyset$.*

Proof. Assume for a contradiction that $\sigma(a) = \emptyset$. Then the resolvent of a is a holomorphic mapping from $\mathbb{C}$ to A and vanishes at ∞, as shown in Proposition 2.2.4(f). In particular, it is bounded and thus constant according to Liouville's theorem 2.3.4. As it vanishes at ∞, it follows that $\mathcal{R}(\lambda, a) = 0$ for all $\lambda \in \mathbb{C}$. But invertible elements of A are always non-zero, as $A \neq \{0\}$ by assumption; so we arrived at a contradiction. ☐

Recall that, on the other hand, in the unital Banach algebra $A = \{0\}$, the element 0 does indeed have empty spectrum.

Definition 2.4.2 (Spectral radius)**.** Let A be a Banach algebra and let $a \in A$. The number

$$\mathrm{r}(a) := \sup\{|\lambda| \mid \lambda \in \sigma(a)\}$$

is called the *spectral radius* of a.[17]

Note that the spectrum of an element a of a Banach algebra A is always compact according to Proposition 2.2.4(e). Hence, if A is non-zero, the supremum in the definition of the spectral radius is actually a maximum.

Remark 2.4.3. Let A be a unital Banach algebra and let $a \in A$. It follows immediately from Proposition 2.2.4(e) that $\mathrm{r}(a) \leq \|a\|$.

Note that one does not have equality between norm and spectral radius, in general. For instance, consider any norm on $\mathbb{C}^{2\times 2}$ which turns this space into a unital Banach algebra. The matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

[16] Which says that every non-constant polynomial over $\mathbb{C}$ has at least one root.

[17] Here, the supremum is meant to be taken within the ordered set $[0, \infty)$, such that the empty set has supremum 0. But according to Theorem 2.4.1 and Remark 2.2.2(a) this comment is only relevant if $A = \{0\}$.

has spectrum $\{0\}$ and thus spectral radius 0, but its norm is non-zero since the matrix is non-zero. However, we can compute that spectral radius of an element a in a unital Banach algebra from the norms of the powers a^n:

Theorem 2.4.4 (Spectral radius formula)**.** *Let A be a unital Banach algebra and $a \in A$. Then one has*

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

Proof. If $A = \{0\}$ there is nothing to prove, so assume that $A \neq \{0\}$. Then there exists a number $\lambda \in \sigma(a)$ such that $|\lambda| = \mathrm{r}(a)$. For each integer $n \geq 1$ we have

$$\lambda^n - a^n = (\lambda - a)\sum_{k=0}^{n-1} \lambda^k a^{n-1-k}.$$

Since $\lambda - a$ is not invertible and both factors on the right hand side commute, it follows from Lemma (c) that $\lambda^n - a^n$ is not invertible, i.e., $\lambda^n \in \sigma(a^n)$.[18] This shows that $\|a^n\| \geq \mathrm{r}(a^n) \geq |\lambda^n| = \mathrm{r}(a)^n$, so

$$\mathrm{r}(a) \leq \inf_{n\in\mathbb{N}} \|a^n\|^{1/n} \leq \liminf_{n\to\infty} \|a^n\|^{1/n}.$$

So it remains to show that

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leq \mathrm{r}(a).$$

Let $D \subseteq \mathbb{C}$ denote the open disk with center 0 and radius $1/\mathrm{r}(a) \in (0,\infty]$. For $\lambda \in D$ the element $1 - \lambda a$ of A is invertible. Indeed, for $\lambda = 0$ this is clear, and for $\lambda \neq 0$ this follows from $\mathrm{r}(a) < |1/\lambda|$ and from $1 - \lambda a = \lambda(1/\lambda - a)$. Consider the mapping $f : D \to A$ that is given by

$$f(\lambda) := (1 - \lambda a)^{-1}$$

for all $\lambda \in U$. For $\lambda \neq 0$ one has $f(\lambda) = \lambda^{-1}\mathcal{R}(\lambda^{-1}, a)$, so f is holomorphic on $D \setminus \{0\}$.[19] But for λ close to 0 one has $|\lambda|\,\|a\| < 1$, so due to the Neumann series,

$$f(\lambda) = \sum_{n=0}^{\infty} \lambda^n a^n.$$

Hence, f is also holomorphic in 0 and satisfies $f^{(n)}(0)/n! = a^n$ for each $n \in \mathbb{N}_0$ according to Theorem 2.3.7(a). Moreover, it follows from the fact that f is holomorphic on the disk D with radius $1/\mathrm{r}(a)$ and from Theorem 2.3.7(b) that

$$\frac{1}{\limsup_{n\to\infty} \|a^n\|^{1/n}} \geq 1/\mathrm{r}(a).$$

This proves the claim. □

[18] By the way, $\lambda^n \in \sigma(a^n)$ is also true for $n = 0$. Why?

[19] Since it is a composition of holomorphic functions.

We discuss a couple of nice consequences of the spectral radius formula in the subsequent corollaries.

Corollary 2.4.5 (Exponential growth)**.** *Let A be a unital Banach algebra and let $a \in A$. Let $r > \mathrm{r}(a)$. Then there exists a number $M \geq 1$ such that*

$$\|a^n\| \leq M r^n \qquad \text{for all } n \in \mathbb{N}_0.$$

Proof. One has $\mathrm{r}(a/r) = \mathrm{r}(a)/r < 1$.[20] Hence, the spectral radius formula shows that there exists an integer $n_0 \geq 0$ such that

$$\left\| \left(\frac{a}{r}\right)^n \right\|^{1/n} \leq 1$$

for all $n \geq n_0$. In other words, $\|a^n\| \leq r^n$ for all $n \geq n_0$. Choosing M as the maximum of 1 and each of numbers $\|a^n\|/r^n$ over $n \in \{0, \dots, n_0 - 1\}$ thus yields the claimed inequality for all $n \geq 0$. □

Corollary 2.4.6 (The Neumann series outside the spectral circle)**.** *Let A be a unital Banach algebra, let $a \in A$, and let $\lambda \in \mathbb{C}$ have modulus $|\lambda| > \mathrm{r}(a)$.*[21]

(a) *The resolvent $\mathcal{R}(\lambda, a)$ is given by the Neumann series*

$$\mathcal{R}(\lambda, a) = \sum_{n=0}^{\infty} \frac{a^n}{\lambda^{n+1}},$$

which converges absolutely.[22]

(b) *Let $|\lambda| > r \geq \mathrm{r}(a)$ and assume that $M \geq 1$ such that $\|a^n\| \leq M r^n$ for all $n \in \mathbb{N}_0$.*[23] *Then the resolvent $\mathcal{R}(\lambda, a)$ satisfies the norm estimate*

$$\|\mathcal{R}(\lambda, a)\| \leq \frac{M}{|\lambda| - r}.$$

Proof. (a) The absolute convergence of the series follows from the root test since, according to the spectral radius formula in Theorem 2.4.4,

$$\lim_{n\to\infty} \left\| \left(\frac{a}{\lambda}\right)^n \right\|^{1/n} = \mathrm{r}(a/\lambda) = \mathrm{r}(a)/|\lambda| < 1.$$

The fact that the series equals the resolvent follows as in the proof of Proposition 1.1.11.

(b) This is an immediate consequence of (a) and of the geometric series formula in $\mathbb{R}$. □

[20] The equality $\mathrm{r}(a/r) = \mathrm{r}(a)/r$ follows from the more general (but still simple) observation that $\sigma(a/r) = \sigma(a)/r$. Alternatively, it also follows from the spectral radius formula.

[21] So $\lambda \in \rho(a)$.

[22] Note that in Proposition 2.2.4(e) this was only shown under the assumption $|\lambda| > \|a\|$ – which turns out to be stronger than the assumption $|\lambda| > \mathrm{r}(a)$ since $\mathrm{r}(a) \leq \|a\|$.

[23] According to Proposition 2.4.5 such an M exists automatically if $r > \mathrm{r}(a)$. In the case $r = \mathrm{r}(a)$ such an M might or might not exist, depending on a.

Corollary 2.4.7 (The distance to the spectrum)**.** *Let A be a unital Banach algebra, let $a \in A$ and $\lambda \in \rho(a)$. Then*[24]

$$\frac{1}{\operatorname{dist}(\lambda, \sigma(a))} = \mathrm{r}(\mathcal{R}(\lambda, a)) \le \|\mathcal{R}(\lambda, a)\|,$$

and for all $\mu \in \mathbb{C}$ that satisfy $|\mu - \lambda| < \operatorname{dist}(\lambda, \sigma(a))$ one has

$$\mathcal{R}(\mu, a) = \sum_{n=0}^{\infty} (\lambda - \mu)^n \mathcal{R}(\lambda, a)^{n+1},$$

where the series converges absolutely.

Proof. We leave this as an exercise on Sheet 4. ☐

2.5 Pseudo resolvents

We will now generalize the concept of resolvents to the notion of so-called *pseudo-resolvents*; these are mappings from a set in the complex plane into a unital Banach algebra which satisfy the resolvent identity. Throughout this section we shall see that many properties of resolvents remain true for pseudo-resolvents.

Definition 2.5.1 (Pseudo resolvents)**.** Let A be a unital Banach algebra and let $\Omega \subseteq \mathbb{C}$ be non-empty. A mapping $\mathcal{R} : \Omega \to A$ is called a *pseudo-resolvent* if it satisfies the *resolvent identity*

$$\mathcal{R}(\lambda) - \mathcal{R}(\mu) = (\mu - \lambda)\mathcal{R}(\mu)\mathcal{R}(\lambda)$$

for all $\lambda, \mu \in \mathbb{C}$.

Note that, if $\mathcal{R}$ is a pseudo-resolvent, then $\mathcal{R}(\lambda)$ and $\mathcal{R}(\mu)$ commute for all λ, μ in the domain of $\mathcal{R}$; this follows readily from the resolvent identity. Note that, for every $\lambda_0 \in \mathbb{C}$, the shifted mapping

$$\mathcal{R}(\,\cdot\, - \lambda_0) : \Omega + \lambda_0 \to A$$

is also a pseudo-resolvent.

In every unital Banach algebra, the resolvent of any given element a is an example of a pseudo-resolvent (that is defined on the open set $\Omega := \rho(a)$). Another simple example of a pseudo-resolvent is the mapping $\mathbb{C} \to A$ that is constantly 0. Here are a few further examples (but we shall meet many more examples during the course):

Examples 2.5.2 (A few pseudo-resolvents)**.**

(a) Let $\lambda_0 \in \mathbb{C}$. The mapping $\mathcal{R} : \mathbb{C} \setminus \{\lambda_0\} \to \mathbb{C}$, $\lambda \mapsto \frac{1}{\lambda - \lambda_0}$ is a pseudo-resolvent in the algebra $\mathbb{C}$. In fact, this mapping is simply the resolvent of the element λ_0 of the algebra $\mathbb{C}$.

[24] Note that this strengthens the first part of Proposition 2.2.4(d).

(b) Let $\mathcal{R} : \mathbb{C} \setminus \{0\} \to \mathbb{C}^{2\times 2}$ be given by

$$\mathcal{R}(\lambda) := \begin{pmatrix} \frac{1}{\lambda} & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{for all } \lambda \in \mathbb{C} \setminus \{0\}.$$

Then $\mathcal{R}$ can readily be checked to be a pseudo-resolvent. However, it is not the resolvent of any element in $\mathbb{C}^{2\times 2}$ since, for instance, $\mathcal{R}(1)$ is not invertible.

(c) There exist constant non-zero pseudo-resolvents. For instance, the mapping $\mathcal{R} : \mathbb{C} \to \mathbb{C}^{2\times 2}$ that is given by

$$\mathcal{R}(\lambda) := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{for all } \lambda \in \mathbb{C}$$

is a pseudo-resolvent.

More generally, if A is a unital Banach algebra and $a \in A$ satisfies $a^2 = 0$, then $\mathcal{R} : \mathbb{C} \to A$, $\lambda \mapsto a$ is a pseudo-resolvent.

In particular, pseudo-resolvent need not converge to 0 at ∞ (in contrast to resolvents of elements of A, see Proposition 2.2.4(f)).

(d) There exist pseudo-resolvents which are defined everywhere on $\mathbb{C}$, but are not constant.[25] In fact, such an examples exist even in finite dimensions, as you will see on Exercise Sheet 4.

Here, we discuss an infinite-dimensional example instead: Let $\mathrm{C}_0((0,1])$ denote the Banach space of all continuous functions $f : (0,1] \to \mathbb{C}$ which satisfy $\lim_{t\downarrow 0} f(t) = 0$,[26] and let $A = \mathcal{L}\big(\mathrm{C}_0((0,1])\big)$. For every $\lambda \in \mathbb{C}$ consider the operator $\mathcal{R}(\lambda) : \mathrm{C}_0((0,1]) \to \mathrm{C}_0((0,1])$ that is given by

$$\big(\mathcal{R}(\lambda) f\big)(t) = e^{-t\lambda} \int_0^t e^{s\lambda} f(s) \,\mathrm{d}s$$

for all $f \in \mathrm{C}_0((0,1])$ and all $t \in (0,1]$. Then $\mathcal{R} : \mathbb{C} \to A$, $\lambda \mapsto \mathcal{R}(\lambda)$ is a pseudo-resolvent. Indeed, for all $\lambda, \mu \in \mathbb{C}$, $f \in \mathrm{C}_0((0,1])$ and $t \in (0,1]$ on has

$$\begin{aligned} \Big((\mu - \lambda)\mathcal{R}(\lambda)\mathcal{R}(\mu) f\Big)(t) &= (\mu - \lambda) e^{-t\lambda} \int_0^t e^{s\lambda} \big(\mathcal{R}(\mu) f\big)(s) \,\mathrm{d}s \\ &= (\mu - \lambda) e^{-t\lambda} \int_0^t e^{s(\lambda - \mu)} \int_0^s e^{r\mu} f(r) \,\mathrm{d}r \,\mathrm{d}s. \end{aligned}$$

With the notation $g(s) := e^{s(\lambda-\mu)}$ and $h(s) := \int_0^s e^{r\mu} f(r) \,\mathrm{d}r$ for all $s \in (0,1]$ we thus obtain

$$\Big((\mu - \lambda)\mathcal{R}(\lambda)\mathcal{R}(\mu) f\Big)(t) = e^{-t\lambda} \int_0^t -g'(s) h(s) \,\mathrm{d}s$$

[25] It follows from Liouville's theorem that such a pseudo-resolvent cannot be bounded since we will show in Proposition 2.5.10 that pseudo-resolvents are always holomorphic.

[26] Note that we can identify $\mathrm{C}_0((0,1])$ with the subspace of $\mathrm{C}([0,1])$ of functions that vanish at the point 0.

$$
\begin{aligned}
&= e^{-t\lambda}\big[g(0)h(0) - g(t)h(t)\big] + e^{-t\lambda}\int_0^t g(s)h'(s)\,\mathrm{d}s \\
&= -e^{-t\mu}\int_0^t e^{r\mu}f(r)\,\mathrm{d}r + e^{-t\lambda}\int_0^t e^{s\lambda}f(s)\,\mathrm{d}s \\
&= \Big(\mathcal{R}(\lambda) - \mathcal{R}(\mu))f\Big)(t).
\end{aligned}
$$

So $\mathcal{R}$ is a pseudo-resolvent as claimed. Moreover, $\mathcal{R}$ is not the resolvent of an element of $A = \mathcal{L}\big(\mathrm{C}_0((0,1])\big)$, as it is defined everywhere on $\mathbb{C}$ and all elements of A have non empty spectrum.[27,28] One can readily check that $\mathcal{R}$ is not constant[29] – so we found an example of a non-constant pseudo-resolvent that is defined on all of $\mathbb{C}$.

Here is a nice auxiliary result that will turn out to be quite useful in the study of pseudo-resolvents:

Lemma 2.5.3. *Let A be a unital Banach algebra, let $r, \tilde{r} \in A$, and let $\lambda, \tilde{\lambda} \in \mathbb{C}$. The following assertions are equivalent:*

(i) *The identity $r - \tilde{r} = (\tilde{\lambda} - \lambda)r\tilde{r}$ holds.*

(ii) *The subsets*

$$
G := \left\{ \begin{pmatrix} ra \\ \lambda ra - a \end{pmatrix} \mid a \in A \right\} \quad \text{and} \quad \tilde{G} := \left\{ \begin{pmatrix} \tilde{r}\tilde{a} \\ \tilde{\lambda}\tilde{r}\tilde{a} - \tilde{a} \end{pmatrix} \mid \tilde{a} \in A \right\}
$$

of $A \times A$ satisfy $\tilde{G} \subseteq G$.

Proof. „(i) $\Rightarrow$ (ii)“ Let $\tilde{w} \in \tilde{G}$. Then there exists $\tilde{a} \in A$ such that

$$
\tilde{w} = \begin{pmatrix} \tilde{r}\tilde{a} \\ \tilde{\lambda}\tilde{r}\tilde{a} - \tilde{a} \end{pmatrix}.
$$

Let us define $a := \tilde{a} + (\lambda - \tilde{\lambda})\tilde{r}\tilde{a} \in A$. Then we have

$$
ra = \big(r + \big(\lambda - \tilde{\lambda}\big)r\tilde{r}\big)\tilde{a} = \tilde{r}\tilde{a}
$$

due to the identity in (i). Thus,

$$
\lambda ra - a = \lambda\tilde{r}\tilde{a} - \tilde{a} - (\lambda - \tilde{\lambda})\tilde{r}\tilde{a} = \tilde{\lambda}\tilde{r}\tilde{a} - \tilde{a}.
$$

[27] More precisely speaking, the resolvent of an element of $T \in A$ cannot be the restriction of the pseudo-resolvent $\mathcal{R}$ to $\rho(T)$ since $\mathcal{R}(\,\cdot\,, T)$ explodes closed to the boundary of $\sigma(T)$, while $\mathcal{R}$ is continuous everywhere on $\mathbb{C}$ (as we will show for all pseudo-resolvents in Proposition 2.5.10).

[28] However, we will see later on that this specific example of a pseudo-resolvent $\mathcal{R}$ is actually the resolvent of a so-called *unbounded operator* on $\mathrm{C}_0((0,1])$ – a topic that we will discuss in detail in Section 4.1.

[29] And it is not difficult to check also that $\|\mathcal{R}(\lambda)\| \to \infty$ as, for instance, λ approaches $-\infty$ from within $\mathbb{R}$.

This shows that

$$\tilde{w} = \begin{pmatrix} \tilde{r}\tilde{a} \\ \tilde{\lambda}\tilde{r}\tilde{a} - \tilde{a} \end{pmatrix} = \begin{pmatrix} ra \\ \lambda ra - a \end{pmatrix} \in G,$$

as claimed.

„(ii) $\Rightarrow$ (i)“ We use the neutral element $1 \in A$, which gives us that $\begin{pmatrix} \tilde{r} \\ \tilde{\lambda}\tilde{r} - 1 \end{pmatrix}$ is an element of $\tilde{G}$. Thus, the same element is also in G, so it can be written as

$$\begin{pmatrix} \tilde{r} \\ \tilde{\lambda}\tilde{r} - 1 \end{pmatrix} = \begin{pmatrix} ra \\ \lambda ra - a \end{pmatrix}$$

for some $a \in A$. By substituting the equality of the first components into the equality of the second components we get $\tilde{\lambda}\tilde{r} - 1 = \lambda\tilde{r} - a$. We now multiply with r from the left and again use that $ra = \tilde{r}$, which yields

$$\tilde{\lambda}r\tilde{r} - r = \lambda r\tilde{r} - \tilde{r},$$

which is precisely the claimed identity. $\square$

Corollary 2.5.4. *Let A be a unital Banach algebra, let $r, \tilde{r} \in A$, and let $\lambda, \tilde{\lambda} \in \mathbb{C}$. Let G and $\tilde{G}$ be defined as in Lemma 2.5.3*(ii)*. The following assertions are equivalent:*

(i) *The elements r and $\tilde{r}$ commute and the identity $r - \tilde{r} = (\tilde{\lambda} - \lambda)r\tilde{r}$ holds.*

(ii) *Both identities $r - \tilde{r} = (\tilde{\lambda} - \lambda)r\tilde{r}$ and $\tilde{r} - r = (\lambda - \tilde{\lambda})\tilde{r}r$ hold.*

(iii) *One has $\tilde{G} = G$.*

Proof. „(i) $\Rightarrow$ (ii)“ This is straightforward.

„(ii) $\Rightarrow$ (i)“ This is easy to see if one distinguishes between the two cases $\lambda = \tilde{\lambda}$ and $\lambda \neq \tilde{\lambda}$.

„(ii) $\Leftrightarrow$ (iii)“ This follows readily from Lemma 2.5.3(ii). $\square$

By using the preceding lemma and corollary we can now prove two nice theorems about pseudo-resolvents.

Theorem 2.5.5 (Uniqueness for pseudo-resolvents)**.** *Let A be a unital Banach algebra, let $\Omega \subseteq \mathbb{C}$ be non-empty, and let $\mathcal{R}_1, \mathcal{R}_2 : \Omega \to A$ be pseudo-resolvents.*

If there exists $\lambda_0 \in \Omega$ such that $\mathcal{R}_1(\lambda_0) = \mathcal{R}_2(\lambda_0)$, then $\mathcal{R}_1 = \mathcal{R}_2$.

Proof. As $\mathcal{R}_1$ is a pseudo-resolvent, it follows from Corollary 2.5.4 that the set

$$G_1(\lambda) := \left\{ \begin{pmatrix} \mathcal{R}_1(\lambda)a \\ \lambda\mathcal{R}_1(\lambda)a - a \end{pmatrix} \mid a \in A \right\}$$

does not depend on λ (for $\lambda \in \Omega$). Similarly, the set

$$G_2(\lambda) := \left\{ \begin{pmatrix} \mathcal{R}_2(\lambda)a \\ \lambda\mathcal{R}_2(\lambda)a - a \end{pmatrix} \mid a \in A \right\},$$

which is also defined for $\lambda \in \Omega$, does not depend on λ since $\mathcal{R}_2$ is a pseudo-resolvent. Now, fix $\lambda \in \Omega$. Then we have

$$G_1(\lambda) = G_1(\lambda_0) = G_2(\lambda_0) = G_2(\lambda),$$

where the equality in the middle follows from the assumption $\mathcal{R}_1(\lambda_0) = \mathcal{R}_2(\lambda_0)$. Now we use the neutral element 1 of A. The tuple

$$\begin{pmatrix} \mathcal{R}_1(\lambda) \\ \lambda\mathcal{R}_1(\lambda) - 1 \end{pmatrix}$$

is in $G_1(\lambda)$ and thus in $G_2(\lambda)$, so there exists an element $a \in A$ such that

$$\begin{pmatrix} \mathcal{R}_1(\lambda) \\ \lambda\mathcal{R}_1(\lambda) - 1 \end{pmatrix} = \begin{pmatrix} \mathcal{R}_2(\lambda)a \\ \lambda\mathcal{R}_2(\lambda)a - a \end{pmatrix}.$$

If we multiply the equality in the first line, $\mathcal{R}_1(\lambda) = \mathcal{R}_2(\lambda)a$, by λ and then substract it from the equality in the second line, we obtain $1 = a$. By substituting this into the equality in the first line again, we finally get $\mathcal{R}_1(\lambda) = \mathcal{R}_2(\lambda)$. □

Next we observe that, in order to have a pseudo-resolvent, it suffices to show the resolvent identity with respect to a fixed pivot point:

Proposition 2.5.6. *Let A be a unital Banach algebra and let $\Omega \subseteq \mathbb{C}$ be non-empty. Let $\mathcal{R} : \Omega \to A$ be a mapping and assume that there exists a number $\lambda_0 \in \Omega$ with the following property: the resolvent identity*

$$\mathcal{R}(\lambda) - \mathcal{R}(\lambda_0) = (\lambda_0 - \lambda)\mathcal{R}(\lambda)\mathcal{R}(\lambda_0)$$

holds for every $\lambda \in \Omega$ and the elements $\mathcal{R}(\lambda)$ and $\mathcal{R}(\lambda_0)$ commute for every $\lambda \in \Omega$.

Then $\mathcal{R}$ is a pseudo-resolvent.

Proof. For every $\lambda \in \Omega$ define

$$G(\lambda) := \left\{ \begin{pmatrix} \mathcal{R}(\lambda)a \\ \lambda\mathcal{R}(\lambda)a - a \end{pmatrix} \mid a \in A \right\}$$

Then it follows from Corollary 2.5.4 that $G(\lambda) = G(\lambda_0)$ for all $\lambda \in \Omega$. Hence one has $G(\lambda) = G(\mu)$ for all $\lambda, \mu \in \Omega$, which in turn proves the resolvent identity between the point λ and μ according to Corollary 2.5.4. □

Corollary 2.5.7. *Let A be a unital Banach algebra and let $\Omega_1, \Omega_2 \subseteq \mathbb{C}$ be two sets with non-empty intersection. Let $\mathcal{R}_j : \Omega_j \to A$ be pseudo-resolvents for $j \in \{1, 2\}$ and assume that there exists $\lambda_0 \in \Omega_1 \cap \Omega_2$ such that $\mathcal{R}_1(\lambda_0) = \mathcal{R}_2(\lambda_0)$.*

Then there exists precisely one pseudo-resolvent $\mathcal{R} : \Omega_1 \cup \Omega_2 \to A$ that extends both $\mathcal{R}_1$ and $\mathcal{R}_2$.

Proof. It follows from Theorem 2.5.5 that $\mathcal{R}_1$ and $\mathcal{R}_2$ coincide when restricted to $\Omega_1 \cap \Omega_2$. Hence, there exists a (obviously unique) function $\mathcal{R} : \Omega_1 \cup \Omega_2 \to A$ which extends both $\mathcal{R}_1$ and $\mathcal{R}_2$. If we fix a point $\lambda_0 \in \Omega_1 \cap \Omega_2$ and apply Proposition 2.5.6 we see that $\mathcal{R}$ is also a pseudo-resolvent. $\square$

We continue with the second theorem that was promised after Corollary 2.5.4.

Theorem 2.5.8 (Largest extension of a pseudo-resolvent)**.** *Let A be a unital Banach algebra, let $\Omega \subseteq \mathbb{C}$ be non-empty, and let $\mathcal{R} : \Omega \to A$ be a pseudo-resolvent.*

Then there exists a largest set $\tilde{\Omega} \subseteq \mathbb{C}$ for which there exists a pseudo-resolvent $\tilde{R} : \tilde{\Omega} \to A$ that satisfies $\tilde{R}|_\Omega = R$. Both $\tilde{\Omega}$ and $\tilde{R}$ are uniquely determined.

Proof. Let $\mathcal{U}$ denote the set of all non-empty subsets of $\mathbb{C}$ which contain Ω and on which there exists a pseudo-resolvent that extends $\mathcal{R}$. Note that for each $U \in \mathcal{U}$ there exists precisely one pseudo-resolvent that extends $\mathcal{R}$; this follows from the uniqueness result in Theorem 2.5.5.

We now set $\tilde{\Omega} := \bigcup \mathcal{U}$. If $U_1, U_2 \in \mathcal{U}$ and $\mathcal{R}_j : U_j \to A$ is, for each $j \in \{1, 2\}$, a pseudo-resolvent that extends $\mathcal{R}$, then the restrictions of both $\mathcal{R}_1$ and $\mathcal{R}_2$ to the intersection $U_1 \cap U_2$ are pseudo-resolvents on $U_1 \cap U_2$ that extend $\mathcal{R}$, and hence those restrictions coincide, again according to Theorem 2.5.5.

Thus, we can glue together all extensions of $\mathcal{R}$ to a mapping $\tilde{\mathcal{R}} : \tilde{U} \to A$ that extends $\mathcal{R}$. By fixing a point $\lambda_0 \in \Omega$ and applying Proposition 2.5.6 we see that $\tilde{\mathcal{R}}$ is a pseudo-resolvent.

By construction, $\tilde{\Omega}$ is the largest set in $\mathbb{C}$ on which there exists a pseuo-resolvent that extends $\mathcal{R}$. Moreover, it follows directly from the property "largest" that $\tilde{\Omega}$ is unique. The uniqueness of $\tilde{\mathcal{R}}$ is, again, due to Theorem 2.5.5. $\square$

The previous theorem enables us to also introduce a meaningful notion of *spectrum* for pseudo-resolvents:

Definition 2.5.9 (Spectrum of a pseudo-resolvent)**.** Let A be a unital Banach algebra, let $\Omega \subseteq \mathbb{C}$ be non-empty, and let $\mathcal{R} : \Omega \to A$ be a pseudo-resolvent.

The set $\tilde{\Omega}$ from Theorem 2.5.8 is called the *domain of the largest extension of* $\mathcal{R}$, and the pseudo-resolvent $\tilde{\mathcal{R}}$ from the same theorem is called the *largest extension of* $\mathcal{R}$. We call the set $\mathbb{C} \setminus \tilde{\Omega}$ the *spectrum of* $\mathcal{R}$.

Proposition 2.5.10 (Properties of pseudo-resolvents)**.** *Let A be a unital Banach algebra, and let $\rho \subseteq \mathbb{C}$ be non-empty, and let $\mathcal{R} : \rho \to A$ be a pseudo-resolvent; assume that ρ is already the domain of the largest extension of $\mathcal{R}$,*[30] *i.e., that $\sigma := \mathbb{C} \setminus \rho$ is the spectrum of $\mathcal{R}$.*

(a) *Let $\lambda_0 \in \rho$ and set*

$$\Omega_{\lambda_0} := \Big\{ \mu \in \mathbb{C} \mid \ (\mu - \lambda_0)\mathcal{R}(\lambda_0) + 1 \text{ is invertible} \Big\}.$$

[30] Otherwise, we can extend it to become this set.

Then Ω_{λ_0} *is open, one has* $\lambda_0 \in \Omega_{\lambda_0} \subseteq \rho$, *and*

$$\mathcal{R}(\mu) = \mathcal{R}(\lambda_0)\Big((\mu - \lambda_0)\mathcal{R}(\lambda_0) + 1\Big)^{-1}.$$

for every $\mu \in \Omega_{\lambda_0}$. *In particular we have the* spectral inclusion theorem *for pseudo-resolvents, i.e.,*[31]

$$\sigma\big(\mathcal{R}(\lambda_0)\big) \setminus \{0\} \supseteq \left\{ \frac{1}{\lambda_0 - \lambda} \mid \lambda \in \sigma \right\}.$$

(b) *The set* ρ *is open and thus, the spectrum* σ *is closed.*

(c) *For every* $\lambda \in \rho$ *one has*

$$\frac{1}{\operatorname{dist}(\lambda, \sigma)} = \mathrm{r}(\mathcal{R}(\lambda)) \leq \|\mathcal{R}(\lambda)\|.$$

If $\mu \in \mathbb{C}$ *and* $|\mu - \lambda| < \operatorname{dist}(\lambda, \sigma)$, *then*

$$\mathcal{R}(\mu) = \sum_{n=0}^{\infty} (\lambda - \mu)^n \mathcal{R}(\lambda)^{n+1},$$

where the series converges absolutely.

(d) *The mapping* $\mathcal{R}$ *is holomorphic (and thus, in particular, continuous) on* ρ.

For every $\lambda \in \rho$ *and every* $n \in \mathbb{N}_0$ *the* n*-th derivative of* $\mathcal{R}$ *at* λ *is given by* $\mathcal{R}^{(n)}(\lambda) = (-1)^n n!\, \mathcal{R}(\lambda)^{n+1}$.

Proof. (a) The set Ω_{λ_0} is open since $\operatorname{Inv}(A)$ is open, and $\lambda_0 \in \Omega_{\lambda_0}$ since the netural element 1 of A is invertible. Moreover,

$$\mathcal{R}_{\lambda_0} : \Omega_{\lambda_0} \ni \mu \mapsto \mathcal{R}(\lambda_0)\Big((\mu - \lambda_0)\mathcal{R}(\lambda_0) + 1\Big)^{-1} \in A$$

is a pseudo-resolvent; indeed, for $\lambda_0 = 0$ this follows from Exercise 5 on Sheet 4, and for general λ_0 one can shift the entire situation by $-\lambda_0$ to reduce it to the case $\lambda_0 = 0$.

Note that $\mathcal{R}(\lambda_0) = \mathcal{R}_{\lambda_0}(\lambda_0)$. Thus, it follows from Corollary 2.5.7 and from the assumption that ρ be the domain of the largest extension of $\mathcal{R}$, that $\Omega_{\lambda_0} \subseteq \rho$. Furthermore, again by the corollary, $\mathcal{R}_{\lambda_0}$ coincides with $\mathcal{R}$ on Ω_{λ_0}. This prove the first part of (a).

The spectral inclusion result can be easily derived from $\Omega_{\lambda_0} \subseteq \rho$.

[31] We will show later on in Theorem 3.3.5 by means of the *Gelfand representation* that this inclusion is actually even an equality. This is important in order to note that, if a pseudo-resolvent $\mathcal{R}$ is actually the resolvent of an element $a \in A$, then the spectrum of $\mathcal{R}$ coincides with $\sigma(a)$, see Corollary 3.3.6.

(b) For every $\lambda_0 \in \rho$ it follows from (a) that ρ contains the open neighbourhood Ω_{λ_0} of λ_0. Hence, ρ is open.

(c) and (d) Let $\lambda \in \rho$. It follows from the Neumann series expansion result in Corollary 2.4.6(a) that the set $\Omega_\lambda \subseteq \rho$ contains the open ball with center λ and radius $\frac{1}{\mathrm{r}(\mathcal{R}(\lambda))}$, and that, for every μ in this ball,

$$\mathcal{R}(\mu) = \mathcal{R}_\lambda(\mu) = \sum_{n=0}^{\infty} (\lambda - \mu)^n \mathcal{R}(\lambda)^{n+1},$$

where the series converges absolutely.

So we showed, in particular, that $\operatorname{dist}(\lambda, \sigma) \geq \frac{1}{\mathrm{r}(\mathcal{R}(\lambda))}$, Moreover, the series expansion that we showed for $\mathcal{R}(\mu)$ in case that $|\mu - \lambda| < \frac{1}{\mathrm{r}(\mathcal{R}(\lambda))}$ yields, according to Theorem 2.3.7(a), that $\mathcal{R}$ is holomorphic and that $\mathcal{R}^{(n)}(\lambda) = (-1)^n n!\, \mathcal{R}(\lambda)^{n+1}$ for each $n \in \mathbb{N}_0$. This proves (d).

Finally, due to the holomorphy we can can use Theorem 2.3.7(b), which tells us that

$$\operatorname{dist}(\lambda, \sigma) \leq \frac{1}{\limsup_{n\to\infty} \|\mathcal{R}(\lambda)^{n+1}\|^{1/n}} = \frac{1}{\mathrm{r}(\mathcal{R}(\lambda))},$$

where the equality at the end follows from the spectral radius formula in Theorem 2.4.4.[32] □

2.6 Continuity properties of the spectrum

Let A be a Banach algebra. In this section we discuss continuity properties of the spectrum $\sigma(a)$ in dependence of the element $a \in A$. Since the spectrum is a set, we need to make sense of the notion of *continuity* of set-valued mappings. To this end we first consider the so-called *Hausdorff distance* between subsets of metric spaces in Definition 2.6.1.

Recall that, for a metric space (M, d), a point $x \in M$, and a set $A \subseteq M$, the *distance of x to A* ist defined as

$$\operatorname{dist}(x, A) := \inf\{\mathrm{d}(x, a) \mid a \in A\} \in [0, \infty].$$

This infimum is understood to be taken within the ordered set $[0, \infty]$, so the empty set has infimum ∞. One has $\operatorname{dist}(x, A) = 0$ if and only if x is an element of the closure $\overline{A}$ of A.

Definition 2.6.1 (Hausdorff distance)**.** Let (M, d) be a metric space.

(a) For all sets $A, B \subseteq M$ we define

$$\mathrm{d}_{\subseteq}(A, B) := \sup\{\operatorname{dist}(a, B) \mid a \in A\} \in [0, \infty].$$

[32] Actually, the spectral radius formula rather tells us that $\limsup_{n\to\infty} \|\mathcal{R}(\lambda)^n\|^{1/n} = \mathrm{r}(\mathcal{R}(\lambda))$; why does it not change anything to consider $\|\mathcal{R}(\lambda)^{n+1}\|^{1/n}$ instead?

(b) For all non-empty sets $A, B \subseteq M$ we define[33]

$$\mathrm{d}_{\mathrm{Haus}}(A, B) := \mathrm{d}_{\subseteq}(A, B) \vee \mathrm{d}_{\subseteq}(B, A) \in [0, \infty];$$

this number is called the *Hausdorff distance between A and B*.

Note that for a subset A and a point x in a metric space, we always have

$$\operatorname{dist}(x, A) = \mathrm{d}_{\subseteq}(\{x\}, A) \leq \mathrm{d}_{\mathrm{Haus}}(\{x\}, A),$$

where the inequality is, in general, not an equality.[34] In the following two propositions we list a number of properties for $\mathrm{d}_{\subseteq}$ and $\mathrm{d}_{\mathrm{Haus}}$.

Proposition 2.6.2 (Properties of $\mathbf{d}_{\subseteq}$). *Let (M, d) be a metric space and let $A, B, C \subseteq M$ be non-empty.*

(a) *One has $\mathrm{d}_{\subseteq}(A, B) = 0$ if and only if $A \subseteq \overline{B}$ (or equivalently, $\overline{A} \subseteq \overline{B}$).*

(b) *Assume that A and B are closed.*[35] *Then*

$$\mathrm{d}_{\subseteq}(A, B) = 0 \qquad \text{if and only if} \qquad A \subseteq B.$$

(c) *One has the triangle inequality $\mathrm{d}_{\subseteq}(A, C) \leq \mathrm{d}_{\subseteq}(A, B) + \mathrm{d}_{\subseteq}(B, C)$.*

Proof. (a) and (b) Assertion (a) readily follows from the properties of distances between points and sets that we listed before Definition 2.6.1, and assertion (b) is an immediate consequence of (a).

(c) For all $a \in A$, $b \in B$, and $c \in C$ one has

$$\mathrm{d}(a, c) \leq \mathrm{d}(a, b) + \mathrm{d}(b, c).$$

Taking the infimum over $c \in C$ on both sides one thus yields

$$\operatorname{dist}(a, C) \leq \mathrm{d}(a, b) + \operatorname{dist}(b, C) \leq \mathrm{d}(a, b) + \mathrm{d}_{\subseteq}(B, C).$$

Next we take the infimum over $b \in B$ on the right hand side and thus obtain

$$\operatorname{dist}(a, C) \leq \operatorname{dist}(a, B) + \mathrm{d}_{\subseteq}(B, C).$$

Finally, let us take with supremum over a on both sides of the previous inequality; this gives

$$\mathrm{d}_{\subseteq}(A, C) \leq \mathrm{d}_{\subseteq}(A, B) + \mathrm{d}_{\subseteq}(B, C),$$

as claimed. □

[33] Here, we use the notation $x \vee y := \max\{x, y\}$ for all $x, y \in [-\infty, \infty]$.
[34] Can you give a counterexample which demonstrates this?
[35] Or, more generally, assume merely that B is closed.

Proposition 2.6.3 (The Hausdorff distance as an extended metric)**.** *Let* (M, d) *be a metric space. Then* $\mathrm{d}_{\mathrm{Haus}}$ *is an* extended metric[36] *on the set of all closed subsets of* M.

Proof. Let $A, B, C \subseteq M$ be closed. It follows immediately from Proposition 2.6.2(b) that $\mathrm{d}_{\mathrm{Haus}}(A, B) = 0$ if and only if $A = B$. Moreover, the definition of the Hausdorff distance readily implies that it is symmetric, i.e., $\mathrm{d}_{\mathrm{Haus}}(A, B) = \mathrm{d}_{\mathrm{Haus}}(B, A)$. Finally, we have

$$\begin{aligned}
\mathrm{d}_{\mathrm{Haus}}(A, C) &= \mathrm{d}_{\subseteq}(A, C) \vee \mathrm{d}_{\subseteq}(C, A) \\
&\leq \big(\mathrm{d}_{\subseteq}(A, B) + \mathrm{d}_{\subseteq}(B, C) \big) \vee \big(\underbrace{\mathrm{d}_{\subseteq}(C, B) + \mathrm{d}_{\subseteq}(B, A)}_{= \mathrm{d}_{\subseteq}(B,A) + \mathrm{d}_{\subseteq}(C,B)} \big) \\
&\leq \mathrm{d}_{\subseteq}(A, B) \vee \mathrm{d}_{\subseteq}(B, A) + \mathrm{d}_{\subseteq}(B, C) \vee \mathrm{d}_{\subseteq}(C, B) \\
&= \mathrm{d}_{\mathrm{Haus}}(A, B) + \mathrm{d}_{\mathrm{Haus}}(B, C),
\end{aligned}$$

where the first inequality follows from Proposition 2.6.2(c), and the second inequality follows from the fact that $(x_1 + x_2) \vee (y_1 + y_2) \leq x_1 \vee y_1 + x_2 \vee y_2$ for all $x_1, x_2, y_1, y_2 \in (-\infty, \infty]$. ☐

Example 2.6.4 (Discontinuity of the spectrum)**.** There exists an operator T and a sequence of operators $(T_n)_{n \in \mathbb{N}}$ in the unital Banach algebra $\mathcal{L}(\ell^2(\mathbb{Z}))$ such that $T_n \to T$ with respect to the operator norm and such that

$$\begin{aligned}
\sigma(T) = \overline{\mathbb{D}} &:= \{\lambda \in \mathbb{C} \mid \ |\lambda| \leq 1\} \quad \text{and} \\
\sigma(T_n) = \mathbb{T} &:= \{\lambda \in \mathbb{C} \mid \ |\lambda| = 1\} \quad \text{for all } n \in \mathbb{N}.
\end{aligned}$$

Hence, $\mathrm{d}_{\mathrm{Haus}}(\sigma(T_n), \sigma(T)) \not\to 0$. We will discuss an explicit example of such operators in the exercises.

In the following theorem we endow the complex plane, as usual, with the Euclidean distance.

Theorem 2.6.5 (Semi-continuity of the spectrum)**.** *Let* A *be a Banach algebra and let* $(a_j)_{j \in J}$ *be a net in* A *that converges to an element* $a \in A$. *Then*

$$\mathrm{d}_{\subseteq} \big(\sigma(a_j), \sigma(a) \big) \to 0$$

Proof. We may assume that A is a unital Banach algebra.[37] Fix $\varepsilon > 0$, and set

$$D_\varepsilon := \{\lambda \in \mathbb{C} \mid \ \mathrm{dist}(\lambda, \sigma(a)) \geq \varepsilon\}.$$

Since the resolvent of a is continuous on the resolvent set and converges to 0 at ∞, it follows that[38]

$$M := \sup_{\lambda \in D_\varepsilon} \|\mathcal{R}(\lambda, a)\| \in [0, \infty).$$

[36] I.e., it satisfies all axioms of a metric, but is allowed to also take the value ∞.
[37] Why?
[38] When can it happen that $M = 0$?

For all sufficiently large j, say $j \succeq j_0$, we have $\|a_j - a\| < 1/M$. Fix such an index j. We are going to show that $D_\varepsilon \subseteq \rho(a_j)$; to this end, let $\lambda \in D_\varepsilon$. Then we have $\lambda \in \rho(a)$ and

$$\|(\lambda - a_j) - (\lambda - a)\| = \|a_j - a\| < \frac{1}{M} \le \frac{1}{\|\mathcal{R}(\lambda, a)\|}.$$

Thus, it follows from Corollary 1.1.12 that $\lambda - a_j$ is also invertible, i.e., $\lambda \in \rho(a_j)$, as claimed.

So every spectral value of a_j is closer to $\sigma(a)$ than ε, i.e., $\mathrm{d}_{\subseteq}(\sigma(a_j), \sigma(a)) \le \varepsilon$.[39] $\square$

Finally we show that the spectrum is continuous for matrices in finite dimensions. In fact, the finite-dimensional case has the particularly nice feature that we can even give a quantitative estimate for the continuity. We need the following lemma.

Lemma 2.6.6 (Upper resolvent estimate in finite dimensions)**.** *Let $d \ge 1$ be an integer and let $\mathbb{C}^{d\times d}$ be endowed with any norm that turns it into a unital Banach algebra.*

(a) *There exists a constant*[40] $\gamma \ge 1$ *such that*

$$\left\|T^{-1}\right\| \le \gamma \frac{\|T\|^{d-1}}{|\det(T)|}.$$

for every invertible $T \in \mathbb{C}^{d\times d}$ and, thus,

$$\|\mathcal{R}(\lambda, T)\| \le \gamma \frac{(\|T\| + |\lambda|)^{d-1}}{\operatorname{dist}(\lambda, \sigma(T))^d}$$

for every $T \in \mathbb{C}^{d\times d}$ and every $\lambda \in \rho(T)$.

(b) *If $p \in [1, \infty)$ and $\mathbb{C}^{d\times d}$ is endowed with the operator norm induced by the p-norm on $\mathbb{C}^d$, then γ can be chosen as d.*[41]

(c) *If $\mathbb{C}^{d\times d}$ is endowed with the operator norm induced by the 2-norm on $\mathbb{C}^d$, then γ can even be chosen as 1.*[42]

Proof. (a) The first estimate in the second displayed formula readily follows from the estimate in the first formula if we substitute $\lambda - T$ for T and use that $\|\lambda - T\| \le$

[39] By the way, this inequality is in fact even strict. Why?

[40] Which depends on the choice of the norm.

[41] For $p = 2$ this estimate is not optimal, see assertion (c). For $p \in (1, \infty) \setminus \{2\}$ I suspect that it is not optimal, but I do not know how to obtain a better estimate in these cases. For $p \in \{1, \infty\}$ I do not have an intuition on whether the estimate is optimal.

[42] Which is particularly interesting since γ does not depend on the dimension d then.

$|\lambda| + \|T\|$ (since we assume the norm to turn $\mathbb{C}^{d\times d}$ into a unital Banach algebra) and that

$$|\det(\lambda - T)| = \prod_{j=1}^{d} |\lambda - \lambda_j| \geq \operatorname{dist}(\lambda, \sigma(T))^d,$$

where $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of T (counted with their algebraic multiplicities). So it suffices to prove that first estimate. Since all norms are equivalent on $\mathbb{C}^{d\times d}$, it thus suffices to prove assertion (b) (or (c)).

(b) We only need to show the first estimate in (a) for $\gamma = d$. To this end we first observe that, for every $M \in \mathbb{C}^{d\times d}$, one has $|\det M| \leq \mathrm{r}(M)^d$; indeed, this follows immediately from the fact that the determinant is the product of the eigenvalues of M.

Now we prove the claimed estimate for $\left\|T^{-1}\right\|$. Cramer's rule tells us that $T^{-1} = S/\det(T)$, where $S \in \mathbb{C}^{d\times d}$ is the matrix whose entryies are given by $S_{j,k} = (-1)^{j+k}\det(T^{(j,k)})$, where $T^{(j,k)} \in \mathbb{C}^{(d-1)\times(d-1)}$ is the matrix that one obtains from T be removing the j-th row and the k-th column. Hence,

$$|S_{j,k}| = \left|\det(T^{(j,k)})\right| \leq \mathrm{r}(T^{(j,k)})^{d-1} \leq \left\|T^{(j,k)}\right\|^{d-1} \leq \|T\|^{d-1};$$

here we also use the norm on $\mathbb{C}^{(d-1)\times(d-1)}$ that is induced by the p-norm on $\mathbb{C}^{(d-1)}$, and the last inequality follows from this choice of the norms on $\mathbb{C}^{(d-1)\times(d-1)}$ and $\mathbb{C}^{d\times d}$, respectively.

Now let $|S| \in \mathbb{C}^{d\times d}$ denote the matrix whose entries are the moduli of the entries of S. Then, again due to our particular choice of the norms,

$$\|S\| \leq \||S|\| \leq \left\|\|T\|^{d-1} E\right\| = \|T\|^{d-1}\|E\|,$$

where $E \in \mathbb{C}^{d\times d}$ denotes the matrix whose entries are all 1.

So it only remains to show that $\|E\| \leq d$. To this end, let $x \in \mathbb{C}^d$ such that $\|x\|_p = 1$. Let $e \in \mathbb{C}^d$ denote the vector whose elements are all 1 and let $p' \in [1,\infty]$ denote the Hölder conjugate of p.[43] Then

$$\begin{aligned}\|Ex\|_p = \Big(\sum_{j=1}^{d} |(Ex)_j|^p\Big)^{1/p} &= d^{1/p}\,|(Ex)_1| \\ &\leq d^{1/p}\,\|e\|_{p'}\,\|x\|_p = d^{1/p'} d^{1/p}\,\|x\|_p = d\,\|x\|_p;\end{aligned}$$

for the inequality between both lines we used Hölder's inequality on $\mathbb{C}^d$. Hence, $\|E\| \leq d$, as claimed.

(c) Again, we only need to show that first estimate in (a) with $\gamma = 1$. This is quite easy if one uses the *polar decomposition* of matrices:[44,45] this decomposition

[43] This means, $\frac{1}{p} + \frac{1}{p'} = 1$.

[44] In case that you are not familiar with the polar decomposition of matrices, you can see the proof of assertion (b) instead, which only relies on Cramer's rule for inverse matrices and which suffices for the proof of Theorem 2.6.7 below.

[45] This argument is taken from [Kat60, Lemma 1 on p. 28 in the Appendix].

result says that there exists a unitary matrix $U \in \mathbb{C}^{d\times d}$ and an hermition matrix $H \in \mathbb{C}^{d\times d}$ such that $T = UH$. As T is invertible, so is H, and $T^{-1} = H^{-1}U^{-1}$. As U and U^{-1} are unitary, they are isometric with respect to the 2-norm on $\mathbb{C}^d$, and the determinant of U has modulus 1, so it follows that

$$\left\|T^{-1}\right\| = \left\|H^{-1}\right\| \qquad \text{and} \qquad \frac{\|T\|^{d-1}}{|\det T|} = \frac{\|H\|^{d-1}}{|\det H|}.$$

So it suffices to prove the claim for the self-adjoint matrix H. But due to the self-adjointness, H can be unitarily diagonalized, and this does not change the norms of H and H^{-1}, nor does it change the determinant of H. So it suffices to prove the claim for diagonal matrices – and for those, it can be checked by an easy computation. $\square$

Now we can show continuity of the spectrum in finite dimensions.

Theorem 2.6.7 (Hölder continuity of the spectrum in finite dimensions)**.** *Let $\mathfrak{C}$ denote the set of all non-empty compact subsets of $\mathbb{C}$, endowed with the Hausdorff metric $\mathrm{d}_{\mathrm{Haus}}$,*[46] *let $d \in \mathbb{N}$ and endow $\mathbb{C}^{d\times d}$ with any norm. Then the mapping*

$$\begin{aligned} \mathbb{C}^{d\times d} &\to \mathfrak{C}, \\ T &\mapsto \sigma(T) \end{aligned}$$

is continuous – and in fact, even locally Hölder continuous with exponent $1/d$.

More precisely, let $\gamma \geq 1$ be as in Lemma 2.6.6(a)*. Let $S, T \in \mathbb{C}^{d\times d}$ and assume that $\delta := (\gamma \|S - T\|)^{1/d} < 1$. Then*[47]

$$\mathrm{d}_{\mathrm{Haus}}\big(\sigma(S), \sigma(T)\big) \leq \delta \max\Big\{1, \frac{2\|T\|}{(1-\delta)}, \frac{2\|S\|}{(1-\delta)}\Big\}.$$

Proof. Let $\delta < 1$. We are going to prove that

$$\mathrm{d}_{\subseteq}\big(\sigma(S), \sigma(T)\big) \leq \delta \max\Big\{1, \frac{2\|T\|}{(1-\delta)}, \frac{2\|S\|}{(1-\delta)}\Big\};$$

by swapping the roles of S and T this implies the claim. So let $\lambda \in \mathbb{C}$ such that

$$\varepsilon := \operatorname{dist}(\lambda, \sigma(T)) > \delta \max\Big\{1, \frac{2\|T\|}{(1-\delta)}, \frac{2\|S\|}{(1-\delta)}\Big\}.$$

It suffices to show that $\lambda \in \rho(S)$. From the choice of ε we get by a brief computation

$$\delta < \varepsilon \qquad \text{and} \qquad \delta < \frac{\varepsilon}{2\|T\| + \varepsilon}.$$

[46] Note that this turns $\mathfrak{C}$ into a metric space according to Proposition 2.6.3.

[47] If one prefers simpler formulas, the following version of the result (which follows immediately from the version in the theorem) is a bit more convenient: if $\delta \leq \frac{1}{2}$, then

$$\mathrm{d}_{\mathrm{Haus}}\big(\sigma(S), \sigma(T)\big) \leq \delta \max\{1, 4\|T\|, 4\|S\|\}.$$

By distinguishing the two cases $2\,\|T\| + \varepsilon \le 1$ and $2\,\|T\| + \varepsilon > 1$ one sees that this implies

$$\delta^d < \frac{\varepsilon^d}{(2\,\|T\| + \varepsilon)^{d-1}}$$

and thus,

$$\|S - T\| < \frac{\varepsilon^d}{\gamma\,(2\,\|T\| + \varepsilon)^{d-1}}.$$

Now we proceed similarly to the proof of Theorem 2.6.5: it follows from Lemma 2.6.6(a) that

$$\begin{aligned} \frac{1}{\|\mathcal{R}(\lambda, T)\|} &\ge \frac{1}{\gamma} \frac{\operatorname{dist}(\lambda, \sigma(T))^d}{\big(\,\|T\| + |\lambda|\,\big)^{d-1}} \ge \frac{1}{\gamma} \frac{\varepsilon^d}{(2\,\|T\| + \varepsilon)^{d-1}} \\ &> \|S - T\| = \|(\lambda - S) - (\lambda - T)\|\,; \end{aligned}$$

for the second inequality we used the definition of ε as well as the estimate $|\lambda| \le \mathrm{r}(T) + \operatorname{dist}(\lambda, \sigma(T)) \le \|T\| + \varepsilon$. Hence, $\lambda \in \rho(S)$ according to Corollary 1.1.12. $\square$

Chapter 3

Commutative Banach Algebras and the Gelfand Representation

3.1 Maximal ideals and characters

The goal of Chapter 3 is to obtain of representation theorem for commutative Banach algebras. We use two important tools on the way to this theorem: *maximal ideals* and *characters*. Let us start with the first of them.

Definition 3.1.1 (Maximal ideals)**.** Let A be an algebra and let $I \subseteq A$ be an ideal.

(a) The ideal I is called *proper* if $I \neq A$.[1]

(b) The ideal I is called *maximal* if it is proper and every proper ideal $J \supseteq I$ is equal to A.[2]

Example 3.1.2 (Maximal ideals in spaces of continuous functions)**.** Let $K \neq \emptyset$ be a compact Hausdorff space and let $\omega_0 \in K$. Then $I_{\{\omega_0\}} := \{f \in \mathrm{C}(K) \mid f(\omega_0) = 0\}$ is an ideal in $\mathrm{C}(K)$ which has co-dimension 1 and is thus maximal.

We will see in the course of this section that maximal ideals have a number of useful properties. First, though, we will discuss whether the existence of such ideals is always ensured. In unital Banach algebras this is indeed the case – a property which we show in the first part of the following proposition.

Proposition 3.1.3 (Existenz and closedness of maximal ideals in the unital case)**.** *Let A be a unital Banach algebra.*

(a) *Every proper ideal in A is contained in a maximal ideal.*

(b) *If A is a unital Banach algebra, then every maximal ideal in A is closed.*

[1] Note that I is allowed to be $\{0\}$.

[2] In other words, I is maximal with respect to set inclusion in the set of all proper ideals in A.

Proof. (a) An ideal in A is proper if and only if it does not contain 1, so the union of a chain of proper ideals is again a proper ideal. Hence, the claim follows from Zorn's lemma.

(b) Let I be a maximal ideal in A. Since the closure $\overline{I}$ is also an ideal, we have either $\overline{I} = I$ or $\overline{I} = A$, so we only need to rule out the second case.

Recall that, due to the Neumann series, every element $a \in A$ that satisfies $\|a - 1\| < 1$ is invertible, and therefore not in I (as I is proper and thus $I \neq A$). Hence, every element in I has distance at least 1 to the neutral element 1. Therefore, $1 \notin \overline{I}$, and we thus conclude that $\overline{I} \neq A$, as claimed. ☐

Now we come to the second important tool that we need to prove Gelfand's representation theorem, namely characters. As we will see in Proposition 3.1.7, they are – in the commutative case – closely related to maximal ideals.

Definition 3.1.4 (Characters)**.** Let A be a Banach algebra.

(a) A *character on* A is a non-zero algebra homomorphism $\tau : A \to \mathbb{C}$.

(b) The set of all characters on A is called the *character space of* A and is denoted by $\Omega(A)$.[3]

Note that there is no character on the trivial algebra $A = \{0\}$, so $\Omega(\{0\}) = \emptyset$. You will see in the exercises that, on non-commutative algebras, one cannot expect the existence of many characters, in general. However, we will see lataer on that there are a lot of characters on commutative Banach algebras, in general – and this will actually be the entire point of the present Chapter 3.

Proposition 3.1.5 (Properties of characters)**.** *Let A be a unital Banach algebra.*

(a) *Every character $\tau \in \Omega(A)$ is continuous.*

(b) *One has $\tau(1) = 1 = \|\tau\|$ for every $\tau \in \Omega(A)$.*

Proof. Fix $\tau \in \Omega(A)$.

(a) The kernel of τ is a proper ideal in A, and has co-dimension 1. Thus, it is a maximal ideal, and therefore closed according to Proposition 3.1.3(b). Hence, τ is continuous.

(b) As τ is non-zero, we have $\tau(1) \neq 0$.[4] Moreover, $\tau(1) = \tau(1 \cdot 1) = \tau(1)\tau(1)$, so we conclude that $\tau(1) = 1$. This also implies that $\|\tau\| \geq |\tau(1)| = 1$. Finally, for every $a \in A$ we have

$$|\tau(a)| = |\tau(a^n)|^{1/n} \leq \|\tau\|^{1/n} \|a^n\|^{1/n} \overset{n\to\infty}{\longrightarrow} \mathrm{r}(a) \leq \|a\|\,,$$

so $\|\tau\| \leq 1$. ☐

[3]Some people also call $\Omega(A)$ the *spectrum of* A (for reasons that will become apparent in Proposition 3.2.3), but we refrain from doing this in this manuscript, to avoid unnecessary terminological confusion.

[4]Why?

The following theorem, which is also interesting on its own right, will be useful in the proof of the subsequent proposition.

Theorem 3.1.6 (Unital Banach algebras with invertible elements only). *Let A be a non-zero unital Banach algebra. If every non-zero element of A is invertible, then A is isomorphic to the unital Banach algebra $\mathbb{C}$.*

Proof. It suffices to prove that every element of A is a scalar multiple of 1, so fix $a \in A$. As $\sigma(a)$ is non-empty according to Theorem 2.4.1, there exists $\lambda \in \mathbb{C}$ such that $a - \lambda$ is non-intertible. But, by assumption, the only non-invertible element of A is 0, so $a - \lambda = 0$, i.e., $a = \lambda$, as claimed. □

Now we come to the main ingredient that we will need for Gelfand's representation theorem in the subsequent section: in commutative unital Banach algebras, maximal ideals and characters are in 1-to-1 correspondence. This will be very useful for the following reasons: on the one hand, characters are very simple examples of algebra homomorphisms[5] and can be used to construct more complicated algebra homomorphisms; but it is not clear at first glance how to even obtain a non-trivial character on a given algebra. On the other hand, maximal ideals exist in abundance, as we easily derived from Zorn's lemma in Proposition 3.1.3(a). Thus, the following 1-to-1 correspondence can be used to translate the fact that there are many maximal ideals into the fact that there are many characters.

Proposition 3.1.7 (Maximal ideals via characters). *Let A be a unital Banach algebra and assume that A is commutative.*

(a) *The mapping*

$$\psi : \tau \mapsto \ker \tau$$

is a bijection between $\Omega(A)$ and the set of all maximal ideals in A.

(b) *If $A \neq \{0\}$, then the character space $\Omega(A)$ is non-empty.*

Proof. (a) *Well-definedness:* As already observed in the proof of Proposition 3.1.5(a), the kernel of every $\tau \in \Omega(A)$ is a maximal ideal in A, so ψ does indeed map from $\Omega(A)$ in the set of maximal ideals.

Injectivity: Assume that two characters $\tau_1, \tau_2 \in \Omega(A)$ have the same kernel, and let $a \in A$. As $\ker \tau_1 = \ker \tau_2$ is a vector subspace of A of co-dimension 1 and does not contain the neutral element 1, there exists an element $b \in \ker \tau_1 = \ker \tau_2$ and a scalar $\lambda \in \mathbb{C}$ such that $a = b + \lambda$. Hence,

$$\tau_1(a) = \tau_1(\lambda) = \lambda = \tau_2(\lambda) = \tau_2(a);$$

for the second and the third equality we used Proposition 3.1.5 (b). Hence, $\tau_1 = \tau_2$.

[5] They are "simple examples" in the sense that they map into the algebra $\mathbb{C}$

Surjectivity: Let $I \subseteq A$ be a maximal ideal. Then I is closed according to Proposition 3.1.3(b), and hence the quotient space A/I is a unital Banach algebra, as shown in Proposition 1.2.5. Moreover, A/I is non-zero since I is a proper ideal.

Now we show that every non-zero element of A/I is invertible: let $b + I \in A/I$ be non-zero. Then $b \notin I$. The set $bA := \{ba \mid \; a \in A\}$ is an ideal in A since A is commutative,and thus, $bA+I$ is also an ideal in A; this ideal contains both I and the element b which is not in I, so it follows from the maximality of I that $bA + I = A$. Hence, there exists $a \in A$ and an element $c \in I$ such that $ba + c = 1$. Thus, in the quotient algebra $A + I$ we have

$$(a + I)(b + I) = (b + I)(a + I) = ba + I = 1 + I,$$

where we again used the commutativity of A for the first equality. Hence, $b+I \in A/I$ is indeed invertible.

As every non-zero element of A/I is invertible, it follows from Theorem 3.1.6 that there exists an algebra isomorphism $j : A/I \to \mathbb{C}$. If $q : A \to A/I$ denotes the quotient mapping, one thus has an algebra homomorphism

$$\tau : A \xrightarrow{q} A/I \xrightarrow{j} \mathbb{C}.$$

Note that τ is surjective (as q is surjective and j is bijective) and thus non-zero, so $\tau \in \Omega(A)$. Moreover, the bijectivity of j implies that $\ker \tau = \ker q = I$.

(b) As A is non-zero, the ideal $\{0\}$ in A is proper; hence, it follows from Proposition 3.1.3(a) that there exists a maximal ideal in A. According to (a) this implies that there exists a character on A. □

The preceding proposition indicates that the set of characters contains a lot of information about any commutative unital Banach algebra. This is the main idea behind *Gelfand's representation theorem* for commutative Banach algebras, which we will discuss in Section 3.2. In order to do this, we need a bit of point set topology; of brief overview of what we need is given in Appendix A.

3.2 The Gelfand representation

By using some fundamental notions from topology, which are described in Appendix A, we can now introduce a topology on the space of characters on a unital Banach algebra. This paves the way to finally show that representation theorem for commutative unital Banach algebras that we are after.

Definition 3.2.1 (Evaluations maps and the topology on the character space)**.** Let A be a unital Banach algebra. For every $a \in A$ define the so-called *evaluation map*

$$\begin{aligned} \hat{a} :\; \Omega(A) &\to \mathbb{C}, \\ \tau &\mapsto \tau(a). \end{aligned}$$

Unless otherwise stated, we always endow the character space $\Omega(A)$ with the initial topology of the family $(\hat{a})_{a \in A}$.

The topology on the character space has two useful properties which complement each other in a nice way: on the one hand, it is sufficiently large to make all the evaluation maps $\hat{a}$ continuous and to be Hausdorff; on the other hand, it is sufficiently small to be compact. This is the content of the following proposition.

Proposition 3.2.2 (Topological properties of the character space). *Let A be a unital Banach algebra.*

(a) *The character space $\Omega(A)$ is compact and Hausdorff.*

(b) *For every $a \in A$ the mapping*

$$\begin{aligned} \hat{a}:\ \Omega(A) &\to \mathbb{C}, \\ \tau &\mapsto \tau(a) \end{aligned}$$

is continuous, i.e., $\hat{a} \in \mathrm{C}\big(\Omega(A)\big)$.

Proof. (a) We first note that it follows from the definition of the weak* topology (Example (b)A.4.2) that a net in $\Omega(A)$ converges to a point $\tau \in \Omega(A)$ with respect to the topology on $\Omega(A)$ (i.e., the topology introduced in Definition 3.2.1) if and only if it converges to τ with respect to the weak* topology in A'; thus it suffices to prove that $\Omega(A)$ is compact and Hausdorff with respect to the weak* topology in A'.

The set $\Omega(A)$ can easily be checked to be a weak*-closed subset of the norm closed unit ball in A', and the latter is weak* compact due to the Banach–Alaoglu theorem A.4.3. As closed subsets of compact sets are again compact, it follows that $\Omega(A)$ is weak* compact. Moreover, as discussed in Example (b)A.4.2, the weak* topology on A' is Hausdorff, and thus it is also Hausdorff on $\Omega(A)$.

(b) This follows from the very definition of the initial topology, see Theorem A.4.1. □

The main goal of this section is to show that, for each element a of a commutative unital Banach algebra A, the continuous function $\hat{a} : \Omega(A) \to \mathbb{C}$ contains a lot of information about a. The following proposition is a large[6] step in this direction.

Proposition 3.2.3 (Description of spectra via characters). *Let A be a unital Banach algebra. If A is commutative, then one has*

$$\sigma(a) = \{\tau(a) \mid \tau \in \Omega(A)\}$$

for every $a \in A$.[7]

Proof. Fix $a \in A$.

[6]Though simple, after all the preparations that we have already made.

[7]In other words, the spectrum $\sigma(a)$ coincides with the image $\hat{a}\big(\Omega(A)\big)$ of the function $\hat{a} : \Omega(A) \to \mathbb{C}$.

„$\supseteq$“ Let $\tau \in \Omega(A)$. Then

$$\tau\big((\tau(a) - a)\big) = \tau(a)\underbrace{\tau(1)}_{=1} - \tau(a) = 0,$$

so $\tau(a) - a \in \ker \tau$. Since $\ker \tau$ is a proper ideal in A, it does not contain any invertible elements, so $\tau(a) \in \sigma(a)$, as claimed.

„$\subseteq$“ Let $\lambda \in \sigma(a)$. The set $(\lambda - a)A$ is an ideal in A since A is commutative; moreover, it does not contain the element 1 since $\lambda - a$ is not invertible,[8] so it is a proper ideal. Hence, there exists a maximal ideal I in A that contains $(\lambda - a)A$; see Proposition 3.1.3(a).

The 1-to-1 correspondence between maximal ideals and kernels of characters that we proved in Proposition 3.1.7(a) gives as a character $\tau \in \Omega(A)$ such that $\ker \tau = I$. In particular, τ vanishes on the element $\lambda - a$ of $(\lambda - a)A \subseteq I$, so

$$\lambda - \tau(a) = \lambda\tau(1) - \tau(a) = \tau(\lambda - a) = 0.$$

So $\lambda = \tau(a)$, as claimed. $\square$

The proof of the inclusion „$\subseteq$“ in the previous proposition illustrates the philoshopy that was indicated before Proposition 3.1.7: for a given spectral value λ of a it is, at first glance, not clear how one could obtain a character τ with the property $\tau(a) = \lambda$. However, we have a way to obtain maximal ideals, and the 1-to-1 correspondence between maximal ideal and characters thus yields the desired τ.

Definition 3.2.4 (The Gelfand representation)**.** Let A be a commutative and unital Banach algebra. For every $a \in A$ let $\hat{a} \in \mathrm{C}(A)$ denote the function

$$\begin{aligned} \Omega(A) &\to \mathbb{C}, \\ \tau &\mapsto \tau(a), \end{aligned}$$

as in Proposition 3.2.2(b). Then the mapping

$$\Theta : A \to \mathrm{C}(\Omega(A)), \qquad a \mapsto \hat{a}$$

is called the *Gelfand representation* of A.

Proposition 3.2.3 shows that, for every element a of a commutative unital Banach algebra, the spectrum $\sigma(a)$ coincides with the image of the function $\Theta(a) = \hat{a} \in \mathrm{C}\left(\Omega(A)\right)$. Let us summarize further important properties of the Gelfand representation in the following theorem; its proof is now a simple synthesis of results that we have already shown.

Theorem 3.2.5 (Properties of the Gelfand representation)**.** *Let A be a unital Banach algebra. If A is commutative and non-zero, then the Gelfand respresentation $\Theta : A \mapsto \mathrm{C}(\Omega(A))$, $a \mapsto \hat{a}$, has the following properties:*

[8] Note that we again used the commutativity of A here.

(a) *The mapping* Θ *is a continuous algebra homomorphism, and one has* $\Theta(1) = \mathbb{1}$.

(b) *For every* $a \in A$ *the range of the function* $\Theta(a) \in \mathrm{C}(\Omega(A))$ *coincides with* $\sigma(a)$.

(c) *One has* $\|\hat{a}\|_\infty = \mathrm{r}(a) \leq \|a\|$ *for all* $a \in A$.[9]

Proof. (a) For all $a, b \in A$ and all $\tau \in \Omega(A)$ one has

$$\big(\Theta(ab)\big)(\tau) = \tau(ab) = \tau(a)\tau(b) = \big(\Theta(a)\big)(\tau)\big(\Theta(b)\big)(\tau) = \big(\Theta(a)\Theta(b)\big)(\tau),$$

where the second equality follows from the fact that τ is algebra homomorphism and thus multiplicative. Hence, $\Theta(ab) = \Theta(a)\Theta(b)$ for all $a, b \in A$. By a similar computation one can check that Θ is linear; so Θ is indeed an algebra homomorphism. Continuity of Θ will be a consequence of (c).

The claim $\Theta(1) = \mathbb{1}$ follows from Proposition 3.1.5(b).

(b) This is precisely the content of Proposition 3.2.3.

(c) This follows immediately from (b). □

3.3 Applications and examples

Example 3.3.1 (The Gelfand homomorphism on spaces of continuous functions is an isomorphism)**.** Let K be a compact Hausdorff space and consider the commutative unital Banach algebra $A := \mathrm{C}(K)$. For every $x \in K$ let $\delta_x \in \Omega(A)$ denote the character $f \mapsto f(x)$.

(a) The mapping

$$\begin{aligned} K &\to \Omega(A), \\ x &\mapsto \delta_x \end{aligned}$$

is a homeomorphism.[10]

(b) The Gelfand representation $\Theta : A \to \mathrm{C}\big(\Omega(A)\big)$ is bijective.

To show the properties claimed in the example we use the following auxiliary result:

Lemma 3.3.2. *Let* K *be a compact Hausdorff space and consider the commutative unital Banach algebra* $A := \mathrm{C}(K)$. *Let* $\tau \in \Omega(A)$.

Then τ *maps real-valued functions into* $\mathbb{R}$, *and* τ *commutes with complex conjugation, i.e.,* $\overline{\tau(f)} = \tau(\overline{f})$ *for every* $f \in A$.

[9]Note that this implies that Θ is continuous and $\|\Theta\| \leq 1$.

[10]I.e., a continuous bijection whose inverse mapping is also continuous.

Proof. First assume that $f \in A$ is real-valued. We have $\tau(f) \in \sigma(f)$ according to Proposition 3.2.3, and $\sigma(f)$ equals the image of f (Example 2.2.6), so it is a subset of $\mathbb{R}$. Hence, $\tau(f) \in \mathbb{R}$.

Now, let $f \in A$ be a general element. Then $\operatorname{Re} f$ and $\operatorname{Im} f$ are in A, too, and they are real-valued. Thus,

$$\overline{\tau(f)} = \overline{\tau(\operatorname{Re} f + \mathrm{i} \operatorname{Im} f)} = \tau(\operatorname{Re} f) - \mathrm{i}\tau(\operatorname{Im} f)) = \tau(\overline{f});$$

for the second equality we used that $\tau(\operatorname{Re} f)$ and $\tau(\operatorname{Im} f)$ are real numbers, which follows from the first parts of the proof. □

Proof of the properties in Example 3.3.1. (a) *Injectivity:* Let $x, y \in K$ such that $x \neq y$. Then the two sets $\{x\}$ and $\{y\}$ are disjoint and, due to the Hausdorff property of K, closed.[11] So we can apply Urysohn's lemma[12] to these two sets, which gives as a function $f \in A$ that satisfies $f(x) \neq f(y)$. Hence, $\langle \delta_x, f\rangle \neq \langle \delta_y, f\rangle$, so the characters δ_x and δ_y are not equal.

Surjectivity: Let $\tau \in \Omega(A)$ be a character; we need to show that there exists a point $x \in K$ such that $\delta_x = \tau$. Then $\ker \tau$ is a maximal ideal, and it follows from Lemma 3.3.2 that $\ker \tau$ is invariant under complex conjugation.[13] Hence, for every $f \in \ker \tau$ we also have $|f|^2 = f\overline{f} \in \ker \tau$.

We now show that there exists a point $x \in K$ such that every $f \in \ker \tau$ vanishes at x. Assume the contrary is the case. For each $x \in K$ we can then find a function $f_x \in A$ such that $f_x(x) \neq 0$. Thus, the open sets $U_x := \{y \in K \mid f_x(y) \neq 0\}$ cover K as x runs through K. Since K is compact by assumption, we can find finitely many points $x_1, \dots, x_n \in K$ such that $K = U_{x_1} \cup \dots \cup U_{x_n}$. Hence, the function $|f_{x_1}|^2 + \dots + |f_{x_n}|^2$ is strictly positive everywhere on K; moreover, it is in $\ker \tau$, as shown in the previous paragraph. But this means that $\ker \tau$ contains an invertible element, which contradicts the fact that $\ker \tau$ is a proper ideal.

So we have, indeed, proved the existence of x such that each $f \in \ker \tau$ satisfies $f(x) = 0$. Hence, the ideal $I_{\{x\}} := \{g \in \mathrm{C}(K) \mid g(x) = 0\}$ is a proper ideal that contains $\ker \tau$. Due to the maximality of $\ker \tau$ we conclude that $\ker \tau = I_{\{x\}}$. For every $f \in A$ we thus have $f - f(x)\,\mathbb{1} \in \ker \tau$, so

$$\tau(f) = f(x)\tau(1) = \delta_x(f),$$

which shows that $\tau = \delta_x$, as claimed.

Continuity: If $(x_j)_{j \in J}$ is a net in K that converges to a point $x \in K$, then we have for every $f \in A$

$$\delta_{x_j}(f) = f(x_j) \to f(x) = \delta_x(f),$$

[11] By the way, the property that all singletons in a topological space X be closed, is the so-called "separation axiom T1"; it is a actually bit weaker than the Hausdorff property.

[12] *Urysohn's lemma* says that, on every compact Hausdorff space X (and more generally, and every so-called *normal* Hausdorff space), for each pair of non-empty disjoint closed sets C_0, C_1 there exists a function $f : X \to [0,1]$ which is 0 on C_0 and 1 on C_1.

[13] I.e., if $f \in \ker \tau$, then also $\overline{f} \in \ker \tau$.

so $\delta_{x_j} \to \delta_x$ in $\Omega(A)$.

Continuity of the inverse mapping: It is a nice exercise in point set topology to show the following general result: if X, Y are topological spaces, X is compact and Y is Hausdorff, then the inverse mapping of every continuous bijection $X \to Y$ is also continuous.[14]

(b) We leave this as an exercise on Sheet 7. $\square$

One very useful application of the Gelfand representation is that it enables us to prove spectral results for commuting elements in non-commutative Banach algebras. To make this work, the following result is useful which relates the spectrum of elements of an algebra to their spectrum in a subalgebra.

Proposition 3.3.3 (The spectrum in subalgebras)**.** *Let A be a unital Banach algebra and let $B \subseteq A$ be a closed subalgebra that contains the neutral element of A. Let $b \in B$. The spectra $\sigma_A(b)$ and $\sigma_B(b)$ in A and B, respectively, are related as follows:*

(a) *One has $\sigma_A(b) \subseteq \sigma_B(b)$.*[15]

(b) *One has $\partial\sigma_A(b) \supseteq \partial\sigma_B(b)$.*[16]

(c) *For the spectral radii of b one has $\mathrm{r}_A(b) = \mathrm{r}_B(b)$.*

Proof. (a) If $\lambda \in \mathbb{C}$ is not in $\sigma_B(b)$, then there exists $c \in B$ such that $c(\lambda - b) = (\lambda - b)c = 1$. As c is also an element of A, it follows that $\lambda - b$ is also invertible in A, so λ is not in $\sigma_A(b)$, either.

(b) Let $\lambda \in \partial\sigma_B(b)$. Then $\lambda - b$ is in the topological boundary of $\mathrm{Inv}(B)$, so it follows from Exercise 4(c) on Sheet 3 that there exists a sequence (b_n) of normalized vectors in B such that $(\lambda - b)b_n \to 0$. The some convergence of course also takes place in A, so $\lambda - b$ is not invertible in A.[17] Hence, $\lambda \in \sigma_A(b)$.

As it follows from (a) that $\big(\sigma_A(b)\big)^\circ \subseteq \big(\sigma_B(b)\big)^\circ$ and as λ is not in the latter set, we conclude that $\lambda \in \partial\sigma_A(b)$, as claimed.

(c) It follows from (a) that $\mathrm{r}_A(b) \le \mathrm{r}_B(b)$. Converse, there exists a number $\lambda \in \partial\sigma_B(b)$ such that $|\lambda| = \mathrm{r}_B(b)$. According to (b) we also $\lambda \in \sigma_A(b)$, so the converse inequality $\mathrm{r}_A(b) \ge \mathrm{r}_B(b)$ holds, too. $\square$

It is not difficult to find two matrices $S, T \in \mathbb{C}^{2\times 2}$ such that

$$\mathrm{r}(S+T) > \mathrm{r}(S) + \mathrm{r}(T).$$

Similarly, one can find two matrices $S, T \in \mathbb{C}^{2\times 2}$ such that

$$\mathrm{r}(ST) > \mathrm{r}(S)\,\mathrm{r}(T).$$

However, this cannot happen if S and T commute, as the following proposition shows.

[14] Hint: Observe that in a Hausdorff space every compact subset is closed, and then use that continuous functions map compact sets to compact sets.

[15] One does not have equality here, in general; see Exercise 5 on Sheet 3.

[16] Here, ∂ denote the topological boundary.

[17] Can you give this argument in more detail?

Proposition 3.3.4 (The spectral radius of commuting elements)**.** *Let A be a unital Banach algebra,*[18] *and let $a, b \in A$ be two commuting elements. Then*

$$\mathrm{r}(a+b) \leq \mathrm{r}(a) + \mathrm{r}(b) \qquad \text{and} \qquad \mathrm{r}(ab) \leq \mathrm{r}(a)\,\mathrm{r}(b).$$

Proof. Let $B \subseteq A$ denote the smallest closed subalgebra of A that contains a, b, and 1. Then B is itself a unital Banach algebra, and it can easily be verified that B is the closure of the closure of the linear span of

$$\{a^m b^n \mid m, n \in \mathbb{N}_0\}.$$

As a and b are assumed to commute, it thus follows that B is commutative.

According to Proposition 3.3.3(c) all elements of B have the same spectral radius in A and in B; so it suffices to show the claim within B.

As B is commutative we can use the Gelfand representation to this end; so let $\Theta : B \to \mathrm{C}\big(\Omega(B)\big)$ the Gelfand representation of B. According to Theorem 3.2.5(c) we have

$$\mathrm{r}(a+b) = \|\Theta(a+b)\|_\infty \leq \|\Theta(a)\|_\infty + \|\Theta(b)\|_\infty = \mathrm{r}(a) + \mathrm{r}(b).$$

Similarly, one obtains

$$\mathrm{r}(ab) = \|\Theta(ab)\|_\infty \leq \|\Theta(a)\|_\infty \|\Theta(b)\|_\infty = \mathrm{r}(a)\,\mathrm{r}(b),$$

so we proved all claims.[19] □

Theorem 3.3.5 (Spectral mapping theorem for pseudo-resolvents)**.** *Let A be a unital Banach algebra, let $\rho \subseteq \mathbb{C}$ be non-empty and open, and let $\mathcal{R} : \rho \to A$ be a pseudo-resolvent. Assume moreover that ρ is already the domain of the largest extension of $\mathcal{R}$, and let $\sigma := \mathbb{C} \setminus \rho$ denote the spectrum of the pseudo-resolvent. Then*

$$\sigma\big(\mathcal{R}(\lambda_0)\big) \setminus \{0\} = \left\{ \frac{1}{\lambda_0 - \lambda} \mid \lambda \in \sigma \right\}.$$

Proof. „$\supseteq$“ We have already seen that this inclusion follows from the properties of pseduo-resolvents that we proved in Proposition 2.5.10(a)

„$\subseteq$“ The proof of this inclusion is an exercise on Sheet 7. □

By combining the preceding theorem with the spectral mapping theorem for resolvents (Theorem 2.2.8), one gets the following result about the largest domain of a pseudo-resolvent which is actually a resolvent:

Corollary 3.3.6 (Spectra of elements vs spectra of pseudo-resolvents)**.** *Let A be a unital Banach algebra and let $a \in A$. Let $\sigma \subseteq A$ denote the spectrum of the (pseudo-)resolvent $\mathcal{R}(\,\cdot\,, a)$ in the sense of Definition 2.5.9. Then $\sigma = \sigma(a)$.*

---

[18] Note that we do not assume A to be commutative.

[19] It is worthwhile to note that the submultiplicativity of the spectral radius can also easily be shown without the Gelfand representation if one uses the spectral radius formula $\mathrm{r}(ab) = \inf_n \|(ab)^n\|^{1/n}$ together with the fact that $(ab)^n = a^n b^n$ for all a, b since a and b commute. The subadditivity, howoever, is apparently not quite so easy to prove without the Gelfand representation.

Part II

Spectral Theory of Linear Operators

Chapter 4

Linear Operators and More

4.1 Unbounded operators

If X is a complex Banach space, then the space $\mathcal{L}(X)$ of bounded linear operators on X is a unital Banach algebra (when endowed with the operator norm), and thus the spectral theory of bounded linear operators can be considered as a special case of the spectral theory in unital Banach algebras;[1] compare Example 2.2.5(b). However, there is a very interesting class of linear operators which does not directly fit into the framework of bounded operators. For instance, consider the complex Banach space $X = \mathrm{C}([0, 1])$; in the study of differential equations it would be useful to consider a linear mapping $L : X \to X$ that sends each function $f \in X$ to its derivative f'; however, this is obviously not possible since their are functions in X that are not differentiable. However, we can define L in the vector subspace $\mathrm{C}^1([0, 1])$ of X that consists of all continuously differentiable functions in X. For this operator we write

$$L : \mathrm{C}([0, 1]) \supseteq \mathrm{C}^1([0, 1]) \to \mathrm{C}([0, 1]), \quad f \mapsto f',$$

The domain of this mapping is $\mathrm{C}^1([0, 1])$; we just use the notation "$\mathrm{C}([0, 1]) \supseteq \mathrm{C}^1([0, 1])$" in the definition of L to indicate that we "would like to keep the surrounding space $C([0, 1])$ in mind". This choice of the surrounding space becomes relevant when we discuss the spectrum of such operators L later on in this section.

Definition 4.1.1 (Linear operators, general case)**.** Let X be a complex Banach space.

(a) A *linear operator on* X is a linear mapping $L : \mathrm{dom}(L) \to X$, where $\mathrm{dom}(L)$ is a vector subspace of X, the *domain* of L.

We often write $L : X \supseteq \mathrm{dom}(L) \to X$ for a linear operator on X.

[1] But this is actually only part of the story for bounded operators: we will see in the subsequent chapters that the spectrum of linear operators carries a very interesting fine structure whose properties are not directly reflected in the spectral theory on unital Banach algebras.

(b) A linear operator $L : X \supseteq \mathrm{dom}(L) \to X$ is called *closed* if its *graph*

$$\mathrm{Gr}(L) := \big\{(x, Lx) \mid\ x \in \mathrm{dom}(L)\big\} \subseteq X \times X$$

is closed in $X \times X$.[2]

(c) Let $L : X \supseteq \mathrm{dom}(L) \to X$ be a linear operator. A norm $\|\cdot\|_L$ on $\mathrm{dom}(L)$ is called a *graph norm* if its equivalent to the norm

$$x \mapsto \|x\|_X + \|Lx\|_X$$

on $\mathrm{dom}(L)$.

(d) A linear operator $L : X \supseteq \mathrm{dom}(L) \to X$ is called *everywhere defined* if $\mathrm{dom}(L) = X$.

(e) A linear operator $L : X \supseteq \mathrm{dom}(L) \to X$ is called *densely defined* if $\mathrm{dom}(L)$ is dense in X.

Note that if a linear operator $L : X \supseteq \mathrm{dom}(L) \to X$ is both closed and everywhere defined, then it is automatically a bounded linear operator on X; this follows from the closed graph theorem. Hence, one often refers to operators with general domain $\mathrm{dom}(L)$ as *unbounded linear operators*.[3]

Proposition 4.1.2 (Characterisation of closedness)**.** *Let $L : X \supseteq \mathrm{dom}(L) \to X$ be a linear operator on a complex Banach space X. The following are equivalent:*

(i) *The operator L is closed.*

(ii) *One (equivalently, every) graph norm of L on* $\mathrm{dom}(L)$ *is complete.*

(iii) *Whenever a sequence (x_n) in* $\mathrm{dom}(L)$ *and $x, y \in X$ such that*

$$x_n \overset{\|\cdot\|_X}{\longrightarrow} x \qquad \textit{and} \qquad Lx_n \overset{\|\cdot\|_X}{\longrightarrow} y,$$

then $x \in \mathrm{dom}(L)$ and $Lx = y$.

Proof. „(i) $\Leftrightarrow$ (ii)“ Endow $X \times X$ with the norm given by $\|(x,y)\|_{X\times X} := \|x\|_X + \|y\|_X$ for all $(x,y) \in X \times X$.[4] Then the mapping $\mathrm{Gr}(L) \to \mathrm{dom}(L)$, $(x,y) \mapsto x$ is bijective and isometric between the normed spaces $(\mathrm{Gr}(L), \|\cdot\|_{X\times X}$ and $(\mathrm{dom}(L), \|\cdot\|_L)$, where we choose $\|\cdot\|_L$ to be the graph norm given by $\|x\|_L := \|x\|_X + \|Lx\|_X$ for all $x \in \mathrm{dom}(L)$. Hence, L is closed if and only if $\mathrm{Gr}(L)$ is closed in $X \times X$ if and only if $(\mathrm{Gr}(L), \|\cdot\|_{X\times X})$ is complete if and only if $(\mathrm{dom}(L), \|\cdot\|_L)$ is complete.

[2] Where $X \times X$ is endowed with the product topology.

[3] And by a – sometime somewhat confusing – abuse of terminology one does so no matter whether L is actually bounded or not. In other words, within this terminological convention, the class of bounded linear operators is contained in the class of unbounded linear operators.

[4] There are various options how one can construct, from the norm on X, a norm on $X \times X$ that induces the product topology.

„(i) $\Rightarrow$ (iii)“ If $(x_n) \subseteq \operatorname{dom}(L)$ converges to $x \in X$ with respect to $\|\cdot\|_X$ and $(Lx_n) \subseteq X$ converges to y (also with respect to $\|\cdot\|_X$), then the sequence (x_n, Lx_n) in $\operatorname{Gr}(L)$ converges to the point $(x, y) \in X \times X$, so it follows from the closedness of $\operatorname{Gr}(L)$ that $(x, y) \in \operatorname{Gr}(L)$. Hence, $x \in \operatorname{dom}(L)$ and $Lx = y$.

„(iii) $\Rightarrow$ (i)“ If $\big((x_n, y_n)\big)$ is a sequence in $\operatorname{Gr}(L)$ that converges to a point $(x, y) \in X \times X$, then (x_n) is a sequence in $\operatorname{dom}(L)$ that converges to x with respect to $\|\cdot\|_X$ and $(y_n) = (Lx_n)$ is a sequence in X that converges to y (also with respect to $\|\cdot\|_X$. Hence it follows from (iii) that $x \in \operatorname{dom}(L)$ and $Lx = y$, i.e., $(x, y) \in \operatorname{Gr}(L)$. ☐

Example 4.1.3 (A differential operator on a space of continuous functions)**.** On the Banach space $X := \mathrm{C}_0((0,1])$ of continuous complex-valued functions f on $(0,1]$ that satisfy $\lim_{t\downarrow 0} f(t)$ (endowed with the sup norm), consider the operator $L : X \supseteq \operatorname{dom}(L) \to X$ that is given by

$$\begin{aligned} \operatorname{dom}(L) &:= \{f \in X \mid \ f \text{ is differentiable and } f' \in X\}, \\ Lf &:= -f' \quad \text{for all } f \in \operatorname{dom}(L). \end{aligned}$$

Then L is a closed and densely defined linear operator on X. A graph norm on $\operatorname{dom}(L)$ is, for instance, given by

$$\|f\|_L := \|f\|_\infty + \left\|f'\right\|_\infty$$

for all $f \in \operatorname{dom}(L)$.

Proof. The operator L is densely defined since, for instance, all C^∞-functions on $(0,1]$ that vanish on a right neighbourhood of 0 are in $\operatorname{dom}(L)$, and the space of those functions can be checked to be dense in $\mathrm{C}_0((0,1])$.[5] The norm $\|\cdot\|_L$ on $\operatorname{dom}(L)$ as defined in the example is clearly a graph norm for L.

In order to show that L is closed it suffices, according to Proposition 4.1.2 to show that the norm $\|\cdot\|_L$ on $\operatorname{dom}(L)$ is complete, so let (f_n) be a Cauchy sequence in $\operatorname{dom}(L)$ with respect to this norm. Then, in particular, the sequence of derivatives (f_n') is Cauchy with respect to the sup norm on $\mathrm{C}_0((0,1])$, so it converges to a continuous and bounded function $g \in \mathrm{C}_0((0,1])$. Now define a function $f \in \mathrm{C}_0((0,1])$ by setting

$$f(t) := \int_0^t g(s) \, \mathrm{d}s \qquad \text{for each } t \in (0,1].$$

Then $f_n' \to g = f'$ with respect to $\|\cdot\|_\infty$, and moreover one also has $f_n \to f$ with respect to $\|\cdot\|_\infty$ since

$$|(f_n - f)(t)| = \left| \int_0^t f_n'(s) - f'(s) \, \mathrm{d}s \right| \le \left\|f_n' - f'\right\|_\infty \qquad \text{for each } t \in (0,1]$$

for every index n. So the function f is in $\operatorname{dom}(L)$ since $-f' = -g \in \mathrm{C}_0((0,1])$, and one has $f_n \to f$ with respect to the norm $\|\cdot\|_L$, which show the claimed completeness of $\operatorname{dom}(L)$ with respect to $\|\cdot\|_L$. ☐

[5] This is a typical PDE argument, so we do not go into details here and leave the proof of this claim for a course in PDE instead.

Proposition 4.1.4 (Continuity with respect to a graph norm). *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X. If $\|\cdot\|_L$ is a graph norm on* $\operatorname{dom}(L)$*, then the mappings*

$$\operatorname{id} : \operatorname{dom}(L) \to X \qquad \text{and} \qquad L : \operatorname{dom}(L) \to X$$

are continuous with respect to the norm $\|\cdot\|_L$ on $\operatorname{dom}(L)$ *and the norm $\|\cdot\|_X$ on X.*

Proof. It suffices to consider the specific graph norm $\|\cdot\|_L$ given by $\|x\|_L := \|x\|_X + \|Lx\|_X$ for all $x \in \operatorname{dom}(L)$. For every $x \in \operatorname{dom}(L)$ one obviously has $\|\operatorname{id}(x)\|_X = \|x\|_X \le \|x\|_L$ and $\|Lx\|_X \le \|x\|_L$, so both mappings $\operatorname{id}, L : \operatorname{dom}(L) \to X$ have norm at most 1 (with respect to the specific graph norm that we are considering). $\square$

Definition 4.1.5 (Dual operators). Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X and assume that L is densely defined. The *dual operator* $L' : X' \supseteq \operatorname{dom}(L') \to X$ is defined by

$$\begin{aligned} \operatorname{dom}(L') &:= \Big\{x' \in X' \mid \exists y' \in X' \text{ such that } \langle x', Lx\rangle = \langle y', x\rangle \text{ for all } x \in \operatorname{dom}(L)\Big\}, \\ L'x' &:= y' \quad \text{for all } x' \in \operatorname{dom}(L'), \end{aligned}$$

where $y' \in X'$ is the uniquely determined[6] element that occcurs in the definition of $\operatorname{dom}(L')$ for a given element $x' \in \operatorname{dom}(L')$.

Definition 4.1.6 (Spectrum and resolvent). Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X.

(a) The *resolvent set* $\rho(L)$ of L is the set of all numbers $\lambda \in \mathbb{C}$ for which the mapping

$$\lambda \operatorname{id} - L : \operatorname{dom}(L) \to X$$

is bijective and the inverse mapping $(\lambda \operatorname{id} - L)^{-1} : X \to \operatorname{dom}(L)$ is continuous from X to X.[7]

For every $\lambda \in \rho(L)$ the bounded linear operator $\mathcal{R}(\lambda, L) := (\lambda \operatorname{id} - L)^{-1} : X \to \operatorname{dom}(L)$ is called the *resolvent* of L at λ.

(b) The *spectrum* $\sigma(L)$ of L is defined as $\sigma(L) := \mathbb{C} \setminus \rho(L)$.

Remarks 4.1.7 (Closedness of operators and boundedness of the resolvent). Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X.

(a) If $\rho(L)$ is non-empty, then L is closed. Indeed, for $\lambda \in \rho(L)$ the resolvent $\mathcal{R}(\lambda, L)$ is continuous from X to X; in particular, it has closed graph. Hence, the set $\{(y, \mathcal{R}(\lambda, L)y) \mid y \in X\}$ is closed in $X \times X$ and thus, so is the set

$$\{(\mathcal{R}(\lambda, L)y, y) \mid y \in X\} = \{(x, (\lambda \operatorname{id} - L)x) \mid x \in \operatorname{dom}(L)\}.$$

From this one can readily derive that the graph of L is also closed in $X \times X$.

[6] Why is y' uniquely determined for each given $x' \in \operatorname{dom}(L')$?

[7] Here, $\operatorname{id} : \operatorname{dom}(L) \to X$ denotes the injection $x \mapsto x$.

(b) If $\lambda \in \rho(L)$, then $\mathcal{R}(\lambda, L)$ is even continuous from X to $\operatorname{dom}(L)$, where we endow $\operatorname{dom}(L)$ with any graph norm. Indeed, this follows from the continuity and injectivity of the embedding $\operatorname{dom}(L) \hookrightarrow X$ together with the closed graph theorem.

(c) If L is closed and $\lambda \in \mathbb{C}$ is such that $\lambda \operatorname{id} - L : \operatorname{dom}(L) \to X$ is bijective, then $\lambda \in \rho(L)$. Indeed, the closedness of L implies that the operator

$$\lambda \operatorname{id} - L : X \supseteq \operatorname{dom}(L) \to X$$

is closed, too, and hence, its inverse $(\lambda \operatorname{id} - L)^{-1} : X \to X$ is also closed. Since the inverse is defined everywhere on X, it is thus even continuous due to the closed graph theorem.

Proposition 4.1.8 (The resolvent is a pseudo resolvent)**.** *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X. If $\rho(L)$ is non-empty, then the resolvent $\mathcal{R}(\,\cdot\,, L) : \rho(L) \to \mathcal{L}(X)$, $\lambda \mapsto \mathcal{R}(\lambda, L)$ satisfies the resolvent identity (and is thus a pseudo-resolvent).*

Proof. For $\lambda, \mu \in \rho(L)$ one has

$$\underbrace{\lambda \operatorname{id} - L}_{X \leftarrow \operatorname{dom}(L)} = \underbrace{(\lambda - \mu) \operatorname{id}}_{X \leftarrow \operatorname{dom}(L)} + \underbrace{\mu \operatorname{id} - L}_{X \leftarrow \operatorname{dom}(L)} = \Big(\underbrace{(\lambda - \mu) \operatorname{id}}_{X \leftarrow \operatorname{dom}(L)} \underbrace{\mathcal{R}(\mu, L)}_{\operatorname{dom}(L) \leftarrow X} + \operatorname{id}_X \Big) \underbrace{(\mu \operatorname{id} - L)}_{X \leftarrow \operatorname{dom}(L)}$$

By composing this with the mapping $\mathcal{R}(\mu, L) : X \to \operatorname{dom}(L)$ from the right we get

$$\underbrace{(\lambda \operatorname{id} - L)}_{X \leftarrow \operatorname{dom}(L)} \underbrace{\mathcal{R}(\mu, L)}_{\operatorname{dom}(L) \leftarrow X} = \underbrace{(\lambda - \mu) \operatorname{id}}_{X \leftarrow \operatorname{dom}(L)} \underbrace{\mathcal{R}(\mu, L)}_{\operatorname{dom}(L) \leftarrow X} + \operatorname{id}_X .$$

Finally, we compose this with the mapping $\mathcal{R}(\lambda, L) : X \to \operatorname{dom}(L)$ from the left and thus obtain

$$\underbrace{\mathcal{R}(\mu, L)}_{\operatorname{dom}(L) \leftarrow X} = (\lambda - \mu) \underbrace{\mathcal{R}(\lambda, L)}_{\operatorname{dom}(L) \leftarrow X} \underbrace{\mathcal{R}(\mu, L)}_{X \leftarrow X} + \underbrace{\mathcal{R}(\lambda, L)}_{\operatorname{dom}(L) \leftarrow X} .$$

This is the resolvent equation, as claimed. ☐

Example 4.1.9 (The spectrum of a differential operator)**.** Consider the differential operator $L : \mathrm{C}_0((0,1]) \supseteq \operatorname{dom}(L) \to \mathrm{C}_0((0,1])$ from Example 4.1.3 that is given by $Lu = -u'$ for all $u \in \operatorname{dom}(L)$.

This closed operator is clearly injective, and for every $f \in \mathrm{C}_0((0,1])$ and every $\lambda \in \mathbb{C}$ the equation

$$(\lambda - L)u = f,$$

which is equivalent to

$$u' = -\lambda u + f$$

is solved by the function $u \in \operatorname{dom}(L)$ that is given by

$$u(t) = e^{-\lambda t} \int_0^t e^{\lambda s} f(s) \, \mathrm{d}s \qquad \text{for all } t \in (0,1].$$

So $\lambda \operatorname{id} - L$ is bijective, and hence $\lambda \in \rho(L)$ according to Remark 4.1.7(c). So $\rho(L) = \mathbb{C}$, and for each $\lambda \in \mathbb{C}$ and each $f \in \mathrm{C}_0((0,1])$ we have

$$\big(\mathcal{R}(\lambda, L) f\big)(t) = e^{-\lambda t} \int_0^t e^{\lambda s} f(s) \, \mathrm{d}s \qquad \text{for all } t \in (0,1].$$

Note that the resolvent $\mathcal{R}(\,\cdot\,, L) : \mathbb{C} \to \mathcal{L}\big(\mathrm{C}_0((0,1])\big)$ is precisely the pseudo-resolvent that we discussed in Example 2.5.2(d).[8]

4.2 An outlook on multi-valued operators

Definition 4.2.1 (Multi-valued operators)**.** Let X be a complex Banach space.

(a) A *multi-valued linear operator* on X is a vector subspace L of $X \times X$.

A multi-valued linear operator on X is called *closed* if it is closed in $X \times X$ (with respect to the product topology).

(b) Let $L \subseteq X \times X$ be a multi-valued linear operator on X. For each all $x \in X$ ones defines

$$Lx := \{ y \in X \mid \ (x,y) \in L \}.$$

(c) Let $L \subseteq X \times X$ be a multi-valued linear operator on X. The vector subspace

$$\begin{aligned} \operatorname{dom}(L) &:= \{ x \in X \mid \ \exists y \in X \text{ such that } (x,y) \in L \} \\ &= \{ x \in X \mid \ Lx \neq \emptyset \} \end{aligned}$$

of X is called the *domain* of X, and the vector subspace

$$\begin{aligned} \operatorname{ran}(L) &:= \{ y \in X \mid \ \exists x \in X \text{ such that } (x,y) \in L \} \\ &= \bigcup_{x \in X} Lx = \bigcup_{x \in \operatorname{dom}(L)} Lx \end{aligned}$$

is called the *range* or *image* of L.

(d) A multi-valued linear operator L on X is said to be *single-valued* if Lx has at most one element for each $x \in X$ (equivalently: if there exists a linear operator $X \supseteq \operatorname{dom}(L) \to X$ with graph L).

[8] Recall that we checked the resolvent identity in Example 2.5.2(d) by means of a concrete computation. Now that we know that this pseudo-resolvent is actually the resolvent of the unbounded operator L, the resolvent identity also follows directly from Proposition 4.1.8.

Equality of two multi-valued operators L_1, L_2 on a Banach space X is simply understood as equalities of the two subsets L_1 and L_2 of $X \times X$ and thus does not need to be defined separately. We note that, for a multi-valued operator L on a Banach space X, one has

$$L = \bigcup_{x \in \operatorname{dom}(L)} \{x\} \times Lx.$$

Definition 4.2.2 (Computations with multi-valued operators)**.** Let X be a complex Banach space, let L, L_1, L_2 be multivalued operators on X and $\alpha \in \mathbb{C}$.

(a) The multivalued operator αL on X is defined is

$$\alpha L := \{(x, \alpha y) \mid (x, y) \in L\}.$$

(b) The sum $L_1 + L_2$ is the multi-valued operator on X that is defined as

$$L_1 + L_2 = \{(x, y_1 + y_2) \mid (x, y_1) \in L_1 \text{ and } (x, y_2) \in L_2\}.$$

(c) The product $L_2 L_1$ is the multi-value operator on X that is defined as

$$L_2 L_1 := \{(x, z) \mid \exists y \in X \text{ such that } (x, y) \in L_1 \text{ and } (y, z) \in L_2\}.$$

(d) The *inverse* L^{-1} is the multi-valued operator on X that is defined as

$$L^{-1} := \{(y, x) \mid (x, y) \in L\}.$$

(e) The *dual* L' of L is the multi-valued operator on X' that is defined as

$$L' := \{(x', y') \mid \langle y', x \rangle = \langle x', y \rangle \text{ for all } (x, y) \in L\}.$$

One can check that if $L : X \supseteq \operatorname{dom}(L) \to X$ is a single-valued and densely defined linear operator, then its dual operator $L' : X' \supseteq \operatorname{dom}(L') \to X'$ in the sense of Definition 4.1.5 coincides with the dual of L in the sense of Definition 4.2.2(e).

In the following proposition we list a number of algebraic properties of multi-valued operators. Many of them are taken from [Haa06, Appendix A.1]. Most of those results follow by simply applying the definitions of the involved operations, so we do not include their proofs.[9] Property (j) in the proposition requires a bit of work with annihilators to be proved; we refer to [Haa06, A.4.2(d)] for details.

Proposition 4.2.3 (Algebraic properties of multi-valued operators)**.** *Let X be a complex Banach space and let L, L_1, L_2 be multi-valued operators on X.*

[9] To demonstrate explicitly how to work with those examples, the proof of assertion (d) was given in the lecture, though.

(a) *One has* $L_1 = L_2$ *if and only if* $\operatorname{dom}(L_1) = \operatorname{dom}(L_2)$ *and* $L_1 x = L_2 x$ *for each* $x \in \operatorname{dom}(L_1) = \operatorname{dom}(L_2)$.

(b) *Multiplication of multi-valued operators on* X *is associative.*

(c) *Addition of multi-valued operators on* X *is commutative and associative.*

(d) *If* $0_X, \operatorname{id}_X \in \mathcal{L}(X)$ *denote the zero and the identity operator on* X,[10] *then*

$$\operatorname{id}_X L = L \operatorname{id}_X = L \qquad \text{and} \qquad L + 0_X = 0_X + L = L.$$

Moreover, one has

$$\begin{aligned} & L\, 0_X = \{(x,z) \mid x \in X \text{ and } z \in L0\} = X \times L0 \\ \text{and} \qquad & 0_X\, L = \{(x,0) \mid x \in \operatorname{dom}(L)\}. \end{aligned}$$

(e) *For every* $\lambda \in \mathbb{C} \setminus \{0\}$ *one has* $\lambda L = (\lambda \operatorname{id}_X) L$ *and* $\lambda(L_2 L_1) = (\lambda L_2) L_1 = L_2(\lambda L_1)$.

(f) *One has* $(L_2 L_1)^{-1} = L_1^{-1} L_2^{-1}$.

(g) *One has* $(L')^{-1} = (L^{-1})'$.

(h) *If* $L_1 \subseteq L_2$, *then also*

$$L_1 L \subseteq L_2 L, \qquad L L_1 \subseteq L L_2, \qquad L_1 + L \subseteq L_2 + L.$$

(i) *One has the* distributivity inclusions

$$\begin{aligned} (L_1 + L_2) L &\subseteq L_1 L + L_2 L \\ L(L_1 + L_2) &\subseteq L L_1 + L L_2. \end{aligned}$$

If L *is single-valued, then the first inclusion is an equality; if* $\operatorname{ran}(L_1) \subseteq \operatorname{ran}(L)$ *or* $\operatorname{ran}(L_2) \subseteq \operatorname{ran}(L)$, *then the second inclusion in an equality.*

In particular, if $L \in \mathcal{L}(X)$ *then both inclusions are equalities.*

(j) *For the* bi-dual $L'' := (L')' \subseteq X'' \times X''$ *of* L *one has* $L'' \cap (X \times X) = \overline{L}$.[11]

Definition 4.2.4 (Spectrum and (pseudo-)resolvent of multi-valued operators)**.** Let L be a multi-valued linear operator on a complex Banach space X.

(a) The *resolvent set* $\rho(L)$ of L consists of those numbers $\lambda \in \mathcal{L}(X)$ for which $(\lambda \operatorname{id}_X - L)^{-1}$ is an element of $\mathcal{L}(X)$.

For $\lambda \in \rho(L)$ the operator $\mathcal{R}(\lambda, L) := (\lambda \operatorname{id}_X - L)^{-1} \in \mathcal{L}(X)$ is called the *resolvent of* L *at* λ.

[10] Note that, as we now identify each operator with its graph, we have $\operatorname{id}_X = \{(x,x) \mid x \in X\}$ and $0_X = \{(x,0) \mid x \in X\}$.

[11] Here, the closure $\overline{L}$ of L is taken in the product topology in $X \times X$, and as usual we identify X with a subspace of X''.

(b) The *spectrum* $\sigma(L)$ of L is defined as $\sigma(L) := \mathbb{C} \setminus \rho(L)$.

Let us now discuss some finite-dimensional examples in order to get a better intution of what is going on.

Examples 4.2.5. (a) Consider the multi-valued operator

$$L := \{(0,x) \mid x \in \mathbb{C}^d\} \subseteq \mathbb{C}^d \times \mathbb{C}^d$$

on $\mathbb{C}^d$. Then $\rho(L) = \mathbb{C}$ and $\mathcal{R}(\lambda, L) = 0$ for each $\lambda \in \mathbb{C}$.

(b) Consider the multi-valued operator

$$L := \left\{ \left(\begin{pmatrix} x \\ 0 \end{pmatrix}, \begin{pmatrix} y \\ x \end{pmatrix} \right) \mid x, y \in \mathbb{C} \right\} \subseteq \mathbb{C}^2 \times \mathbb{C}^2$$

on $\mathbb{C}^2$. For every $\lambda \in \mathbb{C}$ one has

$$\lambda \operatorname{id}_{\mathbb{C}^2} - L = \left\{ \left(\begin{pmatrix} x \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda x - y \\ -x \end{pmatrix} \right) \mid x, y \in \mathbb{C} \right\}$$

and thus

$$\begin{aligned} \left(\lambda \operatorname{id}_{\mathbb{C}^2} - L\right)^2 &= \left\{ \left(\begin{pmatrix} \lambda x - y \\ -x \end{pmatrix}, \begin{pmatrix} x \\ 0 \end{pmatrix} \right) \mid x, y \in \mathbb{C} \right\} \\ &= \left\{ \left(\begin{pmatrix} -\lambda x + y \\ x \end{pmatrix}, \begin{pmatrix} -x \\ 0 \end{pmatrix} \right) \mid x, y \in \mathbb{C} \right\} \\ &= \left\{ \left(\begin{pmatrix} z \\ x \end{pmatrix}, \begin{pmatrix} -x \\ 0 \end{pmatrix} \right) \mid x, z \in \mathbb{C} \right\} = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Hence, one has $\rho(L) = \mathbb{C}$ and

$$\mathcal{R}(\lambda, L) = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$$

for all $\lambda \in \mathbb{C}$.

Motivated by the previous example[12] we present the following theorem which shows how multi-valued operators are, in general, related to pseudo-resolvents. Since we have already seen various related arguments when dealing with spectral properties of pseudo-resolvents, we skip the proof. For further information on the relation between multi-valued operators and pseudo-resolvents, as well as their spectral properties, we refer to [Haa06, Appendix A.2].

Theorem 4.2.6 (Pseudo-resolvents vs. resolvents of multi-valued operators)**.** *Let X be a complex Banach space.*

[12]Which is a bit ironic since, as you can read above, I cannot remember the example right now. But well, at least I remember that I thought it would be a good motivation for Theorem 4.2.6...

(a) *Let L be a multi-valued operator on X assume that $\rho(L) \neq \emptyset$. Then L is closed, its resolvent set $\rho(L)$ is open, and*

$$\mathcal{R}(\,\cdot\,, L) : \rho(L) \to \mathcal{L}(X)$$

is a pseudo resolvent which is equal to its largest extension.

(b) *Conversely, let $\rho \subseteq \mathbb{C}$ be non-empty, and let $\mathcal{R} : \rho \to \mathcal{L}(X)$ be a pseudo-resolvent which already is its largest extension. For each $\lambda_0 \in \rho$ consider the multi-valued operator $L := (-1)\mathcal{R}(\lambda_0)^{-1} + \lambda_0 \operatorname{id}_X$. Then*

$$L = \left\{ \begin{pmatrix} \mathcal{R}(\lambda_0)x \\ \lambda_0 \mathcal{R}(\lambda_0)x - x \end{pmatrix} \mid x \in X \right\}$$

furthermore, the multi-valued operator L does actually not depend on λ_0, and one has $\rho = \rho(L)$ and

$$\mathcal{R}(\lambda) = \mathcal{R}(\lambda, L) \quad \text{for all } \lambda \in \rho = \rho(L).$$

Remark 4.2.7 (Representing pseudo-resolvents as operators)**.** Let $\mathcal{R} : \Omega \to A$ be a pseudo-resolvent on a unital Banach algebra A. In the proof of Theorem 2.5.5 we considered, for $\lambda \in \Omega$, the set $G(\lambda) \subseteq A \times A$ given by

$$G(\lambda) := \left\{ \begin{pmatrix} \mathcal{R}(\lambda)a \\ \lambda \mathcal{R}(\lambda)a - a \end{pmatrix} \mid a \in A \right\}.$$

Similar sets also occurred in Lemma 2.5.3(ii).

Theorem 4.2.6 now gives an intuition why it makes sense to consider those sets. If $A = \mathcal{L}(X)$ for a complex Banach space X, then every pseudo-resolvent on A is the resolvent of a multi-valued operator L on X and Theorem 4.2.6(b) shows how one can obtain L. Considering L in order to study a pseudo-resolvent is a natural choice since L does not depend on λ and determines the pseudo-resolvent uniquely.

For general unital Banach algebras A, though, we do not have a concept of multi-valued operator available. In order to still obtain an object that does not depend on λ and determines a given pseudo-resolvent uniquely, one can proceed as follows: Let us consider the unital Banach algebra A as a subalgebra of the operator space $\mathcal{L}(A)$ by identifying each $a \in A$ with the operator in $\mathcal{L}(A)$ that acts on A as $b \mapsto ab$. If $\mathcal{R} : \Omega \to A$ is a pseudo-resolvent, we thus obtain a pseudoresolvent $\Omega \to \mathcal{L}(A)$, and – Theorem 4.2.6(b) shows – the object $G(\lambda) \subseteq A \times A$ is precisely the multi-valued operator on A which is the given pseudo-resolvent as resolvent. This also explains at a more intuitive level why $G(\lambda)$ does not actually depend on λ.

Proposition 4.2.8 (The resolvent (set) of dual operators)**.** *Let X be a complex Banach space and let L be a closed multi-valued operator on X.*

(a) *One has $L \in \mathcal{L}(X)$ if and only if $L' \in \mathcal{L}(X')$.*

(b) *One has* $\rho(L) = \rho(L')$ *and* $\mathcal{R}(\lambda, L)' = \mathcal{R}(\lambda, L')$ *for each* $\lambda \in \rho(L) = \rho(L')$.

Proof. (a) If $L \in \mathcal{L}(X)$, then L' is single-valued, everywhere defined, and continuous; this is part of the course *Grundlagen der Funktionalanalysis*.

Conversely, assume now that $L' \in \mathcal{L}(X')$.[13] The fact that L' is everywhere defined readily implies that L is single-valued. As L is closed, it suffices to prove that L is everywhere defined, so let $x \in X$.

Since we have already proved the converse implication we know that $L'' \in \mathcal{L}(X'')$, and it follows from Proposition 4.2.3(j) and the closedness of L that $L = \overline{L} = L'' \cap (X \times X)$. If we can show that $L''x \in X$ one thus has $(x, L''x) \in L$ and the claim follows.

In order to show that the vector $L''x$ of X'' is actually an element of X one only has to prove that it is weak* continuous as a mapping from X' to $\mathbb{C}$. [14] In order to check this, it suffices to prove that $\ker(L''x)$ is weak*-closed,[15] and in order to do so it suffices, according to the Krein-Smulian theorem, to check that $\ker(L''x) \cap \mathrm{B}_{\leq r}(0)$ is weak*-closed in X' for every radius $r \geq 0$ (where $\mathrm{B}_{\leq r}(0)$ denotes the norm closed ball of radius r in X'). So let (y'_j) be a net in $\ker(L''x) \cap \mathrm{B}_{\leq r}(0)$ that weak* converges to a point $y' \in \mathrm{B}_{\leq r}(0)$; we may actually assume that this net is universal.[16] Then the net $(L'y'_j)$ is also universal according to Proposition A.3.3. Since L' is bounded, it follows from the Banach–Alaoglu Theorem A.4.3 and Theorem A.3.4 that the net $(L'y'_j)$ is weak*-convergent to a point $z' \in X'$. However, one can easily check that the dual of every multi-valued linear operator is weak*-closed, so we conclude that $z' = L'y'$. Hence,

$$\langle L''x, y' \rangle = \langle x, z' \rangle = \lim_j \langle x, L'y'_j \rangle = \lim_j \langle L''x, y'_j \rangle = 0,$$

so $y' \in \ker(L''x)$, as claimed.

(b) This readily follows from (a) and Proposition 4.2.3(g). $\square$

[13] A simpler argument than the one given in the following was kindly shown to me by Markus Haase: As L' is in $\mathcal{L}(X')$ it follows that $L'' \in \mathcal{L}(X'')$, so L is single-valued and continuous with respect to the norm on X. Moreover, as L' is single-value, L is densely defined. Since L was assumed to be closed, one can conclude that $L \in \mathcal{L}(X)$.

[14] This is a general fact about elements of X'', see for instance [Mur90, Theorem A.2].

[15] In fact, this characterisation of continuity for linear functionals is true in every *locally convex vector space*, see for instance [Mur90, Theorem A.3].

[16] This follows from the general fact that every net has a universal subnet. The notion *subnet* is, however, not discussed in Appendix A.

Chapter 5

Fine Structure of the Spectrum

5.1 Eigenvalues and beyond

Definition 5.1.1 (Eigenvectors, eigenvalues, and the point spectrum)**.** Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X.

(a) A number $\lambda \in \mathbb{C}$ is called an *eigenvalue* of L if $\lambda \operatorname{id}_X - L : \operatorname{dom}(L) \to X$ is not injective, i.e., if there exists a non-zero vector $x \in \operatorname{dom}(L)$ such that $Lx = \lambda x$; in this case, each such vector x is called an *eigenvector* of L for the eigenvalue λ, and the vector subspace

$$\ker(\lambda \operatorname{id} - L) = \{x \in \operatorname{dom}(L) \mid (\lambda \operatorname{id} - L)x = 0\} = \{x \in \operatorname{dom}(L) \mid Lx = \lambda x\}$$

is called the *eigenspace* of L for the eigenvalue λ.

(b) The set $\sigma_{\mathrm{pnt}}(L)$ of all eigenvalues of L is called the *point spectrum* of L.

Clearly every eigenvalue of L is a spectral value, i.e., $\sigma_{\mathrm{pnt}}(L) \subseteq \sigma(L)$.

Proposition 5.1.2 (Eigenspaces of closed linear operators are closed)**.** *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X, and assume that L is closed. Then* $\ker L$ *is closed in X, and hence* $\ker(\lambda \operatorname{id} - L)$ *is closed in X for each $\lambda \in \mathbb{C}$.*

Proof. Let (x_n) be a sequence in $\ker L$ that converges to a point $x \in X$. Then we have $x_n \in \operatorname{dom}(L)$ and $Lx_n = 0$ for each n, so it follows from the closedness of L (see Proposition 4.1.2) that $x \in \operatorname{dom}(L)$ and $Lx = 0$, i.e., $x \in \ker L$. □

Example 5.1.3 (A differential operator)**.** Consider the differential operator $L : \mathrm{C}([0,1]) \subseteq \operatorname{dom}(L) \to \mathrm{C}([0,1])$ that is given by $\operatorname{dom}(L) = \mathrm{C}^1([0,1])$ and $Lu = -u'$ for each $u \in \operatorname{dom}(L)$.

Then every $\lambda \in \mathbb{C}$ is an eigenvalue of L, i.e., $\mathbb{C} = \sigma_{\mathrm{pnt}}(L) = \sigma(L)$. Indeed, let $\lambda \in \mathbb{C}$, and let $u \in \operatorname{dom}(L)$ be given by $u(t) := e^{-\lambda t}$ for each $t \in [0,1]$. Then $u \neq 0$ and $Lu = \lambda u$.

Proposition 5.1.4 (Eigenvalues by means of the dual operators)**.** *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a closed linear operator on a complex Banach space X and assume that L is densely defined.*

(a) *A number $\lambda \in \mathbb{C}$ is an eigenvalue of L' (i.e., $\lambda \operatorname{id} - L' : \operatorname{dom}(L') \to X$ is not injective) if and only if $\lambda \operatorname{id} - L : \operatorname{dom}(L) \to X$ does not have dense range in X.*

(b) *A number $\lambda \in \mathbb{C}$ is an eigenvalue of L (i.e., $\lambda \operatorname{id} - L : \operatorname{dom}(L) \to X$ is not injective) if and only if $\lambda \operatorname{id} - L' : \operatorname{dom}(L') \to X'$ does not have weak* dense range in X'.*

Proof. Throughout the proof we may replace $\lambda \operatorname{id} - L$ with L and $\lambda \operatorname{id} - L'$ with L'.[1]

(a) „$\Rightarrow$“ Assume that the range of L is dense in X and let $x' \in \operatorname{dom}(L')$ such that $L'x' = 0$. For every x in the range of L there exists $v \in \operatorname{dom}(L)$ such that $Lv = x$, and hence $\langle x', x\rangle = \langle L'x', v'\rangle = 0$. Since the range of L is dense in X on x' is continuous, this implies that $\langle x', x\rangle = 0$ for all $x \in X$. Hence, $x' = 0$, so L' is indeed injective.

„$\Leftarrow$“ Assume that the range of L is not dense in X. Then it follows from the Hahn–Banach extension theorem that there exists a non-zero functional $x' \in X'$ that vanishes on the range of L. Thus, one has $\langle x', Lv\rangle = 0 = \langle 0, v\rangle$ for all $v \in \operatorname{dom}(L)$. By the definition of L' (Definition 4.1.5) this implies that $x' \in \operatorname{dom}(L')$ and $L'x' = 0$, so L' is not injective.

(b) „$\Rightarrow$“ Assume that L' has weak* dense range in X' and let $x \in \ker L$. For each x' in the range of L' there exists $v' \in \operatorname{dom}(L')$ such that $L'v' = x'$, so $\langle x', x\rangle = \langle v', Lx\rangle = 0$; so the range of L' is contanied in the weak*-closed subset $\{x' \in X' \mid \langle x', x\rangle = 0\}$ of X'. As the range of L' was assume to be dense in X' we conclude that $\langle x', x\rangle = 0$ for all $x' \in X'$. Due to the Hahn–Banach theorem it thus follows that $x = 0$. So L is indeed injective.

„$\Leftarrow$“ Assume that the range of L' does not have weak* dense range in X'. Then there exists a non-zero vector $x \in X$ such that $\langle x', x\rangle = 0$ for all x' in the range of L'.[2] Hence, for each $v' \in \operatorname{dom}(L')$ one has $\langle L'v', x\rangle = 0 = \langle v', 0\rangle$. So if we interpret L' as a multivalued operator and consider X as a subspace of X'' by means of evaluation, this implies that $(x, 0) \in L''$ (see Definition (e)). Thus it follows from Proposition 4.2.3(j) that $(x, 0) \in L$ since L is closed.[3] In other words, $x \in \ker L$ and $Lx = 0$, which shows that L is not injective. □

From linear algebra you already know *geometric* and *algebraic multiplicities* of eigenvalues of matrices. We will know discuss the same concepts for general linear operators on Banach spaces.[4]

[1] Here we used the the elementary observation that $(\lambda \operatorname{id} - L)' = \lambda \operatorname{id} - L'$.

[2] This follows, for instance, from the Hahn–Banach extension theorem in locally convex vector spaces (since the weak*-topology is locally convex).

[3] In this argument we used that it does not matter whether we compute the dual L' in the sense of Definition 4.1.5 or in the sense of Definition (e); this observation is Exercise 2 on Sheet 8.

[4] The presentation throughout the rest of this section essentially follows [Glü16, Appendix A.2].

Definition 5.1.5 (Generalized eigenspaces and multiplicities)**.** Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X and let $\lambda \in \mathbb{C}$ be an eigenvalue of L.

(a) The number $\dim \ker(\lambda \operatorname{id} - L) \in \mathbb{N} \cup \{\infty\}$ is called the *geometric multiplicity* of λ.

The eigenvalue λ is called *geometrically simple* if it has geometric multiplicity 1.

(b) The vector subspace $\bigcup_{n\in\mathbb{N}} \ker\big((\lambda \operatorname{id} - L)^n\big)$ of L is called the *generalized eigenspace* of L for λ.[5] Its dimension, which is an element of $\mathbb{N}\cup\{\infty\}$, is called the *algebraic multiplicity* of λ.

The eigenvalue λ is called *geometrically simple* if it has algebraic multiplicity 1.

(c) Let $n \in \mathbb{N}$. A vector $x \in X$ is called a *generalized eigenvector of rank n* for the eigenvalue λ of L if $x \in \ker\big((\lambda \operatorname{id} - L)^n\big) \setminus \ker\big((\lambda \operatorname{id} - L)^{n-1}\big)$.

(d) The eigenvalue λ is called *semi-simple* if its generalized eigenspace coincides with its eigenspace.

Obviously a vector $x \in X$ is a generalized eigenvector of rank 1 if and only if it is an eigenvector. If the eigenvalue λ has finite algebraic multiplicity, then it clearly follows that λ is semi-simple if and only if its algebraic and geometric multiplicity coincide.

Example 5.1.6 (Jordan normal form in finite dimensions)**.** Let $T \in \mathbb{C}^{d\times d}$ and let $\lambda \in \mathbb{C}$ be an eigenvalue of T. By considering the Jordan normal form of T one can check that the algebraic multiplicity as introduced in Definition 5.1.5(b) coincides with the multiplicity of λ as a root of the characteristic polynomial $z \mapsto \det(z - T)$. In other words, the definition of algebraic multiplicity that we gave above is consistent with the usual definition for matrices from linear algebra.

Here is a useful criterion to check whether an eigenvalue is semi-simple:

Proposition 5.1.7 (Criterion for the semi-simplicity of eigenvalue)**.** *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X and let $\lambda \in \mathbb{C}$ be an eigenvalue of L.*

(a) *Assume that, for some $m \in \mathbb{N}$, we have* $\ker\big((\lambda \operatorname{id} - L)^m\big) = \ker\big((\lambda \operatorname{id} - L)^{m+1}\big)$. *Then* $\ker\big((\lambda \operatorname{id} - L)^m\big) = \ker\big((\lambda \operatorname{id} - L)^n\big)$ *for each $n \geq m$, so* $\ker\big((\lambda \operatorname{id} - L)^m\big)$ *is the generalized eigenspace of λ.*

[5] Here the powers $(\lambda \operatorname{id} - L)^n$ are defined recursively. Recall that the product of two multi-valued operators was defined in Definition 4.2.2(c); it is not difficult to see that, with this definition, the product of two single-valued operators is again single-valued.

(b) *The eigenvalue λ of L is semi-simple if and only if it does not have a generalized eigenvector of rank* 2.

Proof. (a) It suffices to show that $\ker\big((\lambda \operatorname{id} - L)^{m+1}\big) = \ker\big((\lambda \operatorname{id} - L)^{m+2}\big)$; the claim then follows readily be induction.

The inclusion "$\subseteq$" is clear; to prove the converse inclusion "$\supseteq$", fix a vector $x \in \ker\big((\lambda \operatorname{id} - L)^{m+2}\big)$. Then $(\lambda \operatorname{id} - L)x \in \ker\big((\lambda \operatorname{id} - L)^{m+1}\big) = \ker\big((\lambda \operatorname{id} - L)^{m}\big)$, so it follows that $(\lambda \operatorname{id} - L)^{m+1}x = 0$, as claimed.

(b) The implication "$\Rightarrow$" is clear, and the converse implication "$\Leftarrow$" follows from (a). $\square$

Let us give a nice application of the preceding proposition, which is a generalization of the finite-dimensional observation that a power-bounded matrix can only have semi-simple eigenvalues on the unit circle.[6]

Corollary 5.1.8 (Unimodular eigenvalues of power-bounded operators)**.** *Let X be a complex Banach space and let $T \in \mathcal{L}(X)$ be power-bounded.*[7] *If λ is an eigenvalue of T and has modulus $|\lambda| = 1$, then λ is semi-simple.*

Proof. Let $x \in \ker\big((\lambda - T)^2\big)$. According to Proposition 5.1.7(b) it suffices to show that $x \in \ker(\lambda - T)$.

Let us set $y := (\lambda - T)x$. Then $Ty = \lambda y$, and it follows by induction that

$$T^n x = -n\lambda^{n-1} y + \lambda^n x \qquad \text{for all } n \in \mathbb{N}.$$

If we devide by n, use that $|\lambda| = 1$, and let $n \to \infty$, we thus see that $y = 0$. So $x \in \ker(\lambda - T)$, as claimed. $\square$

5.2 Intermezzo: Vector-valued Laurent series expansion

Definition 5.2.1 (Isolated singularities, poles and order of poles)**.** (a) Let $C \subseteq \mathbb{C}$ be closed. A complex number z_0 is called an *isolated point of C* if $z_0 \in C$ and $C \setminus \{z_0\}$ is also closed.[8]

(b) Let X be a complex Banach space, let $\Omega \subseteq \mathbb{C}$ be non-empty and open, and let $f : \Omega \to X$ be holomorphic. A complex number z_0 is called an *isolated singularity of f* if z_0 is an isolated point of $\mathbb{C} \setminus \Omega$ and f cannot be extended to a holomorphic function $\Omega \cup \{z_0\} \to X$.

[6] This result can be derived by using the Jordan normal form of matrices, and our proof in the infinite-dimensional case essentially resembles the essence of this argument.

[7] Which means that $\sup_{n\in\mathbb{N}} \|T^n\| < \infty$. Note that this implies $\mathrm{r}(T) \leq 1$, so in particular every eigenvalue of T has modulus at most 1.

[8] Equivalently, if $z_0 \in C$ and there exists a number $\varepsilon > 0$ such that the pointed disk with radius ε around z_0 – i.e., the set $\{z \in \mathbb{C} \mid 0 < |z - z_0| < \varepsilon\}$ – does not intersect C.

Theorem 5.2.2 (Laurent series expansion about isolated singularities)**.** *Let X be a complex Banach space, let $\Omega \subseteq \mathbb{C}$ be non-empty and open, and let $f : \Omega \to X$ be holomorphic. Assume that z_0 is an isolated singularity of f. Then there exists a number $r > 0$ and vectors $(a_k)_{k\in\mathbb{Z}}$ in X such that the pointed disk $\{z \in \mathbb{C} \mid 0 < |z - z_0| < r\}$ is contained in Ω and such that, for each z from this pointed disk, one has*

$$f(z) = \sum_{k=-\infty}^{\infty} a_k (z - z_0)^k,$$

where the series converges absolutes.

The coefficients $a_k \in X$ are uniquely determined, and given by

$$a_k = \frac{1}{2\pi i} \oint_\gamma \frac{f(z)}{(z - z_0)^{k+1}} \, \mathrm{d}z \qquad \text{for all } k \in \mathbb{Z},$$

where γ is any circle around z_0 with radius strictly less than r that encircles z_0 counterclockwise. Moreover, there exists at least one index $k < 0$ such that $a_k \neq 0$.

The formula in the preceding theorem is called the *Laurent series expansion* of f about z_0.

Definition 5.2.3 (Poles and essential singularities)**.** In the situation of the previous theorem the number z_0 is called ...

(a) a *pole* of f if only finitely many of the coefficients a_k for $k < 0$ are non-zero. In this case the larges number $q \in \mathbb{N}$ such that $a_{-q} \neq 0$ is called the *order* of the pole z_0.

(b) an *essential singularity* of f if it is not a pole of f.

Proposition 5.2.4 (Characterisation of poles and pole order)**.** *Let X be a complex Banach space, let $\Omega \subseteq \mathbb{C}$ be non-empty and open, and let $f : \Omega \to X$ be holomorphic. Assume that z_0 is an isolated singularity of f.*

(a) *Let $n \in \mathbb{N}$. The number z_0 is a pole of f of order at most n if and only if $\lim_{z\to z_0}(z - z_0)^{n+1} f(z) = 0$ if and only if the limit $\lim_{z\to z_0}(z - z_0)^n f(z)$ exists in X.*

(b) *Let $n \in \mathbb{N}$ and assume that z_0 is a pole of f. Let (z_j) be a sequence in Ω that converges to z_0. Then the pole order of z_0 is most n if and only if $\lim_{j\to\infty}(z_j - z_0)^{n+1} f(z_j) = 0$ if and only if $\lim_{j\to\infty}(z_j - z_0)^n f(z_j)$ exists.*

(c) *Assume that z_0 is a pole of f and let $q \in \mathbb{N}$ denote its order. Then the $-q$-th coefficient of the Laurent series expansion of f about z_0 is equal to $\lim_{z\to z_0}(z - z_0)^q f(z)$.*

Example 5.2.5 (An essential singularity of a scalar-valued function)**.** Consider the holomorphic mapping $f : \mathbb{C}\setminus\{0\} \to \mathbb{C}$, $z \mapsto \exp(-1/z)$. It has an isolated singularity at 0, and its Laurent series expansion about 0 is given by

$$f(z) = \sum_{k=-\infty}^{0} \frac{(-1)^k}{(-k)!} z^k$$

for all $z \in \mathbb{C} \setminus \{0\}$. Hence 0 is an essential singularity of f.

This examples shows that assertion (b) in Proposition 5.2.4 does not remain true if one drops the a priori assumption that z_0 be a pole. Indeed, in the present example the function f is bounded on the right half plan $\{z \in \mathbb{C} \mid \operatorname{Re} z > 0\}$, so

$$\left(\frac{1}{n} - 0\right) f(1/n) \to 0$$

as $n \to \infty$.

5.3 Poles of resolvents

For a closed linear operator $L : X \supseteq \operatorname{dom}(L) \to X$ on a complex Banach space X we use the notation $\operatorname{dom}(L^\infty) := \bigcap_{j\in\mathbb{N}_0} \operatorname{dom}(L^j)$.

Theorem 5.3.1 (Isolated singularities of the resolvent)**.** *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a closed linear operator on a complex Banach space X and let $\lambda_0 \in \sigma(L)$ be an isolated singularity of the resolvent map $\mathcal{R}(\,\cdot\,, L) : \rho(L) \to \mathcal{L}(E)$. Let*

$$\mathcal{R}(\lambda, L) = \sum_{k=-\infty}^{\infty} Q_{k+1}(\lambda - \lambda_0)^k.$$

denote the Laurent series expansion of the resolvent about λ_0.[9] *Then all the operators $Q_k \in \mathcal{L}(X)$ commute, and the following assertions hold:*

(a) *For each $k \in \mathbb{Z}$ one has $Q_k X \subseteq \operatorname{dom}(L)$ and $Q_k L x = L Q_k x$ for all $x \in \operatorname{dom}(L)$.*

(b) *For all $k \geq 1$ one has*

$$Q_k = (-1)^{k+1} Q_1^k \qquad \text{and} \qquad Q_{-k} = (Q_{-1})^k.$$

(c) *For all $k, \ell \geq 1$ one has $Q_k Q_{-\ell} = 0$.*

[9] Note that we enumerated the coefficients Q_k in a slightly unusual manner: with our notation we have

$$\mathcal{R}(\lambda, L) = \cdots + Q_{-1}(\lambda-\lambda_0)^{-2} + Q_0(\lambda-\lambda_0)^{-1} + Q_1(\lambda-\lambda_0)^0 + Q_2(\lambda-\lambda_0)^1 + \ldots;$$

this makes many formulas easier (and arguably also a bit more intuitive).

(d) *The operator Q_0 is a projection. Moreover, it satisfies*

$$Q_0 Q_k = 0 \qquad \textit{and} \qquad Q_0 Q_{-k} = Q_{-k}$$

for all $k \geq 1$.

(e) *For every $k \in \mathbb{Z} \setminus \{1\}$ one has*

$$(\lambda_0 - L) Q_k = -Q_{k-1}.$$

For Q_1 on the other hand, one has

$$(\lambda_0 - L) Q_1 = \operatorname{id} - Q_0.$$

(f) *The operator Q_{-1} has spectral radius 0, and hence the same is true for Q_{-k} for all $k \geq 1$.*

(g) *If $q \geq 1$ and $Q_{-q} = 0$, then λ_0 is a pole of the resolvent, and its pole order is at most q.*

(h) *If $d := \dim(Q_0 X) < \infty$, then λ_0 is a pole of the resolvent of order at most d.*

(i) *For each integer $k \geq 0$ one has $Q_{-k} X \subseteq \operatorname{dom}(L^\infty)$*

Let us discuss an explicit example before we go ahead with the proof of the theorem.

Example 5.3.2. On $X = \mathbb{C}^3$ consider the operator $L \in \mathcal{L}(X)$ given by[10]

$$L = \begin{pmatrix} 1 & -1 & \\ 1 & -1 & \\ & & 1 \end{pmatrix}.$$

A short computation shows that $\sigma(L) = \{0, 1\}$ and that

$$L = \lambda^{-2} \underbrace{\begin{pmatrix} 1 & -1 & \\ 1 & -1 & \\ & & 0 \end{pmatrix}}_{=Q_{-1}} + \lambda^{-1} \underbrace{\begin{pmatrix} 1 & 0 & \\ 0 & 1 & \\ & & 0 \end{pmatrix}}_{=Q_0} + \sum_{k=0}^{\infty} \lambda^k \underbrace{\begin{pmatrix} 0 & 0 & \\ 0 & 0 & \\ & & 1 \end{pmatrix}}_{=Q_{k+1} \text{ for all } k \geq 0}$$

for all $\lambda \in \mathbb{C}$ which satisfy $0 < |\lambda| < 1$. So have have computed the Laurent series expansion of $\mathcal{R}(\,\cdot\,, L)$ about the isolated singularity 0; this singularity is a pole of order 2.

[10]Here we identify L with its representation matrix with respect to te standard basis of $\mathbb{C}^3$.

Proof of Theorem 5.3.1. We first note that, according to Theorem 5.2.2,

$$Q_k = \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{\mathcal{R}(\lambda, L)}{(\lambda - \lambda_0)^k} \, \mathrm{d}\lambda \tag{5.3.1}$$

for all $k \in \mathbb{Z}$, where γ denotes any sufficiently small circle about λ_0 which is directed counter clockwise. Since the resolvent operators of L all commute, this readily implies that the Q_k mutually commute.

(a) Endow $\operatorname{dom}(L)$ with a graph norm $\|\cdot\|_L$. This renders $\operatorname{dom}(L)$ a Banach space since L is closed.

For every $\lambda \in \rho(L)$ the operator $\mathcal{R}(\lambda, L) : X \to \operatorname{dom}(L)$ is continuous as a consequence of the closed graph theorem. Moreover, the mapping $\mathcal{R}(\,\cdot\,, L) : \rho(L) \to \mathcal{L}(X; \operatorname{dom}(L))$ is continuous; this follows from the preceding sentence together with the fact that the mapping is continuous with values in $\mathcal{L}(X)$ and the resolvent identity.

But thus the integral in (5.3.1) also exists as a Riemann integral with values in $\mathcal{L}(X; \operatorname{dom}(L))$, and as the latter space embeds continuously into $\mathcal{L}(X)$, it follows that the integrals in both spaces coincide. Hence, $Q_k \in \mathcal{L}(X; \operatorname{dom}(L))$ and therefore $Q_k X \subseteq \operatorname{dom}(L)$ for each $k \in \mathbb{Z}$.

For every $x \in \operatorname{dom}(L)$ and every $\lambda \in \rho(L)$ one has $L\mathcal{R}(\lambda, L)x = \mathcal{R}(\lambda, L)x$; by using applying the equality (5.3.1) and together with the facts that the integral in this equality can be interpreted as a Riemann integral in $\mathcal{L}(X; \operatorname{dom}(L))$ and that L is continuous from $\operatorname{dom}(L)$ to X, we thus obtain $Q_k L x = L Q_k x$ for every $k \in \mathbb{Z}$.

(b), (c), and (d) Let $k_1, k_2 \in \mathbb{Z}$ and use formula (5.3.1) for two different small circles γ_1, γ_2 with center λ_0, where γ_2 has larger radius than γ_1. Then

$$\begin{aligned}
Q_{k_2} Q_{k_1} =& \frac{1}{(2\pi \mathrm{i})^2} \oint_{\gamma_2} \oint_{\gamma_1} \frac{\mathcal{R}(\lambda_2, L)\mathcal{R}(\lambda_1, L)}{(\lambda_2 - \lambda_0)^{k_2} (\lambda_1 - \lambda_0)^{k_1}} \, \mathrm{d}\lambda_1 \, \mathrm{d}\lambda_2 \\
=& \frac{1}{(2\pi \mathrm{i})^2} \oint_{\gamma_2} \oint_{\gamma_1} \frac{\mathcal{R}(\lambda_2, L) - \mathcal{R}(\lambda_1, L)}{(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_0)^{k_2} (\lambda_1 - \lambda_0)^{k_1}} \, \mathrm{d}\lambda_1 \, \mathrm{d}\lambda_2 \\
=& \frac{1}{(2\pi \mathrm{i})^2} \oint_{\gamma_2} \frac{\mathcal{R}(\lambda_2, L)}{(\lambda_2 - \lambda_0)^{k_2}} \oint_{\gamma_1} \frac{1}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_0)^{k_1}} \, \mathrm{d}\lambda_1 \, \mathrm{d}\lambda_2 \\
&- \frac{1}{(2\pi \mathrm{i})^2} \oint_{\gamma_1} \frac{\mathcal{R}(\lambda_1, L)}{(\lambda_1 - \lambda_0)^{k_1}} \oint_{\gamma_2} \frac{1}{(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_0)^{k_2}} \, \mathrm{d}\lambda_2 \, \mathrm{d}\lambda_1
\end{aligned}$$

By employing the residuu theorem one can compute the above integrals (where one has to distinguish several cases based on the signs of k_1 and k_2) and thus obtains the formulas claimed in (b), (c), and (d).

(e) Fix $k \in \mathbb{Z}$. For every $\lambda \in \rho(L)$ one has $(\lambda_0 - L)\mathcal{R}(\lambda, L) = (\lambda_0 - \lambda)\mathcal{R}(\lambda, L) + \mathrm{id}$. Since the integral in the formula (5.3.1) can be interpreted as a Riemann integral in $\mathcal{L}(X; \operatorname{dom}(L)$ and since $\lambda_0 - L : \operatorname{dom}(L) \to X$ is continuous, it follows that

$$(\lambda_0 - L) Q_k = \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{(\lambda_0 - \lambda)\mathcal{R}(\lambda, L) + \mathrm{id}}{(\lambda - \lambda_0)^k} \, \mathrm{d}\lambda$$

$$= \frac{1}{2\pi \mathrm{i}} \int_\gamma - \frac{\mathcal{R}(\lambda, L)}{(\lambda - \lambda_0)^{k-1}} \,\mathrm{d}\lambda + \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{\mathrm{id}}{(\lambda - \lambda_0)^k} \,\mathrm{d}\lambda;$$

the first summand is $-Q_{k-1}$, and the second summand is equal to 0 if $k \neq 1$, and equal to id if $k = 1$.

(f) Let $\varepsilon > 0$ be sufficiently small, such that (5.3.1) holds for the circle γ with radius ε about λ_0. For every integer $k \geq 1$ it follows from (b) that

$$\left\| Q_{-1}^k \right\| = \|Q_{-k}\| \leq \frac{1}{2\pi} \oint_\gamma \frac{\|\mathcal{R}(\lambda, L)\|}{\varepsilon^{-k}} \,\mathrm{d}\,|\lambda| \leq \sup_{\lambda \in \gamma} \|\mathcal{R}(\lambda, L)\| \, \varepsilon^{k+1}.$$

By taking the k-th root and letting $k \to \infty$ we thus see that $\mathrm{r}(Q_{-1}) \leq \varepsilon$. This shows that $\mathrm{r}(Q_{-1}) = 0$, as claimed.

As $Q_{-k} = Q{-1}^k$ for all $k \geq 1$, it follows from the spectral mapping theorem for polynomials that also $\mathrm{r}(Q_{-k}) = 0$ for all $k \geq 1$.

(g) If $Q_{-q} = 0$ then it follows from (b) that also $Q_{-(q+j)} = Q_{-q} Q_{-j} = 0$ for all $j \geq 0$; this shows the claim.

(h) The operator Q_{-1} commutes with Q_0, so it leaves the range of Q_0 invariant. Moreover, Q_{-1} has spectral radius 0, so its restriction to $Q_0 X$ is nilpotent; more precisely, the d-th poer of this restriction is nilpotent. Hence, $Q_{-d} = Q_{-1}^d = Q_{-1}^d Q_0 = 0$, so according to (g) λ_0 is indeed a pole of order at most d.

(i) Fix an integer $k \leq 0$. We show by induction over n that $Q_{-k} X \subseteq \mathrm{dom}(L^n)$ for each $n \in \mathbb{N}$. In (a) we proved the claim for $n = 1$, so assume now that we proved it for a given number $n \in \mathbb{N}$. Let $x \in X$. It follows from (d) that $Q_{-k} = Q_{-k} Q_0$, and from (a) that $Q_0 x \in \mathrm{dom}(L)$. Hence,

$$L^n Q_{-k} x = L^n Q_{-k} Q_0 x = L^{n-1} Q_{-k} L Q_0 x,$$

where we used the formula from (a) for the second equality (which is possible since $Q_0 x \in \mathrm{dom}(L)$).

The vector $L^{n-1} Q_{-k} L Q_0 x$ is in $\mathrm{dom}(L)$ since Q_{-k} maps X into $\mathrm{dom}(L^n)$ by the induction hypothesis. So we showed that $L^n Q_{-k} x \in \mathrm{dom}(L)$, which means that $Q_{-k} x \in \mathrm{dom}(L^{n+1})$, as claimed. □

Note that property (b) in the theorem above implies that λ_0 is a pole of $\mathcal{R}(\,\cdot\,, L)$ if and only if Q_{-1} is nilpotent. In this case the pole order is the smallest interger $q \geq 1$ such that $(Q_{-1})^q = 0$.

Remark 5.3.3 (The action of $\boldsymbol{Q_{-1}}$)**.** In the situation of Theorem 5.3.1 it follows from (d) and (e) that

$$Q_{-1} Q_0 = Q_{-1} = (L - \lambda_0) Q_0.$$

In other words, on the range of the projection Q_0 the operator Q_{-1} acts as the operator $L - \lambda_0$.

Theorem 5.3.4 (Poles of resolvents)**.** *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a closed linear operator on a complex Banach space X and let $\lambda_0 \in \sigma(L)$ be a pole of the resolvent $\mathcal{R}(\,\cdot\,, L) : \rho(L) \to \mathcal{L}(X)$ of order $q \geq 1$. Let*

$$\mathcal{R}(\lambda, L) = \sum_{n=-q}^{\infty} Q_{k+1}(\lambda - \lambda_0)^k.$$

denote the Laurent series expansion of the resolvent about λ_0.[11]

(a) *One has $\{0\} \neq Q_{-(q-1)}X \subseteq \ker(\lambda_0 - L)$. So in particular, λ_0 is an eigenvalue of L.*

(b) *For all $k \geq q$ the kernel $\ker\big((\lambda_0 \operatorname{id} - L)^k\big)$ does not depend on k and coincides with Q_0X. Hence, Q_0X is the generalized eigenspace of L for the eigenvalue λ_0.*

(c) *For all $k \geq q$ the range of $(\lambda_0 \operatorname{id} - L)^k$ does not depend on k and coincides with $\ker Q_0$.*

(d) *For the eigenvalue λ_0 there exists a generalized eigenvector of rank q, but no generalized eigenvector of rank strictly larger then q.*

In fact, if $x \in X$ satisfies $Q_{-(q-1)}x \neq 0$ then Q_0x is a generalized eigenvector of rank q.

Proof. (a) As λ_0 is a pole of order q we clearly have $Q_{-(q-1)} \neq 0$, so $Q_{-(q-1)}X \neq \{0\}$. Moreover it follows from Theorem 5.3.1(e) that $(\lambda_0 - L)Q_{-(q-1)} = -Q_{-q} = 0$, so $Q_{-(q-1)}$ does indeed map into $\ker(\lambda_0 - L)$.

(b) Fix $k \geq q$. It follows from Theorem 5.3.1(e) that

$$(\lambda_0 - L)^k Q_0 = (-1)^k Q_{-k} = 0,$$

so $Q_0X \subseteq \ker\big((\lambda_0 - L)^k\big)$. Conversely, let $x \in \ker\big((\lambda_0 - L)^k\big)$. Since Q_0 is a projection, so is $\operatorname{id} - Q_0$, and hence it follows from Theorem 5.3.1(e) that

$$(\operatorname{id} - Q_0)x = (\operatorname{id} - Q_0)^k x = Q_1^k(\lambda_0 - L)^k x = 0.$$

This proves that $x \in Q_0X$.[12]

(c) Fix $k \geq q$. It follows from (b) that the range of $(\lambda_0 - L)^k$ is contained in the kernel of Q_0. Conversely, let x be in the kernel of Q_0. Then x is in the range of $\operatorname{id} - Q_0 = (\operatorname{id} - Q_0)^k = (\lambda_0 - L)^k Q_1^k$, thus in the range of $(\lambda_0 - L)^k$.

(d) It follows from (b) that there is no generalized eigenvector of rank strictly larger than q.

[11] Note that $Q_{-(q-1)} \neq 0$ since q is the pole order of λ_0.

[12] Note that this argument could just as well be used to show directly that $\ker\big((\lambda_0 - L)^k\big) \subseteq Q_0X$ for every $k \geq 1$ rather than just for every $k \geq q$; however, this is not important here since the spaces $\ker\big((\lambda_0 - L)^k\big)$ are increasing with respect to k anyway.

Now let $x \in X$ be such that $Q_{-(q-1)}x \neq 0$. For every $k \geq 0$ we have $(\lambda_0 - L)^k Q_0 x = (-1)^k Q_{-k} x$, so $(\lambda_0 - L)^q Q_0 x = (-1)^q Q_{-q} x = 0$, but $(\lambda_0 - L)^{q-1} Q_0 x = Q_{-(q-1)} x \neq 0$. □

We note in passing that, iIn the situation of Theorem 5.3.1, the projection Q_0 is often called the *spectral projection* associated to the spectral value λ_0; its range $Q_0 X$ is called the associated *spectral space*. We will discuss a more general concept of spectral projections later on.

5.4 Approximate eigenvalues

Definition 5.4.1 (Approximate eigenvectors and eigenvalues)**.** Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X.

(a) A number $\lambda \in \mathbb{C}$ is called an *approximate eigenvalue* of L if there exists a sequence (x_n) in $\operatorname{dom}(L)$ which satisfies both $\liminf_{n\to\infty} \|x_n\|_X > 0$ and $(\lambda - L)x_n \to 0$ in X.

In this case, any such sequence (x_n) is called an *approximate eigenvector* of L for the approximate eigenvalue λ.

(b) The set of all approximate eigenvalues of L is called the *approximate point spectrum* of L; we denote it by $\sigma_{\mathrm{appr}}(L)$.

Obviously, every eigenvalue is also an approximate eigenvalue; moreover, it is easy to check that every approximate eigenvalue is a spectral value. So we have

$$\sigma_{\mathrm{pnt}}(L) \subseteq \sigma_{\mathrm{appr}}(L) \subseteq \sigma(L).$$

The condition $\liminf_{n\to\infty} \|x_n\|_X$ in Definition 5.4.1 can be replaced by various similar properties; we list them in the following proposition.

Proposition 5.4.2 (Norm conditions on approximate eigenvectors)**.** *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X, and endow $\operatorname{dom}(L)$ with a graph norm $\|\cdot\|_L$; let $\lambda \in \mathbb{C}$. The following are equivalent:*

(i) *The number λ is an approximate eigenvalue of L, i.e., there exists a sequence (x_n) in $\operatorname{dom}(L)$ such that $\liminf_{n\to\infty} \|x_n\|_X > 0$ and $(\lambda - L)x_n \to 0$ in X.*

(ii) *There exists a sequence (x_n) in $\operatorname{dom}(L)$ such that $\|x_n\|_X = 1$ for each n and such that $(\lambda - L)x_n \to 0$ in X.*

(iii) *There exists a sequence (x_n) in $\operatorname{dom}(L)$ such that $\liminf_{n\to\infty} \|x_n\|_L > 0$ and $(\lambda - L)x_n \to 0$ in X.*

(iv) *There exists a sequence (x_n) in $\operatorname{dom}(L)$ such that $\|x_n\|_L = 1$ for each n and such that $(\lambda - L)x_n \to 0$ in X.*

Proof. „(i) $\Leftrightarrow$ (ii)“ This is straightforward to prove.

„(i) $\Rightarrow$ (iii)“ This follows directly from the fact that the graph norm is stronger than the norm on X.

„(iii) $\Rightarrow$ (i)“ We may assume that the graph norm is given by $\|x\|_L = \|x\|_X + \|Lx\|_X$ for all $x \in \operatorname{dom}(L)$. For every index n we then have

$$\|x_n\|_L \leq (1 + |\lambda|)\, \|x_n\|_X + \|(L - \lambda)x_n\|_X \,,$$

so

$$\|x_n\|_X \geq \frac{1}{1 + |\lambda|}\Big(\|x_n\|_L - \|(L - \lambda)x_n\|_X \Big).$$

So $\liminf_{n \to \infty} \|x_n\|_X > 0$.

„(iii) $\Leftrightarrow$ (iv)“ This is straightforward to prove. $\square$

Examples 5.4.3 (A differential operator)**.** Consider the Banach space $X = \mathrm{C}_0([0, \infty))$ of complex-valued continuous functions on $[0, \infty)$ that vanish at infinity, endowed with the $\|\cdot\|_\infty$-norm. Consider the operator $L : X \supseteq \operatorname{dom}(L) \to X$ that is given by

$$\begin{aligned} \operatorname{dom}(L) &:= \big\{ u \in X \mid\ u \text{ is differentiable and } u' \in X \big\}, \\ Lu &= -u' \quad \text{for every } u \in \operatorname{dom}(L). \end{aligned}$$

Let $\lambda \in \mathbb{C}$. We claim that:

(a) If $\operatorname{Re} \lambda > 0$, then λ is an eigenvalue of L.

(b) If $\operatorname{Re} \lambda = 0$, then λ is an approximate eigenvalue, but not an eigenvalue of L.

(c) If $\operatorname{Re} \lambda < 0$, then $\lambda \in \rho(L)$.

Proof. (a) Consider the function $u : [0, \infty) \to \mathbb{C}$, $t \mapsto e^{-t\lambda}$. This function can be easily checked to be in $\operatorname{dom}(L)$ since $\operatorname{Re} \lambda > 0$, and one has $Lu = \lambda u$. Thus, indeed $\lambda \in \sigma_{\mathrm{pnt}}(L)$.

(b) For every integer $n \geq 1$ consider the function

$$u_n : [0, \infty) \to \mathbb{C}, \qquad t \mapsto e^{-t(\lambda + \frac{1}{n})}.$$

As $\operatorname{Re} \lambda = 0$ one can readily check that each u_n is in $\operatorname{dom}(L)$. Moreover, $\|u_n\|_\infty = 1$ for every n. On the other hand, one has

$$(\lambda - L)u_n = \lambda u_n - (\lambda + \frac{1}{n})u_n = -\frac{1}{n}u_n \to 0$$

with respect to $\|\cdot\|_X$.

To show that $\lambda \notin \sigma_{\mathrm{pnt}}(L)$ we need to check that $\lambda - L$ is injective. Actually, we will even prove this under the more general assumption $\operatorname{Re} \lambda \leq 0$, since we will also need this observation for the proof of (c). If $u \in \operatorname{dom}(L)$ and $(\lambda - L)u = 0$, then $u' = -\lambda u$, so

$$u(t) = e^{-\lambda t} u(0) \quad \text{for all } t \in [0, \infty).$$

As $\operatorname{Re}\lambda \leq 0$ one has $\left|e^{-\lambda t}\right| \geq 1$ for each $t \in [0, \infty)$, and thus

$$|u(0)| \leq \left|e^{-\lambda t} u(0)\right| = |u(t)| \to 0$$

as $t \to \infty$ since $u \in X$. Hence $u(0) = 0$, and thus $u = 0$, as claimed.

(c) We already know from the last argument in the proof of (b) that $\lambda - L : \operatorname{dom}(L) \to X$ is injective. To show that it is also surjective, fix $f \in X$. Consider the function

$$u : [0, \infty) \to X, \quad t \mapsto -e^{-t\lambda} \int_t^\infty e^{\lambda s} f(s) \, \mathrm{d}s.$$

This function is continuously differentiable, and by using that $f(t) \to 0$ as $t \to \infty$ one can check that also $u(t) \to 0$ as $t \to \infty$; so $u \in X$. Moreover, a brief computation shows that $u' = -\lambda u + f$. So one the one hand, $u' \in X$ and thus $u \in \operatorname{dom}(L)$; and on the other hand $(\lambda - L)u = \lambda u + u' = f$. This proves the claimed surjectivity.

The fact that the inverse mapping of $\lambda - L$ is continuous from X to X can either be concluded by showing that L is closed, or by using the explicit formula from the inverse that we derived above and proving its continuity with a simple estimate. □

Recall that a linear operator $T : X \to Y$ between to normed spaces X and Y is said to be *bounded below* if there exists a number $c > 0$ such that $\|Tx\| \geq c \|x\|$ for all $x \in X$.

Theorem 5.4.4 (Approximate eigenvalues and boundedness below)**.** *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a closed linear operator on a complex Banach space X and endow $\operatorname{dom}(L)$ with a graph norm $\|\cdot\|_L$. For every $\lambda \in \mathbb{C}$ the following are equivalent:*

(i) *The number λ is an approximate eigenvalue of L.*

(ii) *The operator $\lambda - L$ from $(\operatorname{dom}(L), \|\cdot\|_L)$ to X is not bounded below.*

(iii) *The number λ is an eigenvalue of L or the operator $\lambda - L : \operatorname{dom}(L) \to X$ does not have closed range in X.*

(iv) *The operator $\lambda - L$ from $(\operatorname{dom}(L), \|\cdot\|_X)$ to X is not bounded below.*

Proof. „(i) ⇔ (ii)“ This equivalence follows from the characterization of approximate eigenvalues in Proposition 5.4.2(iii).

„(ii) ⇒ (iii)“ Assume that (iii) does not hold. Then $\lambda - L : \operatorname{dom}(L) \to X$ is injective and has closed range, say R. This operator is also continuous, $(\operatorname{dom}(L), \|\cdot\|_L)$ is a Banach space due to the closedness of L, and R is a Banach space (with the norm inherited from X) since it is closed in X. Thus the open mapping theorem tells us that the inverse mapping of $\lambda - L$ from R to $(\mathrm{dom}(L), \|\cdot\|_L)$ is continuous; the continuity of the inverse implies that $\lambda - L$ is bounded below, so (ii) does not hold.

„(iii) $\Rightarrow$ (ii)" Assume that (ii) does not hold, i.e., that $\lambda - L$ is bounded below from $(\mathrm{dom}(L), \|\cdot\|_L)$ to X. Then it is also injective, so λ is not an eigenvalue of L. Let us show now that the range R of $\lambda - L$ is complete with respect to the norm inherited from X, and thus closed in X: indeed, let (y_n) be a Cauchy sequence in R, and let $x_n \in \mathrm{dom}(L)$ such that $y_n = (\lambda - L)x_n$ for each n. The boundedness below of $\lambda - L$ implies that (x_n) is a Cauchy sequence in $(\mathrm{dom}(L), \|\cdot\|_L)$, and the latter space is a Banach space as L is assumed to be closed. Hence, (x_n) converges to a vector $x \in \mathrm{dom}(L)$ with respect to $\|\cdot\|_L$. As $\lambda - L$ is continuous from $(\mathrm{dom}(L), \|\cdot\|_L)$ to X it follows that $y_n = (\lambda - L)x_n$ converges to $(\lambda - L)x \in R$ with respect to the norm in X. Thus, R is indeed complete, and hence R is closed in X.

„(i) $\Leftrightarrow$ (iv)" This follows readily from the definition of approximate eigenvalues.

$\square$

Theorem 5.4.5 (The boundary of the spectrum consists of approximate eigenvalues)**.** *Let $L : X \supseteq \mathrm{dom}(L) \to X$ be a closed linear operator on a complex Banach space X. Then $\partial\sigma(L) \subseteq \sigma_{\mathrm{appr}}(L)$.*

Proof. Let $\lambda \in \partial\sigma(L)$. Then there exists a sequence (λ_n) in $\rho(L)$ that converges to λ. We have $\|\mathcal{R}(\lambda_n, L)\| \to \infty$, and thus there exists a sequence (z_n) of normalized vectors in X such that $\alpha_n := \|\mathcal{R}(\lambda_n, L)z_n\| \to \infty$. Let us define

$$x_n := \frac{\mathcal{R}(\lambda_n, L)z_n}{\alpha_n}$$

for every index n. Clearly each x_n is normalized in X and is an element of $\mathrm{dom}(L)$. Moreover, we have

$$(\lambda - L)x_n = (\lambda - \lambda_n)x_n + (\lambda_n - L)x_n = (\lambda - \lambda_n)x_n + \frac{z_n}{\alpha_n} \to 0$$

since all x_n and z_n are normalized and $\alpha_n \to \infty$. $\square$

5.5 Compression spectrum and continuous spectrum

Definition 5.5.1 (Compression spectrum)**.** Let $L : X \supseteq \mathrm{dom}(L) \to X$ be a linear operator on a complex Banach space X. The *compression spectrum* $\sigma_{\mathrm{comp}}(L)$ of L consists of those numbers $\lambda \in \mathbb{C}$ for which the range of $\lambda - L : \mathrm{dom}(L) \to X$ is not dense in X.[13]

[13]This terminology is not generally agreed upon in the literature, and some authors call this set *residual spectrum* rather than compression spectrum.

On the other hand, in those parts of the literature which use the notion compression spectrum (in the sense that we did above) the notion *residual spectrum* typically refers to another subset of the spectrum. So the bottom line is: whenever one reads the notion *residual spectrum* in the literature one should be careful and double-check what the authors actually mean by it.

Clearly, the compression spectrum of a linear operator is a subset of its spectrum. The first assertions in the following proposition is just a rewording of Proposition 5.1.4(a); the second assertion is an immediate consequence of Proposition 5.1.4(b) and of the fact that a subspace of X' which is not weak*-dense cannot be norm dense.

Proposition 5.5.2 (Compression spectrum vs. point spectrum of the dual operator)**.** *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a linear operator on a complex Banach space X and assume that L is densely defined. Then*

$$\sigma_{\text{comp}}(L) = \sigma_{\text{pnt}}(L') \qquad \textit{and} \qquad \sigma_{\text{pnt}}(L) \subseteq \sigma_{\text{comp}}(L').$$

The condition that L be densely defined in the proposition is assumed in order for L' to be single-valued.[14]

The spectrum of any linear operator is covered by the approximate point spectrum and the compression spectrum:

Theorem 5.5.3. *Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a closed linear operator on a complex Banach space X. Then*

$$\sigma(L) = \sigma_{\text{appr}}(L) \cup \sigma_{\text{comp}}(L).$$

Note that the union is not disjoint, in general.[15] Also note that, if L is densely defined, then it follows from the theorem together with Proposition 5.5.2 that $\sigma(L) = \sigma_{\text{appr}}(L) \cup \sigma_{\text{pnt}}(L')$.

Proof of Theorem 5.5.3. „$\supseteq$“ This inclusion is clear.

„$\subseteq$“ Assume that $\lambda \notin \sigma_{\text{appr}}(L)$ and $\lambda \notin \sigma_{\text{comp}}(L)$. The first of those properties implies, according to Theorem 5.4.4(iii), that $\lambda - L : \operatorname{dom}(L) \to X$ is injective and has closed range. The range of this operator is also dense in X since $\lambda \notin \sigma_{\text{comp}}(L)$. Hence, $\lambda - L$ is bijective. Since L was assumed to be closed, it follows that $\lambda \notin \sigma(L)$. ☐

Example 5.5.4 (Fine structure of the spectrum of shift operators)**.** Let $p \in [1, \infty]$ and let $L, R : \ell^p \to \ell^p$ be the *left shift* and the *right shift*, i.e.,

$$\begin{aligned} L &: (x_1, x_2, \dots) \mapsto (x_2, x_3, \dots) \\ \text{and} \qquad R &: (x_1, x_2, \dots) \mapsto (0, x_1, x_2, \dots) \end{aligned}$$

for all $x = (x_1, x_2, \dots) \in \ell^p$. Then one has the following spectral properties:

(a) The left shift L satisfies

$$\sigma(L) = \overline{\mathbb{D}}, \qquad \sigma_{\text{appr}}(L) = \overline{\mathbb{D}},$$
$$\sigma_{\text{pnt}}(L) = \begin{cases} \mathbb{D} & \text{if } p < \infty, \\ \overline{\mathbb{D}} & \text{if } p = \infty, \end{cases} \qquad \sigma_{\text{comp}}(L) = \begin{cases} \emptyset & \text{if } p < \infty, \\ \mathbb{T} & \text{if } p = \infty. \end{cases}$$

[14] Note that we did not discuss the concept *eigenvalue* for multi-valued operators.
[15] Can you give a counterexample?

(b) The right shift R satisfies

$$\sigma(R) = \overline{\mathbb{D}}, \qquad \sigma_{\mathrm{appr}}(R) = \mathbb{T},$$
$$\sigma_{\mathrm{pnt}}(R) = \emptyset, \qquad \sigma_{\mathrm{comp}}(R) = \begin{cases} \mathbb{D} & \text{if } p \in (1,\infty), \\ \overline{\mathbb{D}} & \text{if } p \in \{1,\infty\}. \end{cases}$$

Proof. We first show all claims except for those for the compression spectra of L and R:

(a) The claim for $\sigma_{\mathrm{pnt}}(L)$ can be shown by an explicit computation.[16] Since $\|L\| = 1$ it follows that $\mathrm{r}(L) \le 1$; since the point spectrum contains $\mathbb{D}$ and the spectrum is always closed we conclude that $\sigma(L) = \overline{\mathbb{D}}$.

Moreover we have $\mathbb{T} = \partial\sigma(L) \subseteq \sigma_{\mathrm{appr}}(L)$ according to Theorem 5.4.5, and clearly $\sigma_{\mathrm{pnt}}(L) \subseteq \sigma_{\mathrm{appr}}(L)$, so indeed $\sigma_{\mathrm{appr}}(L) = \overline{\mathbb{D}}$.

(b) One can check by a direct computation that R does not have any eigenvalues, i.e., $\sigma_{\mathrm{pnt}}(R) = \emptyset$.

Next we show that $\sigma(R) = \overline{\mathbb{D}}$. To this end, let $p' \in [1,\infty]$ be the Hölder conjugate of p, i.e., $\frac{1}{p} + \frac{1}{p'} = 1$. If $p < \infty$, then $(\ell^p)'$ can be identified with $\ell^{p'}$ and under this identification we have $R' = L$, where L now denotes the left shift on $\ell^{p'}$. According to Proposition 4.2.8(b) one thus has

$$\sigma(R) = \sigma(R') = \sigma(L) = \overline{\mathbb{D}}.$$

If, on the other hand, $p = \infty$, then ℓ^p can be identified with $(\ell^1)'$, and under this identification one has $L' = R$, where L again denotes the left shift on $\ell^{p'} = \ell^1$. So, again by Proposition 4.2.8(b),

$$\sigma(R) = \sigma(L') = \sigma(L) = \overline{\mathbb{D}}.$$

Next we turn to the approximate point spectrum. First note that $\sigma_{\mathrm{appr}}(R) \supseteq \partial\sigma(R) = \mathbb{T}$ according to Theorem 5.4.5. Moreover, R is isometric, i.e., $\|Rx\| = \|x\|$ for all $x \in \ell^p$; for all $\lambda \in \mathbb{D}$ and all $x \in \ell^p$ this implies

$$\|(\lambda - R)x\| \ge \|Rx\| - \|\lambda x\| = (1 - |\lambda|)\,\|x\|\,;$$

as $|\lambda| < 1$ this shows that $\lambda - R$ is bounded below, so λ is not an approximate eienvalue of R. Hence, $\sigma_{\mathrm{appr}}(R) = \mathbb{T}$.

Finally, let us turn to compression spectra; we start with $\sigma_{\mathrm{comp}}(L)$. If $p < \infty$ the right shift R on $\ell^{p'}$ satisfies $R = L'$; so it follows from Proposition 5.5.2 that

$$\sigma_{\mathrm{comp}}(L) = \sigma_{\mathrm{pnt}}(L') = \sigma_{\mathrm{pnt}}(R) = \emptyset.$$

The case $p = \infty$ will be treated in Exercise Sheet 11.

[16] Which is often done as an exercise in introductory courses to functional analysis.

Lastly, we treat $\sigma_{\mathrm{comp}}(R)$. If $p < \infty$ the left shift L on $\ell^{p'}$ satisfies $L = R'$, so Proposition 5.5.2 shows that

$$\sigma_{\mathrm{comp}}(R) = \sigma_{\mathrm{pnt}}(R') = \sigma_{\mathrm{pnt}}(R),$$

which is $\mathbb{D}$ if $p' < \infty$ (i.e., $p > 1$) and which is $\overline{\mathbb{D}}$ if $p' = \infty$ (i.e., $p = 1$). In the case $p = \infty$ the left shift L on $\ell^{p'} = \ell^1$ satisfies $L' = R$, so the second assertion in Proposition 5.5.2 shows that

$$\sigma_{\mathrm{comp}}(R) = \sigma_{\mathrm{comp}}(L') \supseteq \sigma_{\mathrm{pnt}}(L) = \mathbb{D}.$$

So it only remains to show that $\sigma_{\mathrm{comp}}(R)$ contains the unit circle $\mathbb{T}$; we also do this on Exercise Sheet 11. $\square$

Note that the right shift R on ℓ^p is an *isometry*, i.e., $\|Rx\| = \|x\|$ for each $x \in \ell^p$; however, R is not surjective. Some (though not all) phenomena about the spectrum of R that we observed in Example 5.5.4 are also true for general non-surjective isometries on Banach spaces; this will be discussed in more detail in an exercise.

Remark 5.5.5 (Continuous spectrum)**.** Some authors also defined the so-called *continuous spectrum* of an operator: Let $L : X \supseteq \mathrm{dom}(L) \to X$ be a linear operator on a complex Banach space X. Then its *continuous spectrum* $\sigma_{\mathrm{cont}}(L)$ is defined to be the set of all $\lambda \in \mathbb{C}$ for which the operator $\lambda - L : \mathrm{dom}(L) \to X$ is injective and has dense range, but is not surjective.

It follows from Theorem 5.5.3 that

$$\sigma_{\mathrm{cont}}(L) = \sigma_{\mathrm{appr}}(L) \setminus \Big(\sigma_{\mathrm{pnt}}(L) \cup \sigma_{\mathrm{comp}}(L)\Big),$$

and if L is densely defined, then this means by virtue of Proposition 5.5.2 that

$$\sigma_{\mathrm{cont}}(L) = \sigma_{\mathrm{appr}}(L) \setminus \Big(\sigma_{\mathrm{pnt}}(L) \cup \sigma_{\mathrm{pnt}}(L')\Big).$$

The notion *continuous spectrum* is particularly prevalent in the literature on mathematical physics.

Remark 5.5.6 (The various spectra in finite dimensions)**.** Let $X = \mathbb{C}^d$ and $T \in \mathcal{L}(X)$. Then, clearly,

$$\sigma(T) = \sigma_{\mathrm{pnt}}(T) = \sigma_{\mathrm{appr}}(T) = \sigma_{\mathrm{comp}}(T)$$

and $\sigma_{\mathrm{cont}}(T) = \emptyset$.

Chapter 6

Fredholm Theory and the Essential Spectrum

6.1 Fredholm index and Fredholm operators

Recall that, for a vector subspace V of a vector space X the *codimension* $\operatorname{codim} V$ is defined as $\operatorname{codim} V := \dim(X/V) \in \mathbb{N}_0 \cup \{\infty\}$.

Definition 6.1.1 (Nullity and defect)**.** Let X, Y be vector spaces over the same field and let $T : X \to Y$ be linear.

(a) The number $\operatorname{nul}(T) := \dim \ker T \in \mathbb{N}_0 \cup \{\infty\}$ is called the *nullity* of T.

(b) The number $\operatorname{def} T := \operatorname{codim}(TX) = \dim(Y/TX) \in \mathbb{N}_0 \cup \{\infty\}$ is called the *defect* of T.

Theorem 6.1.2 (Nullity and defect of compositions)**.** *Let $X \xrightarrow{S} Y \xrightarrow{T} Z$ be linear maps between vector spaces.*[1] *Then*

$$\operatorname{nul}(TS) + \operatorname{def}(S) + \operatorname{def}(T) = \operatorname{nul}(S) + \operatorname{nul}(T) + \operatorname{def}(TS).$$

For the proof we need the following observation from linear algebra whose proof we leave as a little exercise: let U, V be vector subspaces of a vector space Y, set $Y_2 := U \cap V$ and let Y_1, Y_3 be vector subspaces of Y such that $U = Y_1 \oplus Y_2$ and $V = Y_2 \oplus Y_3$. Then the sum $Y_1 + Y_2 + Y_3$ in Y is direct.

Proof of Theorem 6.1.2. We decompose X, Y, and Z as direct sums, and the key to the proof is to do this in the right order; we start with Y: define $Y_2 := SX \cap \ker T$. Choose vector subspaces Y_1, Y_3 of Y such that[2] $SX = Y_1 \oplus Y_2$ and $\ker T = Y_2 \oplus Y_3$.

[1] With the same underlying field for all spaces.

[2] It follows from the basis extension theorem that such subspaces exist; we will tacitly use the basis extension theorem several times in the rest of the proof.

As observed before the proof, the sum $Y_1 + Y_2 + Y_3$ is then direct, and we extend this sum by a vector subspace Y_4 of Y to give the entire space Y; hence,

$$Y = \overbrace{Y_1 \oplus \underbrace{Y_2}^{=SX} \oplus Y_3}_{=\ker T} \oplus Y_4.$$

Next we decompose X: define $X_0 := \ker S$, and choose a vector subspace $\tilde{X}$ of X such that $X = X_0 \oplus \tilde{X}$. Then S maps bijectively from $\tilde{X}$ to $SX = Y_1 \oplus Y_2$, so we can decompose $\tilde{X}$ as a direct sum of two vector subspaces X_1 and X_2 such that S maps bijectively from X_1 to Y_1 and bijectively from X_2 to Y_2; in particular, $\dim X_2 = \dim Y_2$. Note that the entire space X is given as

$$X = \underbrace{X_0}_{=\ker S} \oplus X_1 \oplus X_2$$

Finally we decompose Z: let us define $Z_1 := T(Y_1)$ and $Z_4 := T(Y_4)$; we note that $\dim Z_4 = \dim Y_4$ since T maps Y_4 bijectively to Z_4. Since T is injective on $Y_1 \oplus Y_4$ the sum $Z_1 + Z_4$ is direct. To get the entire space Z we extend this direct sum by a further vector subspace Z_5 such that[3]

$$Z = \overbrace{Z_1}^{=TY_1} \oplus \overbrace{Z_4}^{=TY_4} \oplus Z_5.$$

Now we can compute the six numbers in the claimed formula; we start with the three numbers on the left hand side. Clearly, $\operatorname{def}(S) = \dim Y_3 + \dim Y_4$ and $\operatorname{def}(T) = \dim Z_5$. Moreover,

$$\begin{aligned}\ker(TS) &= S^{-1}(\ker T) = S^{-1}(Y_2 \oplus Y_3) = S^{-1}\Big(SX \cap (Y_2 \oplus Y_3)\Big)\\ &= S^{-1}\Big((Y_1 \oplus Y_2) \cap (Y_2 \oplus Y_3)\Big) = S^{-1}(Y_2) = X_0 \oplus X_2.\end{aligned}$$

Therefore, $\operatorname{nul}(TS) = \dim X_0 + \dim X_2$.

Now we compute the three numbers on the right hand side of the claimed equation. Clearly, $\operatorname{nul}(S) = \dim X_0$ and $\operatorname{nul}(T) = \dim Y_2 + \dim Y_3$. Moreover,

$$(TS)X = T(SX) = T(Y_1 \oplus Y_2) = TY_1 = Z_1,$$

so $\operatorname{def}(TS) = \dim Z_4 + \dim Z_5$. So overall we have

$$\begin{aligned}&\operatorname{nul}(TS) + \operatorname{def}(S) + \operatorname{def}(T)\\ =&(\dim X_0 + \underbrace{\dim X_2}_{=\dim Y_2}) + (\dim Y_3 + \underbrace{\dim Y_4}_{=\dim Z_4}) + \dim Z_5\\ =& \dim X_0 + (\dim Y_2 + \dim Y_3) + (\dim Z_4 + \dim Z_5)\\ =& \operatorname{nul}(S) + \operatorname{nul}(T) + \operatorname{def}(ST).\end{aligned}$$ □

[3] To avoid any potential confusion, let us point out explicitly that there are no spaces Z_2 and Z_3 in the proof.

Remark 6.1.3 (Rank theorem in finite dimensions)**.** Let X, Y be finite-dimensional vector spaces over the same field, and let $T \in \mathcal{L}(X;Y)$. Then the so-called *rank theorem* says that

$$\dim X = \dim \ker T + \dim(TX).$$

This can actually be obtained as a consequence of Theorem 6.1.2:

Consider the mappings $X \xrightarrow{T} Y \xrightarrow{q} Y/T(X)$, where q denotes the quotient mapping. Then q is surjective and $qT = 0$, and the theorem gives

$$\underbrace{\mathrm{nul}(qT)}_{=\dim X} + \underbrace{\mathrm{def}(q)}_{=0} + \mathrm{def}(T) = \underbrace{\mathrm{nul}(q)}_{=\dim T(X)} + \mathrm{nul}(T) + \underbrace{\mathrm{def}(qT)}_{=\mathrm{def}(T)};$$

so $\dim X = \dim T(X) + \mathrm{nul}(T)$, as claimed.

Alternatively, one can apply Theorem 6.1.2 to the composition $\ker T \xrightarrow{j} X \xrightarrow{T} Y$, where j denotes the canonical embedding. Then j is injective and $Tj = 0$, and the theorem gives

$$\underbrace{\mathrm{nul}(Tj)}_{=\mathrm{nul}(T)} + \underbrace{\mathrm{def}(T)}_{\dim Y - \dim T(X)} + \underbrace{\mathrm{def}(j)}_{=\dim X - \mathrm{nul}(T)} = \mathrm{nul}(T) + \underbrace{\mathrm{nul}(j)}_{=0} + \underbrace{\mathrm{def}(Tj)}_{=\dim Y},$$

which yields $\dim X = \mathrm{nul}(T) + \dim T(X)$, as claimed.

Definition 6.1.4 ((Semi-)Fredholm operators and Fredholm index)**.** Let X, Y be complex Banach spaces. An operator $T \in \mathcal{L}(X;Y)$ is called...

(a) ... *upper semi-Fredholm* if $\mathrm{nul}(T) < \infty$ and the range TX is closed.

(b) ... *lower semi-Fredholm* if $\mathrm{def}(T) < \infty$ and the range TX is closed.

(c) ... *semi-Fredholm* if it is upper or lower semi-Fredholm. In this case, the *index* of T is defined as

$$\mathrm{ind}(T) := \mathrm{nul}(T) - \mathrm{def}(T) \in \mathbb{Z} \cup \{-\infty, \infty\}.$$

(d) ... *Fredholm* if it is both upper and lower semi-Fredholm.

Note that a semi-Fredholm operator T is upper semi-Fredholm if and only if $\mathrm{ind}(T) < \infty$, lower semi-Fredholm if and only if $\mathrm{ind}(T) > -\infty$, and Fredholm iff $\mathrm{ind}(T) \in \mathbb{Z}$.

Remarks 6.1.5 (Automatic closedness of the range)**.** (a) In the definition of lower semi-Fredholm operators (Definition 6.1.4(b)) the condition that the range TX be closed is actually redundant.[4] This can be seen as follows: if $\mathrm{def}(T) < \infty$,

[4] It is not redundant in the definition of upper semi-Fredholm operators, though – i.e., the range of an operator is not automatically closed if the operator has finite-dimensional kernel. It is part of Exercise Sheet 12 to find a counterexample.

there exists a finite dimensional vector subspace Y_1 of Y such that $Y = TX \oplus Y_1$. As Y_1 is closed necessarily closed due to its finite dimension, the closedness of TX follows from a general result about operator ranges in Proposition B.1.3 in the appendix.

(b) In particular, an operator $T \in \mathcal{L}(X;Y)$ is Fredholm if and only if $\operatorname{nul}(T) < \infty$ and $\operatorname{def}(T) < \infty$.

Theorem 6.1.6 (Characterization of closed range by boundedness below)**.** *Let X, Y be Banach spaces over the same field and let $T \in \mathcal{L}(X;Y)$.*

(a) *The operator T is injective and has closed range if and only if T is bounded below.*

(b) *Then T has closed range if and only if there exists a number $c > 0$ such that*

$$\|Tx\| \geq c \operatorname{dist}(x, \ker T)$$

for all $x \in X$.

Proof. (a) These are the same arguments as in the proof of Theorem 5.4.4.

(b) This follows by applying (a) to the operator $\tilde{T} : X/\ker T \to X$ that is induced by T, since T and $\tilde{T}$ have the same range and since, for each $x \in X$, the number $\operatorname{dist}(x, \ker T)$ is the norm of $x + \ker T$ in $X/\ker T$. □

Examples 6.1.7 (Some Fredholm operators)**.** Let X, Y be complex Banach spaces and $T \in \mathcal{L}(X;Y)$.

(a) If T is bijective, then T is obviously Fredholm and $\operatorname{ind}(T) = 0$.

(b) If X and Y are finite-dimensional, then T is obviously Fredholm and $\operatorname{ind}(T) = \dim X - \dim Y$. Indeed, the rank theorem (see Remark 6.1.3) gives

$$\begin{aligned} \dim X &= \operatorname{nul}(T) + \dim T(X) \\ &= \operatorname{nul}(T) - \underbrace{(\dim Y - \dim T(X))}_{=\operatorname{def}(T)} + \dim Y = \operatorname{ind}(T) + \dim(Y). \end{aligned}$$

(c) So in particular, if X is finite dimensional and $T \in \mathcal{L}(X)$, then T is Fredholm with index 0.

Theorem 6.1.8 (Characterization of upper semi-Fredholm operators)**.** *Let X, Y be complex Banach spaces and let $T \in \mathcal{L}(X;Y)$. The following are equivalent:*

(i) *The operator T is upper semi-Fredholm.*

(ii) *One has $\operatorname{nul}(T) < \infty$ and there exists a closed vector subspace X_1 of X such that $X = \ker T \oplus X_1$ and such that $T|_{X_1}$ is bounded below.*

(iii) *For every bounded sequence* (x_n) *in* X *the following holds: If* (Tx_n) *converges, than* (x_n) *has a convergent subsequence.*

Proof. „(i) $\Rightarrow$ (ii)" The splitting property of X follows from the fact that finite-dimensional subspaces on Banach spaces are always complemented by closed subspaces, and the boundedness below is a consequence of Theorem 6.1.6.

„(ii) $\Rightarrow$ (iii)" Let (x_n) be a sequence as in (iii). As the projection from X onto $\ker T$ along X_1 and its complementary projection are continuous,[5] we can split each x_n as $x_n = u_n + z_n$ for bounded sequences (u_n) in $\ker T$ and (z_n) in X_1. As $\ker T$ is finite dimensional, (u_n) has a convergent subsequence (u_{n_k}). Moreover, as the sequence $(Tx_n) = (T|_{X_1} z_n)$ converges in Y and as $T|_{X_1}$ is bounded below, it follows that (z_n) is a Cauchy sequence in X_1 and thus convergent in X_1 (as X_1 is closed in X and thus complete). Hence, the subsequence

$$\big(x_{n_k}\big) = \big(u_{n_k} + z_{n_k}\big)$$

of (x_n) converges in X, as claimed.

„(iii) $\Rightarrow$ (i)" Due to (iii) every bounded sequence in $\ker T$ has a convergent subsequence, so $\ker T$ is finite dimensional.

In particular, we can split X as $X = \ker T \oplus X_1$ for a closed vector subspace X_1 of X.

To show that TX is closed, we first note that TX coincides with the range of the injective operator $T_1 := T|_{X_1}$. According to Theorem 6.1.6(a) it thus suffices to show that T_1 is bounded below. Assume the contrary. Then there exists a normalized sequence (x_n) in X_1 such that $Tx_n \to 0$. Due to (iii) we can replace (x_n) with a convergent subsequence, say with limit x. Then x also has norm 1 and it is an element of X_1 since X_1 is closed. Due to the continuity of T we have $T_1 x = Tx = 0$; this contradicts the injectivity of T_1. $\square$

In order to make the characterization of upper semi-Fredholm operators in Theorem 6.1.8 also accessible for lower semi-Fredholm operators, we use the following duality result which we quote here without proof. We use the following notation to formulate the theorem: Let X be a Banach space, let $M \subseteq X$ and $N \subseteq X'$. Then the vector subspaces

$$\begin{aligned} M^\perp &:= \{x' \in X' \mid \langle x', x\rangle = 0 \text{ for all } x \in M\} \subseteq X', \\ {}^\perp N &:= \{x \in X \mid \langle x', x\rangle = 0 \text{ for all } x' \in N\} \end{aligned}$$

are called the *annihilators* of M and N in X' and X, respectively. One can show that, if M is a closed vector subsapce of X and N is a weak* closed vector subspace of X', then

$$\dim M^\perp = \dim X/M \qquad \text{and} \qquad \dim {}^\perp N = \dim X'/N.$$

[5] Why?

Theorem 6.1.9 (Closed range theorem)**.** *Let X, Y be Banach spaces over the same field and let $T \in \mathcal{L}(X;Y)$. The following are equivalent:*

(i) *The range TX is closed in Y.*

(ii) *The range $T'Y'$ is closed in X'.*

(iii) *The range $T'Y'$ is weak*-closed in X'.*

(iv) *One has $TX = {}^{\perp}\ker(T')$.*

(v) *One has $T'Y' = \ker(T)^{\perp}$.*

Proof. See for instance [Kab14, Theorem 9.6] for details.[6] □

The theorem has the following useful consequence for semi-Fredholm operators.

Corollary 6.1.10 (Duality of semi-Fredholm operators)**.** *Let X, Y be Banach spaces over the same field and let $T \in \mathcal{L}(X;Y)$. The operator T is semi-Fredholm if and only if $T' \in \mathcal{L}(Y';X')$ is semi-Fredholm; in this case one has*

$$\operatorname{nul}(T) = \operatorname{def}(T'), \qquad \operatorname{nul}(T') = \operatorname{def}(T), \qquad \operatorname{ind}(T') = -\operatorname{ind}(T).$$

Proof. This is a consequence of the Closed range theorem 6.1.9 and of the comments preceding that theorem. □

Corollary 6.1.11 (Composition of (semi-)Fredholm operators)**.** *Let $X \overset{S}{\to} Y \overset{T}{\to} Z$ be bounded linear operators between complex Banach spaces.*

(a) *If S and T are upper semi-Fredholm, then so is TS and* $\operatorname{ind}(TS) = \operatorname{ind}(T) + \operatorname{ind}(S)$.

(b) *If S and T are lower semi-Fredholm, then so is TS and* $\operatorname{ind}(TS) = \operatorname{ind}(T) + \operatorname{ind}(S)$.

(c) *If S and T are Fredholm, then so is TS and* $\operatorname{ind}(TS) = \operatorname{ind}(T) + \operatorname{ind}(S)$.

Proof. (a) It follows from Theorem 6.1.8 that ST is upper semi-Fredholm. The index formula is then an immediate consequence of Theorem 6.1.2.

(b) This follows from (a) and Corollary 6.1.10 by considering the dual operators S' and T'.

(c) This is an immediate consequence of (a) and (b). □

Note that, in general, the composition of two operators with closed range does not need to have closed range; see e.g. [Bou73, p. 362] for a counterexample.

[6] The result in this reference is actually stated in a somewhat more general setting.

Theorem 6.1.12 (The semi-Fredholm operators with fixed index are open)**.** *Let X, Y be complex Banach spaces. For every $j \in \mathbb{Z} \cup \{-\infty, \infty\}$ the set of semi-Fredholm operators with index j is open in $\mathcal{L}(X;Y)$.*

Note that the theorem implies, in particular, that each of the following sets is open in $\mathcal{L}(X;Y)$: (a) the set of semi-Fredholm operators; (b) the set of upper semi-Fredholm operators; (c) the set of lower semi-Fredholm operators; (d) the set of Fredholm operators.

Before we prove the theorem we discuss several consequences of it.

Corollary 6.1.13 (The Fredholm index is continuous)**.** *Let X, Y be complex Banach spaces and endow $\mathbb{Z} \cup \{-\infty, \infty\}$ with the discrete topology.*[7] *Then* ind *is continuous from the set of all Fredholm operators in $\mathcal{L}(X;Y)$ to $\mathbb{Z} \cup \{-\infty, \infty\}$.*

Proof. Theorem 6.1.12 implies that the pre-image of any subset of $\mathbb{Z} \cup \{-\infty, \infty\}$ under ind is open in the set of semi-Fredholm operators. $\square$

Corollary 6.1.14 (Compact perturbations of Fredholm operators)**.** *Let X, Y be complex Banach spaces and let $T \in \mathcal{L}(X;Y)$ be semi-Fredholm. If $K \in \mathcal{L}(X;Y)$ is compact, then $T + K$ is also semi-Fredholm and*

$$\operatorname{ind}(T + K) = \operatorname{ind}(T).$$

Proof. It suffices to consider the case where T is upper semi-Fredholm; the other case then follows by dualization.

By means of Theorem 6.1.8 one can readily check that $T + K$ is also upper semi-Fredholm. Now consider the continuous path $\gamma : [0, 1] \to \mathcal{L}(X;Y)$ that is given by $\gamma(t) = T + tK$ for each $t \in [0, 1]$. According to what we just observed, the path maps into the set of upper semi-Fredholm operators. The composition $\operatorname{ind} \circ \gamma : [0, 1] \to \mathbb{Z} \cup \{-\infty\}$ is continuous (where the co-domain is endowed with the discrete topology) and hence constant. So $T = \gamma(0)$ and $T + K = \gamma(1)$ have the same index, as claimed. $\square$

Now we are going to proof Theorem 6.1.12. We outsource the essence of the argument into the following two lemmas.

Lemma 6.1.15 (Infinite defect)**.** *Let X, Y be complex Banach spaces and let $S, T \in \mathcal{L}(X;Y)$ be bounded below.*[8] *Assume that T has infinite defect. If*

$$\|S - T\| < \frac{1}{2} \underbrace{\inf \Big\{ \|Sx\| \mid x \in X \text{ and } \|x\| = 1 \Big\}}_{=:\gamma},$$

then S has infinite defect, too.

[7] Note that, with respect to the discrete topology, the sequnce $(n)_{n \in \mathbb{N}}$ does not converge to ∞.

[8] In fact, it suffices for the proof if S is bounded below and T has closed range.

Proof. As the range of T is closed and has infinite co-dimension we can, by virtue of Riesz' lemma, find a sequence of normalized vectors $(y_n)_{n\in\mathbb{N}}$ in Y such that, for each $n \in \mathbb{N}_0$, the distance of y_{n+1} to the span of $TX \cup \{y_1, \dots, y_n\}$ is at least $1 - \frac{1}{n}$.

Assume for a contradiction that S has finite defect. As the quotient space Y/SX is thus finite dimensional we may, after replacing (y_n) with a subsequence, assume that the sequence $(y_n + SX)$ in Y/SX converges. So in particular, $\operatorname{dist}(y_{n+1} - y_n, SX) \to 0$. Hence, there exists a sequence (x_n) in X such that

$$\|y_{n+1} - y_n - Sx_n\| \to 0.$$

So in particular, $\gamma \limsup_n \|x_n\| \le \limsup_n \|Sx_n\| \le 2$. Thus,

$$\begin{aligned} 1 = \lim_n \big(1 - \frac{1}{n}\big) &\le \limsup_n \|y_{n+1} - y_n - Tx_n\| \\ &\le \limsup_n \|y_{n+1} - y_n - Sx_n\| + \limsup_n \|(S-T)x_n\| \\ &\le \limsup_n \|x_n\| \, \|S - T\| \le \frac{2}{\gamma} \|S - T\| < 1; \end{aligned}$$

This is a contradiction. $\square$

Lemma 6.1.16 (Topological properties of the set of upper semi-Fredholm operators)**.** *Let X, Y be complex Banach spaces. For every upper semi-Fredholm operator $T \in \mathcal{L}(X;Y)$ there exists $\varepsilon > 0$ with the following property:*

Whenever $S \in \mathcal{L}(X;Y)$ satisfies $\|S - T\| < \varepsilon$, then SX is closed, $\operatorname{nul}(S) \le \operatorname{nul}(T)$,[9] *and* $\operatorname{ind}(S) = \operatorname{ind}(T)$.

Proof. Since $\ker T$ is, by assumption, finite-dimensional, there there exists a closed vector subspace $\tilde{X} \subseteq X$ such that $X = \tilde{X} \oplus \ker T$. The restriction $\tilde{T} := T|_{\tilde{X}} : \tilde{X} \to Y$ is injective and has the same range as T. Thus, its range is closed and we conclude that $\tilde{T}$ is bounded below. Set[10]

$$3\varepsilon := \inf \Big\{ \big\|\tilde{T}\tilde{x}\big\| \ \big| \ \tilde{x} \in \tilde{X} \text{ and } \|\tilde{x}\| = 1 \Big\} > 0.$$

Now let $S \in \mathcal{L}(X;Y)$ and $\|S - T\| < \varepsilon$. For every $\tilde{x} \in \tilde{X}$ of norm 1 one then has

$$\|S\tilde{x}\| \ge \|T\tilde{x}\| - \|(T-S)\tilde{x}\| \ge 3\varepsilon - \|S - T\| \ge 2\varepsilon. \tag{6.1.1}$$

This shows that the restriction $\tilde{S} := S|_{\tilde{X}}$ is bounded below.

In particular $\ker S \cap \tilde{X} = \{0\}$, so the mapping $\ker S \to X/\tilde{X}$, $x \mapsto x + \tilde{X}$ is injective. Therefore,

$$\operatorname{nul}(S) = \dim(\ker S) \le \dim(X/\tilde{X}) = \dim \ker T = \operatorname{nul}(T).$$

[9] So in particular, S is upper semi-Fredholm.

[10] Note however that, in the last step of the proof, we will make ε still smaller.

Moreover, the range $S(\tilde{X}) = \tilde{S}(\tilde{X})$ is closed since $\tilde{S}$ is bounded below. Hence, $S(X) = S(\tilde{X}) + S(\ker T)$ is closed as the sum of a closed and a finite-dimensional subspace.

It remains to show the claim for the index.

First assume that $\operatorname{ind}(T) = -\infty$, i.e., that T has infinite defect. Then $\tilde{T} : \tilde{X} \to Y$ has infinite defect, too (as it has the same range). Moreover, one has

$$\left\| \tilde{S} - \tilde{T} \right\| < \varepsilon \leq \frac{1}{2} \inf \left\{ \left\| \tilde{S}\tilde{x} \right\| \mid \tilde{x} \in \tilde{X} \text{ and } \|\tilde{x}\| = 1 \right\},$$

where the second inequality is (6.1.1). So Lemma 6.1.15 shows that $\tilde{S}$ has infinite defect, too. Hence, $S(X) = \tilde{S}(\tilde{X}) + S(\ker T)$ also has infinite co-dimension, as it is the sum of a space of infinite co-dimension and a finite-dimensional space. This proves that S has infinite defect, too, so $\operatorname{ind}(S) = -\infty$.

Finally assume that $\operatorname{ind}(T) \in \mathbb{Z}$, i.e., that T is Fredholm. Then Y splits as

$$Y = TX \oplus F = \tilde{T}\tilde{X} \oplus F$$

for a finite-dimensional subspace F of Y. Let $q : Y \to Y/F$ denote the quotient mapping. Then $q\tilde{T} : \tilde{X} \to Y/F$ is bijective. The set of all bijective linear operators from $\tilde{X}$ to Y/F is open with respect to the operator norm topology, so by making ε smaller than before if necessary and still assuming that $\|S - T\| < \varepsilon$, we have that $q\tilde{S} : \tilde{X} \to Y/F$ is bijective. Let $j : \tilde{X} \to X$ denote the canonical embedding; then $\tilde{S} = Sj$. Since j, q, and S are upper semi-Fredholm, Corollary 6.1.11(a) yields

$$\begin{aligned} 0 &= \operatorname{ind}(q\tilde{S}) = \operatorname{ind}(qSj) = \operatorname{ind}(q) + \operatorname{ind}(S) + \operatorname{ind}(j) \\ &= \dim F + \operatorname{ind}(S) - \dim \ker T = \operatorname{ind}(S) + \operatorname{def}(T) - \operatorname{nul}(T) = \operatorname{ind}(S) - \operatorname{ind}(T), \end{aligned}$$

which gives the claim. □

Proof of Theorem 6.1.12. For $j \in \mathbb{Z} \cup \{-\infty\}$ the claim follows immediately from Lemma 6.1.16.

For $j = \infty$ we argue by duality: let $T \in \mathcal{L}(X; Y)$ be semi-Fredholm with index ∞. According to Corollary 6.1.10 the dual operator $T' \in \mathcal{L}(Y'; X')$ is semi-Fredholm with index $-\infty$. Hence, by what we have already proved there exists $\varepsilon > 0$ such that all operators in $\mathcal{L}(Y'; X')$ that are closer than ε to T', are semi-Fredholm with index $-\infty$. So if $S \in \mathcal{L}(X; Y)$ satisfies $\|S - T\| < \varepsilon$, then $\|S' - T'\| < \varepsilon$, so S' is semi-Fredholm with index $-\infty$ and hence, again by Corollary 6.1.10, S is semi-Fredholm with index ∞. □

Example 6.1.17 (Nullity and defect need not be continuous)**.** All matrices $A \in \mathbb{C}^{d \times d}$ are Fredholm with index 0 (Example 6.1.7(c)).

Note how this is constistent with Corollary 6.1.13, which tells us that the Fredholm index is continuous (and hence, constant due to the connectedness of $\mathbb{C}^{d \times d}$).

However, nullity and defect need not continuous. For instance, consider the sequence $(\frac{1}{n}\,\mathrm{id})$ which converges to the matrix 0. One has

$$\operatorname{ind}\big(\frac{1}{n}\,\mathrm{id}\big) = \operatorname{nul}\big(\frac{1}{n}\,\mathrm{id}\big) - \operatorname{def}\big(\frac{1}{n}\,\mathrm{id}\big) = 0 - 0 = 0$$

for each n, and

$$\operatorname{ind}(0) = \operatorname{nul}(0) - \operatorname{def}(0) = d - d = 0.$$

This demonstrates nicely how nullity and defect might jump, while there difference remains constant.

In Section 6.2 we will study how Fredholm operators are related to spectral theory. An important part of this connection is due to what is sometimes called *analytic Fredholm theory*; an important part of this theory is summed up in the following theorem. In order to have some time left for Part III of the course, we only include a reference for the theorem instead of giving its proof.

Theorem 6.1.18 (Analytic Fredholm-valued functions)**.** *Let X, Y be complex Banach spaces, let $\Omega \subseteq \mathbb{C}$ be non-empty, open, and connected, and let $F : \Omega \to \mathcal{L}(X;Y)$ be a holomorphic mapping on Ω. Assume that $F(z)$ is a semi-Fredholm operator for every $z \in \Omega$ and that there exists $z_0 \in \Omega$ such that $F(z) : X \to Y$ is bijective. Then the set*

$$D := \{z_0 \in \Omega \mid \; F(z_0) \textit{ is not bijective } \}$$

is at most countable and has no accumulation point inside Ω. Moreover, every $z_0 \in D$ is a pole of F, and the coefficient of $(z - z_0)^k$ in the Laurent series expansion of F about z_0 has finite rank for every $k < 0$.

Proof. We first note that, as $F(z_0)$ is Fredholm with index 0 and Ω is connected, it follows from Corollary 6.1.13 that $F(z)$ is even Fredholm for every $z \in \Omega$. The result can thus be found, for instance, in [GGK90, Corollary 8.4 on p. 203].[11] □

Let us briefly discuss how the preceding theorem together with the preceding results from Fredholm theory can be applied to show the following result about the spectrum of compact operators; you might already know the following from an introductory course to functional analysis. The advantage of the lengthy and theory-intensive part that we have taken so far is that it will allow us to deduce various similar results very easily in the subsequent Section 6.2.

Example 6.1.19 (The spectrum of compact operators)**.** Let X be a complex Banach space and let $K \in \mathcal{L}(X)$ be compact. Then the following assertions hold:

[11] The cited result also contains the following additional property: the coefficient of $(z - z_0)^0$ in the Laurent series is a Fredholm operator of index 0.

(a) The spectrum of K is at most countable, and it has no accumulation points except possibly 0.

(b) Every non-zero spectral value λ_0 of K is an eigenvalue of K of finite algebraic multiplicity[12] and a pole of the resolvent $\mathcal{R}(\,\cdot\,, K)$.

Proof. Consider the connected open set $\Omega := \mathbb{C} \setminus \{0\}$ and the holomorphic mapping $F : \Omega \to \mathcal{L}(X)$, $\lambda \mapsto z \operatorname{id}_X - K$. For every $\lambda \in \Omega$ the operator $\lambda \operatorname{id}_X$ is bijective from X to X, and thus Fredholm with index 0. As K is compact, it thus follows from Corollary 6.1.14 that $F(\lambda)$ is Fredholm (with index 0) for every $\lambda \in \Omega$. Moreover, for every complex number λ of modulus $|\lambda| > \mathrm{r}(K)$,[13] the operator $F(\lambda)$ is bijective. Thus, the assumptions of Theorem 6.1.18 are satisfied. Let us show how this implies both assertions:

(a) The set D from the theorem is, in the situation of the present example, equal to $\sigma(K) \setminus \{0\}$, so the theorem shows that the spectral values of K can only accumulate at 0. This implies, in particular, that $\sigma(K)$ is at most countable.

(b) Also by Theorem 6.1.18 the number λ_0 is a pole of the resolvent $\mathcal{R}(\,\cdot\,, K)$, and all the coefficients in the Laurent series expansion of the resolvent about λ_0 that belong to negative exponents, are finite rank operators. So in particular, in the notation of Theorem 5.3.4, the operator Q_0 has finite rank.

As λ_0 is a pole of the resolvent, it is also an eigenvalue of K according to Theorem 5.3.4(a). Moreover, Theorem 5.3.4(b) says that the space $Q_0 X$ – which we already know to be finite dimensional – is the generalized eigenspace of the eigenvalue λ_0. So the algebraic multiplicity of λ_0 is indeed finite. □

We close this section with two theorems that show in which sense Fredholm operators can be understood as a generalization of bijective operators.

Theorem 6.1.20 (Fredholm operators and invertibility)**.** *Let X, Y be complex Banach spaces and let $T \in \mathcal{L}(X;Y)$. The following are equivalent:*

(i) *The operator T is Fredholm.*

(ii) *There exists an operator $S \in \mathcal{L}(Y;X)$ and finite-rank operators $F_X \in \mathcal{L}(X)$ and $F_Y \in \mathcal{L}(Y)$ such that*

$$ST = \operatorname{id}_X + F_X \qquad \text{and} \qquad TS = \operatorname{id}_X + F_Y$$

(iii) *There exists an operator $S \in \mathcal{L}(Y;X)$ and compact operators $K_X \in \mathcal{L}(X)$ and $K_Y \in \mathcal{L}(Y)$ such that*

$$ST = \operatorname{id}_X + K_X \qquad \text{and} \qquad TS = \operatorname{id}_X + K_Y$$

[12] In the lecture, instead of the finite algebraic multiplicity of the eigenvalue λ_0 an assertion about the coefficients in the Laurent series expansion of the resolvent was made; however, this is equivalent to the assertion about finite algebraic multiplicity here, as follows from Theorem 5.3.4(b) and Theorem 5.3.1(d).

[13] Note that K is bounded, and hence has compact spectrum.

Proof. „(i) $\Rightarrow$ (ii)“ This is not difficult to show if one splits X as $X = \ker T \oplus \tilde{X}$ for a closed vector subspace $\tilde{X}$, and Y as $Y = TX \oplus F$ for a finite-dimensional vector subspace F.

„(ii) $\Rightarrow$ (iii)“ This implication is obvious.

„(iii) $\Rightarrow$ (i)“ Let us first use Theorem 6.1.8(iii) to show that T is upper semi-Fredholm: let (x_n) be a bounded sequence in X and assume that (Tx_n) converges in Y; then (STx_n) converges in X. As K_X is compact, $(K_X x_n)$ has a convergent subsequences, and thus it follows from

$$x_n = STx_n - K_X x_n$$

that (x_n) has a convergent subsequence.

By dualizing the equality $TS = \mathrm{id}_X + K_Y$,[14] and applying the same argument again, we see that also T' is upper semi-Fredholm, and hence T is lower semi-Fredholm according to Corollary 6.1.10. $\square$

Theorem 6.1.21 (Fredholm operators of index **0** and invertibility)**.** *Let X, Y be complex Banach spaces and let $T \in \mathcal{L}(X;Y)$. The following are equivalent:*

(i) *The operator T is Fredholm with index* 0.

(ii) *There exists a finite-rank operators $F \in \mathcal{L}(X;Y)$ such that $T + F$ is bijective.*

(iii) *There exists a compact operator $K \in \mathcal{L}(X;Y)$ such that $T + K$ is bijective.*

Proof. „(i) $\Rightarrow$ (ii)“ Again, we split X and Y: let $X = \ker T \oplus X_1$ and $Y = Y_0 \oplus TX$, where X_1 is a closed vector subspace of X and Y_0 is a finite dimensional vector subspace of Y. The dimensions of $\ker T$ and Y_0 (which are both finite) coincide as $\operatorname{ind} T = 0$; hence, there exists a linear (and continuous) bijection $G : \ker T \to Y_0$.

The restricted operator $T_1 := T|_{X_1} : X_1 \to TX$ is bijective. Let $P : X \to X$ denote the projection onto $\ker T$ along X_1, and define the finite rank operator $F \in \mathcal{L}(X;Y)$ as $F := GP$. Then $T + F : X = \ker T \oplus X_1 \to Y_0 \oplus TX$ acts as $S \oplus T_1$, so it is bijective.

„(ii) $\Rightarrow$ (iii)“ This implication is obvious.

„(iii) $\Rightarrow$ (i)“ As $T + K$ is bijective, it is Fredholm with index 0. Hence, it follows from Corollary 6.1.14 that $T = (T + K) + (-K)$ is also Fredholm with index 0. $\square$

6.2 The essential spectrum

In this section we introduce and study the so-called *essential spectrum* of a linear operator. Ideas and results from various parts of the course culminate here: We use the concept pf Fredholm operators (Section 6.1) to introduce the essential spectrum in Definition 6.2.1. By combining Theorem 6.1.18 about analytic Fredholm valued operators with the insights about poles of the resolvent from Theorem 5.3.4 we

[14]Note that K_Y' is compact since K_Y is so.

show how the essential spectrum is related to so-called *Riesz points* of an operator (Definition 6.2.8, Theorem 6.2.9, and Corollary 6.2.10).

Quotients of Banach algebras (Proposition 1.2.5) become important again since we show in Theorem 6.2.2 that the essential spectrum of a bounded linear operator can be described by means of the Calkin algebra that was introduced in Example 1.2.6. This demonstrates how the spectral theory in general unital Banach algebras (Part I) can be used to study the spectrum of bounded linear operators. This approach is also useful to study the essential spectrum of so-called *power compact* operators in Example 6.2.6; on the same occasion we will also come back to the spectral mapping theorem for polynomials in unital Banach algebras (Exercise 3 on Sheet 4).

Definition 6.2.1 (Essential spectrum and essential spectral radius)**.** Let X be a complex Banach space.

(a) Let $L : X \supseteq \operatorname{dom}(L) \to X$ be a closed linear operator. The set

$$\sigma_{\mathrm{ess}}(L) := \big\{\lambda \in \mathbb{C} \mid \ \lambda - L : \operatorname{dom}(L) \to X \text{ is not Fredholm}\big\}$$

is called the *essential spectrum* of L; here, $\operatorname{dom}(L)$ is endowed with a graph norm.[15] The elements of $\sigma_{\mathrm{ess}}(L)$ are called the *essential spectral values* of L.

(b) Let $T \in \mathcal{L}(X)$. The number $\mathrm{r}_{\mathrm{ess}}(T) := \sup\{|\lambda| \mid \ \lambda \in \sigma_{\mathrm{ess}}(T)\}$ is called the *essential spectral radius* of T.[16]

Note that the essential spectrum is always closed, as the set of Fredholm operators is open in $\mathcal{L}(\operatorname{dom}(L); X)$ is a consequence of Theorem 6.1.12. Moreover, as every bijective bounded linear operator between two Banach spaces is Fredholm, it follows that one always has

$$\sigma_{\mathrm{ess}}(L) \subseteq \sigma(L).$$

Hence, for $T \in \mathcal{L}(X)$, the essential spectrum is compact, and one always has $\mathrm{r}_{\mathrm{ess}}(T) \leq \mathrm{r}(T)$.

Recall from Example 1.2.6 that the *Calkin algebra* of a complex Banach space X is the unital Banach algebra $\mathcal{L}(X)/\mathcal{K}(X)$, where $\mathcal{K}(X) \subseteq \mathcal{L}(X)$ denotes the closed ideal of all compact linear operators on X.

Theorem 6.2.2 (The essential spectrum and the Calkin algebra)**.** *Let X be a complex Banach space and let $T \in \mathcal{L}(X)$. Then the essential spectrum $\sigma_{\mathrm{ess}}(T)$ coincides with the spectrum of the element $T + \mathcal{K}(X)$ in the Calkin algebra $\mathcal{L}(X)/\mathcal{K}(X)$.*

Proof. This can easily be derived from the characterization of Fredholm operators in Theorem 6.1.20. ☐

[15] Which renders $\operatorname{dom}(L)$ a Banach space since L is assumed to be closed; hence, it makes sense to speak about whether the operator $\lambda - L : \operatorname{dom}(L) \to X$ is Fredholm or not.

[16] Here the supremum is taken within the totally ordered set $[0, \infty)$, such that $\sup \emptyset = 0$.

Corollary 6.2.3 (Non-emptyness of the essential spectrum)**.** *Let X be a complex Banach space and let $T \in \mathcal{L}(X)$. Then $\sigma_{\text{ess}}(T)$ is non-empty if and only if X is infinite-dimensional.*

Proof. The Calkin algebra $\mathcal{L}(X)/\mathcal{K}(X)$ is non-zero if and only if X is infinite-dimensional. If the Calkin algebra is zero, its only element has empty spectrum (Remark 2.2.3(b)), and if it is non-zero, then each of its elements has non-empty spectrum (Theorem 2.4.1). Therefore, the claim follows from the description of the essential spectrum of bounded linear operators in Theorem 6.2.2. $\square$

Note that it is important in the preceding corollary that T is a bounded linear operator: as you will see on Exercise Sheet 13, there exist unbounded operators with empty essential spectrum. For instance, as you will prove on this Sheet, every operator with *compact resolvent* has empty essential spectrum; see also Example 6.2.7 below.

We continue with a perturbation result for the essential spectrum:

Proposition 6.2.4 (Stability of the essential spectrum)**.** *Let X be a complex Banach space and let $L : X \supseteq \text{dom}(L) \to X$ be a closed linear operator.*

(a) *If $K \in \mathcal{L}(\text{dom}(L); X)$ is compact,*[17] *then $\sigma_{\text{ess}}(L + K) = \sigma_{\text{ess}}(L)$.*[18]

(b) *If $K \in \mathcal{L}(X)$ is compact, then $\sigma_{\text{ess}}(L + K) = \sigma_{\text{ess}}(L)$.*[19]

Proof. (a) This is an immediate consequence of Corollary 6.1.14.

(b) Since $\text{dom}(L)$, endowed with the graph norm, embeds continuously into X it follows that $K|_{\text{dom}(L)} : \text{dom}(L) \to X$ is also compact. As $L + K = L + K|_{\text{dom}(L)}$, the claim thus follows from (a). $\square$

Example 6.2.5 (Compact operators)**.** Let X be a complex Banach space and let $K \in \mathcal{L}(X)$ be compact. Then $\sigma_{\text{ess}}(K) = \{0\}$ if X is infinite dimensional and $\sigma_{\text{ess}}(K) = \emptyset$ if X is finite dimensional.

Proof. It follows from Corollary 6.1.14 that $\sigma_{\text{ess}}(K) \subseteq \{0\}$. According to Corollary 6.2.3 one has equality if and only if X is infinite dimensional. $\square$

Example 6.2.6 (Compactness of $\boldsymbol{p(T)}$ and power compact operators)**.** Let X be a complex Banach space and let $T \in \mathcal{L}(X)$.

(a) Let p be a polynomial over $\mathbb{C}$ and assume that the operator $p(T)$ is compact. Then every essential spectral value of T is a root of p.

(b) Assume that T is *power compact*, i.e., that T^n is compact for some integer $n \geq 1$. Then $\sigma_{\text{ess}}(T) \subseteq \{0\}$.

[17] Here we again assume $\text{dom}(L)$ to be endowed with the graph norm.

[18] Note that the sum $L + K$ is defined on the domain $\text{dom}(L + K) = \text{dom}(L)$.

[19] Note that, as K is bounded from X to X, the sum $L + K$ is also defined on the domain $\text{dom}(L + K) = \text{dom}(L)$.

Proof. (a) For the sake of easier notation, let $A := \mathcal{L}(X)/\mathcal{K}(X)$ denote the Calkin algebra of X and let $q : \mathcal{L}(X) \to A$ denote the quotient mapping. Then we have

$$p\big(\sigma_{\mathrm{ess}}(T)\big) = p\big(\sigma(q(T)\big) = \sigma\big(p(q(T))\big) = \sigma\big(q(p(T))\big) = \sigma(0) \subseteq \{0\},$$

where the first equality follows from Theorem 6.2.2, the second equality from the polynomial spectral mapping theorem in unital Banach algebras (Exercise 3 on Sheet 4), the third equality from the fact that the quotient mapping is an algebra homomorphism (Proposition 1.2.5), and the last equality from the assumption that $p(T) \in \mathcal{K}(X)$.

(b) This is a special case of (a). $\square$

Example 6.2.7 (Operators with compact resolvent)**.** Let X be a complex Banach space and let $L : X \supseteq \operatorname{dom}(L) \to X$ be a closed linear operator. Assume that $\rho(L) \neq \emptyset$ and that, for some $\lambda \in \rho(L)$, the operator $\mathcal{R}(\lambda, L)$ is compact from X to X.

Then $\sigma_{\mathrm{ess}}(L) = \emptyset$; this will be proved on Exercise Sheet 13.

Observe that, according to Example 6.2.5, the essential spectrum of a compact operator K is contained in $\{0\}$ and that, at the same time, the non-zero spectral values of K have very specific properties according to Example 6.1.19. This is not a coincidence, but rather a special case of the subsequent theorem and its corollary. In order to facilitate the statement of those results, we first introduce the following terminology. Recall again that poles of the resolvent of a linear operator are always eigenvalues (Theorem 5.3.4(a)).

Definition 6.2.8 (Riesz points in the spectrum)**.** Let X be a complex Banach space and let $L : X \subseteq \operatorname{dom}(L) \to X$ be a closed linear operator. A spectral value λ_0 of L is called a *Riesz point* of L if λ_0 is isolated in $\sigma(L)$, is a pole of the resolvent $\mathcal{R}(\,\cdot\,, L)$, and has finite algebraic multiplicity as an eigenvalue of L.

With this terminology, Example 6.1.19(b) says that every non-zero spectral value of a compact operator $K \in \mathcal{L}(X)$ is a Riesz point of K.

Theorem 6.2.9 (Non-essential spectral values vs. Riesz points)**.** *Let X be a complex Banach space and let $L : X \subseteq \operatorname{dom}(L) \to X$ be a closed linear operator. For every spectral value $\lambda_0 \in \sigma(L)$ the following are equivalent:*

(i) *The number λ_0 is a Riesz point of L.*

(ii) *One has $\lambda_0 \notin \sigma_{\mathrm{ess}}(L)$, and the connected component of the open set $\mathbb{C} \setminus \sigma_{\mathrm{ess}}(L)$ that contains λ_0 intersects the resolvent set $\rho(L)$.*

Proof. „(i) $\Rightarrow$ (ii)“ We need to show that $\lambda_0 - L : \operatorname{dom}(L) \to X$ has finite nullity and defect;[20] then it is a Fredholm operator according to Remark 6.1.5.

[20] Note that $\operatorname{dom}(L)$ is again endowed with a graph norm here.

Let λ_0 be a Riesz point. Then λ_0 is an eigenvalue and its algebraic multiplicity is finite; hence, its geometric multiplicity is finite, too, i.e., $\ker(\lambda_0 - L)$ is finite dimensional. In other words, $\operatorname{nul}(\lambda_0 - L) < \infty$.

Since λ_0 is a Riesz point, it is also a pole of the resolvent. Let Q_0 be as in Theorem 5.3.4. Then, according to part (c) of that Theorem, $\ker Q_0$ coincides with the range of $(\lambda_0 - L)^q$, where $q \in \mathbb{N}$ denote the pole order of λ_0. As the operator Q_0 is a projection (Theorem 5.3.1(d)), its range is closed and X splits as $X = Q_0X \oplus \ker Q_0$, so the codimension of the range of $(\lambda_0 - L)^q$ is equal to the dimension of Q_0X. This dimension is finite since Q_0X is the generalized eigenspace of the eigenvalue λ_0 according to Theorem 5.3.4(b); so the codimesion of the range of $(\lambda_0 - L)^q$ is finite. As the range of $\lambda_0 - L$ contains the range of $(\lambda - L)^q$, the former has even smaller codimension; so $\operatorname{def}(\lambda_0 - L) < \infty$.

So $\lambda_0 - L$ is indeed a Fredholm operator, i.e., $\lambda_0 \notin \sigma_{\mathrm{ess}}(L)$. Let $\Omega \subseteq \mathbb{C}$ denote denote the connected component of $\mathbb{C} \setminus \sigma_{\mathrm{ess}}(L)$ that contains λ_0. Since Ω is open and contains λ_0, it also contains a small ball with center λ_0. As λ_0 is a Riesz point of L it is isolated in $\sigma(L)$, so all points sufficiently close to λ_0 are in $\rho(L)$. This shows that Ω intersects $\rho(L)$.

„(ii) $\Rightarrow$ (i)“ Let $\lambda_0 \notin \sigma_{\mathrm{ess}}(L)$ and let $\Omega \subseteq \mathbb{C}$ denote the connected component of $\mathbb{C} \setminus \sigma_{\mathrm{ess}}(L)$ that contains λ_0. Let $F : \Omega \to \mathcal{L}(\operatorname{dom}(L); X)$ be given by

$$F(\lambda) := \lambda - L$$

for all $\lambda \in \Omega$. Then F is analytic and takes values in the Fredholm operators only. Moreover, as Ω was assumed to intersect $\rho(L)$, there exists a point $\lambda \in \Omega$ for which the operator $F(\lambda) : \operatorname{dom}(L) \to X$ is bijective. Hence, all assumptions of Theorem 6.1.18 are satisfied. The set D from this theorem is, in the present situation, equal to $\Omega \cap \sigma(L)$, so the theorem says that every point in this set – in particular the point λ_0 – is isolated in $\sigma(L)$ and a pole of the resolvent; moreover, the assertion about the Laurent series coefficients in the theorem shows that the operator Q_0 (as in the proof of the converse implication we use the notation from Theorem 5.3.4) has finite rank. But the range Q_0X is the generalized eigenspace of λ_0 (Theorem 5.3.4(b)), so the algebraic multiplicity of the eigenvalue λ_0 is indeed finite. Hence, λ_0 is a Riesz point, as claimed. □

Corollary 6.2.10 (Non-essential spectral values vs. Riesz points, more concretely)**.** *Let X be a complex Banach space and let $L : X \supseteq \operatorname{dom}(L) \to X$ be a closed linear operator.*

(a) *If $\rho(L) \neq \emptyset$ and if the open set $\mathbb{C} \setminus \sigma_{\mathrm{ess}}(L)$ is connected, then the Riesz points of L are precisely the elements of $\sigma(L) \setminus \sigma_{\mathrm{ess}}(L)$.*

(b) *A number $\lambda_0 \in \partial\sigma(L)$ is a Riesz point of L if and only if $\lambda_0 \notin \sigma_{\mathrm{ess}}(L)$.*

Proof. (a) By assumption that set $\mathbb{C} \setminus \sigma_{\mathrm{ess}}(L)$ consists of only one connected component, and it obviously intersects $\rho(L)$ since $\rho(L) \neq \emptyset$. So the claim follows immediatly from Corollary 6.2.9.

(b) Let $\lambda_0 \in \partial\sigma(L)$.

„$\Rightarrow$" This implication is immediate from Theorem 6.2.9.

„$\Leftarrow$" Let $\lambda_0 \notin \sigma_{\mathrm{ess}}(L)$. As λ_0 is a boundary point of $\sigma(L)$, every ball centered at λ_0 intersects $\rho(L)$. Hence, the connected component of $\mathbb{C} \setminus \sigma_{\mathrm{ess}}(L)$ also intersects $\rho(L)$. Therefore, this implication also follows from Theorem 6.2.9. $\square$

Example 6.2.11 (Compactness of $\boldsymbol{p(T)}$, again)**.** As in Example 6.2.6(a) let X be a complex Banach space, $T \in \mathcal{L}(X)$, and assume that p is a complex polynomial such that $p(T)$ is compact. Assume in addition that $p \neq 0$.

We know from Example 6.2.6(a) that $\sigma_{\mathrm{ess}}(T)$ is contained in the set of roots of p, and the latter set is finite as $p \neq 0$. Hence, $\sigma_{\mathrm{ess}}(T)$ is finite, so its complement in $\mathbb{C}$ is connected. So Corollary 6.2.10(a) shows that every spectral value λ of T that satisfies $p(\lambda) \neq 0$ is a Riesz point of T.

Part III

Functional Calculi and Spectral Representation

Chapter 7

Holomorphic Functional Calculus

7.1 Intermezzo: The Cauchy integral formula

We need to recall a number of concepts from complex analysis. In order to do so, *paths* and *cycles* will be important:

Definition 7.1.1 (Paths and cycles). (a) A *closed C^1-path in* $\mathbb{C}$ is a continuously differentiable mapping $\gamma : [0,1] \to \mathbb{C}$ such that $\gamma(0) = \gamma(1)$.

(b) A closed C^1-cycle Γ in $\mathbb{C}$ is a tuple $\Gamma = (\gamma_1, \dots, \gamma_n)$ of finitely many closed C^1-paths in $\mathbb{C}$.

Based on paths and cycles one can define the following integral notion: Let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be open, let X be a complex Banach space, and let $f : \Omega \to X$ be continuous, and let γ be a closed C^1-path which is contained in Ω. Then one sets

$$\oint_\gamma f(z)\mathrm{d}z := \int_0^1 f(\gamma(t))\dot{\gamma}(t)\ \mathrm{d}t,$$

where the latter integral is understood as a vector-valued Riemann integral in the sense of Definition 1.3.10.[1] Moreover, if $\Gamma = (\gamma_1, \dots, \gamma_n)$ is a closed C^1-cycle and all the paths $\gamma_1, \dots, \gamma_n$ are contained in Ω, then we define

$$\oint_\Gamma f(z)\mathrm{d}z := \sum_{k=1}^n \oint_{\gamma_k} f(z)\ \mathrm{d}z.$$

If Γ is a closed C^1-cycle in $\mathbb{C}$ and $z_0 \in \mathbb{C}$ is not located on any of the paths of which Γ consists, then the number

$$\mathrm{n}(\Gamma; z_0) := \frac{1}{2\pi\mathrm{i}} \oint_\Gamma \frac{1}{z - z_0}\ \mathrm{d}z$$

[1] In fact such an integral already occurred in Theorem 5.2.2 where we used it to compute the coefficients of Laurent series expansions.

is called the *index* or *winding number* of Γ around z_0. It is a standard result in complex analysis that the winding number is always an element of $\mathbb{Z}$; see for instance [Con73, Section IV.4]. Intuitively, the winding number $\mathrm{n}(\Gamma; z_0)$ describes how often the cycle Γ encircles the point z_0 counterclockwise.

Theorem 7.1.2 (Cauchy's theorem and Cauchy's integral formula)**.** *Let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be open, let X be a complex Banach space, and let $f : \Omega \to X$ be holomorphic, and let Γ be a closed C^1-cycle in $\mathbb{C}$ that is contained in Ω. Assume moreover that $\mathrm{n}(\Gamma; c) = 0$ for every $c \in \mathbb{C} \setminus \Omega$.*

(a) *One has* Cauchy's integral theorem*, i.e.,*

$$\oint_\Gamma f(z) \,\mathrm{d} = 0.$$

(b) *For every $z_0 \in \Omega$ that is not located on any of the paths in Γ one has* Cauchy's integral formula

$$\oint_\Gamma \frac{f(z)}{z - z_0} \,\mathrm{d}z = \mathrm{n}(\Gamma; z_0) f(z_0).$$

Proof. For scalar-valued functions those are standard results in complex analysis; see for instance [Con73, Section IV.5]. The vector-valued case can be deduced from the scalar-valued case by testing against bounded linear functionals. □

Remark 7.1.3 (Independence of the path of integration)**.** Let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be open, let X be a complex Banach space, and let $f : \Omega \to X$ be holomorphic, and let $\Gamma, \tilde{\Gamma}$ be a closed C^1-cycle in $\mathbb{C}$ that is contained in Ω. Assume that $\mathrm{n}(\Gamma, c) = \mathrm{n}(\tilde{\Gamma}, c)$ for all $c \in \mathbb{C} \setminus C$.

Then Theorem 7.1.2(a) implies that

$$\oint_\Gamma f(z) \,\mathrm{d}z = \oint_{\tilde{\Gamma}} f(z) \,\mathrm{d}z.$$

We leave the details of how to derive this from Theorem 7.1.2(a) as an exercise.

In order to construct the holomorphic functional calculus the following existence result for cycles is important:

Proposition 7.1.4 (Existence of cycles)**.** *Let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be open and let $K \subseteq \Omega$ be compact. Then there exists a closed C^1-cycle Γ which is contained in $\Omega \setminus K$ and satisfies the properties*

$$\mathrm{n}(\Gamma; z_0) = 1 \qquad \textit{and} \qquad \mathrm{n}(\Gamma; c) = 0$$

for every $z_0 \in K$ and every $c \in \mathbb{C} \setminus \Omega$.

Proof. See for instance [Con95, Proposition 1.8 on p. 4 in Section 13.1]. □

7.2 The holomorphic functional calculus

The following notation is convenient: for a non-empty open set $\Omega \subseteq \mathbb{C}$ we denote by $\mathrm{Hol}(\Omega)$ the algebra of all holomorphic functions $\Omega \to \mathbb{C}$, where addition, scalar multiplication, and multiplication, are defined pointwise.

Definition 7.2.1 (Holomorphic functional calculus)**.** Let A be a unital Banach algebra, let $a \in A$, let $\Omega \subseteq \mathbb{C}$ be an open set that contains $\sigma(a)$, and let $f : \Omega \to \mathbb{C}$ be holomorphic. Then we define $f(a) \in A$ as

$$f(a) := \frac{1}{2\pi \mathrm{i}} \oint_\Gamma f(z)\mathcal{R}(z,a)\mathrm{d}z,$$

where Γ is any closed C^1-cycle in $\mathbb{C}$ that is contained in Ω and that satisfies $\mathrm{n}(\Gamma; c) = 0$ for all $c \in \mathbb{C} \setminus \Omega$ and $\mathrm{n}(\Gamma; z_0) = 1$ for all $z_0 \in \sigma(a)$. [2]

Remark 7.2.2 (The functional calculus does not depend on the cycle nor on the surrounding set)**.** In the setting of Definition 7.2.1, the definition of $f(a)$ does not depend on the choice of Γ; this follows from Remark 7.1.3.

Consequently, the element $f(a)$ does not change, either, if we restrict f to a smaller open set that still contains $\sigma(a)$.

Theorem 7.2.3 (The holomorphic functional calculus is an algebra homomorphism)**.** *Let a be an element of a unital Banach algebra A, let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be an open set that contains $\sigma(a)$, Then the mapping*

$$\begin{aligned} \mathrm{Hol}(\Omega) &\to A \\ f &\mapsto f(a) \end{aligned}$$

is an algebra homomorphism which map the constant function $\mathbb{1}$ to the neutral element 1 of A; moreover, it maps the function $(z \mapsto z)$ to a.

Proof. Linearity of the map is clear, and the proof of multiplicativity is an exercise on Sheet 14. To show that $\mathbb{1}(a) = 1$, one nots that $\mathbb{1}$ is holomorphic on all of $\mathbb{C}$, so one can use a circle with center 0 and radius r for any $r > \mathrm{r}(a)$ as the $\Gamma = \Gamma_r$ over which ones integrates in the definition of $\mathbb{1}(a)$. Thus,

$$2\pi\mathrm{i}\, \mathbb{1}(a) = \oint_{\Gamma_r} \mathcal{R}(z,a)\mathrm{d}z = \oint_{\Gamma_r} \frac{(z-a)+a}{z}\mathcal{R}(z,a)\mathrm{d}z = \oint_{\Gamma_r} \frac{1}{z}\mathrm{d}z + \oint_{\Gamma_r} \frac{a}{z}\mathcal{R}(z,a)\mathrm{d}z;$$

the first summand is equal to $2\pi\mathrm{i} \in A$, and the second one converges to 0 as $r \to \infty$ as follows from the Neumann series expansion of $\mathcal{R}(z,a)$. Hence, $\mathbb{1}(a) = 1$, is claimed.

Finally, let $f : \Omega \to \mathbb{C}$ be given by $f(z) = z$ for all $z \in \Omega$. Then, for an appropriately chosen cycle Γ,

$$2\pi\mathrm{i} f(a) = \oint_\Gamma (z-a)\mathcal{R}(z,a)\mathrm{d}z + \oint_\Gamma a\mathcal{R}(z,a)\mathrm{d}z = \oint_\Gamma 1\mathrm{d}z + 2\pi\mathrm{i}a\, \mathbb{1}(a) = 2\pi\mathrm{i}a,$$

which proves the claim. □

[2] Note that such a cycle exists according to Proposition 7.1.4.

Theorem 7.2.4 (The spectral mapping theorem and composition of holomorphic functional calculi)**.** *Let a be an element of a unital Banach algebra A, let $\emptyset \neq \Omega \subseteq \mathbb{C}$ be an open set that contains $\sigma(a)$, and let $f : \Omega \to \mathbb{C}$ be holomorphic.*

(a) *One has the* spectral mapping theorem $\sigma\big(f(a)\big) = f\big(\sigma(a)\big)$.

(b) *The holomorphic functional calculus respect the composition of functions, i.e.:*

if $\tilde{\Omega} \subseteq \mathbb{C}$ is an open set that contains $f(\Omega)$ and $g : \tilde{\Omega} \to \mathbb{C}$ is holomorphic, then $(g \circ f)(a) = g\big(f(a)\big)$.

For a proof of this theorem we refer, for instance, to [HP57, Theorems 5.3.1 and 5.3.2 on p. 171] or [Yos80, Corollaries 1 and 2 in Section VIII.7 on p. 227] (the results in the latter reference are only formulated on the Banach algebra $\mathcal{L}(X)$ for a Banach space X, but the same arguments work in general unital Banach algebras).

Examples 7.2.5 (Powers and the resolvent)**.** Let A be a unital Banach algebra and let $a \in A$.

(a) For a polynomial function $p : \mathbb{C} \to \mathbb{C}$ the element $p(a)$ that is defined by means of the holomorphic functional calculus coincides with the element that is defined by simply substituting a into the standard representation of p as a linear combination of monomials. This follows from the properties of the holomorphic functional calculus given in Theorem 7.2.4.

(b) Let $\lambda \in \mathbb{C} \setminus \sigma(a)$ and let $f : \mathbb{C} \setminus \{\lambda\} \to \mathbb{C}$ be given by $f(z) = \frac{1}{\lambda - z}$ for all $z \in \mathbb{C} \setminus \{\lambda\}$. Then the domain of f contains the $\sigma(a)$ and thus, $f(a)$ is well-defined by means of the holomorphic functional calculus.

As one would intuitively expect, one has $f(a) = \mathcal{R}(\lambda, a)$. This can be shown be considering the element $(\lambda - a)f(a)$ and using that the holomorphic functional calculus is multiplicative.

Proposition 7.2.6 (The holomorphic functional calculus via the Taylor series)**.** *Let a be an element of a unital Banach algebra A, let $z_0 \in \mathbb{C}$ and let $r \in (0, \infty]$ such that the open ball $\mathrm{B}_{<r}(z_0)$ in $\mathbb{C}$ contains $\sigma(a)$. If $f : \mathrm{B}_{<r}(z_0) \to \mathbb{C}$ is holomorphic, then*

$$f(a) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(a - z_0)^k,$$

where the series converges absolutely in A.

The proof of this proposition will be part of Exercise Sheet 15.

Example 7.2.7 (The exponential function)**.** Let a be an element of a unital Banach algebra A, and let $\exp : \mathbb{C} \to \mathbb{C}$ be the complex exponential function. The element

$\exp(a)$ that ones obtains by means of the holomorphic functional calculus is then, according to Proposition 7.2.6, given by

$$\exp(a) = \sum_{n=0}^{\infty} \frac{a^n}{n!}.$$

Hence, the definition of $\exp(a)$ by means of the holomorphic functional calculus is consistent with the definition of $\exp(a)$ that we gave in Section 1.4.

7.3 Spectral projections

Definition 7.3.1 (Spectral projections and spectral spaces)**.** Let X be a complex Banach space and let $T \in \mathcal{L}(X)$. Let $\sigma_1 \subseteq \sigma(T)$ be a closed subset of $\sigma(T)$ which is isolated within $\sigma(T)$, meaning that $\sigma_2 := \sigma(T) \setminus \sigma_1$ is also closed.

Then there are non-empty and disjoint open sets Ω_1 and Ω_2 in $\mathbb{C}$ that contains σ_1 and σ_2, respectively. The indicator function $\mathbb{1}_{\Omega_1} : \Omega_1 \cup \Omega_2 \to \mathbb{C}$ is holomorphic, and the operator $P_1 := \mathbb{1}_{\Omega_1}(T) \in \mathcal{L}(X)$ is called the *spectral projection* of T associated to σ_1. The range of P_1 is called the *spectral space* of T associated with σ_1.

In the situation of the preceding definition note that, according to Remark 7.2.2, the spectral projection (and hence the spectral space) of σ_1 does not depend on the choice of the open sets Ω_1 and Ω_2. Moreover, since the functional calculus is an algebra homomorphism (Theorem 7.2.3), it is easy to checkt that the spectral projection is indeed a projection (see Exercise 2 on Sheet 14 for details).

Remark 7.3.2 (Isolated singularities of the resolvent)**.** Let X be a complex Banach space and let $T \in \mathcal{L}(X)$. If $\lambda_0 \in \sigma(T)$ is an isolated point in $\sigma(T)$ and

$$\mathcal{R}(\lambda, T) = \sum_{k=-\infty}^{\infty} Q_{k+1}(\lambda - \lambda_0)^k$$

for λ close to λ_0 is the Laurent series expansion of $\mathcal{R}(\,\cdot\,, T)$ about λ_0, then Q_0 is the spectral projection of T associated to $\{\lambda_0\}$.

Theorem 7.3.3 (Splitting via spectral projections)**.** *Assume that we are in the situation of Definition 7.3.1 and set $P_2 := \mathrm{id}_X - P_1$.*

(a) *The projection P_2 is the spectral projection of T associated to σ_2.*

(b) *The proctions P_1, P_2 commute with T, so T leaves P_1X and P_2X invariant.*

(c) *For each $k \in \{1, 2\}$ the restriction $T_2 := T|_{P_kX} \in \mathcal{L}(P_kX)$ has spectrum σ_k.*

Proof. We use the notation from Definition 7.3.1, and we set $\Omega := \Omega_1 \cup \Omega_2$.

(a) One has $\mathbb{1}_{\Omega_1}(a) + \mathbb{1}_{\Omega_2}(a) = \mathbb{1}_{\Omega}(a) = 1$, so $\mathbb{1}_{\Omega_2}(a) = 1 - P_1 = P_2$.

(b) It is clear that P_1 and P_2 commute with T; moreover, one can readily check that two commuting operators leave their images invariant.

(c) We first note that, according to (b), T is the direct sum of the operators T_1 and T_2; hence, $\sigma(T) = \sigma(T_1) \cup \sigma(T_2)$.

Now let $\lambda \in \sigma_2$; we claim that $\lambda \notin \sigma(T_1)$. To see this, consider the holomorphic function $f : \Omega \to \mathbb{C}$ that is given by

$$f(z) = \begin{cases} \frac{1}{\lambda - z} & \text{for all } z \in \Omega_1, \\ 0 & \text{for all } z \in \Omega_2. \end{cases}$$

Then $f(T)$ also commutes with P_1 and P_2 and thus leaves their image invariant. Moreover, $f(T)(\lambda - T) = P_1$, so $\lambda - T_1$ is a bijective operator on P_1X with inverse $f(T)|_{P_1X}$. So $\lambda \notin \sigma(T_1)$, as claimed.

By exchanging the roles of T_1 and T_2 we see that, similarly, a number $\lambda \in \sigma_1$ cannot be an element of $\sigma(T_2)$. As $\sigma(T_1) \cup \sigma(T_2) = \sigma(T)$ is the disjoint union of σ_1 and σ_2, we conclude that $\sigma(T_1) = \sigma_1$ and $\sigma(T_2) = \sigma_2$, as claimed. □

Remark 7.3.4 (Spectral projections for unbounded operators)**.** In this section, we only discussed spectral projections for bounded operators since, for this case, they can be obtained from the holomorphic functional calculus in Banach algebras discussed in Section 7.2.

However, for bounded isolated parts of the spectrum of an unbounded linear operator on a Banach space, one can also define spectral projection by using the same approach. See for instance [EN00, pp. 244–246] for more information.

Similarly, it is possible to define a holomorphic functional calculus for certain unbounded operators, too. For a classical account of this – which only works for a rather restricted set of functions, though – we refer to [TL80, Section V.8, in particular from page 314 on]. More recent results on the functional calculus of a class of unbounded operators can be found in [Haa06].

Chapter 8

C^*-Algebras and the Continuous Functional Calculus

8.1 C^*-algebras

Definition 8.1.1 (**C^***-algebras)**.** (a) Let A be an algebra. An *involution* on A is a mapping $\cdot^* : A \to A$ which satisfies

$$(\lambda a)^* = \overline{\lambda} a^*, \qquad (a+b)^* = a^* + b^*,$$
$$a^{**} := (a^*)^* = a, \qquad (ab)^* = b^* a^*$$

for all $a, b \in A$ and all $\lambda \in \mathbb{C}$.

(b) A *C^*-algebra* is a Banach algebra A together with an involution $\cdot^* : A \to A$ that satisfies the additional condition $\|a\|^2 \le \|a^* a\|$ for all $a \in A$.

A *unital C^*-algebra* is a C^*-algebra whose underlying Banach algebra is unital.

Note that if A is a unital C^*-algebra, then $1^* = 1$. Indeed one has $1^* = 1\,1^*$ and, by taking the involution on both sides, also $1 = (1\,1^*)^* = 1\,1^*$.

The condition $\|a\|^2 \le \|a^* a\|$ in the definition of a C^*-algebra turns out to automatically imply the following – formally stronger – properties of the norm.[1]

Proposition 8.1.2 (Properties of the norm in a **C^***-algebra)**.** *Let A be a C^*-algebra. Then we have*

$$\|a\| = \|a^*\| \qquad \text{and} \qquad \|a\|^2 = \|a^* a\|$$

for each $a \in A$. In particular, the involution is isometric and hence continuous.

Proof. For all $a \in A$ we have

$$\|a\|^2 \le \|a^* a\| \le \|a^*\|\,\|a\| \tag{8.1.1}$$

[1] Thus, one can just as well us those stronger conditions for the definition of a C^*-algebra, as is for instance done in [Mur90, Section 2.1].

where the first inequality was required directly in the definition of a C^*-algebra and the second inequality follows from the submultiplicativity of the norm. Hence, $\|a\| \le \|a^*\|$ for all $a \in A$. By replacing a with a^* in this inequality we also obtain $\|a^*\| \le \|a\|$, and hence, $\|a\| = \|a^*\|$ for all $a \in A$.

By substituting $\|a^*\|$ with $\|a\|$ on the right of the inequality (8.1.1) we obtain $\|a\|^2 \le \|a^*a\| \le \|a^2\|$ and hence, $\|a\|^2 = \|a^*a\|$, as claimed. $\square$

The two most standard examples of C^*-algebras are the following:

Example 8.1.3 (Spaces of continuous functions)**.** Let $K \neq \emptyset$ be a compact Hausdorff space. Then $\mathrm{C}(K)$, together with the involution that is given by pointwise complex conjugation, is a unital C^*-algebra. This is straightforward to check.

Example 8.1.4 (The bounded linear operators on a Hilbert space)**.** Let H be a complex Hilbert space. Then the mapping $\mathcal{L}(H) \ni T \mapsto T^* \in \mathcal{L}(H)$, which maps each operator T to its Hilbert space adjoint, turns $\mathcal{L}(H)$ into a unital C^*-algebra.

Indeed, for every $T \in \mathcal{L}(H)$ and all $x \in H$ of norm 1 one has

$$\|Tx\|^2 = |(Tx, Tx)| = |(T^*Tx, x)| \le \|T^*T\| .$$

So by taking the supremum over all normalized x and y we obtain $\|T\|^2 \le \|T^*T\|$, which is the C^*-algebra property.

An example of a non-unital C^*-algebra is the space of all compact linear operators on an infinite dimensional Hilbert space. This will be discussed on Exercise Sheet 15.

8.2 Self-adjoint and normal elements

Definition 8.2.1 (Self-adjoint elements)**.** Let A be a C^*-algebra.

(a) An element $a \in A$ is called *self-adjoint* if $a^* = a$.

(b) For each element $a \in A$ the self-adjoint elements

$$\operatorname{Re} a := \frac{1}{2}(a + a^*) \qquad \text{and} \qquad \operatorname{Im} a := \frac{1}{2\mathrm{i}}(a - a^*)$$

of A are called the *real part* and *imaginary part* of a.

Note that for all elements a of a C^*-algebra A one has $a = \operatorname{Re} a + \mathrm{i} \operatorname{Im} a$.

Definition 8.2.2 (Normal and unitary elements)**.** Let A be a C^*-algebra.

(a) An element $a \in A$ is called *normal* if it commutes with a^*.

(b) Assume now that A is unital. Then an element $u \in A$ is called *unitary* if $u^*u = uu^* = 1$.[2]

[2]Equivalently, if u is invertible and $u^{-1} = u^*$.

Obivously, every self-adjoint element of a C^*-algebra is normal, and every unitary element of a unital C^*-algebra is normal.

Example 8.2.3 (Unitary images of the exponential function)**.** Let A be a C^*-algebra and let $a \in A$. If a is self-adjoint, then $e^{\mathrm{i}a}$ is unitary.

Proof. The continuity of the involution and the definition of the exponential function yield

$$\exp(\mathrm{i}a)^* = \exp\big((\mathrm{i}a)^*\big) = \exp(-\mathrm{i}a),$$

where the last equality uses the self-adjointness of a and the anti-linearity of the involution. Since $\mathrm{i}a$ and $-\mathrm{i}a$ commute, Proposition 1.4.2(b) yields

$$\exp(\mathrm{i}a)^* \exp(\mathrm{i}a) = \exp(\mathrm{i}a)\exp(\mathrm{i}a)^* = \exp(\mathrm{i}a - \mathrm{i}a) = \exp(0) = 1,$$

so $\exp(\mathrm{i}a)$ is indeed a unitary. □

Proposition 8.2.4 (The spectral radius of normal elements)**.** *Let A be a unital C^*-algebra and let $a \in A$.*

(a) *One has $\|a\| = \mathrm{r}(a^*a)^{1/2}$.*

(b) *If a is self-adjoint or, more generally, normal, then $\mathrm{r}(a) = \|a\|$.*

Proof. We first show (b) in the special case where a is self-adjoint: Proposition 8.1.2 then gives $\|a^2\| = \|a\|^2$ since $a^* = a$. By iterating this equality (which is possible since all powers of a are self-adjoint, too) we obtain $\|a^{2^n}\| = \|a\|^{2^n}$ for all $n \in \mathbb{N}$. Hence, the spectral radius formula from Theorem 2.4.4 gives

$$\mathrm{r}(a) = \lim_{n\to\infty} \left\|a^{2^n}\right\|^{1/2^n} = \|a\| .$$

Now we can proof both assertions of the proposition:

(a) Let $a \in A$ be arbitrary. Then a^*a is self-adjoint, so it follows from what we have just shown that $\mathrm{r}(a^*a) = \|a^*a\|$. But the latter number is equal to $\|a\|^2$ according to Proposition 8.1.2.

(b) Let $a \in A$ be normal. For every $n \in \mathbb{N}$ one then has

$$\|a^n\|^2 = \mathrm{r}\left((a^n)^*a^n\right) = \mathrm{r}\left((a^*a)^n\right) = \mathrm{r}(a^*a)^n = \|a\|^{2n} ,$$

where the first equality follows from (a) (applied to the element a^n), the second equality follows from the normality of a, the third equality follows from the spectral mapping theorem for polynomials (Exerise 3 on Sheet 4), and the last equality follows again from (a). Thus, $\|a^n\| = \|a\|^n$, and hence the spectral radius formula from Theorem 2.4.4 yields $\mathrm{r}(a) = \|a\|$, as claimed. □

Corollary 8.2.5 (Spectrum and norm of unitary elements)**.** *Let $A \neq \{0\}$ be a unital C^*-algebra and let $u \in A$ be unitary. Then $\|u\| = 1$ and $\mathrm{r}(u) \subseteq \mathbb{T}$.*

Proof. It follows from Proposition 8.2.4(a) that $\|u\| = \mathrm{r}(u^*u)^{1/2} = \mathrm{r}(1)^{1/2} = 1$, where the last equality uses that $A \neq \{0\}$.

So in particular, $\mathrm{r}(u) \le \|u\| = 1$. But $u^{-1} = u^*$ is unitary, too, so it follows that $\mathrm{r}(u^{-1}) \le 1$, as well. So every spectral value λ of u satisfies $|\lambda| \le 1$ and $\left|\lambda^{-1}\right| \le 1$, and hence, $|\lambda| = 1$. □

8.3 Gelfand representation for commutative C^*-algebras

Proposition 8.3.1 (The spectrum of self-adjoint elements)**.** *Let A be a unital C^*-algebra and let $a \in A$ be self-adjoint. Then $\sigma(a) \subseteq \mathbb{R}$.*

Proof. As a is self-adjoint, it is clearly normal. Moreover, $\exp(\mathrm{i}a)$ is unitary according to Example 8.2.3 and hence, $\sigma\big(\exp(\mathrm{i}a)\big) \subseteq \mathbb{T}$ by Corollary 8.2.5. So the spectral mapping theorem for the holomorphic functional calculus, Theorem 7.2.4(a), shows that $\sigma(\mathrm{i}a) \subseteq \mathrm{i}\mathbb{R}$ and hence, $\sigma(a) \subseteq \mathbb{R}$. □

Corollary 8.3.2 (Characters preserve adjoints)**.** *Let A be a unital C^*-algebra and let $\tau \in \Omega(A)$. Then $\tau(a^*) = \overline{\tau(a)}$ for all $a \in A$.*

Proof. Let $a \in A$. The we have $a = \operatorname{Re} a + \mathrm{i} \operatorname{Im} a$, where $\operatorname{Re} a$ and $\operatorname{Im} a$ are self-adjoint elements of A. Hence, it follows from the preceding Proposition 8.3.1 that $\operatorname{Re} a$ and $\operatorname{Im} a$ have real spectrum and that, thus, $\tau(\operatorname{Re} a)$ and $\tau(\operatorname{Im} a)$ are in $\mathbb{R}$. Thus,

$$\overline{\tau(a)} = \overline{\tau(\operatorname{Re} a) + \mathrm{i}\tau(\operatorname{Im}(a))} = \tau(\operatorname{Re} a) - \mathrm{i}\tau(\operatorname{Im} a) = \tau(a^*),$$

as claimed. □

Theorem 8.3.3 (Gelfand representations of commutative $\boldsymbol{C^*}$-algebras)**.** *Let $A \neq \{0\}$ be a unital C^*-algebra and assume that A is commutative. Then the Gelfand homomorphism*

$$\Theta : A \to \mathrm{C}\left(\Omega(A)\right), \qquad a \mapsto \hat{a}$$

is isometric, bijective, and preserves adjoints.

Proof. Since A is commutative, every element of A is normal. Hence, we have

$$\|\Theta(a)\|_\infty = \mathrm{r}(a) = \|a\|$$

for each $a \in A$, where the first equality follows from Theorem 3.2.5(c) and the second equality from Proposition 8.2.4(b). So Θ is indeed isometric, and hence it is, in particular, injective.

Every every $\tau \in \Omega(A)$ and every $a \in A$ one has

$$\Theta(a^*)(\tau) = \tau(a^*) = \overline{\tau(a)} = \overline{\Theta(a)}(\tau),$$

where the second equality follows from Corollary 8.3.2. So Θ does indeed preserve adjoints.

Finally, let us show surjectivity of Θ. As Θ is an algebra homomorphism and maps 1 to $\mathbb{1}$, the image B of Θ is a subalgebra of $\mathrm{C}\left(\Omega(A)\right)$ that contains $\mathbb{1}$. Moreover, B is invariant under taking ajdoints since Θ preserves adjooints. If $\tau_1, \tau_2 \in \Omega(A)$ are two distinct characters, then there exists $a \in A$ such that $\tau_1(a) \neq \tau_2(a)$. Thus, $\Theta(a)(\tau_1) \neq \Theta(a)(\tau_2)$, so B seperates the points of $\Omega(A)$. Thus, the Stone–Weierstraß approximation theorem 1.1.8 implies that B is dense in $\mathrm{C}\left(\Omega(A)\right)$. But as Θ is isometric, B is closed, so we conclude that $B = \mathrm{C}\left(\Omega(A)\right)$. ☐

8.4 The continuous functional calculus

For normal elements of C^*-algebras (rather than for general elements of general Banach algebras) one can define a functional calculus even for continuous function (rather than only for holomorphic ones), and it even suffices if the function is defined merely on the spectrum $\sigma(a)$ (rather than in a neighbourhood of $\sigma(a)$.

To make this precise, we need the following definition and the subsequent result:

Definition 8.4.1 (***C****-subalgebras)**.** Let A be a C^*-algebra. A *C^*-subalgebra* of A is a subalgebra B of A which is closed and which is invariant with respect to the involution $\cdot^*$.

Note that is C^*-subalgebra is again a C^*-algebra in its own right.

Proposition 8.4.2 (Functional calculus via the Gelfand representation)**.** *Let A be a unital C^*-algebra and let $a \in A$.*

(a) *The smallest closed subalgebra B of A that contains* 1*,* a*, and* a^* *is a C^*-subalgebra of A.*[3] *It is commutative if and only if a is normal.*

(b) *Let $f \in \mathrm{C}\left(\sigma(a)\right)$ and let B be a commutative C^*-algebra of A that contains 1 and a.*[4] *Then the element $f(a) := \Theta_B^{-1}\left(f \circ \Theta_B(a)\right)$ of $B \subseteq A$, where $\Theta_B : B \to \mathrm{C}\left(\Omega(B)\right)$ is the Gelfand representation of B, is well-defined and does not depend on the choice of B.*

For the proof we need the following auxiliary result:

Lemma 8.4.3 (The spectrum in ***C****-subalgebras)**.** *Let A be a unital C^*-algebra and let $B \subseteq A$ be a C^*-subalgebra that contains* 1*. For every $b \in B$ one has $\sigma_A(b) = \sigma_B(b)$ for the spectra of b within the Banach algebras A and B.*

Proof. „$\subseteq$“ This implication is clear (and is true is general Banach algebras rather than only C^*-algebras, see Proposition 3.3.3(a)).

[3] And clearly, it is the smallest C^*-subalgebra of A that contains 1 and a.
[4] Note that such a B only exists if a is normal.

„$\supseteq$" First assume that b is self-adjoint. Then both sets $\sigma_A(b)$ and $\sigma_B(b)$ are a subset of $\mathbb{R}$ according to Proposition 8.3.1 Thus, each of those two spectra coincides with its topological boundary within $\mathbb{C}$ and hence, the inclusion follows from Proposition 3.3.3(b).

Now we consider general elements $b \in B$. It suffices to show that if b is invertible in A, then it is also invertible in A. So assume that there exists $a \in A$ such that $ab = ba = 1$. Then we have $b^*a^* = a^*b^* = 1^* = 1$ and thus $bb^*a^*a = 1$. Thus, the self-adjoint element bb^* of B is invertible in A. As we have already treated the self-adjoint case separately above, we hence know that bb^* is invertible in B, i.e., there exists $c \in B$ such that $bb^*c = 1$. By multiplying with a from the left we get $b^*c = a$ and hence $a \in B$, which shows that b is invertible in B, as claimed. ☐

Proof of Proposition 8.4.2. (a) One readily checks that B is the closed linear span of all elements of the form

$$a^{j_1}(a^*)^{k_1} \dots a^{j_n}(a^*)^{k_n}$$

for integers $n \in \mathbb{N}_0$ and $j_1, \dots, j_n, k_1, \dots, k_n \in \mathbb{N}_0$. From this one can easily that B is a C^*-subalgebra of A. Moreover, this representation of B also show that B is commutativ if a is normal; the converse implication is obvious.

(b) *Well-definedness:* According to Theorem 3.2.5(b) the range of the function $\Theta_B(a) \in \mathrm{C}\big(\Omega(B)\big)$ is contained in $\sigma_B(a)$ and according to Lemma 8.4.3 this set coincides with $\sigma_A(a)$. Since $f \in \mathrm{C}\big(\sigma_A(a)\big)$ by assumption, the composition $f \circ \Theta_B(a)$ is well-defined and an element of $\mathrm{C}\big(\Omega(B)\big)$, so we can apply Θ_B^{-1} to it.

Independence of B: Due to the Stone–Weierstraß approximation theorem 1.1.8 f can, on $\sigma(a)$, be approximated with respect to the sup norm $\|\cdot\|_{\mathrm{C}(\sigma_A(a))}$ by a sequence of functions $q_n : \mathbb{C} \ni z \mapsto q_n(z) \in \mathbb{C}$ that are polynomials in z and z^*. So we have

$$\|f \circ \Theta_B(a) - q_n \circ \Theta_B(a)\|_{\mathrm{C}(\Omega(A))} = \|f - q_n\|_{\mathrm{C}(\sigma_A(a))} \to 0.$$

As Θ is isometric, so is Θ^{-1} and hence,

$$\Theta_B^{-1}\big(f \circ \Theta_B(a)\big) = \lim_{n\to\infty} \Theta_B^{-1}\big(q_n \circ \Theta_B(a)\big) = \lim_{n\to\infty} q_n\big(\Theta_B^{-1}\Theta_B(a)\big) = \lim_{n\to\infty} q_n(a).$$

But the latter term does not depend on B, which shows the claim. ☐

Definition 8.4.4 (The continuous functional calculus)**.** Let A be a unital C^*-algebra and let $a \in A$ be normal. The mapping

$$\begin{aligned} \mathrm{C}\big(\sigma(a)\big) &\to A \\ f &\mapsto f(a), \end{aligned}$$

where $f(a)$ is defined as in Proposition 8.4.2(b) for any commutative C^*-subalgebra B of A that contains both 1 and a,[5] is called the *continuous functional calculus of a*.

[5]Note that such a B exists according to Proposition 8.4.2(a).

Proposition 8.4.5 (The continuous functional calculus is an algebra homomorphism)**.** *Let A be a unital C^*-algebra and let $a \in A$ be normal. The continuous functional calculus*

$$\begin{aligned} \mathrm{C}\left(\sigma(a)\right) &\to A \\ f &\mapsto f(a), \end{aligned}$$

is an algebra homomorphism that maps $\mathbb{1}$ to 1 and that respects the involution $\cdot^*$.

Proof. This follows readily from the definition of the continuous functional calculus if one has that the Gelfand representation is an algebra isomorphism, sends 1 to $\mathbb{1}$, and respects the involution. □

Remark 8.4.6 (Approximation by polynomials)**.** Let A be a unital C^*-algebra, let $a \in A$ be normal, and let $f \in \mathrm{C}\left(\sigma(a)\right)$. Let (q_n) be a sequence functions $q_n : \mathbb{C} \ni z \mapsto q_n(z) \in \mathbb{C}$ that are polynomials in z and z^* $\|p_n - f\|_{\mathrm{C}(\sigma(a))} \to 0$.[6] Then the proof of Proposition 8.4.2(b) shows that $f(a) = \lim_{n\to\infty} q_n(a)$.

Remark 8.4.7 (The continuouos and the holomorphic functional calculus are consistent)**.** Let A be a unital C^*-algebra and let $a \in A$ be normal. Let $\Omega \subseteq \mathbb{C}$ be an open set that contains $\sigma(a)$ and let $f : \Omega \to \mathbb{C}$ be holomorphic. Then, clearly, $f|_{\sigma(a)}$ is in $\mathrm{C}\left(\sigma(a)\right)$. One has

$$f(a) = f|_{\sigma(a)}(a),$$

where the left hand side is understood in the sense of the holomorphic functional calculus and the right hand side is understood in the sense of the continuous functional calculus.

Indeed, this is clearly true if f is a polynomial. For general holomorphic $f : \Omega \to \mathbb{C}$ this can be shown by approximation by polynomials.

Example 8.4.8 (The modulus of elements of a $\boldsymbol{C^*}$-algebra)**.** Let A be a unital C^*-algebra and let $a \in A$. Then a^*a is self-adjoint and one can show that always $\sigma(a^*a) \subseteq [0, \infty)$, see for instance [Mur90, Theorem 2.2.4 on p. 46]. As the mapping $f : [0, \infty) \to [0, \infty) \subseteq \mathbb{C}$, $t \mapsto \sqrt{t}$, is continuous, it follows that the element

$$|a| := \sqrt{a^*a}$$

is well-defined by means of the continuous functional calculus.

[6] Recall again that such a sequence always exists due to the Stone–Weiterstraß approximation theorem.

Appendices

Appendix A

A Few Notions from Point Set Topology

In this appendix we give a very brief overview over some concepts from point set topology that we need in the course.[1] We leave out most of the proofs, though.

A.1 Topological spaces

Metric spaces are often used in analysis as a quite general framework where the concepts of continuity and convergence can be defined and studied. There are situations, though (for instance when studying the character space of a commutative Banach algebra, see Section 3.2), where metric space do not suffice to describe all interesting situations where convergence and continuity play a role.

Thus, one introduces an even more general theoretic framework, namely *topological spaces*.

Definition A.1.1 (Topologies, topological spaces and open sets)**.**

(a) Let X be a set. A *topology* on X is a subset τ of the power set 2^X of X with the following properties:[2]

  (I) We have $\emptyset \in \tau$ and $X \in \tau$.

  (II) The set τ is stable with respect to finite intersections, i.e.: for each $n \in \mathbb{N}$ and all $U_1, \dots, U_n \in \tau$ we have $U_1 \cap \dots \cap U_n \in \tau$.[3]

  (III) The set τ is stable with respect to arbitrary unions, i.e.: for each index set I and each family $(U_i)_{i \in I}$ such that $U_i \in \tau$ for each i, we have have $\bigcup_{i \in I} U_i \in \tau$.

[1] The appendix is mainly based on, and partially copied from, lecture notes that I wrote for a course in *Topology* at the University of Passau during the winter term 2020/21.

[2] Throughout the course, we use the notation 2^X for the *power set* of a set X, i.e., for the set of all subsets of X.

[3] Note that, by a simple induction argument, this property is equivalent to requiring that $U_1 \cap U_2 \in \tau$ for all $U_1, U_2 \in \tau$.

(b) A *topological space* is a pair (X, τ) such that X is a set and τ is a topology on X.

(c) Let (X, τ) be a topological space. A subset $U \subseteq X$ is called *open* iff $U \in \tau$.

(d) Let (X, τ) be a topological space. As subset $C \subseteq X$ is called *closed* if its complement $X \setminus C$ is open.

A few very simple examples are useful to get a first taste of how general the concept of a topological space is:

Examples A.1.2 (Discrete and indiscrete topology)**.** Let X be a set.

(a) The power set 2^X itself is a topology on X. It is called the *discrete topology* on X.

(b) The set $\{\emptyset, X\}$ is a topology on X. It is called the *indiscrete topology* on X.

Example A.1.3 (Metric spaces as topological spaces)**.** Let (M, d) be a metric space, and let $\tau \subseteq 2^M$ consist of all sets $U \subseteq M$ with the following property: for every $x \in U$ there exists a number $\varepsilon > 0$ such that the ball

$$\mathrm{B}_{<\varepsilon}(x) := \{y \in M \mid \mathrm{d}(y, x) < \varepsilon\}$$

is contained in U. Then τ is a topology on M, and its elements are precisely the usual open sets in M.

The indiscrete topology on a set X is an example of a topological that behaves quite weird when one is used to metric spaces. The following notion is often useful to rule out particularly strange kinds of behaviour.[4]

Definition A.1.4 (Topological Hausdorff spaces)**.** A topological space (X, τ) is called *Hausdorff* if the following holds: for any two distinct points $x_1, x_2 \in X$ there exist disjoint open sets U_1, U_2 such that $x_1 \in U_1$ and $x_2 \in U_2$.

For instance, if X has at least two points and τ is the indiscrete topology on X, then (X, τ) is not Hausdorff. If, on the other hand, τ is the discrete topology on a set X, then (X, τ) is Hausdorff. Moreover, every metric space is Hausdorff with respect to its usual topology.

[4]The words "particularly strange" in this sentence are written from an analytic perspective. In more algebraic topics, non-Hausdorff topologies appear quite often.

A.2 Continuity and convergence

As indicated at the beginning of the previous subsection, topological spaces mainly become interesting through their interaction with two other concepts: continuity and convergence.[5]

From now on we often use the following (slightly imprecise, but very convenient and very common) convention: when talking about a topological space (X, τ) we supress τ in the notation, and thus only say that "X is a topological space". Of course, the topology τ still needs to be there, we just do not always mention it explicitly.

Definition A.2.1 (Continuous mappings)**.** Let (X, τ_X) and (Y, τ_Y) be topological spaces. A mapping $f : X \to X$ is called *continuous* if $f^{-1}(V)$ is open in X for every open set $V \subseteq Y$.

Definition A.2.2 (Convergence in topological spaces)**.** Let X be a topological space, let $x \in X$ and let $(x_j)_{j \in J}$ be a net in X. We say that $(x_j)_{j\in J}$ *converges to* x if the following holds: for every open set U that contains x there exists an index $j_0 \in J$ such that $x_j \in U$ for all $j \succeq j_0$.

We sometimes write $x_j \to x$ to say that $(x_j)_{j\in J}$ converges to x, and we call x a *limit* of $(x_j)_{j\in J}$ if $(x_j)_{j\in J}$ converges to x.

In topological spaces, limits need not be unique, in general. For instance, if we endow a set X with the indiscrete topology, then every net in X converges to every point in X. In fact, uniqueness of limits is characterized by the Hausdorff property that we introduced before:

Theorem A.2.3 (Limits are unique iff the space is Hausdorf)**.** *Let X be a topological space. The following are equivalent:*

(i) *The space X is Hausdorff.*

(ii) *Every net in X converges to at most one point in X.*

Nets are, in a sense, the appropriate generalizations of sequences to the setting of topological spaces. One can, for instance, use them to characterize closedness of sets and continuity of functions:

Proposition A.2.4 (Closedness and continuity via nets)**.** *Let X, Y be topological spaces.*

(a) *A subset $C \subseteq X$ is closed if and only if the following holds: whenever a net $(x_j)_{j\in J}$ in C converges to a point $x \in X$, then $x \in C$.*

(b) *A mapping $f : X \to Y$ is continuous if and only if the following holds: whenever a net $(x_j)_{j\in J}$ in X converges to a point $x \in X$, then the net $\big(f(x_j)\big)_{j\in J}$ in Y converges to the point $f(x)$.*

[5] From an algebraic viewpoint, one might argue that it is really continuity which is the important concept. From a functional analytic perspective, though, convergence also takes an important role.

A.3 Compactness and universal nets

A very important and useful concept in topology is *compactness* of topological spaces; this notion is defined as follows:

Definition A.3.1 (Compactness spaces)**.** A subset S of a topological space X is called *compact* if the following holds: for every family $(U_j)_{j\in J}$ of open sets that satisfies $\bigcup_{j\in J} U_j \supseteq S$, there exists a finite subset $F \subseteq J$ such that $\bigcup_{j\in F} U_j \supseteq S$.[6]

It follows right from the definition of compact sets and continuous maps that the image of a compact set under a continuous map is again compact.

One very useful characterization of compact sets is via so-called *universal nets*.

Definition A.3.2 (Universial nets)**.** Let X be a set and let $(x_j)_{j\in J}$ be a net in X.

(a) Let $S \subseteq X$. We say that the net $(x_j)_{j\in J}$ is *eventually in* S if there exists an index $j_0 \in J$ such that $x_j \in S$ for all $j \geq j_0$.

We say that the net $(x_j)_{j\in J}$ is *eventually constant* if there exists an index $j_0 \in J$ and a point $x \in X$ such that $x_j = x$ for all $j \succeq j_0$ (in other words, if there exists $x \in X$ such that the net is eventually in $\{x\}$).

(b) The net $(x_j)_{j\in J}$ is called *universal* if, for every set $S \subseteq X$, the net is eventually in S or eventually in $X \setminus S$.

For instance, every constant net and, more generally, every eventually constant net is universal. Universal nets that are not eventually constant are very non-concrete and somewhat weird objects,[7] and it is not obvious that such objects even exist. However, one can use Zorn's lemma to show the existence of so-called *ultrafilters*, and from this one can derive that universal nets exist in abundance.

It is not difficult to see that whether a net is universal does not depend on the surrounding set – i.e., if $(x_j)_{j\in J}$ is a net in a set X and $S \subseteq X$ contains this net, then $(x_j)_{j\in J}$ is universal in X if and only if it is universal in S. On a related note, the following is easy to show:

Proposition A.3.3 (Images of universal nets are universal)**.** *Let X, Y be sets and let $f : X \to Y$ be a mapping. If $(x_j)_{j\in J}$ is a universal net in X, then the net $(f(x_j))_{j\in J}$ in Y is universal, too.*

The following theorem is one reason why universal nets are very useful:

Theorem A.3.4 (Compactness via universal nets)**.** *Let C be a subset of a topological space X. The following are equivalent:*

(i) *The set C is compact.*

[6] In other words: Every open cover of S has a finite subcover.

[7] For instance, one can check that a sequence that is not eventually constant, cannot be a universal net.

(ii) *Every universal net in C converges to a point in C.*[8]

A.4 Initial topologies

It is a general principle when building mathematical theories that one would like to have tools to construct a new mathematical object from given ones. The following result gives an example of how to construct a new topology from given topological spaces and mappings.

Theorem A.4.1 (Initial topology)**.** *Let X be a set, let $(X_h)_{h\in H}$ be a family of topological spaces, and let $f_h : X \to X_h$ be a mapping for each $h \in H$.*

Among all topologies on X which make all the functions f_h continuous[9] *there exists a smallest*[10] *one; it is called the* initial topology *of $(f_h)_{h\in H}$.*[11]

The initial topology has the following properties:

(a) *A net $(x_j)_{j\in J}$ converges to a point $x \in X$ if and only if, for each $h \in H$, the net $(f_h(x_j))_{j\in J}$ in X_h converges to $f_h(x)$.*

(b) *A map $g : W \to X$ from a topological space W to X is continuous if and only if the map $f_h \circ g : W \to X_h$ is continuous for every $h \in H$.*

(c) *If X_h is Hausdorff for each $h \in H$ and if for every every pair of distinct points $x, y \in X$ there exists $h \in H$ such that $f_h(x) \neq f_h(y)$, then X is Hausdorff, too.*

Note that assertion (c) readily follows from the definitions of the Hausdorff property and contintuity of the maps f_h.

The following examples of initial topologies are quite prominent in function analysis:

Examples A.4.2 (Weak and weak* topologies on Banach spaces)**.** Let X be a Banach space.

(a) The initial topology on X of the family $(x')_{x'\in X'}$ is called the *weak topology* on X.

It follows from Theorem A.4.1(a) that a net $(x_j)_{j\in J}$ in X converges to a point $x \in X$ with respect to the weak topology if and only if $\langle x', x_j\rangle \to \langle x', x\rangle$ for all $x' \in X'$.

[8] Note however that this assertion does not say that all limits of a universal net in C are also in C; in fact, compact subsets of topological spaces need to be closed in general. However, if X is Hausdorff, then one can check that every compact subset of X is indeed closed.

[9] Why does there exists any such topology on X, after all?

[10] With respect to set inclusion between topologies.

[11] Note that this is slightly imprice, since the initial topology does not only depend the mappings f_h, but also on the topologies on the X_h.

More, the weak topology is Hausdorff due to Theorem A.4.1(c) and the Hahn–Banach theorem.[12]

(b) For each element $x \in X$ consider the element $\hat{x}$ in the bi-dual space X'' that is given by $\langle \hat{x}, x' \rangle := \langle x', x \rangle$ for all $x' \in X'$. Then the initial topology on X' of the family $(\hat{x})_{x \in X}$ is called the *weak* topology* on X'.

It follows from Theorem A.4.1(a) that a net $(x'_j)_{j \in J}$ in X' converges to a point $x' \in X'$ with respect to the weak* topology if and only if $\langle x'_j, x \rangle \to \langle x', x \rangle$ for all $x \in X$.

The weak* topology is Hausdorff due to Theorem A.4.1(c).

Let us now illiustrate the usefulness of universal nets by the following proof of the following important theorem about the weak*-topology:

Theorem A.4.3 (Banach–Alaoglu: The dual unit ball is weak* compact)**.** *Let X be a Banach space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Then the unit ball* $\mathrm{B}_{\leq 1}(0)$ *in X' is compact with respect to the weak* topology.*

Proof. Let $(x'_j)_{j \in J}$ be a universal net in X' such that $\|x'_j\| \leq 1$ for all $j \in J$. For every $x \in X$ the net $(\langle x'_j, x \rangle)_{j \in J}$ is a universal net in the closed – and thus compact – disk with radius $\|x\|$ in $\mathbb{K}$. Thus, it follows from Theorem A.3.4 that this net converges to a number $\alpha_x \in \mathbb{K}$ of modulus $|\alpha_x| \leq \|x\|$.

Consider the mapping $x' : X \to \mathbb{K}$, $x \mapsto \alpha_x$. One can readily check that x' is linear; moreover, it follows from the estimate $|\alpha_x| \leq \|x\|$ for each $x \in X$ that x' is continuous – i.e., $x' \in X'$ – and $\|x'\| \leq 1$. Hence, x' is an element of the unit ball $\mathrm{B}_{\leq 1}(0)$ in X'. Moreover, the net $(x'_j)_{j \in J}$ converges weak* to x' due to the characterization of weak* convergence in Example A.4.2(b).

So every universal net in the ball $\mathrm{B}_{\leq 1}(0)$ in X' converges with respect to the weak* topology to a point in the ball. According to Theorem A.3.4 this proves compactenss of $\mathrm{B}_{\leq 1}(0)$. ☐

Another important instance of an initial topology is the *product topology*, which is a topology on a cartesian product that is constructed from topologies on its components:

Example A.4.4 (Product topology)**.** Let H be a non-empty set, and for each $h \in H$ let X_h be a topological space. We set $X := \prod_{h \in H} X_h$, and for each $h_0 \in H$ we consider the h_0-th coordinate mapping

$$\begin{aligned} p_{h_0} &: X \to X_{h_0} \\ x = (x_h)_{h \in H_0} &\mapsto x_{h_0}. \end{aligned}$$

The initial topology of the family $(p_h)_{h \in H}$ is called the *product topology* of the family of spaces $(X_h)_{h \in H}$.

[12]More precisely, the version of the Hahn–Banach theorem which says that for every $x \in X$ there exists $x' \in X'$ such that $\langle x', x \rangle \neq 0$.

The following famous result is surprisingly easy to prove by means of universal nets:

Theorem A.4.5 (Tychonoff's theorem on the compactness of product spaces)**.** *Let H be a non-empty set, and for each $h \in H$ let X_h be a compact topological space. Then the space $X := \prod_{h\in H} X_h$, endowed with the product topology, is compact, too.*

Proof. Let $(x_j)_{j\in J} = \Big((x_j^{(h)})_{h\in H}\Big)_{j\in J}$ be a universal net in X. Then for each $h \in H$ the net $(x_j^{(h)})_{j\in J}$ in X_h is universal; this follows by applying Proposition A.3.3 to the h-th coordinate map $X \to X_h$.

As X_h as compact, Theorem A.3.4 shows that this net converges to an element $y^{(h)} \in X_h$. Hence, the net $(x_j)_{j\in J}$ converges coordinatewise – and thus with respect to the product topology – to the point $(y^{(h)})_{h\in H}$ in X. So we showed that every universal net in X converges, which implies compactness of X, again according to Theorem A.3.4. □

We note in passing that it is not difficult to derive the Banach–Alaoglu theorem A.4.3 from Tychonoff's theorem A.4.5. When one uses the technology of universal nets, though, it seems a bit more natural to derive both results directly from the characterization of compact sets via univeral nets, as we have done above.

Appendix B

Some Concepts from Functional Analysis

In this appendix we discuss a few concepts from functional analysis that are used in the course. Some of the material might not be covered in introductory text books on functional analysis.

B.1 Operator ranges

Definition B.1.1 (Operator range)**.** Let X be a real or complex Banach space. An *operator range* in X is a vector subspace U of X such that there exists a Banach space W (over the same field as X) and a bounded linear operator $T : W \to X$ with the property $TW = U$.

Proposition B.1.2 (Characterisation of operator ranges)**.** *Let U be a vector subspace of a Banach space X. The following are equivalent:*

(i) *The subspace U is an operator range in X.*

(ii) *There exists a norm $\|\cdot\|_U$ on U that turns U into a Banach space and that makes the embedding $(U, \|\cdot\|_U) \hookrightarrow (X, \|\cdot\|_X)$ continuous.*

Proof. „(i) $\Rightarrow$ (ii)“ Let W be a Banach space and $T \in \mathcal{L}(W; X)$ with range V. Then T induces an injective operator $\tilde{T} \in \mathcal{L}(W/\ker T; X)$ with the same range. By replacing W with $W/\ker T$ and T with $\tilde{T}$ we may assume that T is injective, and thus bijective from W to V.

By transporting the norm on W to V via T we thus obtain a complete norm $\|\cdot\|_V$ on V which makes the embedding of V into X continuous.

„(ii) $\Rightarrow$ (i)“ The embedding $U \hookrightarrow X$ is a continuous linear operator between the Banach spaces $(U, \|\cdot\|_U)$ and $(X, \|\cdot\|_X)$ and has range U, so U is an operator range. $\square$

Clearly, every closed vector subspace of a Banach space is an operator range. We discuss further examples of operator ranges in Examples B.1.5 below.

The following result is very useful in the theory of (lower semi-)Fredholm operators, see Remark 6.1.5.

Proposition B.1.3 (Direct sums of operator ranges)**.** *Let X be a Banach spaces and, for some $n \in \mathbb{N}$, let $V_1, \dots, V_n$ be operator ranges in X such that X is the algebraically direct sum of them – meaning that every $x \in X$ be be written as $x = v_1 + \dots + v_n$ for uniquely determined vectors $v_1 \in V_1, \dots, v_n \in V_n$.*

The each of the operator ranges $V_1, \dots, V_n$ is closed.

Proof. According to Proposition B.1.2 we can, for each $k \in \{1, \dots, n\}$, endow V_k with a complete norm $\|\cdot\|_{V_k}$ that makes the embedding of V_k into X continuous.

The space $V := V_1 \times \dots \times V_n$ (with componentwise addition and scalar multiplication) is a Banach space when endowed with the norm $\|\cdot\|_V$ given by

$$\|(v_1, \dots, v_n)\|_V := \|v_1\|_{V_1} + \dots + \|v_n\|_{V_n}$$

for all $(v_1, \dots, v_n) \in V$. Moreover, the mapping

$$\begin{aligned} \Phi : \qquad\qquad V &\to X, \\ (v_1, \dots, v_n) &\mapsto v_1 + \dots + v_n \end{aligned}$$

is bijective as X is the direct sum of the subspaces $V_1, \dots, V_n$, and it is continuous since the spaces $(V_k, \|\cdot\|_{V_k})$ embed continuously into $(X, \|\cdot\|_X)$.

It thus follows from the continuous inverse theorem that the inverse mapping $\Phi^{-1} : X \to V$ is continuous, too. This implies that, for each $k \in \{1, \dots, n\}$, the norms $\|\cdot\|_X$ and $\|\cdot\|_{V_k}$ are equivalent on the space V_k, and thus, V_k is also complete with respect to the norm $\|\cdot\|_X$. Hence, V_k is closed in $(X, \|\cdot\|_X)$, as claimed. □

Remark B.1.4 (Uniqueness of the norm)**.** Let V be an operator range in a Banach space X. All complete norms on V that make the embedding of V into X continuous, are equivalent. This follows from the closed graph theorem.

Examples B.1.5 (Simple examples of operator ranges)**.**

(a) Every closed vector subspace of a Banach space X is an operator range in X; so in particular, every finite-dimension vector subspace of X is an operator range in X.

(b) Let (Ω, μ) be a finite measure space, and let $1 \le p \le q \le \infty$. Then $L^q(\Omega, \mu)$ is an operator range in $L^p(\Omega, \mu)$.

(c) Let $1 \le p \le q \le \infty$. Then ℓ^p is an operator range in ℓ^q.

(d) For every $p \in [1, \infty]$ the space $C([0, 1])$ is an operator range in $L^p([0, 1])$.

In a certain sense, one can think of operator spaces as a generalization of closed subspaces. For instance, they have the following stability properties. We omit the proof, but it is a nice exercise if you are interested to learn more about operator ranges.

Proposition B.1.6 (Stability properties of operator ranges)**.** *Let X, Y be Banach spaces over the same field.*

(a) *Let $T \in \mathcal{L}(X;Y)$ and let $U \subseteq X$ and $V \subseteq Y$ be operator ranges. Then TU and $T^{-1}V$ are operator ranges in Y and X, respectively.*

(b) *If $V_1, \dots, V_n \subseteq X$ are operator ranges for some $n \in \mathbb{N}$, then $V_1 + \dots + V_n$ and $V_1 \cap \dots \cap V_n$ are operator ranges in X, too.*

Let us note that the intersection of infinitely many operator ranges is not an operator range, in general;[1] this is an important distinction between operator ranges and closed subspaces.

On the other hand, a result for closed subspaces that remains true for operator ranges, is the following consequence of Baire's theorem:

Theorem B.1.7 (Countable unions of operator ranges)**.** *Let X be a Banach space and for every $n \in \mathbb{N}$, let V_n be an operator range in X. If $X = \bigcup_{n \in \mathbb{N}} V_n$, then there exists $n_0 \in \mathbb{N}$ such that $X = V_{n_0}$.*

Proof. It follows from that assumption $X = \bigcup_{n \in \mathbb{N}} V_n$ and from Baire's theorem that at least one of the subspaces V_n, say V_{n_0}, is not meagre in X. Hence, by a version of the open mapping theorem – see for instance [Rud91, Theorem 2.11 on p. 48] – we have $V_{n_0} = X$. $\square$

As a funny consequence of the previous theorem one can, e.g., give a non-constructive proof of the existence of a function $f \in L^1\big([0,1]\big)$ which is not in $L^p\big([0,1]\big)$ for any $p > 1$. (But on it is not difficult either to construct an explicit example of such a function, without appealing to Theorem B.1.7.)

[1] Can you find a counterexample?

Bibliography

[ABHN11] Wolfgang Arendt, Charles J. K. Batty, Matthias Hieber, and Frank Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monogr. Math., Basel.* Basel: Birkhäuser, 2nd ed. edition, 2011.

[Bou73] Richard Bouldin. The product of operators with closed range. *Tohoku Math. J. (2)*, 25:359–363, 1973.

[Con73] John B. Conway. *Functions of one complex variable*, volume 11 of *Grad. Texts Math.* Springer, Cham, 1973.

[Con95] John B. Conway. *Functions of one complex variable. II*, volume 159 of *Grad. Texts Math.* New York, NY: Springer-Verlag, 1995.

[EN00] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Grad. Texts Math.* Berlin: Springer, 2000.

[GGK90] Israel Gohberg, Seymour Goldberg, and Marinus A. Kaashoek. *Classes of linear operators. Vol. I*, volume 49 of *Oper. Theory: Adv. Appl.* Basel etc.: Birkhäuser Verlag, 1990.

[Glü16] Jochen Glück. *Invariant Sets and Long Time Behaviour of Operator Semigroups.* PhD thesis, Universität Ulm, 2016. DOI: 10.18725/OPARU-4238.

[Haa06] Markus Haase. *The functional calculus for sectorial operators*, volume 169 of *Operator Theory: Advances and Applications.* Birkhäuser Verlag, Basel, 2006.

[HP57] Einar Hille and Ralph S. Phillips. *Functional analysis and semi-groups*, volume 31 of *Colloq. Publ., Am. Math. Soc.* American Mathematical Society (AMS), Providence, RI, 1957.

[Kab14] Winfried Kaballo. *Aufbaukurs Funktionalanalysis und Operatortheorie.* Springer Spektrum, Berlin, 2014. Distributionen—lokalkonvexe Methoden—Spektraltheorie [Distributions, locally convex methods, spectral theory].

[Kat60] Tosio Kato. Estimation of iterated matrices, with application to the von Neumann condition. *Numer. Math.*, 2:22–29, 1960.

[MN91] Peter Meyer-Nieberg. *Banach lattices.* Universitext. Berlin etc.: Springer-Verlag, 1991.

[Mur90] Gerard J. Murphy. *C^*-algebras and operator theory.* Academic Press, Inc., Boston, MA, 1990.

[Ped89] Gert K. Pedersen. *Analysis now*, volume 118 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989.

[Rud91] Walter Rudin. *Functional analysis.* New York, NY: McGraw-Hill, 2nd ed. edition, 1991.

[TL80] Angus Ellis Taylor and David C. Lay. Introduction to functional analysis. 2nd ed. New York - Chichester - Brisbane: John Wiley & Sons., 1980.

[Yos80] Kosaku Yosida. *Functional analysis. 6th ed*, volume 123 of *Grundlehren Math. Wiss.* Springer, Cham, 1980.