25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

November 13, 2021

Lecture 1

A survey on operator theory I

Along the first lecture of this Internet Seminar, we are going to consider usually a complex Banach space X, a linear operator T with domain $D(T)$ and range $Rg(T)$, both contained in X, and we will study the operator $\lambda - T$, where λ is a complex number.

If λ is such that $\lambda - T : D(T) \to X$ is bijective and $\lambda - T$ has continuous inverse, then λ is said to belong to the resolvent set $\rho(T)$ of T, otherwise λ belongs to the spectrum $\sigma(T)$ of T.

The main aim of Spectral Theory is the systematic analysis of the properties of T and $(\lambda - T)^{-1}$, through the study of the sets $\rho(T)$ and $\sigma(T)$.

If X is a finite-dimensional space, then we can represent T as a matrix, and in that case spectral analysis reduces to the study of the eigenvalues of the matrix, since if $\lambda - T$ does not have an inverse, then there exists $x \in X$, with $x \neq 0$ such that $Tx = \lambda x$. It is well known that the spectrum of a matrix contains at least one complex eigenvalue.

Dealing with the infinite dimensional case, the picture gets more diversified, since, e.g., the spectrum could be empty and spectral values need not to be eigenvalues.

This week we start by providing the basic definitions and tools to deal with operators on Banach spaces and their spectra.

Notation: Given a Banach space $(X, \|\cdot\|)$ over the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, we denote by $B_X := \{\, x \in X \mid \|x\| \leq 1\}$, the closed unit ball of X and by $I : X \to X$ the identity operator.

1.1 Main definitions and properties of bounded operators on a Banach space

If X and Y are Banach spaces, we denote by $\mathcal{L}(X, Y)$ the space of linear and continuous operators $T : X \to Y$. We recall the following result.

Proposition 1.1.1. *If X and Y are Banach spaces and $T : X \to Y$ is a linear operator, then the following properties are equivalent:*

(i) *T is continuous on X;*

(ii) *T is continuous at $x = 0$;*

(iii) *T is bounded (i.e., $T(B_X)$ is a bounded set in Y).*

Consequently, for every $T \in \mathcal{L}(X,Y)$ the following definition of *operator norm* of T is meaningful

$$\|T\| := \sup_{x \in B_X} \|Tx\|_Y.$$

Clearly $\|Tx\|_Y \le \|T\|\|x\|_X$ for every $x \in X$. Moreover, the space $(\mathcal{L}(X,Y), \|\cdot\|)$ is a Banach space.

If $X = Y$, then we set $\mathcal{L}(X) := \mathcal{L}(X,X)$. If $Y = \mathbb{K}$, then the Banach space $\mathcal{L}(X,\mathbb{K})$ is called *topological dual space* of X and is denoted by X'. In this case, the operator norm is denoted by $\|\cdot\|'_{X'}$ and simply by $\|\cdot\|'$ when no confusion may arise.

If Z is another Banach space, $S \in \mathcal{L}(Y,Z)$ and $T \in \mathcal{L}(X,Y)$, then it is easy to prove that the composition operator $ST := S \circ T$ belongs to $\mathcal{L}(X,Z)$ and

$$\|ST\| \le \|S\| \cdot \|T\| \ . \tag{1.1}$$

If $T \in \mathcal{L}(X,Y)$, then the dual operator $T' : Y' \to X'$ is defined by

$$(T'y')(x) := y'(Tx), \qquad \forall y' \in Y', \ \forall x \in X.$$

The operator T' belongs to $\mathcal{L}(Y',X')$. Indeed, it is straightforward to prove that T' is linear. Moreover, for any fixed $y' \in Y'$, it holds that

$$|(T'y')(x)| = |y'(Tx)| \le \|y'\|'_{Y'} \cdot \|Tx\|_Y \le \|y'\|'_{Y'} \cdot \|T\| \cdot \|x\|_X, \qquad \forall x \in X.$$

Hence $T'y' \in X'$ and

$$\|T'y'\|'_{X'} = \sup_{x \in B_X} |T'y'(x)| \le \|T\|\|y'\|'_{Y'}.$$

By the arbitrariness of y', we get that $T' \in \mathcal{L}(Y',X')$ and the inequality

$$\|T'\| = \sup_{y' \in B_{Y'}} \|T'y'\|'_{X'} \le \|T\| \tag{1.2}$$

holds true. Actually, it holds that

$$\|T\| = \|T'\|. \tag{1.3}$$

Indeed, fix $x_0 \in X$. By the Hahn–Banach-theorem, see Corollary A.1.2, there exists $y'_0 \in Y'$ such that $\|y'_0\|'_{Y'} = 1$ and $y'_0(Tx_0) = \|Tx_0\|_Y$. Consequently, $T'y'_0(x_0) = \|Tx_0\|_Y$, so that

$$\|Tx_0\|_Y = (T'y'_0)(x_0) \le \|T'\| \cdot \|y'_0\|'_{Y'} \cdot \|x_0\|_X = \|T'\| \cdot \|x_0\|_X.$$

It follows that $\|T\| \le \|T'\|$.

If M is a subspace of X and N is a subspace of X', set

$$M^\perp := \{\, y' \in X' \mid y'(x) = 0 \text{ for every } x \in M \,\},$$
$${}^\perp N := \{\, x \in X \mid y'(x) = 0 \text{ for every } y' \in N \,\}.$$

For every $T \in \mathcal{L}(X)$, it holds that

$$\ker(T') = T(X)^\perp, \qquad \ker(T) = {}^\perp T'(X'). \tag{1.4}$$

(See Exercise 1.4.1)

1.2 Closed operators

Definition 1.2.1. A linear operator $T : D(T) \to Y$, where $D(T)$ is a subspace of X, is said to be

(i) *densely defined* if $D(T)$ is dense in X;

(ii) *closed* if its graph $\mathcal{G}(T) := \{ (x, Tx) \mid x \in D(T) \}$ is a closed subspace of the product Banach space $X \times Y$ endowed with the norm $\|(x, y)\| := \|x\|_X + \|y\|_Y$ for $(x, y) \in X \times Y$.

Since X, Y are Banach spaces, from the above definition the following result follows.

Proposition 1.2.2. *Let $T : D(T) \subseteq X \to Y$ be a linear operator. Then, the following properties are equivalent:*

(i) *T is a closed operator;*

(ii) *for every sequence $(x_n)_{n\in\mathbb{N}} \subset D(T)$ such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} Tx_n = y$ for some $(x, y) \in X \times Y$, it holds that $x \in D(T)$ and $y = Tx$;*

(iii) *the space $D(T)$, endowed with the graph norm*

$$\|x\|_T = \|x\| + \|Tx\| \,, \qquad x \in D(T),$$

is a Banach space.

Clearly if $T \in \mathcal{L}(X, Y)$, then it is closed. If the domain of T is the whole space X, then the converse holds by the Closed Graph Theorem.

Theorem 1.2.3 (Closed Graph Theorem)**.** *Let X and Y be two Banach spaces and let $T : X \to Y$ be a closed operator. Then, T is continuous.*

Proof. A linear operator T is always continuous with respect to its graph norm $\|\cdot\|_T$ on X. Since, by Proposition 1.2.2, $\|\cdot\|_X \le \|\cdot\|_T$ and $(X, \|\cdot\|_X)$, $(X, \|\cdot\|_T)$ are both Banach spaces, it follows that the norms $\|\cdot\|_X$ and $\|\cdot\|_T$ are equivalent (see Exercise 1.4.3). Thus, T is continuous with respect to the original norm. □

Given two linear operators $S : D(S) \subseteq X \to Y$ and $T : D(T) \subseteq X \to Y$, T is said to be an extension of S, and we write in this case $S \subset T$, if $D(S) \subseteq D(T)$ and $Sx = Tx$ for every $x \in D(S)$.

Definition 1.2.4. An operator $T : D(T) \subset X \to Y$ is said to be closable if it admits a closed extension.

A straightforward reasoning gives the next result, see Exercise 1.4.2.

Proposition 1.2.5. *Let X be a Banach space and let $T : D(T) \subset X \to Y$ be a linear operator. Then the following properties hold:*

(i) *the operator T is closable if and only if for any sequence $(x_n)_{n\in\mathbb{N}} \subset D(T)$ such that $\lim_{n\to\infty} x_n = 0$ and $\lim_{n\to\infty} Tx_n = y$, it holds $y = 0$;*

(ii) *if T is a closable operator, then there exists the smallest closed extension of T. It is denoted by $\overline{T}$ and it is called closure of T. Moreover, it holds that*

$$\overline{\mathcal{G}(T)} = \mathcal{G}(\overline{T}).$$

Example 1.2.6. Let $T : C^1([0,1]) \subset L^2((0,1)) \to \mathbb{C}$ be defined by $Tf := f'(0)$. Then, T is not closable. Indeed, consider the sequence $(f_n)_{n\in\mathbb{N}} \subset C^1([0,1])$ such that $f_n(x) = \frac{1}{n}\sin(nx)$ for every $x \in [0,1]$. Then, $\lim_{n\to\infty} f_n = 0$ in $L^2((0,1))$, but $Tf_n = 1$ for every $n \in \mathbb{N}$, so that $\lim_{n\to\infty} Tf_n \neq 0$.

1.3 Spectrum and resolvent

Let $(X, \|\cdot\|)$ be a Banach space over $\mathbb{C}$. In the sequel, if $T : D(T) \subseteq X \to X$ is a linear operator on X and $\lambda \in \mathbb{C}$, then $\lambda \pm T$ stands for the operator $\lambda I \pm T$, with domain $D(T)$.

Definition 1.3.1. Let $T : D(T) \subseteq X \to X$ be a linear operator. The set

$$\rho(T) := \{\lambda \in \mathbb{C} : \lambda - T : D(T) \to X \text{ is bijective and } (\lambda - T)^{-1} \in \mathcal{L}(X)\}$$

is called resolvent set of T. If $\rho(T) \neq \varnothing$ and $\lambda \in \rho(T)$, then the operator

$$R(\lambda, T) := (\lambda - T)^{-1}$$

is said resolvent of T at λ. The set $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is called spectrum of T. Moreover, the set $\sigma_p(T) := \{\lambda \in \sigma(T) \mid \ker(\lambda - T) \neq \{0\}\}$ is called point spectrum of T. Every scalar $\lambda \in \sigma_p(T)$ is said eigenvalue of T and every $0 \neq x \in X$ such that $(\lambda - T)x = 0$ is said eigenvector of T corresponding to the eigenvalue λ.

Proposition 1.3.2. *If $T : D(T) \subseteq X \to X$ is a closed operator and $\lambda - T$ is bijective, then $(\lambda - T)^{-1} \in \mathcal{L}(X)$.*

Proof. By the assumption $\mathcal{G}(\lambda - T)$ is a closed subspace of $X \times X$, hence

$$\mathcal{G}((\lambda - T)^{-1}) = \{ (\lambda x - Tx, x) \mid x \in D(T) \}$$

is closed too. The Closed Graph Theorem 1.2.3 yields the assertion. □

Remark 1.3.3. If $T : D(T) \subseteq X \to X$ is a linear operator such that $\rho(T)$ is not empty, then T is closed. Indeed, for any sequences $(x_n) \subset D(T)$ such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} Tx_n = y$ for some x, $y \in X$, we have

$$\begin{aligned} R(\lambda, T)y &= \lim_{n\to\infty} R(\lambda, T)Tx_n \\ &= \lim_{n\to\infty} (\lambda R(\lambda, T)x_n - x_n) \\ &= \lambda R(\lambda, T)x - x. \end{aligned}$$

Thus, $x \in D(T)$ and $Tx = y$, which proves the closedness of T.

Lemma 1.3.4. *Let*

$$\mathcal{I}(X) := \{ S \in \mathcal{L}(X) \mid S \ \ \textit{is bijective} \}.$$

Then, the following properties hold:

(1) *if* $S \in \mathcal{L}(X)$, *with* $\|S\| < 1$, *then* $I - S \in \mathcal{I}(X)$ *and its inverse is given by the so-called Neumann-series*

$$(I - S)^{-1} = \sum_{n=0}^{\infty} S^n;$$

(2) *if* $\rho \in \mathbb{C}$, $S \in \mathcal{L}(X)$, *with* $\|S\| < |\rho|$, *then* $\rho \pm S \in \mathcal{I}(X)$;

(3) *if* $T \in \mathcal{I}(X)$ *and* $S \in \mathcal{L}(X)$, *then* $T + S \in \mathcal{I}(X)$ *if and only if* $I + T^{-1}S \in \mathcal{I}(X)$;

(4) $\mathcal{I}(X)$ *is an open subset of* $\mathcal{L}(X)$.

Proof. (1) Consider the series

$$\sum_{n=0}^{\infty} S^n \tag{1.5}$$

in the Banach space $\mathcal{L}(X)$.
Since $\|S^n\| \le \|S\|^n$ and $\|S\| < 1$, the series $\sum_{n=0}^{\infty} \|S^n\|$ converges and therefore the series in (1.5) converges in $\mathcal{L}(X)$. Now, observe that

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

and, analogously,

$$\left(\sum_{n=0}^{\infty} S^n \right) (I - S) = I.$$

The previous identities ensure that $I - S$ is invertible and

$$(I - S)^{-1} = \sum_{n=0}^{\infty} S^n.$$

(2) Since $\|\rho^{-1} S\| < 1$, by property (1) we get that $I \pm \rho^{-1} S \in \mathcal{I}(X)$ and so $\rho \pm S = \rho(I \pm \rho^{-1} S) \in \mathcal{I}(X)$.

(3) Since T is invertible, it follows that $T + S = T(I + T^{-1}S)$. Hence, $T + S \in \mathcal{I}(X)$ if and only if $I + T^{-1}S \in \mathcal{I}(X)$.

(4) Let $T \in \mathcal{I}(X)$ and $S \in \mathcal{L}(X)$ be such that $\|S\| < \|T^{-1}\|^{-1}$. Then,

$$\|T^{-1}S\| \le \|T^{-1}\| \|S\| < 1.$$

Combining properties (1) and (3), we get that $I + T^{-1}S \in \mathcal{I}(X)$, and therefore $T + S \in \mathcal{I}(X)$. This means that the open ball with center T and radius $\|T^{-1}\|^{-1}$ is contained in $\mathcal{I}(X)$. By the arbitrariness of $T \in \mathcal{I}(X)$, it follows that $\mathcal{I}(X)$ is an open subset of $\mathcal{L}(X)$. □

Proposition 1.3.5. *Let* $T : D(T) \subseteq X \to X$ *be a linear operator, Then, the following properties hold.*

(1) *Let* $\lambda_0 \in \rho(T)$. *If* $\lambda \in \mathbb{C}$ *and* $|\lambda - \lambda_0| < \|R(\lambda_0, T)\|^{-1}$, *then* $\lambda \in \rho(T)$ *and*

$$R(\lambda, T) = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0, T)^{n+1}, \tag{1.6}$$

where the series in (1.6) *converges in the operator norm.*

(2) $\rho(T)$ *is open.*

(3) *If* $\rho(T) \neq \varnothing$, *then the mapping* $R(\cdot, T) : \rho(T) \to \mathcal{L}(X)$ *is analytic.*

(4) (Resolvent identity) *For every* $\lambda, \mu \in \rho(T)$

$$R(\lambda, T) - R(\mu, T) = (\mu - \lambda) R(\lambda, T) R(\mu, T).$$

In particular, $R(\lambda, T)$ *and* $R(\mu, T)$ *commute.*

(5) *Let* $(\lambda_n)_{n \in \mathbb{N}} \subseteq \rho(T)$ *be a sequence converging to some* $\lambda_0 \in \mathbb{C}$. *Then,* $\lambda_0 \in \sigma(T)$ *if and only if* $\lim_{n \to \infty} \|R(\lambda_n, T)\| = \infty$.

Proof. (1) Observe that

$$\lambda - T = \lambda - \lambda_0 + \lambda_0 - T = [I + (\lambda - \lambda_0) R(\lambda_0, T)](\lambda_0 - T), \qquad \lambda \in \mathbb{C}. \tag{1.7}$$

If $|\lambda - \lambda_0| < \|R(\lambda_0, T)\|^{-1}$, namely if $\|(\lambda - \lambda_0) R(\lambda_0, T)\| < 1$, then, by applying Lemma 1.3.4(1), we conclude that $I + (\lambda - \lambda_0) R(\lambda_0, T) \in \mathcal{I}(X)$. Combining this fact with the identity (1.7) yields that the operator $\lambda - T$ is bijective with bounded inverse operator given by

$$\begin{aligned} R(\lambda, T) =& R(\lambda_0, T)[I - (\lambda_0 - \lambda) R(\lambda_0, T)]^{-1} \\ =& R(\lambda_0, T) \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0, T)^n \\ =& \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0, T)^{n+1}. \end{aligned}$$

Properties (2) and (3) follow by property (1). In particular, the series representation of the resolvent (1.6) gives that $R(\cdot, T)$ is analytic on the open set $\rho(T)$ when $\rho(T) \neq \varnothing$.

(4) Fix $\lambda, \mu \in \rho(T)$. Then,

$$\begin{aligned} R(\lambda, T) =& R(\lambda, T)(\mu - T) R(\mu, T) = R(\lambda, T)[(\mu - \lambda) + (\lambda - T)] R(\mu, T) \\ =& (\mu - \lambda) R(\lambda, T) R(\mu, T) + R(\mu, T). \end{aligned}$$

Thus

$$R(\lambda, T) - R(\mu, T) = (\mu - \lambda) R(\lambda, T) R(\mu, T).$$

(5) Assume that $\lambda_0 \in \sigma(T)$. For every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that, for every $n > n_0$, it holds $|\lambda_n - \lambda_0| < \varepsilon$. By property (1),

$$\varepsilon > |\lambda_n - \lambda_0| \geq \frac{1}{\|R(\lambda_n, T)\|}$$

for every $n > n_0$. Hence $\lim_{n \to \infty} \|R(\lambda_n, T)\| = \infty$.

Vice versa, assume that $\lambda_0 \in \rho(T)$. Then, the function $R(\cdot, T)$ is clearly bounded on the compact set $\{\lambda_n \mid n \in \mathbb{N} \cup \{0\}\} \subseteq \rho(T)$, contradicting the assumption that $\lim_{n \to \infty} \|R(\lambda_n, T)\| = \infty$. ☐

Proposition 1.3.5(2) yields that $\sigma(T)$ is a closed subset of $\mathbb{C}$. In general situations, nothing more can be said on the spectrum and the resolvent sets, as the following example shows.

Example 1.3.6. Let $X = C([0,1],\mathbb{C})$ be the space of complex-valued continuous functions on $[0,1]$, endowed with the sup-norm. Then, the operators $T_i f = f'$, $i = 1,2$, with domains $D(T_1) = C^1([0,1],\mathbb{C})$ and $D(T_2) = \{ f \in C^1([0,1],\mathbb{C}) \mid f(1) = 0\}$, are closed in $C([0,1],\mathbb{C})$, (see Exercise 1.4.4).

Fix $\lambda \in \mathbb{C}$ and consider the function f_λ, defined by $f_\lambda(x) := e^{\lambda x}$ for every $x \in [0,1]$. Then, $f_\lambda \in D(T_1)$ and

$$(\lambda - T_1)(f_\lambda) = \lambda f_\lambda - \lambda f_\lambda = 0;$$

this means that $\lambda - T_1$ is not injective. By the arbitrariness of λ, we get that $\sigma(T_1) = \mathbb{C}$ and $\rho(T_1) = \varnothing$.

For any $\lambda \in \mathbb{C}$, consider now the operator S_λ, defined by

$$S_\lambda g(x) = \int_x^1 e^{\lambda(x-s)} g(s)ds, \qquad 0 \le x \le 1, \;\; g \in C([0,1],\mathbb{C}).$$

Clearly, S_λ belongs to $\mathcal{L}(X)$. Moreover, for every $g \in C([0,1],\mathbb{C})$ the function $S_\lambda g$ is the unique solution of the Cauchy problem $\lambda f - f' = g$, $f(1) = 0$. It follows that $(\lambda - T_2)S_\lambda = S_\lambda(\lambda - T_2) = I$, namely, $\lambda \in \rho(T_2)$. Again, by the arbitrariness of λ, we get that $\rho(T_2) = \mathbb{C}$ and $\sigma(T_2) = \varnothing$.

These easy examples highlight how the spectrum and the resolvent sets of an operator are sensitive to the domain of the operator.

The following result will be useful.

Proposition 1.3.7. *Let $(X, \|\cdot\|)$ be a complex Banach space, $T : D(T) \to X$ be a closed linear operator and $\lambda \in \mathbb{C}$. If $\lambda \in \sigma(T)$ and $\mathrm{rg}(\lambda - T)$ is not closed, then there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ such that $\|x_n\| = 1$ and $\lim_{n\to\infty} \|Tx_n - \lambda x_n\| = 0$.*

Proof. If $\lambda \in \sigma_p(T)$, the result is obvious. Assume then that $\ker(\lambda - T) = \{0\}$. Hence, the operator $(\lambda - T)^{-1}\colon \mathrm{rg}(\lambda - T) \to X$ is well defined and not bounded. Indeed, if one assumes, by contradiction, that there is $C \ge 0$ such that $\|x\| \le C\|(\lambda - T)x\|$ for all $x \in D(T)$, then $\mathrm{rg}(\lambda - T)$ is closed, since T is closed.

Thus, there exists a sequence $(y_n)_{n\in\mathbb{N}} \subset \mathrm{rg}(\lambda - T)$ such that

$$\|(\lambda - T)^{-1} y_n\| \ge n\|y_n\|, \qquad n \in \mathbb{N}.$$

By linearity, we can assume that $y_n = (\lambda - T)x_n$ with $\|x_n\| = 1$ for every $n \in \mathbb{N}$, thus getting

$$1 = \|x_n\| = \|(\lambda - T)^{-1} y_n\| \ge n\|(\lambda - T)x_n\|, \qquad n \in \mathbb{N}.$$

$(x_n)_{n\in\mathbb{N}}$ is the sequence that we were looking for. □

Proposition 1.3.7 inspires the following definition:

Definition 1.3.8. Let $(X, \|\cdot\|)$ be a complex Banach space, $T : D(T) \to X$ be a closed linear operator and $\lambda \in \mathbb{C}$. If $\lambda \in \sigma(T)$ and $\lambda - T$ is not injective or $\mathrm{rg}(\lambda - T)$ is not closed, then λ is called an *approximate eigenvalue* of the operator T. The set of all approximate eigenvalues of T will be denoted by $\sigma_a(T)$.

We end this section with a description of the spectrum of a resolvent operator.

Proposition 1.3.9. *Let $(X, \|\cdot\|)$ be a complex Banach space. Let $T : D(T) \subseteq X \to X$ be a linear operator on X and fix $\lambda \in \rho(T)$. Then, the following properties are satisfied.*

(1) $\sigma(R(\lambda, T)) \setminus \{0\} = \{\frac{1}{\lambda-\mu} \mid \mu \in \sigma(T)\}$.

(2) $\sigma_p(R(\lambda, T)) = \{\frac{1}{\lambda-\mu} \mid \mu \in \sigma_p(T)\}$.

Proof. First, we prove that, for every $\mu \in \rho(T) \setminus \{\lambda\}$, $(\lambda-\mu)^{-1}$ belongs to the resolvent set of the operator $R(\lambda, T)$. For this purpose, we introduce the operator $S := (\lambda-\mu)(\lambda-T)R(\mu, T)$ and observe that

$$\begin{aligned}\left(\frac{1}{\lambda-\mu} - R(\lambda, T)\right) S =&(\lambda - T)R(\mu, T) - (\lambda-\mu)R(\mu, T)\\ =&(\mu - T)R(\mu, T) = I,\\ S\left(\frac{1}{\lambda-\mu} - R(\lambda, T)\right) =&(\lambda - T)R(\mu, T) - (\lambda-\mu)R(\mu, T)\\ =&(\mu - T)R(\mu, T) = I.\end{aligned}$$

From these identities it follows that the operator $(\lambda-\mu)^{-1}I - R(\lambda, T)$ is invertible and

$$\left(\frac{1}{\lambda-\mu} - R(\lambda, T)\right)^{-1} = S = (\lambda-\mu)(\lambda-T)R(\mu, T) \in \mathcal{L}(X). \tag{1.8}$$

Therefore, $(\lambda-\mu)^{-1}$ belongs to $\rho(R(\lambda, T))$.

We can now show property (1). Let $\nu \in \sigma(R(\lambda, T)) \setminus \{0\}$. If $\nu \neq \frac{1}{\lambda-\mu}$ for all $\mu \in \sigma(T)$, then the complex number $\lambda - \frac{1}{\nu}$ does not belong to $\sigma(T)$. Therefore, $\lambda - \frac{1}{\nu} \in \rho(T)$. Since $\lambda - \frac{1}{\nu} \neq \lambda$, from the first part of the proof we can conclude that $\nu \in \rho(R(\lambda, T))$ and (1.8) shows that

$$(\nu - R(\lambda, T))^{-1} = \frac{1}{\nu}(\lambda - T)R\left(\lambda - \frac{1}{\nu}, T\right).$$

This is a contradiction.

Vice versa, let $\nu = \frac{1}{\lambda-\mu}$ with $\mu \in \sigma(T)$. Let us suppose that $\nu \in \rho(R(\lambda, T))$ and consider the operator $S_1 := \nu R(\lambda, T)(\nu - R(\lambda, T))^{-1}$. Then,

$$\begin{aligned}(\mu - T)S_1 =&(\mu - T)\nu R(\lambda, T)(\nu - R(\lambda, T))^{-1}\\ =&(\mu - \lambda + \lambda - T)\nu R(\lambda, T)(\nu - R(\lambda, T))^{-1}\\ =&(-R(\lambda, T) + \nu)(\nu - R(\lambda, T))^{-1} = I\end{aligned}$$

and, similarly, using the fact that $R(\lambda, T)$ and $(\nu - R(\lambda, T))^{-1}$ commute, we can show that

$$S_1(\mu - T) = \nu R(\lambda, T)(\nu - R(\lambda, T))^{-1}(\mu - T) = I\,.$$

This means that $\mu \in \rho(T)$, thereby obtaining a contradiction.

The proof of property (2) follows in a similar way by taking into account the definition of the point spectrum of an operator and taking into account that if $\lambda \in \rho(T)$, then $R(\lambda, T)$ is injective and hence $0 \notin \sigma_p(R(\lambda, T))$. □

Remark 1.3.10. Let $T : D(T) \to X$ be a densely defined linear operator on X such that $D(T) \neq X$ and $\rho(T) \neq \varnothing$. Then, for any $\lambda \in \rho(T)$, $R(\lambda, T)$ cannot admit an inverse belonging to $\mathcal{L}(X)$ and, hence, $0 \in \sigma(R(\lambda, T))$. Thus,

$$\sigma(R(\lambda, T)) = \{0\} \cup \left\{ \frac{1}{\lambda - \mu} \mid \mu \in \sigma(T) \right\}.$$

1.4 Exercises

Exercise 1.4.1. Let X be a Banach space. If M is a subspace of X and N is a subspace of X', then set

$$M^{\perp} := \{ y' \in X' \mid y'(x) = 0 \text{ for every } x \in M \},$$
$${}^{\perp}N := \{ x \in X \mid y'(x) = 0 \text{ for every } y' \in N \}.$$

Prove that, for every $T \in \mathcal{L}(X)$, it holds that

$$\ker(T') = T(X)^{\perp}, \qquad \ker(T) =^{\perp} T'(X').$$

Exercise 1.4.2. Prove Proposition 1.2.5.

Exercise 1.4.3. Let X be a vector space and let $\|\cdot\|_1$, $\|\cdot\|_2$ be two norms in X such that $(X, \|\cdot\|_1)$, $(X, \|\cdot\|_2)$ are both Banach spaces and there exists a constant $C > 0$ such $\|x\|_1 \leq C\|x\|_2$ for all $x \in X$. Prove that $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.

[**Hint:** Use Proposition A.2.2.]

Exercise 1.4.4. On $X = C([0,1], \mathbb{C})$, the space of complex-valued continuous functions on $[0,1]$, endowed with the sup-norm, consider the operators $T_i f = f'$, $i = 1, 2$, with domains $D(T_1) = C^1([0,1], \mathbb{C})$ and $D(T_2) = \{f \in C^1([0,1], \mathbb{C}) : f(1) = 0\}$. Prove that T_1 and T_2 are closed operators on X.

Exercise 1.4.5. On $X = C([0,1], \mathbb{C})$ define the operators

$$Tf := f(0) - \int_0^1 f(x)\, dx, \qquad T_n f = f(0) - \int_{\frac{1}{n}}^1 f(x)\, dx, \qquad f \in X,\ n \in \mathbb{N}.$$

1. Prove that T, $T_n \in X'$ and compute $\|T\|'$ and $\|T_n\|'$.
2. Prove that
$$\lim_{n\to\infty} \|T_n - T\|' = 0.$$

Exercise 1.4.6. To any $f \in L^1((0,1))$ associate the sequence of its moments $(f_n^{\#})_n$ given by

$$f_n^{\#} = \int_0^1 t^n f(t)\, dt.$$

1. Prove that $T : f \mapsto (f_n^{\#})_{n\in\mathbb{N}}$ belongs to $\mathcal{L}(L^1((0,1)), c_0)$, where c_0 is the space of null sequences.

2. Compute its dual operator $T' : \ell^1 \to L^\infty((0,1))$.

Exercise 1.4.7. On $X = C([0,1])$, endowed with the sup-norm, consider the operator T defined by

$$Tf(s) := \int_0^1 2stf(t)\,dt, \qquad f \in X,\ s \in [0,1].$$

1. Prove that $T \in \mathcal{L}(X)$ and compute $\|T\|$.

2. Solve the integral equation

$$f(s) - \int_0^1 2stf(t)\,dt = \sin(\pi s).$$

Exercise 1.4.8. Let X be a Banach space and let T and S be two linear operators from X into itself. Prove that if T, $S \in \mathcal{L}(X)$, then $\sigma(ST) \setminus \{0\} = \sigma(TS) \setminus \{0\}$.

[**Hint:** Compute $(I + TR(\lambda, ST)S)(\lambda I - TS)$ for $\lambda \in \rho(ST)$.]

Notes

For further details, we refer the reader to the classical monographs [1–4].

Appendix A

Functional Analytic Tools

In this appendix we collect some fundamental results from Functional Analysis which we use in this edition of the Internet Seminar.

A.1 Hahn-Banach Theorem

Given a Banach space X, one of the most important questions one may ask about its dual space X' is the following: are there 'enough' elements in X'? For example, are there enough elements to separate points? This is answered using the Hahn–Banach theorem and its consequences that we state here below.

Theorem A.1.1 (Hahn-Banach theorem for normed spaces)**.** *Let X be a real or complex normed space and Y be a non-trivial subspace of X. Then for any $f \in Y'$ there exists $\tilde{f} \in X'$ such that $\tilde{f}(y) = f(y)$ for all $y \in Y$ and $\|\tilde{f}\|' = \|f\|'$.*

Many useful results follow from the Hahn–Banach theorem A.1.1. For example:

Corollary A.1.2 (Separation)**.** *Let X be a real or complex normed space. Then for any $x \in X$ there exists a linear functional $f \in X'$ such that $f(x) = \|x\|$ and $\|f\|' = 1$. Hence, if $x \neq y \in X$ then there exists $f \in X'$ such that $f(x) \neq f(y)$.*

Corollary A.1.3. *Let X be a Banach space and D be a subspace of X. If*

$$\forall f \in X' : \quad (f(x) = 0 \; \forall x \in D) \Longrightarrow \; f = 0,$$

then D is dense in X.

A.2 Open Mapping Theorem

We recall that a continuous map between two topological spaces has the property that the pre-image of any open set is open, but in general the image of an open set is not open. In the case of bounded linear maps between Banach spaces this special property is satisfied under the assumption of the surjectivity of the map.

Theorem A.2.1 (Open Mapping theorem)**.** *Let X and Y be two Banach spaces and let $T : X \to Y$ be a surjective and bounded linear operator. Then, T is an open operator, i.e., T maps open sets in X onto open sets in Y.*

The assumption that T maps X onto Y is essential. Consider, for example, the projection $(x, y) \mapsto (x, 0)$ defined on $\mathbb{R}^2$. This map is not open.

As an application of Theorem A.2.1 a general property of inverse maps can be deduced, i.e.:

Proposition A.2.2. *Let X and Y be two Banach spaces and let T be a bijective bounded linear operator from X to Y. Then the inverse T^{-1} is also a bounded linear operator from Y to X.*

A.3 Quotient spaces, topologically complemented subspaces and projections

Given a vector space E, a projection is a linear map $P : E \to E$ such that $P^2 = P$. In this case, setting $E_1 = P(E)$ and $E_2 = \ker(P)$, the pair (E_1, E_2) is a complemented pair of subspaces of E, namely $E_1 + E_2 = E$ and $E_1 \cap E_2 = E$. This is equivalent to say that every $x \in E$ has a unique decomposition as $x = x_1 + x_2$ with $x_i \in E_i$ $(i = 1, 2)$.

Vice versa, given a complemented pair (E_1, E_2) of subspaces of E, setting $Px = x_1$, when $x = x_1 + x_2$ with $x_i \in E_i$, $(i = 1, 2)$, it is immediate to prove that P is a projection and that $E_1 = P(E)$ and $E_2 = \ker P$.

If X is a Banach space and P is a continuous projection, then the pair of complemented subspaces associated with P is said to be topologically complemented. According to the definition, the subspaces forming a pair of topologically complemented subspaces are closed, being kernels of the continuous maps P and $I - P$.

In general a pair of closed subspaces which are algebraically complemented are not topologically complemented.

Nevertheless, closed subspaces of Hilbert spaces and finite dimensional subspaces in any Banach space are topologically complemented.

Let X be a Banach space and M be a closed subspace of X. Further, let $\varphi : X \to X/M$ be the quotient mapping defined by $\varphi(x) = x + M$ for every $x \in X$. The space X induces a norm on X/M, which is defined by

$$\|\varphi(x)\|_{X/M} = \inf\{\|x + y\| : y \in M\}.$$

X/M, endowed with the norm $\|\cdot\|_{X/M}$, is a Banach space and φ is a continuous and open mapping.

Moreover, the space $(X/M, \|\cdot\|_{X/M})'$ is isometric to $(M^\perp, \|\cdot\|')$, where

$$M^\perp := \{\, y' \in X' \mid y'(x) = 0 \text{ for every } x \in M \,\}.$$

If codim$M$:=dim(X/M) is finite, then it can be proved that M is a topologically complemented subspace in X.

For a proof of the previous results, we refer the reader to [3, 1.40-42 and 4.8-9] or to [**?**, 4.2]

A.4 Uniform Boundedness

We recall the following useful result to establish the uniform boundedness of a family of continuous linear operators between two normed spaces.

Theorem A.4.1 (Banach–Steinhaus theorem)**.** *Let X be a Banach space and let Y be a normed space. Let $\{T_\alpha\}_{\alpha\in A}$ be a family of bounded linear operators from X to Y. Suppose that for each $x \in X$, the set $\{T_\alpha x\}_{\alpha\in A}$ is a bounded subset of Y. Then the set $\{\|T_\alpha\|\}_{\alpha\in A}$ is bounded, i.e., $\sup_{\alpha\in A}\|T_\alpha\| < \infty$.*

A consequence of the theorem above is the following fact.

Corollary A.4.2. *Let X be a Banach space and Y be a normed space. If a sequence $(T_n)_{n\in\mathbb{N}} \subset \mathcal{L}(X,Y)$ is strongly convergent (i.e., $T_n x$ converges for every $x \in X$), then there exists an operator $T \in \mathcal{L}(X,Y)$ such that $(T_n)_{n\in\mathbb{N}}$ is strongly convergent to T.*

A.5 Riesz-Fréchet Theorem

A complex vector space H is said to be *an inner product space* (or *unitary space*) if there exists an *inner product* (or *scalar product*), i.e. a map $\langle\cdot,\cdot\rangle : H \times H \to \mathbb{C}$ with the following properties:

(a) $\langle y, x\rangle = \overline{\langle x, y\rangle}$ for every $x, y \in H$;

(b) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ and $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ for every $x, y, z \in H$ and $\alpha \in \mathbb{C}$;

(c) $\langle x, x\rangle \geq 0$ for every $x \in H$;

(d) $\langle x, x\rangle = 0$ if and only if $x = 0$.

By the properties (a), (b), (c) and (d) above, we can define a *norm* on H by setting $\|x\|^2 := \langle x, x\rangle$, for $x \in H$. The following famous inequality holds true:

$$|\langle x, y\rangle| \leq \|x\|\|y\| \quad \text{(Cauchy-Schwarz Inequality)} \tag{A.1}$$

for all $x, y \in H$.

If the normed space $(H, \|\cdot\|)$ is complete, i.e., if every Cauchy sequence converges in H, then H is called a Hilbert space.

If $(H, \|\cdot\|)$ is a Hilbert space, then it is easy to show that the map $H \ni x \mapsto \langle x, y\rangle \in \mathbb{C}$ is a continuous linear functional on H for all $y \in H$. Indeed, for a fixed $y \in H$, set $f_y(x) := \langle x, y\rangle$ for every $x \in H$. Then, by the Cauchy-Schwarz Inequality it follows that

$$|f_y(x)| = |\langle x, y\rangle| \leq \|x\|\|y\|$$

for all $x \in H$. This means that $f_y \in H'$ and $\|f\|' \leq \|y\|$. Actually, $\|f\|' = \|y\|$ as $f_y(y) = \langle y, y\rangle = \|y\|^2$.

The Riesz-Fréchet Theorem shows that all continuous linear functionals on a Hilbert space H are of this type.

Theorem A.5.1 (Riesz-Fréchet Theorem)**.** *Let H be a Hilbert space and let $f \in H'$. Then, there exists a unique $y \in H$ such that $f(x) = \langle x, y\rangle$ for all $x \in H$. Moreover, $\|f\|' = \|y\|$.*

Notes

For the proof of all the above results we refer the reader to [3, 4].

Bibliography

[1] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[2] M. Reed, B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[3] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[4] W. Rudin, Real and Complex Analysis, 3rd ed., McGraw-Hill, 1986.

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

October 28, 2021

Lecture 2

A survey on operator theory II

In this lecture, we keep on our survey on the main properties of linear operators and their spectra, focusing on the class of bounded operators. In this situation, the spectrum turns to be a non-empty compact subset of the complex plane and a formula allows to calculate the radius of the smallest disk containing the spectrum.

Moreover, the multiplication operator, whose relevance will become clear in the next lectures, is introduced and its spectral properties are studied. Finally the main features of normal and selfadjoint operators are discussed.

2.1 More on spectrum and resolvent

In this section, we keep the same notation used in Section 1.3. To begin with, we observe that, if the operator T is bounded, then the resolvent function $R(\cdot, T)$ and the spectrum of T satisfy further interesting properties.

Proposition 2.1.1. *Let $(X, \|\cdot\|)$ be a complex Banach space and $T \in \mathcal{L}(X)$. Then, the following properties hold.*

(1) *If $\lambda \in \mathbb{C}$ is such that $|\lambda| > \|T\|$, then $\lambda \in \rho(T)$ and*

$$R(\lambda, T) = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}. \tag{2.1}$$

Consequently, $\rho(T) \neq \varnothing$.

(2) $\sigma(T) \neq \varnothing$.

(3) $\sigma(T) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq \|T\|\}$.

(4) *$\sigma(T)$ is a compact subset of $\mathbb{C}$.*

Proof. (1) Fix $\lambda \in \mathbb{C}$ such that $|\lambda| > \|T\|$. By Lemma 1.3.3(2), $\lambda - T \in \mathcal{I}(X)$, namely $\lambda \in \rho(T)$, and

$$R(\lambda, T) = \frac{1}{\lambda}(I - \lambda^{-1}T)^{-1} = \frac{1}{\lambda}\sum_{n=0}^{\infty} \frac{T^n}{\lambda^n} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}.$$

(2) Assume that $\sigma(T) = \varnothing$, or, equivalently, $\rho(T) = \mathbb{C}$. Then, the function $\lambda \in \mathbb{C} \to R(\lambda, T)$ is entire and satisfies the following inequality:

$$\|R(\lambda,T)\| \le \frac{1}{|\lambda|}\frac{1}{1-\frac{\|T\|}{|\lambda|}} = \frac{1}{|\lambda|-\|T\|}, \qquad \lambda \in \mathbb{C},\ |\lambda| > \|T\|, \tag{2.2}$$

by the series representation of the resolvent (2.1). Hence $\|R(\lambda,T)\|$ tends to zero as $|\lambda| \to \infty$.

Now fix $x \in X$ and $f \in X'$, and define the entire function $\varphi : \mathbb{C} \to \mathbb{C}$ by setting

$$\varphi(\lambda) = (f \circ R(\lambda,T))x, \qquad \lambda \in \mathbb{C}.$$

By (2.2), the function φ satisfies the inequality

$$|\varphi(\lambda)| \le \frac{\|f\|'\|x\|}{|\lambda|-\|T\|}, \qquad \lambda \in \mathbb{C},\ |\lambda| > \|T\|.$$

It follows that φ is a bounded function in $\mathbb{C}$ and $\lim_{|\lambda|\to\infty}\varphi(\lambda) = 0$. Therefore, by Liouville's theorem, $\varphi = 0$. By the arbitrariness of f and x, we get that $R(\lambda,T) = 0$ for every $\lambda \in \mathbb{C}$, thus getting a contradiction, since $X \neq \{0\}$.

(3) By property (1), $\{\lambda \in \mathbb{C} \mid |\lambda| > \|T\|\} \subseteq \rho(T)$. The assertion follows by taking the complementary sets.

(4) By combining Proposition 1.3.4(2) and property (3), we get that $\sigma(T)$ is a bounded closed set, and therefore is compact in $\mathbb{C}$. □

Definition 2.1.2. Let $(X, \|\cdot\|)$ be a complex Banach space and $T \in \mathcal{L}(X)$. The spectral radius of T is defined as

$$r(T) := \sup\{\,|\lambda| \mid \lambda \in \sigma(T)\}.$$

Since $\sigma(T)$ is a compact subset of $\mathbb{C}$ for every $T \in \mathcal{L}(X)$, it holds that

$$r(T) = \max\{|\lambda| \mid \lambda \in \sigma(T)\}.$$

This means that the spectrum of a bounded operator T touches at least at one point the smallest closed disk which is centered at the origin and contains $\sigma(T)$. Moreover, the following result holds.

Theorem 2.1.3. *Let $(X, \|\cdot\|)$ be a complex Banach space and let $T \in \mathcal{L}(X)$. Then, the following formula holds:*

$$r(T) = \lim_{n\to\infty}\|T^n\|^{\frac{1}{n}}.$$

Moreover, $r(T) \le \|T\|$.

Proof. Set $r := \limsup_{n\to\infty}\|T^n\|^{\frac{1}{n}}$ and observe that the series

$$\sum_{n=0}^{\infty}\mu^n T^n, \qquad \mu \in \mathbb{C},$$

converges with respect to the operator norm if $|\mu| < r^{-1}$ and does not converge if $|\mu| > r$. It follows that the series

$$\sum_{n=0}^{\infty} \frac{T^n}{\lambda^n}, \qquad 0 \neq \lambda \in \mathbb{C},$$

converges in the operator norm if $|\lambda| > r$. Moreover, the following identity holds for any $|\lambda| > r$:

$$\frac{\lambda - T}{\lambda} \sum_{n=0}^{\infty} \frac{T^n}{\lambda^n} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^n} \frac{\lambda - T}{\lambda} = I.$$

Hence, $\{\lambda \in \mathbb{C} \mid |\lambda| > r\} \subseteq \rho(T)$ and

$$R(\lambda, T) = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}, \qquad \lambda \in \mathbb{C},\ |\lambda| > r. \tag{2.3}$$

So $\sigma(T) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq r\}$. By the definition of spectral radius we get that $r(T) \leq r$.

Let us prove that $r \leq r(T)$. For this purpose, we begin by observing that the series in (2.3) converges uniformly with respect to the operator norm on every circle with center at the origin and radius $\rho > r$. Thus, by a term-by-term integration, (cf. Appendix B), we get

$$\begin{aligned}\frac{1}{2\pi i}\int_{|\lambda|=\rho} \lambda^n R(\lambda, T) d\lambda =& \frac{1}{2\pi i} \sum_{k=0}^{\infty} \left(\int_{|\lambda|=\rho} \lambda^{n-k-1} d\lambda \right) T^k \\ =& \frac{1}{2\pi} \sum_{k=0}^{\infty} \left(\int_0^{2\pi} \rho^{n-k} e^{i(n-k)t} dt \right) T^k = T^n.\end{aligned} \tag{2.4}$$

Since the function $\lambda \mapsto \lambda^n R(\lambda, T)$ is analytic in $\{\lambda \in \mathbb{C}, |\ |\lambda| > r(T)\}$, the integral in (2.4) does not change if we integrate on a circle with center at 0 and radius $\rho > r(T)$. So, by applying (2.4), we get that

$$\|T^n\| \leq \rho^{n+1} \max_{|\lambda|=\rho} \|R(\lambda, T)\|$$

for every $\rho > r(T)$ (observe that the maximum is achieved since the function $R(\cdot, T)$ is continuous on $\rho(T)$). This yields

$$r = \limsup_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leq \rho, \qquad \rho > r(T),$$

and therefore, by the arbitrariness of ρ,

$$r = \limsup_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leq r(T).$$

Let us prove now that

$$\limsup_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}.$$

Fix $n \in \mathbb{N}$, and let $m \in \mathbb{N}$ be such that $m > n$. Then, there exist and are unique $k, h \in \mathbb{N}$ such that $m = kn + h$ with $0 \leq h < n$. Hence,

$$\|T^m\|^{\frac{1}{m}} = \|T^{h+kn}\|^{\frac{1}{m}} \leq \|T\|^{\frac{h}{m}} \cdot \|T^n\|^{\frac{k}{m}} = \|T\|^{\frac{h}{m}} \left(\|T^n\|^{\frac{1}{n}}\right)^{\frac{kn}{m}}.$$

Since

$$0 \leq \frac{h}{m} < \frac{n}{m}, \qquad \frac{m-n}{m} < \frac{kn}{m} \leq 1,$$

it holds that $h/m \to 0$ and $kn/m \to 1$ as $m \to \infty$. Then, for every $n \in \mathbb{N}$, we can estimate

$$\limsup_{m\to\infty} \|T^m\|^{\frac{1}{m}} \leq \|T^n\|^{\frac{1}{n}} \leq \|T\|.$$

By the arbitrariness of n it follows that

$$\limsup_{m\to\infty} \|T^m\|^{\frac{1}{m}} \leq \liminf_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leq \|T\|,$$

namely $\lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}$ exists and is less than or equal to $\|T\|$. □

2.2 A fundamental example: multiplication operators

In this section we will test the concepts that we have introduced so far on a model operator, namely the multiplication operator. As it will be made clear, this operator plays an outstanding role in spectral theory.

Let (Ω, μ) be a measure space, μ being a σ-finite measure, and let $m : \Omega \to \mathbb{C}$ be a μ-measurable function. We recall that the essential range of m is

$$m_{\text{ess}}(\Omega) = \{\, \omega \in \mathbb{C} \mid \mu(\{x \in \Omega \mid |m(x) - \omega| < \varepsilon\}) > 0,\ \forall \varepsilon > 0\}.$$

Fix $1 \leq p < \infty$. The multiplication operator associated with m on $L^p(\Omega, \mu)$ is defined by:

$$D(M_m) = \{f \in L^p(\Omega, \mu) \mid mf \in L^p(\Omega, \mu)\}$$
$$M_m(f) = mf, \qquad f \in D(M_m).$$

Proposition 2.2.1. *Under the previous assumptions:*

(1) $(M_m, D(M_m))$ *is densely defined and closed.*

(2) $D(M_m) = L^p(\Omega, \mu)$ *and* M_m *is bounded if and only if* m *is essentially bounded, i.e., the set* $m_{\text{ess}}(\Omega)$ *is bounded in* $\mathbb{C}$. *In this case,*

$$\|M_m\| = \|m\|_\infty := \sup\{|\omega| : \omega \in m_{\text{ess}}(\Omega)\}.$$

(3) M_m *is boundedly invertible if and only if* $0 \notin m_{\text{ess}}(\Omega)$.

(4) $\sigma(M_m) = m_{\text{ess}}(\Omega)$.

Proof. (1) Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $D(M_m)$ converging in $L^p(\Omega\mu)$ to some function f and such that mf_n converges to a function g in $L^p(\Omega,\mu)$. Then, there exists a subsequence $(f_{k_n})_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} f_{k_n}(x) = f(x)$ for almost every (shortly a.e.) x in Ω. So $\lim_{n\to\infty} m(x)f_{k_n}(x) = m(x)f(x)$ a.e. x in Ω and therefore $g = mf$. We have thus proved that $(M_m, D(M_m))$ is closed.

In order to prove that $D(M_m)$ is dense in $L^p(\Omega,\mu)$, assume first that μ is finite and consider the sequence $(E_n)_{n\in\mathbb{N}}$ of subsets of Ω defined by

$$E_n := \{\, x\in\Omega \mid |m(x)| < n \,\}.$$

Clearly $\bigcup_{n\in\mathbb{N}} E_n = \Omega$ and $\lim_{n\to\infty}\mu(\Omega\setminus E_n) = 0$. Set $u_n = u\chi_{E_n}$, where χ_{E_n} is the characteristic function of E_n. For every $n\in\mathbb{N}$, u_n belongs to $D(M_m)$. By the absolute continuity of the integral, for every $\varepsilon > 0$ there exists $\delta > 0$ such that if $\mu(E) < \delta$, then $\int_E |u|^p d\mu < \varepsilon$. Let $\nu\in\mathbb{N}$ be such that $\mu(\Omega\setminus E_n) < \delta$ for every $n > \nu$. Hence, if $n > \nu$:

$$\int_\Omega |u-u_n|^p d\mu = \int_{\Omega\setminus E_n} |u|^p d\mu < \varepsilon.$$

If $\mu(\Omega) = \infty$, then there exists a family $(\Omega_i)_{i\in\mathbb{N}}$ of subsets of Ω with finite measure such that $\bigcup_{i\in\mathbb{N}}\Omega_i = \Omega$. By repeating the previous arguments for every i and taking the limit, we get the assertion, see Exercise 2.4.1.

It is immediate to verify that, if m is essentially bounded, then $D(M_m) = L^p(\Omega,\mu)$ and M_m is a bounded operator. Vice versa, assume that m is not essentially bounded and introduce the disjoint measurable subsets of Ω

$$F_n := \{x\in\Omega : n \le |m(x)| < n+1\}.$$

Since m is not essentially bounded, there must exist infinitely many sets F_n with positive measure. Take any F_{n_k} with $\mu(F_{n_k}) > 0$ for all $k\in\mathbb{N}$, and define the function

$$f(x) := \begin{cases} \dfrac{1}{k^\nu \mu(F_{n_k})^{1/p}}, & x\in F_{n_k} \\ 0, & x\notin F, \end{cases}$$

where $F = \bigcup_{k\in\mathbb{N}} F_{n_k}$ and $\nu\in(\frac{1}{p},\frac{1}{p}+1]$. Thus,

$$\int_\Omega |f|^p d\mu = \sum_{k=1}^{\infty}\int_{F_{n_k}} \frac{1}{k^{\nu p}\mu(F_{n_k})} d\mu = \sum_{k=1}^{\infty}\frac{1}{k^{\nu p}} < \infty$$

and

$$\int_\Omega |mf|^p d\mu \ge \sum_{k=1}^{\infty}\int_{F_{n_k}} \frac{k^p}{k^{\nu p}\mu(F_{n_k})} d\mu = \sum_{k=1}^{\infty}\frac{1}{k^{p(\nu-1)}} = \infty,$$

since $\nu\in(\frac{1}{p},\frac{1}{p}+1]$, which is a contradiction. If m is essentially bounded, then it is easy to see that $\|M_m\| \le \|m\|_\infty$. In order to prove the other inequality, set for every $\varepsilon > 0$,

$$\Omega_\varepsilon = \{\, x\in\Omega \mid |m(x)| > \|m\|_\infty - \varepsilon \,\}.$$

Then $\mu(\Omega_\varepsilon) > 0$. Choosing $u_\varepsilon \in L^p(\Omega,\mu)$ such that $\|u_\varepsilon\|_p = 1$ and $u_\varepsilon = 0$ on $\Omega \setminus \Omega_\varepsilon$, we get

$$\|M_m u_\varepsilon\|_p^p = \int_{\Omega_\varepsilon} |m u_\varepsilon|^p d\mu \geq (\|m\|_\infty - \varepsilon) \int_{\Omega_\varepsilon} |u_\varepsilon|^p d\mu = (\|m\|_\infty - \varepsilon).$$

So, $\|M_m\| \geq \|m\|_\infty - \varepsilon$ for every $\varepsilon > 0$ and thus $\|M_m\| \geq \|m\|_\infty$.

(3) If $0 \notin m_{\text{ess}}(\Omega)$, then there exists $\varepsilon > 0$ such that $|m(x)| > \varepsilon$ for a.e. $x \in \Omega$. As a consequence, the function $\frac{1}{m}$ is essentially bounded and, by property (2), $M_{\frac{1}{m}}$ is a bounded operator on $L^p(\Omega,\mu)$. Moreover, it is clear that $M_{\frac{1}{m}} = M_m^{-1}$. Vice versa, let $C = \|M_m^{-1}\| > 0$. If $0 \in m_{\text{ess}}(\Omega)$, then $\mu(\{\, x \in \Omega \mid |m(x)| < \frac{1}{C} \,\}) > 0$ and it would exist $\delta \in]0, \frac{1}{C}[$ such that the measure of the set $E = \{\, x \in \Omega \mid \delta < |m(x)| < \frac{1}{C} \,\}$ is 0. By setting $u(x) = \frac{1}{m(x)\mu(E)^{\frac{1}{p}}}\chi_E$ and $v = M_m u$, we would get that $\|v\| = 1$ and

$$\|M_m^{-1}\| \geq \|M_m^{-1} v\|_p = \|u\|_p = \left(\int_E \frac{1}{|m(x)|^p \mu(E)} d\mu \right)^{\frac{1}{p}} > C.$$

This gives a contradiction.

(4) By the very definition, $\lambda \in \sigma(M_m)$ if and only if $\lambda - M_m = M_{\lambda - m}$ is not invertible, or, equivalently, by property (3), $0 \notin (\lambda - m)_{\text{ess}}(\Omega)$, i.e. $\lambda \notin m_{\text{ess}}(\Omega)$. □

2.3 Normal and self-adjoint operators

Along this section, $(H, \|\cdot\|)$ will be a complex Hilbert space endowed with the scalar product $\langle \cdot, \cdot \rangle$. If $T \in \mathcal{L}(H)$, then for every $y \in H$ the operator $x \mapsto \langle Tx, y\rangle$ is linear and bounded. By Riesz–Fréchet Theorem, there exists a unique element $T^*y \in H$ such that

$$\langle Tx, y\rangle = \langle x, T^*y\rangle, \qquad x, y \in H.$$

It is immediate to prove that $T^* \in \mathcal{L}(H)$ and that $\|T^*\| = \|T\|$.

Definition 2.3.1. The operator T^* is called *adjoint operator* of T.

Lemma 2.3.2. *Let $T, S \in \mathcal{L}(H)$ and $\alpha \in \mathbb{C}$. Then, the following properties hold:.*

(1) $T^{**} = T$.

(2) $(\alpha T)^* = \overline{\alpha} T^*$.

(3) $(T + S)^* = T^* + S^*$.

(4) $(TS)^* = S^* T^*$.

(5) *If T is invertible, then T^* is invertible too and $(T^*)^{-1} = (T^{-1})^*$.*

(6) $\|T^*T\| = \|T\|^2$.

Proof. Properties (1)–(5) easily follow from the definition. In order to prove property (6), it suffices to observe that $\|T^*T\| \leq \|T\| \cdot \|T^*\| = \|T\|^2$ and

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*(Tx)\rangle \leq \|x\|\|T^*(Tx)\| \leq \|T^*T\|.$$

for every $x \in H$ with $\|x\| = 1$. □

Definition 2.3.3. An operator $T \in \mathcal{L}(H)$ is called self-adjoint if $T = T^*$, while it is said to be normal if $TT^* = T^*T$.

Example 2.3.4. Let (Ω, μ) be a measure space, with a σ-finite measure μ, and let m be a complex-valued function in $L^\infty(\Omega, \mu)$. Consider the multiplication operator M_m on $L^2(\Omega, \mu)$. Then, $M_m^* = M_{\overline{m}}$. Indeed, if $f \in L^2(\Omega, \mu)$, then it holds that

$$\langle mh, f\rangle = \int_\Omega mh\overline{f}d\mu = \int_\Omega h\overline{(\overline{m}f)}d\mu = \langle h, \overline{m}f\rangle, \qquad \forall h \in L^2(\Omega, \mu).$$

It follows immediately that M_m is a normal operator and that M_m is self-adjoint if and only if m is real-valued.

Proposition 2.3.5. *Let $T \in \mathcal{L}(H)$ be a normal operator. Then, the following properties hold true:*

(1) $\|Tx\| = \|T^*x\|$ *for every* $x \in H$*;*

(2) *If* $Tx = \lambda x$ *for* $x \in H$ *and* $\lambda \in \mathbb{C}$*, then* $T^*x = \overline{\lambda}x$*;*

(3) $\|T^2\| = \|T\|^2$*;*

(4) $r(T) = \|T\|$.

Proof. (1) Fix $x \in H$ and observe that

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

(2) Since $(T - \lambda I)^* = T^* - \overline{\lambda}I$ and $T - \lambda I$ is normal, we can apply property (1) to the operator $T^* - \overline{\lambda}I$, getting that

$$\|T^*x - \overline{\lambda}x\| = \|Tx - \lambda x\| = 0.$$

(3) By applying property (1), we get that $\|T^2x\| = \|T^*Tx\|$ for every $x \in H$. From Lemma 2.3.2, it follows that $\|T^2\| = \|T^*T\| = \|T\|^2$.

(4) By induction, (3) implies that $\|T\|^{2^m} = \|T^{2^m}\|$ for every $m \in \mathbb{N}$. Indeed, it is sufficient to observe that $T^{2^{m-1}}$ is normal and to proceed as in the proof of property (3). Hence $\|T\| = \|T^{2^m}\|^{2^{-m}}$ for every $m \in \mathbb{N}$. Letting $m \to \infty$ and recalling that $r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$, we get that $r(T) = \|T\|$. □

The following proposition reveals that the spectrum of a normal operator consists only of approximate eigenvalues (see Definition 1.3.7).

Proposition 2.3.6. *Let $T \in \mathcal{L}(H)$ be a normal operator. Then, $\lambda \in \sigma(T)$ if and only if there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in H, with $\|x_n\| = 1$, such that* $\lim_{n\to\infty} \|Tx_n - \lambda x_n\| = 0$.

Proof. Let $\lambda \in \mathbb{C}$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in H, with $\|x_n\| = 1$, such that $\lim_{n\to\infty} \|Tx_n - \lambda x_n\| = 0$. Then, clearly $\lambda - T$ is not boundedly invertible and then $\lambda \in \sigma(T)$. Conversely, assume that $\lambda \in \sigma(T)$ and, by contradiction, that a sequence $(x_n)_{n\in\mathbb{N}}$, such that $\|x_n\| = 1$, for every $n \in \mathbb{N}$, and $\lim_{n\to\infty} \|Tx_n - \lambda x_n\| = 0$, does not exist. Then, there exists $\delta > 0$ such that

$$\|Tx - \lambda x\| \geq \delta\|x\| \tag{2.5}$$

for every $x \in H$; hence $\lambda - T$ is injective. But $\lambda - T$ is also normal and

$$(\lambda - T)^* = \overline{\lambda} - T^*.$$

Therefore,

$$\|\overline{\lambda}x - T^*x\| = \|\lambda x - Tx\| \geq \delta\|x\|$$

for every $x \in H$. Thus $\overline{\lambda} - T^*$ is injective too. Let us prove that $\mathrm{rg}(\lambda - T)$ is dense in H. For this purpose, let $y \in H$ be such that $\langle \lambda x - Tx, y\rangle = 0$ for every $x \in H$. Then, $\langle x, \overline{\lambda}y - T^*y\rangle = 0$ for every $x \in H$, namely $\overline{\lambda}y - T^*y = 0$, so that $y = 0$ since $\overline{\lambda} - T^*$ is injective. Thus, $\mathrm{rg}(\lambda - T)$ is dense in H. Combining the density of the range and using (2.5) to prove that the range is closed in H, it follows that $\lambda \in \rho(T)$. □

In the case of self-adjoint operators, we can add further important information on the spectrum.

Proposition 2.3.7. *Let $T \in \mathcal{L}(H)$ be self-adjoint. Then, $\sigma(T) \subset \mathbb{R}$.*

Proof. Fix $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then, for every $0 \neq x \in H$ it holds that

$$\begin{aligned} 0 <&|\lambda - \overline{\lambda}|\|x\|^2 = |\langle Tx - \lambda x, x\rangle - \langle Tx - \overline{\lambda}x, x\rangle| \\ =&|\langle Tx - \lambda x, x\rangle - \langle x, Tx - \lambda x\rangle| \leq 2\|Tx - \lambda x\|\|x\|. \end{aligned}$$

Therefore, if $(x_n)_{n\in\mathbb{N}}$ is a sequence in H with $\|x_n\| = 1$, then

$$\|Tx_n - \lambda x_n\| > \frac{|\lambda - \overline{\lambda}|}{2} > 0.$$

Hence, λ cannot be an approximate eigenvalue, so by Proposition 2.3.6, $\lambda \notin \sigma(T)$. □

We conclude the section with an important subclass of normal operators.

Definition 2.3.8. An operator $T \in \mathcal{L}(H)$ is said to be unitary if $T^*T = TT^* = I$.

Proposition 2.3.9. *The following properties are satisfied.*

(1) *Every unitary operator is normal.*

(2) *An operator T is unitary if and only if it is invertible and $T^{-1} = T^*$.*

(3) *If T is a unitary operator, then T^{-1} and T^* are unitary.*

(4) *Any unitary operator is an isometry.*

Proof. Properties (1) and (2) follow immediately from the definition.

(3) If T is unitary, then

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

Analogously, one proves that $T^{-1}(T^{-1})^* = I$ and, therefore, T^{-1} is unitary. Since $T^* = T^{-1}$, the operator T^* is unitary too.

(4) If T is unitary, then

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

so that T is an isometry. □

Remark 2.3.10. It is worth recalling that, thanks to the polarization identity

$$\langle x,y\rangle = \frac{1}{4}[\langle x+y,x+y\rangle - \langle x-y,x-y\rangle + i\langle x+iy,x+iy\rangle - i\langle x-iy,x-iy\rangle], \tag{2.6}$$

any isometry in a Hilbert space preserves the scalar product, namely $\langle Tx,Ty\rangle = \langle x,y\rangle$ for every $x,y \in H$.

Examples 2.3.11. (1) Let $\ell^2(\mathbb{Z})$ be the Hilbert space of complex valued sequences $x = (x_n)_{n\in\mathbb{Z}}$ such that $\|x\|^2 = \sum_{n\in\mathbb{Z}} |x_n|^2 < \infty$. Consider the right translation operator $T\colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ defined by

$$T(x_n)_{n\in\mathbb{Z}} := (x_{n-1})_{n\in\mathbb{Z}}, \qquad (x_n)_{n\in\mathbb{Z}} \in \ell^2(\mathbb{Z})\,.$$

Then, T is unitary. Indeed, T is clearly invertible and

$$\langle Tx,y\rangle = \sum_{n\in\mathbb{Z}} x_{n-1}\overline{y}_n = \sum_{n\in\mathbb{Z}} x_n\overline{y}_{n+1} = \langle x, T^{-1}y\rangle, \qquad x,y\in\ell^2(\mathbb{Z}),$$

i.e., $T^* = T^{-1}$.

(2) Let $H = L^2((0,1))$ and let $T \in \mathcal{L}(H)$ be the operator defined by

$$(Tf)(x) := f(1-x), \qquad \text{a.e. } x\in(0,1),\ \ f\in L^2((0,1)).$$

The operator T is clearly invertible. Moreover $T = T^* = T^{-1}$. Hence, T is a unitary operator.

2.4 Exercises

Exercise 2.4.1. On $L^p(\Omega,\mu)$, $p\in[1,\infty)$, where (Ω,μ) is a measure space with μ a σ-finite measure and $\mu(\Omega)=\infty$, we consider the multiplication operator

$$\begin{aligned} &D(M_m) = \{f\in L^p(\Omega,\mu) \mid mf\in L^p(\Omega,\mu)\} \\ &M_m(f) = mf, \qquad f\in D(M_m) \end{aligned}$$

associated with a μ-measurable function $m:\Omega\to\mathbb{C}$. Prove that $D(M_m)$ is dense in $L^p(\Omega,\mu)$.

Exercise 2.4.2. Let $\varphi:[0,1]\to[0,1]$ be a bijective and continuous function. Consider the operator $T: L^2((0,1))\to L^2((0,1))$ defined by

$$(Tf)(x) := \int_0^{\varphi(x)} f(s)\,ds.$$

(1) Prove that $T\in\mathcal{L}(L^2((0,1)))$ and compute T^* (distinguishing the case in which φ is increasing from that in which φ is decreasing).

(2) If φ is increasing, then prove that T cannot be self-adjoint.

(3) Find a decreasing function φ such that T is self-adjoint. Is such a φ unique?

Exercise 2.4.3. Let X be a Banach space and $T\in\mathcal{L}(X)$ be an isometry, i.e., $\|Tx\| = \|x\|$ for all $x\in X$. Prove the following properties.

(1) $\sigma_a(T) \cap \{\lambda \in \mathbb{C} : |\lambda| < 1\} = \varnothing$.

(2) If $0 \in \rho(T)$, then $\sigma(T^{-1}) = \{\frac{1}{\mu} : \mu \in \sigma(T)\}$ and $\sigma(T) \subseteq \{\lambda \in \mathbb{C} : r(T^{-1})^{-1} \leq |\lambda| \leq r(T)\}$.

(3) If $0 \in \rho(T)$, then $\sigma(T) \subseteq \{\lambda \in \mathbb{C} : |\lambda| = 1\}$.

Exercise 2.4.4. On $X = C_0(\mathbb{R}, \mathbb{C}) = \{f \in C(\mathbb{R}, \mathbb{C}) : \lim_{|t|\to\infty} f(t) = 0\}$ endowed with the sup-norm, consider the operator

$$(Tf)(t) = f(t+1), \qquad t \in \mathbb{R},\ f \in X.$$

(1) Prove that T is an isometry and

$$\sigma(T) = \{\xi\lambda : |\xi| = 1,\ \lambda \in \sigma(T)\}.$$

[**Hint:** Use the transformation $(S_\theta f)(t) := e^{i\theta t} f(t)$ for $t \in \mathbb{R}$ and $f \in X$.]

(2) Show that

$$\sigma(T) = \{\lambda \in \mathbb{C} : |\lambda| = 1\}.$$

[**Hint:** Show that $1 \in \sigma(T)$ and use (1) to prove that $\{\lambda \in \mathbb{C} : |\lambda| = 1\} \subseteq \sigma(T)$.]

Exercise 2.4.5. Let X be a Banach space and $T \in \mathcal{L}(X)$. Prove that $\sigma(T) = \sigma(T')$.

Exercise 2.4.6. Let X be a Banach space and $T \in \mathcal{L}(X)$ be an operator such that $\|T\| \in \sigma(T)$. Prove that

$$\|I + T\| = 1 + \|T\|.$$

Exercise 2.4.7. Let X be a Banach space and let T and S be two linear operators from X into itself. Prove that if $ST - TS = I$, then S or T is not bounded.

Exercise 2.4.8. Let H be a complex Hilbert space and let $T \in \mathcal{L}(H)$ be a self-adjoint operator. Prove that

(1) $\|(T \pm iI)x\|^2 = \|Tx\|^2 + \|x\|^2$ for every $x \in H$;

(2) $U_T := (T + iI)(T - iI)^{-1}$ is a unitary operator on H. U_T is called the *Cayley transform* of T;

(3) $1 \in \rho(U_T)$ and $T = -i(I + U_T)(I - U_T)^{-1}$.

Appendix B

Differential and integral calculus for functions with values in a Banach space

In this appendix we define the Riemann integral for vector-valued functions. For more details and for the proof of the results that we present here, we refer the reader to [1, Chapter 2], [2, Chapter 3] and [3, Chapter 3].

Definition B.0.1. A bounded function $f : [a,b] \to X$ $(-\infty < a < b < +\infty)$ is said to be *integrable* on $[a,b]$ if there exists $x \in X$ with the following property: for each $\varepsilon > 0$ there exists $\delta > 0$ such that, for every partition $\mathcal{P} = \{a = t_0 < t_1 < \ldots < t_n = b\}$ of $[a,b]$, with $\max_{i=1,\ldots,n}(t_i - t_{i-1}) < \delta$, and for every choice of the points $\xi_i \in [t_{i-1}, t_i]$, we have

$$\left\| x - \sum_{i=1}^{n} f(\xi_i)(t_i - t_{i-1}) \right\| < \varepsilon.$$

In this case we define

$$\int_a^b f(t)dt = x.$$

Arguing as in the real-valued case, one can easily check that the set of integrable functions on $[a,b]$ is a vector space and that the integral over $[a,b]$ is a linear operator on this vector space. In particular, if $f : [a,b] \to X$ is continuous, then it is integrable. Moreover, if f is integrable on $[a,b]$, then the map $t \mapsto \|f(t)\|_X$ is integrable on $[a,b]$ as well. Finally, if f is integrable on $[a,b]$, then it is integrable on every $[c,d] \subset [a,b]$ and

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt$$

for every $c \in (a,b)$.

As in the real-valued case, the definition of the Riemann integral can be easily extended to the case of unbounded intervals or unbounded functions.

Definition B.0.2. Let $I \subset \mathbb{R}$ be an interval with endpoints a and b $(-\infty \le a < b \le +\infty)$ with a and b not necessarily in I. Moreover, let $f : I \to X$ be Riemann integrable on $[c,d]$

for every $a < c < d < b$. We say that f admits an improper integral on I if, for each $t_0 \in I$, the limits

$$\lim_{c \to a^+} \int_c^{t_0} f(t)dt, \qquad \lim_{d \to b^-} \int_{t_0}^d f(t)dt$$

exist in X. In this case we set

$$\int_I f(x)dx = \lim_{c \to a^+} \int_c^{t_0} f(t)dt + \lim_{d \to b^-} \int_{t_0}^d f(t)dt.$$

Note that the previous definition is independent of t_0.

To simplify the notation in the following proposition we simply denote by I the interval in which the function f is defined. When we say that "f is integrable on I" we mean indifferently, that f is Riemann integrable on I or that it admits improper integral on I.

Proposition B.0.3. *Let $f : I \to X$ be integrable on I. Then, the following properties are satisfied:*

(i) *for each bounded operator $T \in L(X, Y)$, the function Tf is Riemann integrable on $[a, b]$ and*

$$T \int_I f(t)dt = \int_I Tf(t)dt;$$

(ii) *if $A : D(A) \subset X \to Y$ is a closed operator such that $f(t) \in D(A)$ for every $t \in I$ and Af is integrable on I, then*

$$\int_I f(t)dt \in D(A) \qquad \textit{and} \qquad A \int_I f(t)dt = \int_I Af(t)dt.$$

Proof. We limit ourselves to proving the last part of the proposition, since the first one is similar and even simpler.

Proof of (ii). We first assume that $I = [a, b]$ for some $a, b \in \mathbb{R}$, $a < b$ and f is Riemann integrable on $[a, b]$. Set $x = \int_a^b f(t)dt$. We fix $n \in \mathbb{N}$ and consider the partition $\mathcal{P}_n = \{a = t_0 < \ldots < t_n = b\}$ where $t_k = a + (b - a)k/n$ for every $k = 0, \ldots, n$. Moreover, for every $k = 0, \ldots, n - 1$ fix $\xi_i \in [t_i, t_{i+1}]$ and set

$$S_n = \sum_{i=1}^n f(\xi_i)(t_i - t_{i-1}), \qquad n \in \mathbb{N}.$$

Of course S_n belongs to $D(A)$ for each $n \in \mathbb{N}$ and

$$AS_n = \sum_{i=1}^n Af(\xi_i)(t_i - t_{i-1}), \qquad n \in \mathbb{N}.$$

Since both f and Af are integrable, S_n and AS_n converge, respectively, to x and $y = \int_a^b Af(t)dt$ as n tends to $+\infty$. Since A is closed, it follows that x belongs to $D(A)$ and $Ax = y$.

Now, suppose that f admits improper integral on I. To fix the ideas we assume that $I = [a, +\infty)$ and that f is Riemann integrable on $[a, b]$ for each $b > a$. Then,

$$A \int_a^b f(t)dt = \int_a^b Af(t)dt$$

By hypothesis

$$\lim_{b\to+\infty} \int_a^b Af(t)dt = \int_a^{+\infty} Af(t)dt \qquad \text{and} \qquad \lim_{b\to+\infty} \int_a^b f(t)dt = \int_a^{+\infty} f(t)dt.$$

Since A is closed,

$$\int_a^{+\infty} f(t)dt \in D(A) \qquad \text{and} \qquad A\int_a^{+\infty} f(t)dt = \int_a^{+\infty} Af(t)dt. \qquad \square$$

Next we recall the *fundamental theorem of calculus* for X-valued functions. For this purpose we first recall the definition of *Fréchet derivative*. Let $I \subset \mathbb{R}$ be an (open) interval and let $t_0 \in I$. The function $f : I \to X$ is *Fréchet differentiable* at $t_0 \in I$ if the limit

$$\lim_{t\to t_0} \frac{f(t) - f(t_0)}{t - t_0}$$

exists in X. Such a limit, when existing, is denoted by $f'(t_0)$ and is called the Fréchet derivative of f at t_0. In an analogous way the right- and left-derivative can be defined.

Theorem B.0.4. *Let $f : [a, b] \to X$ be continuous. Then, the integral function $F : [a, b] \to X$ defined by*

$$F(t) = \int_a^t f(s)ds, \qquad t \in [a, b],$$

is (Fréchet) differentiable, and $F'(t) = f(t)$ for each $t \in [a, b]$.

Now, we recall the definition of the integral of vector-valued functions of a complex variable, along a smooth curve γ.

Definition B.0.5. Let Ω be an open subset of $\mathbb{C}$, $f : \Omega \to X$ be a continuous function and $\gamma : [a, b] \to \Omega$ be a piecewise C^1-curve. The integral of f along γ is defined as follows:

$$\int_\gamma f(z)dz = \int_a^b f(\gamma(t))\gamma'(t)dt.$$

As in the case of vector-valued functions defined on a real interval, we can define the *improper complex integrals* in an obvious way.

Definition B.0.6. Let $\Omega \subset \mathbb{C}$ be a (possibly) unbounded open set. Moreover, let $I = (a, b)$ be a (possibly unbounded) interval and $\gamma : I \to \mathbb{C}$ be a (piecewise) C^1 curve in Ω. We say that the continuous function $f : \Omega \to X$ admits an improper integral along γ if for each $t_0 \in (a, b)$ the limits

$$\lim_{s\to a^+} \int_s^{t_0} f(\gamma(\tau))\gamma'(\tau)d\tau \qquad \text{and} \qquad \lim_{s\to b^-} \int_{t_0}^b f(\gamma(\tau))\gamma'(\tau)d\tau$$

exist in X. In such a case, we set

$$\int_\gamma f(z)dz = \lim_{s\to a^+} \int_s^{t_0} f(\gamma(\tau))\gamma'(\tau)d\tau + \lim_{s\to b^-} \int_{t_0}^s f(\gamma(\tau))\gamma'(\tau)d\tau.$$

Note that the definition of the improper integral is independent of the choice of t_0. Moreover, if I is bounded and the integral of f along γ exists, then f admits an improper integral along γ and the two integrals coincide.

Bibliography

[1] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[2] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[3] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

November 3, 2021

Lecture 3

Compact operators and Riesz-Schauder theory

In this lecture, we introduce the compact operators and explore their main properties. For instance, we show that compact operators form a closed subspace of the set of bounded operators between Banach spaces and that the adjoint of a compact operator is compact too. One fundamental example of compact operators is given by the integral operators between L^p-spaces.

While in a finite dimensional setting each bounded operator is automatically compact, this is not the case in a infinite dimensional Banach space, since, for instance, the identity operator is not compact.

It is well known that the spectrum of a bounded linear operator on a finite dimensional Banach space consists of eigenvalues only. On the other hand, in the previous lectures we have seen that the structure of the spectrum of a bounded linear operator defined on a infinite dimensional Banach space could be much more complicated. The picture is easier when one deals with compact operators. Indeed, 0 is always in the spectrum of a compact operator and all the other elements in the spectrum are eigenvalues whose corresponding eigenspaces are finite dimensional. Another important feature of compact operators is the so-called *Fredholm alternative* which states the following: if $T : X \to X$ is a compact operator and $\lambda \in \mathbb{C} \setminus \{0\}$, then either the equation $\lambda x - Tx = y$ is uniquely solvable for every $y \in X$ or y belongs to the image of the operator $\lambda I - T$ if and only if it satisfies n "compatibility" conditions, where n denotes the dimension of the eigenspace associated to λ.

3.1 Compact operators: definitions, examples and Schauder theorems

Definition 3.1.1. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two Banach spaces and let $T\colon X \to Y$ be a linear operator. T is called compact operator if $T(B_X)$ is a relatively compact subset of Y.

It is easy to show that a linear operator $T\colon X \to Y$ is compact if and only if for every sequence $(x_n)_{n\in\mathbb{N}} \subset B_X$, $(Tx_n)_{n\in\mathbb{N}}$ admits a convergent subsequence (see Exercise 3.3.1).

Example 3.1.2. Let $\Omega \subset \mathbb{R}^N$ be a bounded open set and let $K : \overline{\Omega} \times \overline{\Omega} \to \mathbb{R}$ be a continuous function. Further, let $T : L^p(\Omega) \to L^p(\Omega)$ be the integral operator defined by

$$Tf(x) = \int_\Omega K(x,y)f(y)dy, \qquad x \in \Omega,\ \ f \in L^p(\Omega).$$

From the dominated convergence theorem it easily follows that $T(L^p(\Omega)) \subseteq C(\overline{\Omega})$. Let us prove that T is a compact operator. For this purpose, we observe that, applying Hölder's inequality, it can be easily checked that

$$\|Tf\|_\infty \le \sup_{x\in\Omega} \|K(x,\cdot)\|_{p'}\|f\|_p \le \|K\|_\infty |\Omega|^{1/p'}$$

for every $f \in L^p(\Omega)$ with $\|f\|_p \le 1$, where $|\Omega|$ denotes the Lebesgue measure of Ω. Thus, $T(B_{L^p(\Omega)})$ is an equibounded subset of $C(\overline{\Omega})$, i.e. $(T(B_{L^p(\Omega)}))(x)$ is a bounded subset of $\mathbb{R}$, for each $x \in \overline{\Omega}$. Moreover, since K is a uniformly continuous function, for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$y \in \Omega,\ \ x_1, x_2 \in \Omega : \|x_1 - x_2\| < \delta \ \Rightarrow\ |K(x_1,y) - K(x_2,y)| < \varepsilon|\Omega|^{-1/p'}.$$

Therefore, applying again Hölder's inequality we conclude that

$$|Tf(x_1) - Tf(x_2)| \le \int_\Omega |K(x_1,y) - K(x_2,y)||f(y)|dy \le \varepsilon$$

for every $f \in B_{L^p(\Omega)}$ and $x_1, x_2 \in \Omega$ such that $\|x_1 - x_2\| < \delta$. This means that $T(B_{L^p(\Omega)})$ is a equi-uniformly continuous subset of $C(\overline{\Omega})$. Applying Arzelà-Ascoli theorem, see Theorem A.7.1, we can infer that $T(B_{L^p(\Omega)})$ is a relatively compact subset of $C(\overline{\Omega})$. Finally, recalling that $C(\overline{\Omega})$ is continuously embedded into $L^p(\Omega)$, we obtain that $T(B_{L^p(\Omega)})$ is relatively compact in $L^p(\Omega)$.

Since relatively compact subsets of a Banach space are also bounded sets, it follows that each compact operator T is bounded. (cf. Proposition 1.1.1).

Remark 3.1.3. Given two Banach spaces X and Y, denote by $\mathcal{K}(X,Y)$ the space of all compact linear operators and by $\mathcal{F}(X,Y)$ the space of all bounded linear operators with finite dimensional range. Recalling that every bounded subset of a finite dimensional space is relatively compact, we easily conclude that

$$\mathcal{F}(X,Y) \subset \mathcal{K}(X,Y) \subset \mathcal{L}(X,Y)\,.$$

Moreover, given three Banach spaces X, Y, Z,

$$S \circ R \circ T \in \mathcal{K}(X,W)$$

for every $S \in \mathcal{L}(Z,W)$, $T \in \mathcal{L}(X,Y)$ and $R \in \mathcal{K}(Y,Z)$. Indeed, since $T \in \mathcal{L}(X,Y)$, $T(B_X)$ is a bounded subset of Y and, hence, there exists $\lambda > 0$ such that $T(B_X) \subset \lambda B_Y$. It follows that $R(T(B_X)) \subset \lambda R(B_Y)$, where $R(B_Y)$ is a relatively compact subset of Z, since $R \in \mathcal{K}(Y,Z)$. Noticing that $S \in \mathcal{L}(Z,W)$, we can infer that $S(R(T(B_X)))$ is a relatively compact subset of W.

The following characterization holds true.

Theorem 3.1.4 (Schauder theorem). *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two Banach spaces and let $T : X \to Y$ be a linear operator. Then, T is compact if and only if the dual operator T' is compact.*

Proof. Let $T : X \to Y$ be a compact operator and let us denote by K the closure of $T(B_X)$ in Y. Since T is a bounded operator, there exists $M > 0$ such that $\|y\|_Y \le M$ for every $y \in K$. Therefore,

$$\|f\|_K := \sup\{|f(y)| \mid y \in K\} \le M,$$

for every f in the unit ball $B_{Y'}$ of Y' and, consequently, we can identify $B_{Y'}$ with a bounded subset of the Banach space $C(K)$ consisting of all the continuous functions over the compact set K, endowed with the sup-norm

$$\|g\|_K := \sup\{|g(y)| \mid y \in K\} \qquad g \in C(K). \tag{3.1}$$

Moreover, for every $f \in B_{Y'}$ and $y, z \in K$ it holds that

$$|f(y) - f(z)| \le \|y - z\|_Y.$$

This implies that $B_{Y'}$ is an equi-continuous subset of $C(K)$ and, since it is bounded, it is relatively compact due to Arzelà-Ascoli theorem. Thus, from every sequence $(f_n)_{n\in\mathbb{N}} \subset B_{Y'}$ we can extract a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ which is a Cauchy sequence with respect to the norm (3.1). Since

$$\begin{aligned}\|T'f_{n_k} - T'f_{n_j}\|_{X'} &= \sup_{x\in B_X} |(T'f_{n_k} - T'f_{n_j})(x)| = \sup_{x\in B_X} |(f_{n_k} - f_{n_j})(Tx)| \\ &= \|f_{n_k} - f_{n_j}\|_K\end{aligned}$$

for every $j, k \in \mathbb{N}$, $(T'f_{n_k})_{k\in\mathbb{N}}$ turns out to be a Cauchy sequence in X', which is a Banach space, and, therefore, it converges to an element of X'. We have so proved that $T'(B_{Y'})$ is a relatively compact subset of X' and, hence, T' is a compact operator.

Vice versa, let us suppose that T' is a compact operator. Arguing as above, we can show that the operator $T''\colon X'' \to Y''$ is compact. If we denote by $j_X\colon X \to X''$ (by $j_Y\colon Y \to Y''$, respectively) the canonical embedding defined by $j_X(x)(f) := f(x)$ for every $x \in X$ and $f \in X'$ (similarly, for j_Y), then j_X and j_Y are isometries and

$$j_Y \circ T = T'' \circ j_X.$$

So, $j_Y(T(B_X)) = T''(j_X(B_X))$ is a relatively compact subset of Y'', since $T'' \circ j_X$ is a compact operator (see Remark 3.1.3) Taking into account that j_Y is an isometry from Y into Y'', it follows that $T(B_X)$ is a relatively compact set of Y. This shows that T is a compact operator. □

Proposition 3.1.5. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two Banach spaces. Then, $\mathcal{K}(X, Y)$ is a closed subspace of $\mathcal{L}(X, Y)$.*

Proof. The proof that $\mathcal{K}(X, Y)$ is a linear subspace of $\mathcal{L}(X, Y)$ is left to the reader (see Exercise 3.3.2).

To prove that $\mathcal{K}(X,Y)$ is closed, fix a sequence $(T_n)_{n\in\mathbb{N}} \subset \mathcal{K}(X,Y)$ converging to an operator $T : X \to Y$ in $(\mathcal{L}(X,Y), \|\cdot\|)$ as n tends to ∞. Further, fix $\varepsilon > 0$ and let $n_0 \in \mathbb{N}$ be such that $\|T_{n_0} - T\| < \varepsilon$. Then,

$$T(B_X) \subset (T - T_{n_0})(B_X) + T_{n_0}(B_X) \subset \varepsilon B_Y + T_{n_0}(B_X). \tag{3.2}$$

Note that $T_{n_0}(B_X)$ is a relatively compact subset of Y, since $T_{n_0} \in \mathcal{K}(X,Y)$. We can thus determine a finite number of elements in B_Y, let us say $y_1, y_2, \ldots, y_k$, such that

$$T_{n_0}(B_X) \subset \bigcup_{i=1}^{k} (y_i + \varepsilon B_Y). \tag{3.3}$$

Using (3.2) and (3.3) it is immediate to conclude that

$$T(B_X) \subset \bigcup_{i=1}^{k} (y_i + 2\varepsilon B_Y)\,.$$

Due to the arbitrariness of ε, this means that $T(B_X)$ is a totally bounded subset of Y and, hence, it is relatively compact, see Exercise 3.3.1. □

Remark 3.1.6. From Proposition 3.1.5 it follows that, if $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are two Banach spaces, and $(T_n)_{n\in\mathbb{N}}$ is a sequence of operators from X into Y, with finite dimensional range, converging in $\mathcal{L}(X,Y)$ to some operator T, then T is a compact operator.

The question whether each compact operator between Banach spaces is the limit, in the operator norm, of a suitable sequence of operators with finite dimensional range, was an open problem for a long time, the so-called "*approximation problem*". In 1972, Enflo solved this problem showing that the answer to the previous question is negative. Anyway it is worth to mention that, if Y is a Hilbert space (or more generally, if Y admits a Schauder basis), then the closure of the space $\mathcal{F}(X,Y)$ coincides with $\mathcal{K}(X,Y)$ (cf. e.g., [1]).

Remark 3.1.7. Let $(X, \|\cdot\|)$ be an infinite dimensional Banach space and let T belong to $\mathcal{K}(X) := \mathcal{K}(X,X)$. Then, $0 \in \sigma(T)$. Indeed, if 0 were in the resolvent set of operator T, then, by Remark 3.1.3, $I = T^{-1} \circ T$ would be a compact operator and, hence, B_X would be a compact set. By Theorem A.6.1, this would imply that X is finite dimensional, which is a contradiction.

3.2 Riesz-Schauder theory for compact operators

In what follows, we denote by $(X, \|\cdot\|)$ an infinite dimensional Banach space over $\mathbb{C}$. Moreover, for every $T \in \mathcal{L}(X)$ and $\lambda \in \mathbb{C}$, we denote by T_λ the operator $\lambda - T$. If $0 \neq \lambda \in \sigma_p(T)$, where $\sigma_p(T)$ is the set of all the eigenvalues of T, then we denote by $N(T_\lambda) := \ker T_\lambda$ the *eigenspace* associated with the eigenvalue λ. In general, if T is a bounded operator, then $N(T_\lambda)$ is an infinite dimensional space. Assuming that T is compact, we can prove the following.

Proposition 3.2.1. *Fix $T \in \mathcal{K}(X)$. Then, for every $\lambda \in \mathbb{C} \setminus \{0\}$ and $n \in \mathbb{N}$, it holds that $\dim N(T_\lambda^n) < \infty$.*

Proof. We first prove the proposition in the case $n = 1$. For this purpose, we set $B := B_X \cap N(T_\lambda)$ (i.e., B is the closed unit ball of $N(T_\lambda)$). Clearly, $T_\lambda(B) = \{0\}$ and, consequently, $\lambda B = T(B)$. Since $T(B)$ is a relatively compact subset of X, B is itself a relatively compact subset of X and also of $N(T_\lambda)$. Therefore, $N(T_\lambda)$ is finite dimensional and the assertion follows in the case $n = 1$.

The proof in the case $n > 1$ follows from the case $n = 1$ since

$$T_\lambda^n = (\lambda - T)^n = \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k}(-1)^k T^k = \lambda^n - S$$

where $S = \sum_{k=1}^{n} \binom{n}{k}(-1)^{k+1}\lambda^{n-k}T^k \in \mathcal{K}(X)$ in view of Remark 3.1.3. □

Proposition 3.2.2. *For every $T \in \mathcal{K}(X)$, $\lambda \in \mathbb{C} \setminus \{0\}$ and $n \in \mathbb{N}$, the set $T_\lambda^n(X)$ is a closed subspace of X.*

Proof. Due to Proposition 3.2.1, we know that $\dim N(T_\lambda) < \infty$. This implies that there exists a bounded linear operator $P\colon X \to X$ such that $P(X) = N(T_\lambda)$ and $P^2 = P$. Set $Y := \ker P$. Then, we can split X into the direct sum

$$X = N(T_\lambda) \oplus Y \tag{3.4}$$

and conclude that $T_\lambda(X) = T_\lambda(Y)$. Next, we fix $z \in \overline{T_\lambda(X)}$ and take a sequence $(y_n)_{n\in\mathbb{N}} \subset Y$ such that $T_\lambda(y_n)$ converges to z in X. By contradiction, let us assume that the sequence $(y_n)_{n\in\mathbb{N}}$ is not bounded. Then, there exists a subsequence $(y_{n_k})_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} \|y_{n_k}\| = \infty$. Set $v_k := \frac{y_{n_k}}{\|y_{n_k}\|} \in Y$ for every $k \in \mathbb{N}$. Clearly, $\|v_k\| = 1$ for every $k \in \mathbb{N}$, so that $(v_k)_{k\in\mathbb{N}}$ is a bounded sequence. Moreover, $T_\lambda(v_k) = \frac{1}{\|y_{n_k}\|}T_\lambda(y_{n_k})$ converges to zero as k tends to ∞ and

$$v_k = \lambda^{-1}(T_\lambda(v_k) + T(v_k)) \tag{3.5}$$

for every $k \in \mathbb{N}$. Since the sequence $(v_k)_{k\in\mathbb{N}}$ is bounded and T is a compact operator, there exists a subsequence of $(T(v_k))_{k\in\mathbb{N}}$ which converges. Without loss of generality, we assume that the whole sequence $(T(v_k))_{k\in\mathbb{N}}$ converges to some element w of X. From formula (3.5) it follows that v_k converges to $\lambda^{-1}w =: v \in X$ as k tends to ∞, so that $\|v\| = 1$. Further, the boundedness of T and the uniqueness of the limit imply that $T_\lambda(v_k)$ converges to $T_\lambda(v) = 0$ as k tends to ∞. This yields that $v \in N(T_\lambda)$ and, therefore, $v = P(v) = \lim_{k\to\infty} P(v_k) = 0$. This is a contradiction since $\|v\| = 1$. Hence, the sequence $(y_n)_{n\in\mathbb{N}}$ is bounded.

Arguing as we did for the sequence $(v_k)_{k\in\mathbb{N}}$, by writing $y_n = \lambda^{-1}(T_\lambda(y_n) + Ty_n)$ and using the compactness of T, we can infer that, possibly up to a subsequence, $(y_n)_{n\in\mathbb{N}}$ converges to some $y \in X$ as n tends to ∞. The boundedness of T_λ and the uniqueness of the limit show that $T_\lambda(y_n)$ converges to $T_\lambda(y) = z$, so that $z \in T_\lambda(X)$. We have so proved that $T_\lambda(X)$ is a closed subspace of Y.

Since $T(X_1) \subset X_1$ and $X_1 := T_\lambda(X)$ is closed, it follows that $T|_{X_1} \in \mathcal{K}(X_1)$. Thus, from the first part of the proof, we have $T_\lambda^2(X) = T_\lambda(X_1)$ is a closed subspace of X_1 and hence of X. By induction we obtain the closedness of $T_\lambda^n(X)$ for every $n \in \mathbb{N}$. □

Corollary 3.2.3. *For every $T \in \mathcal{K}(X)$, $\lambda \in \mathbb{C} \setminus \{0\}$ and $n \in \mathbb{N}$, it holds that*

$$\operatorname{codim} T_\lambda^n(X) < \infty\,.$$

Proof. Since $T'_\lambda = \lambda - T'$ and $T' \in \mathcal{K}(X')$ due to Theorem 3.1.4, we can apply Proposition 3.2.1 and infer that $\dim N(T'_\lambda) < \infty$. On the other hand, Proposition 3.2.2 shows that $T_\lambda(X)$ is a closed subspace of X. Hence, the quotient space $X/T_\lambda(X)$, endowed with the norm $\|\cdot\|_{X/T_\lambda(X)}$ is a Banach space. In particular, its topological dual space is isomorphic to $(T_\lambda(X)^\perp, \|\cdot\|')$, see Appendix A.3. Consequently, thanks to the first equality in (1.4), to $N(T'_\lambda)$. It follows that the dual space of $X/T_\lambda(X)$ and, consequently, $X/T_\lambda(X)$ itself, is a finite dimensional space, i.e., $\operatorname{codim} T_\lambda(X) < \infty$.

The proof in the case $n > 1$ is left to the reader as an exercise (see Exercise 3.3.4). □

The following lemma plays a crucial role in what follows.

Lemma 3.2.4 (Riesz's lemma). *Let $(X, \|\cdot\|)$ be a real or complex Banach space and let M be a closed proper subspace of X. Then, for every $\delta \in (0,1)$ there exists $x \in X$ such that $\|x\| = 1$ and $d(x, M) > \delta$, where*

$$d(x, Y) = \inf\{\|x - y\| \ \mid \ y \in Y\}.$$

Proof. Since M is a proper closed subspace of X, Hahn–Banach theorem shows that there exists a closed hyperplane H such that $M \subset H$. Let $f \in X'$, with $\|f\|' = 1$, be such that $H = \{y \in X \ \mid \ f(y) = 0\}$.

Fix $\delta \in (0,1)$. Then, there exists $x \in X$ such that $\|x\| = 1$ and $|f(x)| > \delta$. Consequently,

$$\delta < |f(x)| = |f(x-y)| \le \|f\|' \cdot \|x - y\| = \|x - y\|$$

for every $y \in H$, so that

$$\delta < |f(x)| \le \inf\{\|x - y\| \ \mid \ y \in H\} = d(x, H) \le d(x, M). \tag{3.6}$$

□

Remark 3.2.5. Let H be a closed hyperplane of X, i.e. $H = \{y \in X \ \mid \ f(y) = 0\} = \ker f$ for some $f \in X'$ with $\|f\|' = 1$. Fix $x \in X$ with $\|x\| = 1$. Then, $d(x, H) = 1$ if and only if $|f(x)| = 1$. Indeed, if $|f(x)| = 1$, then, using (3.6), we conclude that

$$1 \le d(x, H) \le \|x\| = 1.$$

Vice versa, let us suppose that $d(x, H) = 1$. For every fixed $z \in X \setminus H$, with $\|z\| = 1$, we set $y := x - \frac{f(x)}{f(z)} z$. Then, $f(y) = f(x) - \frac{f(x)}{f(z)} f(z) = 0$, i.e. $y \in H$ and

$$1 = d(x, H) \le \|x - y\| = \left| \frac{f(x)}{f(z)} \right|.$$

It follows that $|f(z)| \le |f(x)|$. The arbitrariness of z, with $\|z\| = 1$, yields that $1 = \|f\|' \le |f(x)| \le \|f\|'\|x\| = 1$, i.e. $|f(x)| = 1$.

Examples 3.2.6. We show by two examples that, in general, $f \in X'$ does not achieve its operator norm at any point of ∂B_X.

(1) Let $X = c_0$ and $f \in (c_0, \|\cdot\|_\infty)'$ be the functional defined by $f((y_n)_{n\in\mathbb{N}}) := \sum_{n=1}^\infty 2^{-n} y_n$ for every $(y_n)_{n\in\mathbb{N}} \in c_0$. Then,

$$|f(y)| \le \sum_{n=1}^\infty 2^{-n}|y_n| \le \|y\|_\infty \sum_{n=1}^\infty 2^{-n} = \|y\|_\infty$$

for every $y = (y_n)_{n\in\mathbb{N}} \in c_0$, This implies that $\|f\|' \le 1$. On the other hand, for every $k \in \mathbb{N}$ the sequence $y^{(k)} := (y_n^{(k)})_{n\in\mathbb{N}}$, where $y_n^{(k)} = 1$ if $n \le k$ and $y_n^{(k)} = 0$ otherwise, belongs to c_0 and

$$0 \le f(y^{(k)}) = \sum_{n=1}^{k} 2^{-n} \le \|f\|'\|y^{(k)}\|_\infty = \|f\|' .$$

Taking the limit as k tends to ∞, it follows that

$$1 = \sum_{n=1}^{\infty} 2^{-n} \le \|f\|'.$$

Consequently, $\|f\|' = 1$.

Let us now suppose that there exists $x = (x_n)_{n\in\mathbb{N}} \in c_0$, with $\|x\| = 1$, such that $|f(x)| = 1$. Since $x \in c_0$, there exists $k \in \mathbb{N}$ such that $|x_n| < 1/2$ for every $n \ge k$. It follows that

$$\begin{aligned} 1 =&|f(x)| = \left|\sum_{n=1}^{\infty} 2^{-n} x_n\right| \le \sum_{n=1}^{k-1} 2^{-n}|x_n| + \sum_{n\ge k} 2^{-n}|x_n| \\ &\le \sum_{n=1}^{k-1} 2^{-n} + \frac{1}{2}\sum_{n\ge k} 2^{-n} < \sum_{k=1}^{\infty} 2^{-k} = 1 , \end{aligned}$$

which is a contradiction.

(2) Let $X = C([0,1])$ and $f \in (C([0,1]), \|\cdot\|_\infty)'$ be functional defined by

$$f(x) = \int_0^{1/2} x(t)dt - \int_{1/2}^{1} x(t)dt$$

for every $x \in C([0,1])$. Arguing as in (1), it can be easily checked that there exists no $x \in C([0,1])$, with $\|x\|_\infty = 1$, such that $|f(x)| = 1$.

Remark 3.2.7. Let $(X, \|\cdot\|)$ be a reflexive Banach space and let $f \in X'$ be a functional with $\|f\|' = 1$. Then, there exists $x \in X$, with $\|x\| = 1$, such that $|f(x)| = 1$. Indeed, since $\|f\|' = \sup_{\|x\|=1} |f(x)| = 1$, we can determine a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ such that $\|x_n\| = 1$ and $\lim_{n\to\infty} |f(x_n)| = 1$. Recalling that the unit ball of X is weakly compact, up to a subsequence, $(x_n)_{n\in\mathbb{N}}$ weakly converges to some $x_0 \in X$. This implies that $\lim_{n\to\infty} |f(x_n)| = |f(x_0)| = 1$ and $\|x_0\| = 1$.

Actually, the above property characterizes the reflexivity of the Banach space X as the next theorem shows (for a proof we refer the reader to e.g., [4, p. 84]).

Theorem 3.2.8 (James' theorem). *Let $(X, \|\cdot\|)$ be a real or complex Banach space. Then, X is reflexive if and only if for every $f \in X'$, with $\|f\|' = 1$, there exists $x \in X$ such that $\|x\| = 1$ and $|f(x)| = 1$.*

We now introduce the ascent and descent indices of a bounded linear operator.
Let $S \in \mathcal{L}(X)$, where $(X, \|\cdot\|)$ is an infinite dimensional Banach space. It is easy to check that

$$\{0\} = N(S^0) \subseteq N(S) \subseteq N(S^2) \subseteq \ldots \subseteq N(S^k) \subseteq \ldots ,$$

where $S^0 = I$.

Definition 3.2.9. If $N(S^k) = N(S^{k+1})$ for some $k \in \mathbb{N} \cup \{0\}$, then

$$a(S) := \min\{k \in \mathbb{N} \cup \{0\} \ \mid \ N(S^k) = N(S^{k+1})\}$$

is called the *ascent index* of S.

Similarly, it holds that

$$X = S^0(X) \supseteq S(X) \supseteq S^2(X) \supseteq \ldots \supseteq S^n(X) \supseteq \ldots .$$

Definition 3.2.10. If $S^k(X) = S^{k+1}(X)$ for some $k \in \mathbb{N} \cup \{0\}$, then

$$d(S) := \min\{k \in \mathbb{N} \cup \{0\} \ \mid \ S^k(X) = S^{k+1}(X)\}$$

is called the *descent index* of S.

Remark 3.2.11. (1) If $a(S) = k_0$, then $N(S^k) = N(S^{k+1})$ for every $k \geq k_0$. Indeed, for every $k \geq k_0$ and $x \in N(S^{k+1})$, it holds that $0 = S^{k+1}(x) = S^{k_0+1}(S^{k-k_0}(x))$, so that $S^{k-k_0}(x) \in N(S^{k_0+1}) = N(S^{k_0})$. Consequently, $S^k(x) = S^{k_0}(S^{k-k_0}(x)) = 0$, i.e., $x \in N(S^k)$.

(2) If $d(S) = k_0$, then $S^k(X) = S^{k+1}(X)$ for every $k \geq k_0$. Indeed, if $k \geq k_0$ and $y \in S^k(X)$, then there exists $x \in X$ such that $y = S^k(x) = S^{k-k_0}(S^{k_0}(x))$, where $S^{k_0}(x) \in S^{k_0}(X) = S^{k_0+1}(X)$. Therefore, we can determine $x' \in X$ such that $S^{k_0}(x) = S^{k_0+1}(x')$. This implies that $y = S^k(x) = S^{k-k_0}(S^{k_0+1}(x')) = S^{k+1}(x') \in S^{k+1}(X)$.

Proposition 3.2.12. *Let $S \in \mathcal{L}(X)$ be such that $S^n(X)$ is a closed subspace of X for all $n \in \mathbb{N}$. If $a := a(S) < \infty$ and $d := d(S) < \infty$, then $a = d$. Moreover, the Banach space X splits into the* (*topological*) *direct sum*

$$X = N(S^m) \oplus S^m(X)\,, \tag{3.7}$$

where $m = a(S) = d(S)$.

Proof. Set $m := \max\{a, d\} \in \mathbb{N} \cup \{0\}$. Then, by Remark 3.2.11 we know that

$$S^k(X) = S^d(X) \qquad \text{and} \qquad N(S^k) = N(S^a) \tag{3.8}$$

for every $k \geq m$. Moreover, $N(S^m) \cap S^m(X) = \{0\}$. Indeed, if $x \in N(S^m) \cap S^m(X)$, then $S^m(x) = 0$ and $x = S^m(y)$ for some $y \in X$. It follows that $S^{2m}(y) = S^m(S^m(y)) = S^m(x) = 0$, i.e., $y \in N(S^{2m}) = N(S^m)$, by (3.8), so that $x = S^m(y) = 0$. Hence, $S^m_{|S^m(X)}$ is a one to one bounded linear operator mapping $S^m(X)$ into $S^m(X)$. This operator is also surjective since its image is $S^{2m}(X)$ which coincides with $S^m(X)$ (recall that $m \geq d$). Note that, by assumptions, $S^m(X)$ is a closed subset of X and hence $S^m_{|S^m(X)}$ is also an open linear operator from $S^m(X)$ onto $S^m(X)$ by the open mapping theorem.

Set $R := (S^m_{|S^m(X)})^{-1} \in \mathcal{L}(S^m(X))$. Then, the mapping $R \circ S^m \colon X \to S^m(X)$ satisfies the condition

$$(R \circ S^m)^2 = R \circ (S^m \circ R) \circ S^m = R \circ I \circ S^m = R \circ S^m\,,$$

i.e., $R \circ S^m$ is a continuous projection onto $S^m(X)$. We can thus split the Banach space X into the topological direct sum

$$X = N(R \circ S^m) \oplus (R \circ S^m)(X) = N(S^m) \oplus S^m(X) = N(S^a) \oplus S^d(X)\,, \tag{3.9}$$

where $N(R \circ S^m) = N(S^m)$ since R is one to one. From (3.9) we obtain that

$$S^a(X) = S^{a+d}(X) = S^d(X)\,,$$

and this implies that $a \geq d$.

Let us now prove that $d \geq a$. Since we have proved that $a \geq d$, it follows that $N(S^d) \subseteq N(S^a)$, so it remains to prove only that $N(S^a) \subseteq N(S^d)$. For this purpose, we fix $x \in N(S^a)$. Clearly $S^d(x) \in N(S^a) \cap S^d(X)$, hence $S^d(x) = 0$ by (3.9) and therefore $x \in N(S^d)$. □

Proposition 3.2.13. *For every $T \in \mathcal{K}(X)$ and $\lambda \in \mathbb{C} \setminus \{0\}$, it holds that $a(T_\lambda) = d(T_\lambda) < \infty$.*

Proof. Fix $\lambda \in \mathbb{C} \setminus \{0\}$. In view of Propositions 3.2.2 and 3.2.12 we just need to prove that $a(T_\lambda) < \infty$ and $d(T_\lambda) < \infty$.

By contradiction, let us suppose that $N(T_\lambda^{n-1})$ is a proper subspace of $N(T_\lambda^n)$ for every $n \in \mathbb{N}$ (where $T_\lambda^0 = I$). Applying Riesz's lemma, we can infer that for every $n \in \mathbb{N}$ there exists $x_n \in N(T_\lambda^n)$ such that $\|x_n\| = 1$ and $d(x_n, N(T_\lambda^{n-1})) \geq \frac{1}{2}$.

Note that

$$\begin{aligned} |\lambda|^{-1}\|T(x_n) - T(x_m)\| =& |\lambda|^{-1}\|\lambda x_n - T_\lambda(x_n) + T_\lambda(x_m) - \lambda x_m\| \\ =& \|x_n - (\lambda^{-1}T_\lambda(x_n) - \lambda^{-1}T_\lambda(x_m) + x_m)\| \geq \frac{1}{2} \end{aligned}$$

for every $n, m \in \mathbb{N}$, with $n > m$, since $\lambda^{-1}T_\lambda(x_n) - \lambda^{-1}T_\lambda(x_m) + x_m \in N(T_\lambda^{n-1})$. This implies that no convergent subsequence of $(T(x_n))_{n\in\mathbb{N}}$ exists, which is a contradiction, since $T \in \mathcal{K}(X)$ and $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence. We thus conclude that $a(T_\lambda) < \infty$.

In a completely similar way, we can show that $d(T_\lambda) < \infty$. □

By using the above result one obtains that any non zero element of the spectrum of a compact operator is an eigenvalue of that operator. This is the aim of the following result.

Proposition 3.2.14. *For every $T \in \mathcal{K}(X)$, the spectrum $\sigma(T)$ is the union of $\sigma_p(T)$ and $\{0\}$. Moreover, $\sigma(T)$ is at most countable and it admits at most zero as limit point.*

Proof. Fix $\lambda \in \mathbb{C} \setminus \{0\}$ such that $\lambda \notin \sigma_p(T)$. Then, $N(T_\lambda) = \{0\}$ and, consequently, $N(T_\lambda^k) = \{0\}$ for every $k \geq 0$. This means that $a(T_\lambda) = 0$. On the other hand, Proposition 3.2.13 shows that $d(T_\lambda) < \infty$. So, by Proposition 3.2.12 we infer that $d(T_\lambda) = a(T_\lambda) = 0$ and we conclude that $T_\lambda(X) = X$. So, by Proposition 1.3.2, one deduces that $\lambda \in \rho(T)$.

Let us now prove that $\sigma_p(T)$ is at most countable and that admits at most 0 as limit point. To prove these two properties, we will show that for every $\delta > 0$ the set $\{\lambda \in \sigma_p(T) \mid |\lambda| \geq \delta\}$ is finite. By contradiction, let us assume that $\{\lambda \in \sigma_p(T) \mid |\lambda| \geq \delta\}$ contains an infinite number of elements for some $\delta > 0$. Then, there exists a sequence $(\lambda_n)_{n\in\mathbb{N}} \subset \sigma_p(T)$ such that $\lambda_n \neq \lambda_m$ and $|\lambda_n| \geq \delta$ for every $n, m \in \mathbb{N}$. For every $n \in \mathbb{N}$ let us fix $x_n \in N(T_{\lambda_n})$ with $\|x_n\| = 1$. Note that the elements $x_1, \ldots, x_n$ are linearly independent for every $n \in \mathbb{N}$. Suppose that this is not the case. Then, there exists $(\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n \setminus \{0\}$ such that $\alpha_1 x_1 + \ldots + \alpha_n x_n = 0$. Up to reordering the elements $x_1, \ldots, x_n$, if necessary, we may assume that $x_1 = \beta_2 x_2 + \ldots + \beta_n x_n$, so that

$$\lambda_1 x_1 = T(x_1) = T(\beta_2 x_2 + \ldots \beta_n x_n) = \beta_2 \lambda_2 x_2 + \ldots + \beta_n \lambda_n x_n\,,$$

or, equivalently,

$$\beta_2(\lambda_2-\lambda_1)x_2+\ldots+\beta_n(\lambda_n-\lambda_1)x_n=0.$$

Iterating this procedure and possibly reordering the elements, we would deduce that $x_{n-1}=\gamma x_n$, with $\gamma\neq 0$ so that

$$\lambda_{n-1}x_{n-1}=T(x_{n-1})=\gamma T(x_n)=\gamma\lambda_n x_n=\lambda_n x_{n-1}\,,$$

which is a contradiction, since $\lambda_{n-1}\neq\lambda_n$ and $x_{n-1}\neq 0$. We have so proved that $x_1,\ldots,x_n$ are linearly independent.

For every $n\in\mathbb{N}$, let us set $M_n=\operatorname{span}\{x_1,\ldots,x_n\}$. By the first part of the proof, M_{n-1} is a proper closed subspace of M_n and, hence, in view of Riesz's lemma, we can find $y_n\in M_n$ such that $\|y_n\|=1$ and $d(y_n,M_{n-1})\geq\frac{1}{2}$. Consequently, $\|y_n-y_{n-1}\|\geq\frac{1}{2}$ for every $n\in\mathbb{N}$. Now, we observe that

$$\begin{aligned}\|T(y_n)-T(y_m)\|=&\|T(y_n)+\lambda_n y_n-\lambda_n y_n-T(y_m)\|\\=&\|\lambda_n y_n-(T_{\lambda_n}(y_n)-T(y_m))\|\end{aligned}\tag{3.10}$$

for every $n,m\in\mathbb{N}$, with $n>m$. Since $y_m\in M_m$, it follows that $y_m=\sum_{k=1}^m\alpha_k x_k$ so that

$$T(y_m)=\sum_{k=1}^m\alpha_k T(x_k)=\sum_{k=1}^m\alpha_k\lambda_k x_k\in M_m\subseteq M_{n-1}.$$

Moreover,

$$T_{\lambda_n}(y_n)=\sum_{k=1}^n\alpha_k T_{\lambda_n}(x_k)=\sum_{k=1}^n\alpha_k(\lambda_n-\lambda_k)x_k\in M_{n-1}.$$

Therefore, $T_{\lambda_n}(y_n)-T(y_m)\in M_{n-1}$ and from (3.10) we obtain that

$$\|T(y_n)-T(y_m)\|=|\lambda_n|\|y_n-\lambda_n^{-1}(T_{\lambda_n}(y_n)-T(y_m))\|\geq\frac{1}{2}\delta.$$

We have so proved that $(T(y_n))_{n\in\mathbb{N}}$ does not admit any convergent subsequence. This is clearly a contradiction since $T\in\mathcal{K}(X)$ and $(y_n)_{n\in\mathbb{N}}$ is a bounded sequence. □

The following remark will be used in the proof of Proposition 3.2.16.

Remark 3.2.15. Let X and Y be two finite dimensional Banach spaces, with dimensions m ed n, respectively. Further, let $\{x_1,\ldots,x_m\}$ be a basis of X, with $\|x_k\|_X=1$ for every $k=1,\ldots,m$, let $\{f_1,\ldots,f_m\}$ be the dual basis[1] of $\{x_1,\ldots,x_m\}$ and let $\{y_1,\ldots,y_n\}$ be a basis of Y. Denote by $L\colon X\to Y$ the linear operator defined by

$$L(x)=\sum_{k=1}^m f_k(x)y_k,\quad x\in X,$$

where we set $y_j=0$ for every $j=n+1,\ldots,m$, if $m>n$. The operator L belongs to $\mathcal{L}(X,Y)$ since

$$\|L(x)\|_Y\leq\|x\|_X\cdot\sum_{k=1}^m\|f_k\|_X'\|y_k\|_Y$$

for every $x\in X$. Moreover, since $L(X)=\operatorname{span}(y_1,\ldots,y_m)$ it follows that

[1] i.e., $f_i(x_j)=\delta_{ij}$ for every $i,j=1,\ldots,m$, where δ_{ij} is the Kronecker delta.

- if $m \le n$ then L is one to one,
- if $m \ge n$ then L is surjective.

Proposition 3.2.16. *Let $T \in \mathcal{K}(X)$. Then, for every $\lambda \in \mathbb{C} \setminus \{0\}$ it holds that*

$$\dim N(T_\lambda) = \operatorname{codim} T_\lambda(X) < \infty\,.$$

Proof. By Proposition 3.2.1, we know that $\dim N(T_\lambda) < \infty$. This guarantees that there exists a bounded linear projection $P : X \to X$ such that $P(X) = N(T_\lambda)$ so that we can split X into the direct sum

$$X = N(T_\lambda) \oplus Y\,, \tag{3.11}$$

where $Y = \ker P$. On the other hand, Propositions 3.2.2 and 3.2.3, show that $T_\lambda(X)$ is a closed subspace of X with finite codimension. Therefore, we can determine a linear and bounded projection $Q\colon X \to X$ such that $Q(X) = T_\lambda(X)$ and, consequently, we can split X into the sum

$$X = Z \oplus T_\lambda(X)\,, \tag{3.12}$$

where $Z = \ker Q$ is a closed subspace of X, with $\dim Z = \operatorname{codim} T_\lambda(X)$ (see Appendix A.3).

Let us consider the operator L in Remark 3.2.15, with $X = N(T_\lambda)$ and $Y = Z$. Since Z is a finite dimensional space, the operator L is compact as well as the operator $S = L \circ P$.

Note that the operator $A := T + S$ is compact and $A_\lambda = \lambda - (T + S) = T_\lambda - S$. In particular, the decomposition (3.11) shows that

$$A_\lambda(y) = T_\lambda(y) - S(y) = T_\lambda(y) - L(P(y)) = T_\lambda(y) - L(0) = T_\lambda(y)$$

for every $y \in Y$. This implies that $(A_\lambda)_{|Y} = (T_\lambda)_{|Y}\colon Y \to T_\lambda(X)$ is an isomorphism. Moreover, (3.11) also shows that

$$A_\lambda(x) = T_\lambda(x) - S(x) = -S(x) = -L(Px) = -L(x)$$

for every $x \in N(T_\lambda)$ so that $(A_\lambda)_{|N(T_\lambda)} = -L$.

In view of Remark 3.2.15, the operator L is injective or surjective. We will prove that, actually, L is both injective and surjective, so that the spaces $N(T_\lambda)$ and Z have the same dimension.

Let us assume that L is injective. Then, the operator A_λ is injective as well, since $N(A_\lambda) \subseteq N(L) \cap N(T_\lambda)$. Indeed, if $x \in N(A_\lambda)$ then $T_\lambda(x) = S(x) = L(P(x)) \in Z \cap T_\lambda(X)$. Applying the decomposition (3.12) we can infer that $T_\lambda(x) = L(P(x)) = 0$, i.e., $x \in N(T_\lambda)$ and $x = P(x) \in N(L)$. Therefore, $x \in N(L) \cap N(T_\lambda)$. Recalling that A is a compact operator, we deduce that $d(A_\lambda) = a(A_\lambda) = 0$, so that A_λ is also surjective. This implies that $(A_\lambda)_{|N(T_\lambda)} = -L$ is an isomorphism from $N(T_\lambda)$ into Z.

Let us now assume that L is surjective. Then, the operator A_λ is surjective as well since $A_\lambda(Y) = T_\lambda(X)$ and $A_\lambda(N(T_\lambda)) = L(N(T_\lambda)) = Z$ and $X = Z \oplus T_\lambda(X)$. The compactness of the operator A implies that $a(A_\lambda) = d(A_\lambda) = 0$, so that A_λ is also injective. In particular, $(A_\lambda)_{|N(T_\lambda)} = -L$ is injective and, consequently, is an isomorphism from $N(T_\lambda)$ into Z. □

Corollary 3.2.17. *Let $T \in \mathcal{K}(X)$. Then, for every $\lambda \in \mathbb{C} \setminus \{0\}$ and $n \in \mathbb{N}$ it holds that*

$$\operatorname{codim} T^n_\lambda(X) = \dim N((T^n)'_\lambda) = \operatorname{codim} (T^n)'_\lambda(X') = \dim N(T^n_\lambda)\,.$$

Proof. The assertion when $n = 1$ follows from observing that the dual operator T' is compact as well due to Schauder's Theorem and

$$(X/T_\lambda(X))' \text{ is isometric to } N(T'_\lambda),$$

Since the above spaces have finite dimension, we deduce that

$$\operatorname{codim} T_\lambda(X) = \dim(X/T_\lambda(X)) = \dim(X/T_\lambda(X))' = \dim N(T'_\lambda)\,.$$

The other equalities follow from Proposition 3.2.16.

To prove the assertion when $n > 1$, it suffices to observe that $T_\lambda^n = \lambda^n - S$, where $S \in \mathcal{K}(X)$. □

The previous proposition shows that every compact operator satisfies the *Fredholm alternative*, i.e.,

Theorem 3.2.18 (Fredholm alternative). *Let $T : X \to X$ be a compact operator. Then, for every $\lambda \in \mathbb{C} \setminus \{0\}$ one and only one of the following conditions is satisfied:*

(i) *for every $y \in X$ the equation $\lambda x - Tx = y$ is solvable with a unique solution $x \in X$;*

(ii) *the equation $\lambda x - Tx = 0$ admits a finite number of linearly independent solutions. In this latter case, the equation $\lambda x - Tx = y$ is solvable if and only if $x'_j(y) = 0$ for every $j = 1, \dots, n$, where $\{x'_1, \dots, x'_n\}$ is a basis of $N(T'_\lambda)$.*

Proof. If condition (i) is not satisfied, then λ belongs to the spectrum of T, i.e. it belongs to $N(T_\lambda)$, which is a finite dimensional subspace of X (see Propositions 3.2.1 and 3.2.14). Hence, the first part of property (ii) follows. On the other hand, $T_\lambda(X) = {}^\perp N(T'_\lambda)$, where $N(T'_\lambda)$ is finite dimensional too. Therefore, the equation $\lambda x - Tx = y$ is solvable if and only if $x'(y) = 0$ for every $x' \in N(T'_\lambda)$, and the second part of property (ii) follows. □

Notes

For further details, we refer the reader to the classical monographs [3, 7, 8, 10].

3.3 Exercises

Exercise 3.3.1. Let X and Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. Prove that the following assertions are equivalent.

(1) $T \in \mathcal{K}(X, Y)$.

(2) For every sequence $(x_n)_{n\in\mathbb{N}} \subset B_X$, $(T(x_n))_{n\in\mathbb{N}}$ admits a convergent subsequence.

(3) $T(B_X)$ is totally bounded, i.e. for any $\varepsilon > 0$, there is a finite subset M of Y such that $T(B_X) \subseteq \bigcup_{y\in M}(y + \varepsilon B_Y)$.

Exercise 3.3.2. Prove that, if $S, T : X \to Y$ are compact operators and $\lambda \in \mathbb{K}$, then $S + T$ and λS are compact operators as well.

Exercise 3.3.3. Prove that a Banach space X is finite dimensional if and only if its dual space X' is finite dimensional.

Exercise 3.3.4. Complete the proof of Corollary 3.2.3.

Exercise 3.3.5. Complete the proof of Proposition 3.2.13, showing that $d(T_\lambda) < \infty$.

Exercise 3.3.6. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two Banach spaces and $T\colon X \to Y$ be a compact linear operator. Show that if $(x_n)_n \subset X$ is a weakly convergent sequence to $x \in X$, then $(Tx_n)_n$ converges to Tx in $(Y, \|\cdot\|_Y)$.

Exercise 3.3.7. Let $(a_{j,k})_{j,k\in\mathbb{N}} \subset \mathbb{C}$ satisfy $\sum_{j,k=1}^{\infty} |a_{j,k}|^2 < \infty$. Show that the linear operator $T\colon \ell^2 \to \ell^2$ defined by

$$Tx := \left(\sum_{k=1}^{\infty} a_{j,k} x_k \right)_{j\in\mathbb{N}}, \qquad x = (x_k)_{k\in\mathbb{N}} \in \ell^2,$$

is compact.

Exercise 3.3.8. Let X be a Banach space. Prove that $T(X) = {}^{\perp}N(T')$ for every operator $T \in \mathcal{L}(X)$.

Exercise 3.3.9.

1. Let $p, q \in (1, \infty)$ be such that $\frac{1}{q} + \frac{1}{p} = 1$ and let $\Omega \subseteq \mathbb{R}^n$ be a measurable set. For $k \in L^q(\Omega \times \Omega)$ define the linear operator $(Tf)(x) := \int_\Omega k(x,y) f(y)\, dy$ for any $f \in L^p(\Omega)$. Prove that T is a bounded and compact operator from $L^p(\Omega)$ into $L^q(\Omega)$.

2. Here we assume that $p = 2$, $\Omega = (0,1)$ and $k(x,y) = \min\{x,y\}$.

   (i) Prove that T is a self-adjoint operator and
   $$\|T\|_{\mathcal{L}(L^2(0,1))} \le \|k\|_{L^2((0,1)\times(0,1))} = \frac{1}{\sqrt{6}}.$$

   (ii) Deduce, for any $g \in L^2((0,1))$, the integral equation
   $$f(x) - 2\int_0^1 \min\{x,y\} f(y)\, dy = g(x), \qquad \text{for a.e. } x \in (0,1)$$
   has a unique solution $f \in L^2((0,1))$.

   (iii) Verify that $f(x) = \sin\left(\frac{\pi}{2}x\right)$ is an eigenfunction of T and compute the corresponding eigenvalue λ_0.

   (iv) Show that $r(T) = \lambda_0$ and deduce that $\|T\|_{\mathcal{L}(L^2(0,1))} < \|k\|_{L^2((0,1)\times(0,1))}$.
   [**Hint:** Use Proposition 2.3.5 to infer that $r(T) = \|T\|_{\mathcal{L}(L^2((0,1))}$ and observe that, if $\lambda \in \mathbb{C}$ is an eigenvalue of T, then there exists $f \in C^2([0,1])$ such that $\lambda f'' = -f$ and $f(0) = f'(1) = 0$.]

Appendix A

Functional Analytic Tools

In this appendix we collect some fundamental results from Functional Analysis which we use in this edition of the Internet Seminar.

A.1 Hahn-Banach Theorem

Given a Banach space X, one of the most important questions one may ask about its dual space X' is the following: are there 'enough' elements in X'? For example, are there enough elements to separate points? This is answered using the Hahn–Banach theorem and its consequences that we state here below.

Theorem A.1.1 (Hahn-Banach theorem for normed spaces)**.** *Let X be a real or complex normed space and Y be a non-trivial subspace of X. Then for any $f \in Y'$ there exists $\tilde{f} \in X'$ such that $\tilde{f}(y) = f(y)$ for all $y \in Y$ and $\|\tilde{f}\|' = \|f\|'$.*

Many useful results follow from the Hahn–Banach theorem A.1.1. For example:

Corollary A.1.2 (Separation)**.** *Let X be a real or complex normed space. Then for any $x \in X$ there exists a linear functional $f \in X'$ such that $f(x) = \|x\|$ and $\|f\|' = 1$. Hence, if $x \neq y \in X$ then there exists $f \in X'$ such that $f(x) \neq f(y)$.*

Corollary A.1.3. *Let X be a Banach space and D be a subspace of X. If*

$$\forall f \in X' : \quad (f(x) = 0 \ \forall x \in D) \Longrightarrow \ f = 0,$$

then D is dense in X.

A.2 Open Mapping Theorem

We recall that a continuous map between two topological spaces has the property that the pre-image of any open set is open, but in general the image of an open set is not open. In the case of bounded linear maps between Banach spaces this special property is satisfied under the assumption of the surjectivity of the map.

Theorem A.2.1 (Open Mapping theorem)**.** *Let X and Y be two Banach spaces and let $T : X \to Y$ be a surjective and bounded linear operator. Then, T is an open operator, i.e., T maps open sets in X onto open sets in Y.*

The assumption that T maps X onto Y is essential. Consider, for example, the projection $(x, y) \mapsto (x, 0)$ defined on $\mathbb{R}^2$. This map is not open.

As an application of Theorem A.2.1 a general property of inverse maps can be deduced, i.e.:

Proposition A.2.2. *Let X and Y be two Banach spaces and let T be a bijective bounded linear operator from X to Y. Then the inverse T^{-1} is also a bounded linear operator from Y to X.*

A.3 Quotient spaces, topologically complemented subspaces and projections

Given a vector space E, a projection is a linear map $P : E \to E$ such that $P^2 = P$. In this case, setting $E_1 = P(E)$ and $E_2 = \ker(P)$, the pair (E_1, E_2) is a complemented pair of subspaces of E, namely $E_1 + E_2 = E$ and $E_1 \cap E_2 = E$. This is equivalent to say that every $x \in E$ has a unique decomposition as $x = x_1 + x_2$ with $x_i \in E_i$ $(i = 1, 2)$.

Vice versa, given a complemented pair (E_1, E_2) of subspaces of E, setting $Px = x_1$, when $x = x_1 + x_2$ with $x_i \in E_i$, $(i = 1, 2)$, it is immediate to prove that P is a projection and that $E_1 = P(E)$ and $E_2 = \ker P$.

If X is a Banach space and P is a continuous projection, then the pair of complemented subspaces associated with P is said to be topologically complemented. According to the definition, the subspaces forming a pair of topologically complemented subspaces are closed, being kernels of the continuous maps P and $I - P$.

In general a pair of closed subspaces which are algebraically complemented are not topologically complemented.

Nevertheless, closed subspaces of Hilbert spaces and finite dimensional subspaces in any Banach space are topologically complemented.

Let X be a Banach space and M be a closed subspace of X. Further, let $\varphi : X \to X/M$ be the quotient mapping defined by $\varphi(x) = x + M$ for every $x \in X$. The space X induces a norm on X/M, which is defined by

$$\|\varphi(x)\|_{X/M} = \inf\{\|x + y\| : y \in M\}.$$

X/M, endowed with the norm $\|\cdot\|_{X/M}$, is a Banach space and φ is a continuous and open mapping.

Moreover, the space $(X/M, \|\cdot\|_{X/M})'$ is isometric to $(M^\perp, \|\cdot\|')$, where

$$M^\perp := \{\, y' \in X' \mid y'(x) = 0 \text{ for every } x \in M \,\}.$$

If $\operatorname{codim} M{:=}\dim(X/M)$ is finite, then it can be proved that M is a topologically complemented subspace in X.

For a proof of the previous results, we refer the reader to [8, 1.40-42 and 4.8-9] or to [6, 4.2]

A.4 Uniform Boundedness

We recall the following useful result to establish the uniform boundedness of a family of continuous linear operators between two normed spaces.

Theorem A.4.1 (Banach–Steinhaus theorem)**.** *Let X be a Banach space and let Y be a normed space. Let $\{T_\alpha\}_{\alpha\in A}$ be a family of bounded linear operators from X to Y. Suppose that for each $x \in X$, the set $\{T_\alpha x\}_{\alpha\in A}$ is a bounded subset of Y. Then the set $\{\|T_\alpha\|\}_{\alpha\in A}$ is bounded, i.e., $\sup_{\alpha\in A}\|T_\alpha\| < \infty$.*

A consequence of the theorem above is the following fact.

Corollary A.4.2. *Let X be a Banach space and Y be a normed space. If a sequence $(T_n)_{n\in\mathbb{N}} \subset \mathcal{L}(X,Y)$ is strongly convergent (i.e., $T_n x$ converges for every $x \in X$), then there exists an operator $T \in \mathcal{L}(X,Y)$ such that $(T_n)_{n\in\mathbb{N}}$ is strongly convergent to T.*

A.5 Riesz-Fréchet Theorem

A complex vector space H is said to be *an inner product space* (or *unitary space*) if there exists an *inner product* (or *scalar product*), i.e. a map $\langle\cdot,\cdot\rangle : H \times H \to \mathbb{C}$ with the following properties:

(a) $\langle y, x\rangle = \overline{\langle x, y\rangle}$ for every $x, y \in H$;

(b) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ and $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ for every $x, y, z \in H$ and $\alpha \in \mathbb{C}$;

(c) $\langle x, x\rangle \geq 0$ for every $x \in H$;

(d) $\langle x, x\rangle = 0$ if and only if $x = 0$.

By the properties (a), (b), (c) and (d) above, we can define a *norm* on H by setting $\|x\|^2 := \langle x, x\rangle$, for $x \in H$. The following famous inequality holds true:

$$|\langle x, y\rangle| \leq \|x\|\|y\| \quad \text{(Cauchy-Schwarz Inequality)} \tag{A.1}$$

for all $x, y \in H$.

If the normed space $(H, \|\cdot\|)$ is complete, i.e., if every Cauchy sequence converges in H, then H is called a Hilbert space.

If $(H, \|\cdot\|)$ is a Hilbert space, then it is easy to show that the map $H \ni x \mapsto \langle x, y\rangle \in \mathbb{C}$ is a continuous linear functional on H for all $y \in H$. Indeed, for a fixed $y \in H$, set $f_y(x) := \langle x, y\rangle$ for every $x \in H$. Then, by the Cauchy-Schwarz Inequality it follows that

$$|f_y(x)| = |\langle x, y\rangle| \leq \|x\|\|y\|$$

for all $x \in H$. This means that $f_y \in H'$ and $\|f\|' \leq \|y\|$. Actually, $\|f\|' = \|y\|$ as $f_y(y) = \langle y, y\rangle = \|y\|^2$.

The Riesz-Fréchet Theorem shows that all continuous linear functionals on a Hilbert space H are of this type.

Theorem A.5.1 (Riesz-Fréchet Theorem)**.** *Let H be a Hilbert space and let $f \in H'$. Then, there exists a unique $y \in H$ such that $f(x) = \langle x, y\rangle$ for all $x \in H$. Moreover, $\|f\|' = \|y\|$.*

A.6 Characterization of finite dimensional spaces

Using Riesz's lemma, see Lemma 3.2.4, we can characterize finite dimensional vector spaces.

Theorem A.6.1. *Let X be a real or complex Banach space. Then, the following assertions are equivalent:*

(1) $\dim X < \infty$.

(2) *Every bounded and closed subset of X is compact.*

(3) *B_X is compact.*

Proof. (1) $\Longrightarrow$ (2): Let $\{e_1, e_2, \dots, e_m\}$ be a basis of X and $K \subset X$ be bounded and closed. Let $x_n := \sum_{j=1}^{m} x_{j,n} e_j \in K$ with $x_{j,n} \in \mathbb{K}$ for $1 \leq j \leq m$ and $n \in \mathbb{N}$, where $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$ depending on the fact that X is a real or complex Banach space. Since all norms in X are equivalent, it is sufficient to show that there exists a subsequence (x_{n_k}) such that $\lim_{k\to\infty} \|x_{n_k} - x\|_\infty = 0$ for some $x \in X$. Since the sequence $(|x_{j,n}|)_{n\in\mathbb{N}}$ is bounded in $\mathbb{K}$ for all $1 \leq j \leq m$, there exists a subsequence $(x_{n_k})_k$ with $\lim_{k\to\infty} x_{j,n_k} = x_j$ for all $j = 1, \dots m$. Hence, $\lim_{k\to\infty} \|x_{n_k} - x\|_\infty = 0$, where $x := \sum_{j=1}^{m} x_j e_j$.

(2) $\Longrightarrow$ (3): Trivial.

(3) $\Longrightarrow$ (1): Assume by contradiction that $\dim X = \infty$. Then, for every $n \in \mathbb{N}$ there exists a subspace X_n of X such that

$$X_1 \subsetneq X_2 \subsetneq \dots \subsetneq X \text{ with } \dim X_n = n.$$

Since X_n is a finite dimensional space, it follows that X_n is a closed (proper) subspace of X_{n+1}. Applying Riesz's lemma, see Lemma 3.2.4, we conclude that there exists $x_n \in X_{n+1}$ with $\|x_n\| = 1$ and $d(x_n, X_n) \geq \frac{1}{2}$. This implies that the sequence $(x_n)_{n\in\mathbb{N}} \subset B_X$ satisfies the estimate $\|x_n - x_m\| \geq \frac{1}{2}$ for all $n > m$. Thus, $(x_n)_{n\in\mathbb{N}}$ cannot admit any convergent subsequence, which is a contradiction. □

A.7 Arzelà-Ascoli Theorem

The following theorem was obtained in 1895 by Arzelà for subsets of $C([a,b])$. A weaker form had been discovered earlier by Ascoli. For this reason, the result is often referred to as the *Arzelà–Ascoli theorem.*

Theorem A.7.1 (Arzelà–Ascoli Theorem)**.** *Let K be a compact space and let $(C(K), \|\cdot\|_\infty)$ denote the Banach space of all complex valued continuous functions on K. Let $\Phi \subset C(K)$ be a equibounded and equicontinuous family, i.e., Φ satisfies the following properties:*

(a) $\sup\{|f(x)| : f \in \Phi\} < \infty$ *for every $x \in K$;*

(b) *for every ε and $x \in K$ there exists a neighborhood V of x such that $|f(x) - f(y)| < \varepsilon$ for all $y \in V$ and $f \in \Phi$.*

Then, Φ is a relatively compact subset of $C(K)$, i.e., every sequence in Φ contains a uniformly convergent subsequence.

Notes

For the proof of all the above results we refer the reader to [8, 9].

Bibliography

[1] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[2] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[3] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[4] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[5] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[6] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981

[7] M. Reed, B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[8] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[9] W. Rudin, Real and Complex Analysis, 3rd ed., McGraw-Hill, 1986.

[10] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

November 10, 2021

Lecture 4

Spectral representation theorem for bounded operators I

One of the main results in Linear Algebra states that a linear self-adjoint operator on a finite dimensional vector space can be diagonalized, namely there exists a basis that allows to represent the operator as a diagonal operator. With this formulation, the result does not have an immediate counterpart in the general contest of infinite dimensional spaces. But if we give a different interpretation, namely we look at the finite dimensional space as a space $L^2(\mu)$, where μ is the counting measure on a finite set and at a diagonal operator as a multiplication operator, then the previous result can be restated in the following way: a linear self-adjoint operator on a finite dimensional space is unitarily equivalent to a multiplication operator.
With this perspective, we can extend the spectral theorem for self-adjoint operators to the infinite dimensional case.
We will first focus on compact operators, that somehow represent the link between the finite dimensional and the infinite dimensional case, since compact self-adjoint operators are proved to be unitarily equivalent to a diagonal operator on ℓ^2.
Further, we will concentrate on bounded self-adjoint operators. By combining Weierstrass theorem and Riesz Representation theorem, we will first give a meaning to the expression $f(T)$, where T is a self-adjoint bounded operator and f is a continuous function acting on the spectrum of T, and we introduce the so-called spectral measures associated with T. They will allow to prove that a self-adjoint operator acting on a complex separable Hilbert space is unitarily equivalent to a multiplication operator defined on a space $L^2(M, \mu)$, where (M, μ) is a measure space and μ is finite.

Throughout the lecture, $(H, \|\cdot\|)$ is a Hilbert space over $\mathbb{C}$ with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$.

4.1 Spectral decomposition theorem for self-adjoint compact operators

In this section, we study the behaviour of self-adjoint compact operators defined on Hilbert spaces.

Let us begin with the following fundamental remark.

Remark 4.1.1. Let $T \in \mathcal{L}(H)$ be a self-adjoint compact operator. Then, $\sigma(T) = \sigma_p(T) \cup \{0\} \subset \mathbb{R}$ due to Propositions 2.3.5 and 3.2.14. Since $r(T) = \max\{|\lambda| \mid \lambda \in \sigma(T)\} = \|T\|$, we can infer that $\|T\| \in \sigma_p(T)$ or $-\|T\| \in \sigma_p(T)$.

Theorem 4.1.2 (Spectral representation theorem 1)**.** *Let $T \in \mathcal{L}(H)$ be a self-adjoint compact operator. Then, there exists a closed and separable subspace H_0 of H, a complete orthonormal system $(x_n)_{n\in J} \subset H_0$, where $J = \mathbb{N}$ or J is a finite subset of $\mathbb{N}$, and $(\lambda_n)_{n\in J}$, which is an infinitesimal sequence if $J = \mathbb{N}$, such that*

$$Tx = \begin{cases} \sum_{n\in J} \lambda_n \langle x, x_n\rangle x_n & \text{if } x \in H_0, \\ 0, & \text{if } x \in H_0^{\perp}. \end{cases} \tag{4.1}$$

Proof. Since $T \in \mathcal{K}(H)$ is self-adjoint, we can apply Proposition 3.2.14 and Remark 4.1.1 to infer that $\sigma(T) = \sigma_p(T) \cup \{0\} \subset \mathbb{R}$ and $\sigma_p(T) = (\alpha_n)_{n\in J}$, where J is either $\mathbb{N}$ or a finite subset of $\mathbb{N}$. In the first case, $(\alpha_n)_{n\in\mathbb{N}}$ belongs to c_0.

To fix the ideas, we assume that $J = \mathbb{N}$. The proof in the other case is completely similar.

Set $H_n := N(T_{\alpha_n})$ for every $n \in \mathbb{N}$. Then, $\langle x, y\rangle = 0$ for every $x \in H_n$ and $y \in H_m$, with $n \neq m$, since

$$\alpha_n\langle x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \alpha_m\langle x, y\rangle\,,$$

and $\alpha_n \neq \alpha_m$. Moreover, $H_n \neq \{0\}$ and it is finite dimensional, due to Proposition 3.2.1. Therefore, we can determine a finite orthonormal basis $\mathcal{B}_n = \{x_1^n, \ldots x_{j_n}^n\}$ in H_n. As a byproduct, the countable set $\bigcup_{n\in\mathbb{N}} \mathcal{B}_n$ is a complete orthonormal system of $H_0 := \overline{\operatorname{span}(\bigcup_{n\in\mathbb{N}} \mathcal{B}_n)}$.

Next, we observe that $T(H_0^{\perp}) \subset H_0^{\perp}$. Indeed, if $x \in H_0^{\perp}$, then

$$\langle x_k^n, Tx\rangle = \langle Tx_k^n, x\rangle = \alpha_n\langle x_k^n, x\rangle = 0$$

for every $n \in \mathbb{N}$ and every $k = 1, \ldots, j_n$, and hence $Tx \in H_0^{\perp}$. Therefore, $T_{|H_0^{\perp}}$ is a self-adjoint compact operator, mapping $H_0^{\perp}$ into $H_0^{\perp}$. We claim that $T_{|H_0^{\perp}} = 0$. If this were not the case, Remark 4.1.1 would guarantee the existence of $\lambda \in \mathbb{R} \setminus \{0\}$, $x_0 \in H_0^{\perp}$, $x_0 \neq 0$, such that $Tx_0 = \lambda x_0$. We are led to a contradiction since all the eigenvectors of T belong to H_0.

Finally, we observe that, if $x \in H_0$, then $x = \sum_{n=1}^{\infty}\sum_{k=1}^{j_n}\langle x, x_k^n\rangle x_k^n$ so that

$$Tx = \sum_{n=1}^{\infty}\sum_{k=1}^{j_n}\langle x, x_k^n\rangle Tx_k^n = \sum_{n=1}^{\infty}\sum_{k=1}^{j_n}\alpha_n\langle x, x_k^n\rangle x_k^n.$$

This completes the proof, choosing $\{x_n \mid n \in \mathbb{N}\} = \bigcup_{k\in\mathbb{N}} \mathcal{B}_k$ and $\lambda_n = \alpha_k$ if $x_n \in \mathcal{B}_k$. □

Remark 4.1.3. From formula (4.1) it follows immediately that $(\lambda_n)_{n\in\mathbb{N}} \subset \sigma_p(T)$ and $x_n \in N(T_{\lambda_n})$ for every $n \in \mathbb{N}$.

If H is a separable Hilbert space, then the spectral representation theorem 4.1.2 can be reformulated as follows (see Exercise 4.3.1).

Theorem 4.1.4. *Let H be a separable Hilbert space over $\mathbb{C}$ and let $T \in \mathcal{L}(H)$ be a self-adjoint compact operator. Then, there exist a complete orthonormal system $(e_n)_{n\in\mathbb{N}} \subset H$ and an infinitesimal sequence $(\mu_n)_{n\in\mathbb{N}} \subset \mathbb{R}$, such that*

$$Tx = \sum_{n\in\mathbb{N}} \mu_n\langle x, e_n\rangle e_n \tag{4.2}$$

for every $x \in X$.

Remark 4.1.5. If H is separable, then the operator T is equivalent to a diagonal operator on the space ℓ^2. Indeed, if we denote by U the operator

$$U\colon H \to \ell^2, \qquad x \mapsto (\langle x, e_n\rangle)_{n\in\mathbb{N}}\,,$$

then $T = U^{-1}MU$, where M is the diagonal operator on ℓ^2 defined by $M(\xi) = (\mu_n \xi_n)_{n\in\mathbb{N}}$ for every $\xi = (\xi_n)_{n\in\mathbb{N}} \in \ell^2$.

The same spectral representation can be obtained for more general compact operators on Hilbert spaces as the following result shows.

Theorem 4.1.6 (Spectral representation theorem 2)**.** *Let $T \in \mathcal{L}(H)$ be a compact operator. Then, there exist a closed and separable subspace H_0 of H, a complete orthonormal system $(x_n)_{n\in J} \subset H_0$, where $J = \mathbb{N}$ or J is a finite subset of $\mathbb{N}$, an orthonormal system $(y_n)_{n\in J} \subset H$ and $(\lambda_n)_{n\in J}$, which is an infinitesimal sequence, if $J = \mathbb{N}$, such that*

$$Tx = \begin{cases} \sum_{n\in J} \lambda_n \langle x, x_n\rangle y_n, & \text{if } x \in H_0,\\ 0, & \text{if } x \in H_0^{\perp}\,. \end{cases} \tag{4.3}$$

Proof. Observe that the operator T^*T is compact and self-adjoint. Hence, the spectral representation theorem 4.1.2 guarantees the existence of a separable closed subspace H_0 of H, of a set J which is either $\mathbb{N}$ or a finite subset of $\mathbb{N}$, of a complete orthonormal system $(x_n)_{n\in J} \subset H_0$ and of $\{\mu_n\}_{n\in J} \subset \mathbb{R}$, which is an infinitesimal sequence if $J = \mathbb{N}$, such that

$$T^*Tx = \begin{cases} \sum_{n\in J} \mu_n \langle x, x_n\rangle x_n, & \text{if } x \in H_0,\\ 0, & \text{if } x \in H_0^{\perp}\,, \end{cases} \tag{4.4}$$

where $\sigma_p(T^*T) \supset (\mu_n)_{n\in J}$ and $x_n \in N((T^*T)_{\mu_n})$ for $n \in J$.

It follows that $T_{|H_0^{\perp}} = 0$. Indeed, from the proof of Theorem 4.1.2 we know that T^*T identically vanishes on $H_0^{\perp}$. Hence, $\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle = \langle 0, x\rangle = 0$ for every $x \in H_0^{\perp}$, i.e., $Tx = 0$.

Next, we observe that, if $x \in N((T^*T)_{\mu_n})$, then

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

which yields that $\mu_n \geq 0$. Thus, we can set $\lambda_n = \sqrt{\mu_n}$. Clearly, $(\lambda_n)_{n\in J} \in c_0$ if $J = \mathbb{N}$. If $x \in H_0$, then $x = \sum_{n\in J} \langle x, x_n\rangle x_n$ and, therefore,

$$Tx = \sum_{n\in J} \langle x, x_n\rangle Tx_n = \sum_{n=1}^{\infty} \lambda_n \langle x, x_n\rangle y_n$$

where $y_n = \lambda_n^{-1} Tx_n$ if $\lambda_n \neq 0$ and $y_n = 0$ if $\lambda_n = 0$. To conclude, we just need to show that the set of all $y_n \neq 0$ is an orthonormal system. For this purpose, it suffices to observe that

$$\begin{aligned} \langle y_n, y_m\rangle =& \lambda_n^{-1}\lambda_m^{-1} \langle Tx_n, Tx_m\rangle = \lambda_n^{-1}\lambda_m^{-1} \langle T^*Tx_n, x_m\rangle \\ =& \mu_n \lambda_n^{-1}\lambda_m^{-1} \langle x_n, x_m\rangle = \lambda_n \lambda_m^{-1} \langle x_n, x_m\rangle = \delta_{m,n} \end{aligned}$$

for every $n, m \in J$ such that $y_n, y_m \neq 0$, where $\delta_{m,n}$ is the Kronecker delta. □

4.2 Spectral representation theorem for bounded and self-adjoint operators on a separable Hilbert space

In order to state and prove the spectral representation theorem for bounded and self-adjoint operators it is necessary to introduce an appropriate functional calculus, which will allow us to define $f(T)$ when $T \in \mathcal{L}(H)$ and f is a continuous function defined on the spectrum of T.

When f is a polynomial, the operator $f(T)$ is defined in the obvious way as the following definition shows.

Definition 4.2.1. Let $(X, \|\cdot\|)$ be a complex Banach space and $T \in \mathcal{L}(X)$. If $P(z) = \sum_{n=0}^{N} a_n z^n$ is a complex polynomial, then we define $P(T) := \sum_{n=0}^{N} a_n T^n$.

Lemma 4.2.2 (Spectral mapping theorem). *Let $(X, \|\cdot\|)$ be a complex Banach space, $T \in \mathcal{L}(X)$ and P be a complex polynomial. Then*

$$\sigma(P(T)) = \{P(\lambda) \mid \lambda \in \sigma(T)\} = P(\sigma(T)) \,. \tag{4.5}$$

Proof. Fix $\lambda \in \sigma(T)$. Then, λ is a root of the polynomial $P - P(\lambda)$ and hence, by the fundamental theorem of algebra, $P(\lambda) - P(x) = (\lambda - x)Q(x) = Q(x)(\lambda - x)$ for every $x \in \mathbb{C}$, Q being a suitable polynomial. Accordingly, $P(\lambda) - P(T) = (\lambda - T)Q(T) = Q(T)(\lambda - T)$. Since $\lambda - T$ does not admit an inverse, it follows that $P(\lambda) - P(T)$ does not admit an inverse too. Otherwise, $Q(T)(P(\lambda) - P(T))^{-1}$ would be a bounded inverse of $\lambda - T$ (see Exercise 4.3.3). Hence, $P(\lambda) \in \sigma(P(T))$.

Vice versa, fix $\mu \in \sigma(P(T))$ and let $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ be the roots of the polynomial $P - \mu$, some of which possibly repeated according to its multiplicity. Then, $P(z) - \mu = a(z - \lambda_1)\cdots(z - \lambda_n)$ and, hence,

$$P(T) - \mu I = a(T - \lambda_1)\cdots(T - \lambda_n).$$

If $\lambda_1, \ldots, \lambda_n \notin \sigma(T)$, then

$$(P(T) - \mu)^{-1} = a^{-1}(T - \lambda_n)^{-1}\cdots(T - \lambda_1)^{-1},$$

which is a contradiction with the assumption $\mu \in \sigma(P(T))$. So, $\lambda_i \in \sigma(T)$ for some $i \in \{1, \ldots, n\}$, that is $\mu = P(\lambda_i)$ and, hence, $\mu \in P(\sigma(T))$. □

Lemma 4.2.3. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator and P be a complex polynomial. Then,*

$$\|P(T)\| = \sup_{\lambda \in \sigma(T)} |P(\lambda)| \,.$$

Proof. Since T is a self-adjoint operator, $P(T)$ is a normal operator (see Exercise 4.3.2). So, we can apply Proposition 2.3.5 to conclude that $\|P(T)\| = r(P(T))$. Now, from Lemma 4.2.2 it follows that

$$\begin{aligned}\|P(T)\| =& r(P(T)) = \sup\{|\lambda| : \lambda \in \sigma(P(T))\}\\ =& \sup\{|P(\lambda)| : \lambda \in \sigma(T)\}\\ =& \sup_{\lambda \in \sigma(T)} |P(\lambda)| \,.\end{aligned}$$

□

Lemma 4.2.3 allows us to extend the functional calculus from polynomials to the space $C(\sigma(T),\mathbb{C})$ of all $\mathbb{C}$-valued continuous functions on the spectrum of T. We recall that an operator $T \in \mathcal{L}(H)$ is said to be positive if $\langle Tx, x\rangle \ge 0$ for all $x \in H$.

Theorem 4.2.4. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator. Then, there exists a unique linear mapping*

$$\Phi\colon C(\sigma(T),\mathbb{C}) \to \mathcal{L}(H)$$

satisfying the following properties for every $f, g \in C(\sigma(T),\mathbb{C})$ and $\lambda \in \mathbb{C}$:

(1) $\Phi(fg) = \Phi(f)\Phi(g)$, $\Phi(\lambda f) = \lambda\Phi(f)$;

(2) $\Phi(1) = I$, $\Phi(\overline{f}) = \Phi(f)^*$;

(3) $\|\Phi(f)\| = \|f\|_\infty$;

(4) *if* $f = id_{\sigma(T)}$, *then* $\Phi(f) = T$;

(5) $\sigma(\Phi(f)) = \{f(\lambda) : \lambda \in \sigma(T)\}$;

(6) *if* $f \ge 0$, *then* $\Phi(T) \ge 0$;

(7) *if* $B \in \mathcal{L}(H)$ *commutes with* T, *then* B *also commutes with* $\Phi(f)$.

Proof. For each complex polynomial P we define $\Phi(P) := P(T)$. Since by Lemma 4.2.2 the identity $\|P(T)\| = \|P\|_{C(\sigma(T),\mathbb{C})}$ holds, Φ is an isometric linear mapping from the space $(\mathcal{P}_{|\sigma(T)}, \|\cdot\|_\infty)$ of the restrictions to $\sigma(T)$ of all complex polynomials into the space $(\mathcal{L}(H), \|\cdot\|)$. Therefore, Φ extends uniquely to an isometric linear mapping $\tilde{\Phi}$ from the completion of $(\mathcal{P}_{|\sigma(T)}, \|\cdot\|_\infty)$ into $(\mathcal{L}(H), \|\cdot\|)$. By the Weierstrass approximation theorem (see Theorem C.1.13), the completion of $(\mathcal{P}_{|\sigma(T)}, \|\cdot\|_\infty)$ is the space $(C(\sigma(T),\mathbb{C}), \|\cdot\|_\infty)$ because $\sigma(T) \subseteq \mathbb{R}$. To enlighten the notation, we denote such an extension again by Φ.

Properties (1) and (2) are clearly satisfied whenever f and g are polynomials. Hence, they easily extend to the case when f and g belong to $C(\sigma(T),\mathbb{C})$ by using a density argument. Property (4) is obvious.

We now show property (5), which is a proper analogue of Lemma 4.2.3. So, fix $\mu \in f(\sigma(T))$. Then $\mu = f(\lambda)$ for some $\lambda \in \sigma(T)$ and the following identity is satisfied

$$\mu - \Phi(f) = \Phi(f(\lambda)) - \Phi(f) = \Phi(f(\lambda) - f)\,.$$

For a fixed $\varepsilon > 0$, we choose a function $g_\varepsilon \in C(\sigma(T),\mathbb{C})$ such that $g_\varepsilon(\lambda) = 1$, $\|g_\varepsilon\|_\infty = 1$ and

$$|[f(\eta) - f(\lambda)]g_\varepsilon(\eta)| < \varepsilon$$

for all $\eta \in \sigma(T)$, so that $\|(f - f(\lambda))g_\varepsilon\|_\infty \le \varepsilon$ (see Exercise 4.3.4. On the other hand, since $\|\Phi(g_\varepsilon)\| = \|g_\varepsilon\|_\infty = 1$, there exists $x_\varepsilon \in H$ with $\|x_\varepsilon\| \le 1$ such that $\|\Phi(g_\varepsilon)(x_\varepsilon)\| \ge 1/2$. If we set $y_\varepsilon = \Phi(g_\varepsilon)(x_\varepsilon)$, then $\|y_\varepsilon\| \ge 1/2$ and

$$\begin{aligned}\|(\mu - \Phi(f))(y_\varepsilon)\| =& \|\Phi(f(\lambda) - f)\Phi(g_\varepsilon)(x_\varepsilon)\|\\ =& \|\Phi((f(\lambda) - f)g_\varepsilon)(x_\varepsilon)\|\\ \le& \|\Phi((f(\lambda) - f)g_\varepsilon)\|\|x_\varepsilon\| \le \varepsilon.\end{aligned}$$

This means that the operator $\mu I - \Phi(f)$ is not invertible in $\mathcal{L}(H)$ and, hence, $\mu = f(\lambda) \in \sigma(\Phi(T))$.

Vice versa, fix $\mu \notin f(\sigma(T))$. Then, the function $h(\lambda) := (\mu - f(\lambda))^{-1}$, for $\lambda \in \sigma(T)$, is clearly continuous on $\sigma(T)$. Moreover, by properties (1) and (2) it follows that

$$\begin{aligned}(\mu - \Phi(f))\Phi(h) =&\Phi(\mu - f)\Phi(h) = \Phi(h)\Phi(\mu - f)\\ =&\Phi(h(\mu - f)) = \Phi(id_{\sigma(T)}) = I\,,\end{aligned}$$

thereby implying that the operator $\mu - \Phi(f)$ is invertible in $\mathcal{L}(H)$ with inverse $(\mu - \Phi(f))^{-1} = \Phi(h)$. Thus, $\mu \notin \sigma(\Phi(f))$. Since $\mu \notin f(\sigma(T))$ is arbitrary, we can conclude that $(f(\sigma(T)))^c \subseteq (\sigma(\Phi(f))^c$, that is $\sigma(\Phi(f)) \subseteq f(\sigma(T))$.

In order to show property (6), we observe that if $f \geq 0$ on $\sigma(T)$, then $f = g^2$ for some real-valued continuous function g. By properties (1) e (2) it follows that $\Phi(f) = \Phi(g^2) = \Phi(g)^2$, where $\Phi(g)$ is a self-adjoint operator due to property (2). Therefore, for every $x \in H$, it holds that

$$\langle \Phi(f)(x), x\rangle = \langle \Phi(g)^2(x), x\rangle = \langle \Phi(g)(x), \Phi(g)(x)\rangle = \|\Phi(g)(x)\|^2 \geq 0\,,$$

i.e., $\Phi(f) \geq 0$.

Property (7) is immediate to prove. Indeed, if B commutes with T, then B clearly commutes with $P(T)$ for every complex polynomial P. By density it follows that B also commutes with $\Phi(f)$.

It remains to show the uniqueness of the mapping Φ. To this end it suffices to observe that if there would exist another linear mapping $\Psi\colon C(\sigma(T), \mathbb{C}) \to \mathcal{L}(H)$ satisfying the properties (1), (2), (3) e (4), then $\Psi(P) = \Phi(P)$ for every complex polynomial P and, hence, $\Psi \equiv \Phi$ by density. □

In the rest of the section, we will use also the notation $f(T)$ instead of $\Phi(f)$ when we want to point out the dependence on T.

Proposition 4.2.5. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator. Then, for each pair of vectors $x, y \in H$ there exists a unique positive Borel measure $\mu_{x,y}$ on the compact set $\sigma(T)$ such that*

$$\langle f(T)(x), y\rangle = \int_{\sigma(T)} f d\mu_{x,y}$$

for every $f \in C(\sigma(T), \mathbb{C})$. The family $\{\mu_{x,y}\}_{x,y\in H}$ satisfies the following properties for every $x, y \in H$:

(1) $|\mu_{x,y}|(\sigma(T)) \leq \|x\|\|y\|$;

(2) $\mu_y := \mu_{y,y}$ *is a finite positive Borel measure on* $\sigma(T)$ *with* $\mu_y(\sigma(T)) = \|y\|^2$;

(3) $\mu_{x,y} = \overline{\mu_{y,x}}$;

(4) $\mu_{f(T)(x),y} = f\mu_{x,y} = \mu_{x,\overline{f}(T)(y)}$ *for all* $f \in C(\sigma(T), \mathbb{C})$;

(5) $\mu_{\alpha x_1 + x_2, y} = \alpha\mu_{x_1,y} + \mu_{x_2,y}$ *and* $\mu_{x,\alpha y_1 + y_2} = \overline{\alpha}\mu_{x,y_1} + \mu_{x,y_2}$ *for all* $x, y, x_i, y_i \in H$ *for* $i = 1, 2$ *and* $\alpha \in \mathbb{C}$.

The measure μ_y is said to be the spectral measure associated with the vector y.

Proof. For fixed $x, y \in H$, consider the functional $L_{x,y}\colon C(\sigma(T), \mathbb{C}) \to \mathbb{C}$ defined by setting

$$L_{x,y}(f) := \langle f(T)(x), y\rangle$$

for every $f \in C(\sigma(T), \mathbb{C})$. Then, $L_{x,y}$ is clearly linear because of the linearity of x and the inner product. Moreover, using the fact that x is an isometry, it follows that

$$|L_{x,y}(f)| = |\langle f(T)x, y\rangle| \leq \|f(T)\|\|x\|\|y\| = \|f\|_\infty\|x\|\|y\|$$

for all $f \in C(\sigma(T), \mathbb{C})$. Hence, $L_{x,y}$ is a continuous linear functional on $C(\sigma(T), \mathbb{C})$ with $\|L_{x,y}\|' \leq \|x\|\|y\|$. By the Riesz representation theorem there exists a unique regular Borel measure $\mu_{x,y}$ on $\sigma(T)$ such that

$$L_{x,y}(f) = \int_{\sigma(T)} f d\mu_{x,y}$$

for every $f \in C(\sigma(T), \mathbb{C})$, with $|\mu_{x,y}|(\sigma(T)) = \|L_{x,y}\|' \leq \|x\|\|y\|$. So, (1) is proved.

For each $y \in H$ we can estimate

$$|\mu_y|(\sigma(T)) \leq \|y\|^2 = \langle\Phi(1)y, y\rangle = \int_{\sigma(T)} 1 d\mu_y = \mu_y(\sigma(T)) \leq |\mu_y|(\sigma(T)),$$

because $\Phi(1) = I$. This implies that $\mu_y(\sigma(T)) = |\mu_y|(\sigma(T)) = \|y\|^2$. Since μ_y and its total variation agree on $\sigma(T)$, μ_y is a finite positive measure on $\sigma(T)$. So, also property (2) is established.

To show property (3) we observe that

$$\int_{\sigma(T)} f d\mu_{x,y} = \langle f(T)(x), y\rangle = \overline{\langle \overline{f}(T)(y), x\rangle} = \overline{\int_{\sigma(T)} \overline{f} d\mu_{y,x}} = \int_{\sigma(T)} f d\overline{\mu_{y,x}}$$

for every $f \in C(\sigma(T), \mathbb{C})$. By uniqueness it follows that $\mu_{x,y} = \overline{\mu_{y,x}}$. So, (3) is proved.

To prove property (4) we proceed as follows. We fix $f \in C(\sigma(T), \mathbb{C})$ and observe that

$$\int_{\sigma(T)} g d\mu_{f(T)(x),y} = \langle g(T)f(T)(x), y\rangle = \langle (gf)(T)(x), y\rangle = \int_{\sigma(T)} gf d\mu_{x,y}$$

for every $g \in C(\sigma(T), \mathbb{C})$. Again, by uniqueness, it follows that $\mu_{f(T)(x),y} = f\mu_{x,y}$. This fact together with property (3) yields that

$$\mu_{x,\overline{f}(T)(y)} = \overline{\mu_{\overline{f}(T)(y),x}} = \overline{\overline{f}\mu_{y,x}} = f\mu_{x,y}.$$

It remains to establish property (5). For fixed $\alpha \in \mathbb{C}$ and $x_i \in H$ $(i = 1, 2)$ and every $f \in C(\sigma(T), \mathbb{C})$ we have that

$$\begin{aligned}\int_{\sigma(T)} f d\mu_{\alpha x_1+x_2,y} =&\langle f(T)(\alpha x_1 + x_2), y\rangle = \alpha\langle f(T)(x_1), y\rangle + \langle f(T)(x_2), y\rangle \\ =&\alpha \int_{\sigma(T)} f d\mu_{x_1,y} + \int_{\sigma(T)} f d\mu_{x_2,y} = \int_{\sigma(T)} f d(\alpha\mu_{x_1,y} + \mu_{x_2,y}).\end{aligned}$$

By uniqueness, it follows that $\mu_{\alpha x_1+x_2,y} = \alpha\mu_{x_1,y} + \mu_{x_2,y}$. In a similar way one shows that $\mu_{x,\alpha y_1+y_2} = \overline{\alpha}\mu_{x,y_1} + \mu_{x,y_2}$. □

Remark 4.2.6. Let $T \in \mathcal{L}(H)$ be a self-adjoint operator and let $\{\mu_{x,y}\}_{x,y\in H}$ be the family of Borel measures on $\sigma(T)$ constructed in Proposition 4.2.5. Then, the family $\{\mu_{x,y}\}_{x,y\in H}$ allows us to extend the functional calculus also to the class of all bounded Borel measurable functions on $\sigma(T)$. Indeed, fix a Borel and bounded function $f : \sigma(T) \to \mathbb{C}$ and consider the functional $a_f \colon H \times H \to \mathbb{C}$ defined by setting

$$a_f(x,y) := \int_{\sigma(T)} f d\mu_{x,y}$$

for every $x, y \in H$. Then, by property (5) of Proposition 4.2.5 the form a_f is sesquilinear. Moreover, by property (1) in Proposition 4.2.5

$$|a_f(x,y)| \le \|f\|_\infty |\mu_{x,y}|(\sigma(T)) \le \|f\|_\infty \|x\|\|y\|$$

for all $x, y \in H$. This shows that a_f is a bounded sesquilinear form on H with a bound less than or equal to $\|f\|_\infty$.

Now, the Riesz Fréchet theorem permits to construct $f(T)$. Indeed, for a fixed $x \in H$, the operator $H \ni y \mapsto a_f(x,y) \in \mathbb{C}$ is a continuous anti-linear functional on H. So, there exists a unique $\xi_x \in H$, with $\|\xi_x\| = \|a_f(x,\cdot)\|'$, such that

$$a_f(x,y) = \langle \xi_x, y\rangle$$

for every $y \in H$. Therefore, we can define $f(T)(x) := \xi_x$.

By repeating this process for each $x \in H$, we can define an operator $f(T) \colon H \to H$, by setting $f(T)(x) = \xi_x$ for every $x \in H$. The operator $f(T)$ is linear. Indeed, for fixed $\alpha \in \mathbb{C}$, $x_i \in H$ $(i = 1, 2)$ and every $y \in H$ we can write

$$\begin{aligned}\langle f(T)(\alpha x_1 + x_2), y\rangle =& \langle \xi_{\alpha x_1 + x_2}, y\rangle = a_f(\alpha x_1 + x_2, y)\\ =&\alpha a_f(x_1,y) + a_f(x_2,y) = \alpha\langle \xi_{x_1}, y\rangle + \langle \xi_{x_2}, y\rangle\\ =&\alpha\langle f(T)(x_1), y\rangle + \langle f(T)(x_2), y\rangle = \langle \alpha f(T)(x_1) + f(T)(x_2), y\rangle;\end{aligned}$$

hence, $f(T)(\alpha x_1 + x_2) = \alpha f(T)(x_1) + f(T)(x_2)$. Moreover, for every $x \in H$ we can estimate

$$\|f(T)(x)\| = \|\xi_x\| = \|a_f(x,\cdot)\|' = \sup_{\|y\|\le 1} |a_f(x,y)| \le \|f\|_\infty \|x\|.$$

This means that $f(T)$ is also continuous with operator norm $\|f(T)\| \le \|f\|_\infty$. Therefore, $f(T) \in \mathcal{L}(H)$.

We point out that, by construction,

$$\langle f(T)(x), y\rangle = \int_{\sigma(T)} f d\mu_{x,y}$$

for all $x, y \in H$.

The functional calculus defined above continues to satisfy the same properties listed in Theorem 4.2.4 (see e.g. [8, Theorem VII.2]).

Definition 4.2.7. Fix $T \in \mathcal{L}(H)$. A vector $x \in H$ is said to be a *cyclic vector* for T if $\overline{\text{span}\{T^n(x) :\ n \in \mathbb{N} \cup \{0\}\}} = H$.

Lemma 4.2.8. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator. If there exists a cyclic vector x for T, then there exists a unitary operator $U\colon H \to L^2(\sigma(T), \mu_x)$ such that*

$$(UTU^{-1})(f)(\lambda) = \lambda f(\lambda) \quad \mu_x - a.e.$$

for all $f \in L^2(\sigma(T), \mu_x)$.

Proof. For each $f \in C(\sigma(T), \mathbb{C})$ we set

$$U\Phi(f)(x) := f\,,$$

where Φ is the mapping constructed in the proof of Theorem 4.2.4. Then, U is well defined in the space $\{\Phi(f)(x) : f \in C(\sigma(T), \mathbb{C})\}$. Indeed, if $f, g \in C(\sigma(T), \mathbb{C})$ are two functions for which $\Phi(f)(x) = \Phi(g)(x)$, it follows via properties (1), (2) and (4) in Theorem 4.2.4 that

$$\begin{aligned}\Phi(f)T^n(x) =& \Phi(f)\Phi(\lambda^n)(x) = \Phi(f\lambda^n)(x) = \Phi(\lambda^n f)(x)\\ =& \Phi(\lambda^n)\Phi(f)(x) = \Phi(\lambda^n)\Phi(g)(x) = \Phi(g)T^n(x)\,,\end{aligned}$$

i.e., $\Phi(f) = \Phi(g)$ on the dense subspace $\text{span}\{T^n(x) : n \in \mathbb{N} \cup \{0\}\}$ of H. Thanks to the continuity of $\Phi(f)$ and of $\Phi(g)$ we deduce that $\Phi(f) = \Phi(g)$ on the whole space H. So, it follows that

$$0 = \|\Phi(f) - \Phi(g)\| = \|\Phi(f - g)\| = \|f - g\|_\infty\,,$$

i.e., $f \equiv g$. Moreover, by properties (1) and (2) of Theorem 4.2.4 it follows that

$$\begin{aligned}\|\Phi(f)(x)\|^2 =& \langle x, \Phi(f)^*\Phi(f)(x)\rangle = \langle x, \Phi(\overline{f}f)(x)\rangle\\ =& \langle \Phi(\overline{f}f)(x), x\rangle = \int_{\sigma(T)} |f|^2 d\mu_x\end{aligned}$$

for every $f \in C(\sigma(T), \mathbb{C})$. This means that U is an isometry from the space $(\{\Phi(f)(x) : f \in C(\sigma(T), \mathbb{C})\}, \|\cdot\|)$ into $L^2(\sigma(T), \mu_x)$.

Since the space $\{\Phi(f)(x) : f \in C(\sigma(T), \mathbb{C})\}$ is dense in H, we can extend U to an isometry from H into $L^2(\sigma(T), \mu_x)$. On the other hand, the fact that $C(\sigma(T), \mathbb{C})$ is a dense subspace of $L^2(\sigma(T), \mu_x)$ implies that U is also surjective. At this point, we observe that

$$(UTU^{-1})(f)(\lambda) = (UT\Phi(f)(x))(\lambda) = (U\Phi(\lambda f)(x))(\lambda) = \lambda f(\lambda)$$

for all $f \in C(\sigma(T), \mathbb{C})$. This identity continues to be hold for all $f \in L^2(\sigma(T), \mu_x)$ because $C(\sigma(T), \mathbb{C})$ is a dense subspace of $L^2(\sigma(T), \mu_x)$. □

In order to extend such a result to any bounded and self-adjoint operator on a complex separable Hilbert space, we need an auxiliary result. Before we state it, we recall that, if $(H_n)_{n\in J}$ is a countable family of Hilbert spaces, then by $X = \bigoplus_{n\in J} H_n$ we mean the set of all the elements $x = (x_n)_{n\in J}$ such that $x_n \in H_n$ for every $n \in J$ and

$$\|x\|_X^2 := \sum_{n\in J} \|x_n\|_{H_n}^2 < \infty.$$

Clearly, if J is finite set, then X coincides with the cartesian product of the spaces H_n.

Lemma 4.2.9. *Let H be a complex separable Hilbert space and let $T \in \mathcal{L}(H)$ be a self-adjoint operator. Then, there exists a family $\{H_n\}_{n\in J}$ of closed subspaces of H, with $J \subseteq N$ finite or infinite, such that*

(1) $H = \bigoplus_{n\in J} H_n$;

(2) *if* $x \in H_n$, *then* $T(x) \in H_n$;

(3) *for every* $n \in J$ *the operator* $T_{|H_n}$ *admits a cyclic vector* $x_n \in H_n$.

Proof. Let $\{e_n\}_{n\in\mathbb{N}}$ be a complete orthonormal system of H. If we set $x_1 := e_1$ and $H_1 := \overline{\text{span}\{T^n(x_1) : n \in \mathbb{N} \cup \{0\}\}}$, then H_1 is a T-invariant subspace of H and x_1 is a cyclic vector for $T_{|H_1}$.

If $e_n \in H_1$ for all $n \in \mathbb{N}$, then $H_1 = H$ and, hence, the proof is complete. Otherwise, let n_1 be the first index for which $e_{n_1} \notin H_1$; this means that $e_n \in H_1$ for every $n < n_1$. Then, we denote by $P_{H_1^\perp}$ the orthogonal projection from H onto the closed subspace $H_1^\perp$ and we set $x_2 := P_{H_1^\perp}(e_{n_1})$. We observe that $x_2 \neq 0$ because $e_{n_1} \notin H_1$. Moreover, since T is a self-adjoint operator and T maps H_1 into itself, T also maps $H_1^\perp$ into itself. Indeed, for a fixed $h \in H_1^\perp$, we have that $\langle T(h), k\rangle = \langle h, T(k)\rangle = 0$ for all $k \in H_1$. This implies that $T(h) \in H_1^\perp$. If we set $H_2 := \overline{\text{span}\{T^n(x_2) : n \in \mathbb{N} \cup \{0\}\}}$, then it follows that $H_2 \subset H_1^\perp$. In particular, H_2 is a T-invariant subspace of H and x_2 is a cyclic vector for $T_{|H_2}$.

If $H = H_1 \oplus H_2$, then the proof is complete. Otherwise, let n_2 be the first index for which $e_{n_2} \notin H_1 \oplus H_2$. Denoted by $P_{(H_1\oplus H_2)^\perp}$ the orthogonal projection from H onto the closed subspace $(H_1 \oplus H_2)^\perp$, we set $x_3 := P_{(H_1\oplus H_2)^\perp}(e_{n_2})$ and we proceed as before. After a finite number of steps, we say N, we could get $H = H_1 \oplus H_2 \oplus \ldots H_N$, where for every $i = 1, \ldots, N$ the closed subspace H_i is T-invariant and $T_{|H_i}$ admits a cyclic vector. Otherwise, we will have a family $\{H_i\}_{i\in\mathbb{N}}$ of T-invariant and orthogonal closed subspaces of H such that every $T_{|H_i}$ admits a cyclic vector x_i. In each case, by construction, $e_n \in H_1$ for every $n < n_1$, $e_n \in H_1 \oplus H_2$ for every $n_1 \leq n < n_2$, and so on. This implies that $(e_n)_{n\in\mathbb{N}} \subset \bigoplus_{i\in\mathbb{N}} H_i$ from which it follows that $H = \bigoplus_{i\in\mathbb{N}} H_i$. □

Thanks to the previous lemmas, we can now prove the following result.

Theorem 4.2.10. *Let H be a complex separable Hilbert space and let $T \in \mathcal{L}(H)$ be a self-adjoint operator. Then, there exist a finite or infinite set $J \subseteq \mathbb{N}$, a family $\{\mu_n\}_{n\in J}$ of measures on $\sigma(T)$ and an unitary operator*

$$U : H \to \bigoplus_{n\in J} L^2(\sigma(T), d\mu_n)$$

such that

$$(UTU^{-1}(\psi))_n(\lambda) = \lambda\psi_n(\lambda) \qquad \mu_n - a.e.$$

for all $\psi = (\psi_n)_{n\in J} \in \bigoplus_{n\in J} L^2(\sigma(T), d\mu_n)$ *and for all* $n \in J$.

Proof. The result follows by applying first Lemma 4.2.9 to find the decomposition of H and then Lemma 4.2.8 to each component of the decomposition. So, the n-th measure μ_n is the spectral measure associated with the n-th cyclic vector. In particular, μ_n is a measure on $\sigma(T_{|H_n})$, but we can extend it to the whole compact set $\sigma(T)$ by setting $\mu_n \equiv 0$ on $\sigma(T) \setminus \sigma(T_{|H_n})$. □

We now can state and prove the spectral theorem in its classical formulation.

Theorem 4.2.11 (Spectral theorem for bounded self-adjoint operators)**.** *Let H be a complex separable Hilbert space and $T \in \mathcal{L}(H)$ be a self-adjoint operator. Then, there exist a measurable space (M, μ) with finite positive measure, a bounded and measurable function $m\colon M \to \mathbb{R}$ and an unitary operator $U\colon H \to L^2(M, d\mu)$ such that*

$$(UTU^{-1}(f))(\eta) = m(\eta)f(\eta) \quad \mu - a.e.$$

for all $f \in L^2(M, d\mu)$.

Proof. By Lemma 4.2.9 we can write $H = \bigoplus_{n\in J} H_n$ with $J \subseteq \mathbb{N}$ finite or infinite, where $\{H_n\}_{n\in J}$ is an orthogonal family of closed T-invariant subspaces of H, such that the operator $T_{|H_n}$ admits a cyclic vector x_n for all $n \in J$. We can always suppose that $\|x_n\| = 2^{-n}$ for all $n \in J$. We now denote by μ_n the spectral (positive) measure on $\sigma(T)$ associated with x_n. As in the proof of Theorem 4.2.10, we observe that μ_n is a measure on $\sigma(T_{|H_n})$, but we can extend it to the whole compact set $\sigma(T)$ by setting $\mu_n \equiv 0$ on $\sigma(T) \setminus \sigma(T_{|H_n})$. On the other hand, by Lemma 4.2.8, for every $n \in J$ there exists an unitary operator $U_n\colon H_n \to L^2(\sigma(T), \mu_n)$ such that

$$(U_n T U_n^{-1}(\psi_n))(\lambda) = \lambda\psi_n(\lambda) \quad \mu_n - \text{a.e.}$$

for all $\psi_n \in L^2(\sigma(T), \mu_n)$.

Set $M := J \times \sigma(T)$. Then, we say that $E \subseteq M$ is μ–measurable if $E_n := \{\lambda \in \sigma(T) : (n, \lambda) \in E\}$ is μ_n–measurable for all $n \in J$. In such a case, we define $\mu(E) := \sum_{n\in J} \mu_n(E_n)$. We observe that $\mu(M) = \sum_{n\in J} \mu_n(\sigma(T)) = \sum_{n\in J} \|x_n\|^2 = \sum_{n\in J} 2^{-n} < \infty$. This implies that the positive measure μ constructed above is finite. Moreover, if $f \in L^2(M, d\mu)$, then

$$\int_M |f|^2 d\mu = \sum_{n\in J} \int_{\sigma(T)} |f(n, \lambda)|^2 d\mu_n(\lambda)\,.$$

Now, let us consider the operator

$$U\colon H = \bigoplus_{n\in J} H_n \to L^2(M, d\mu)$$

defined by

$$g = \sum_{n\in J} g_n \mapsto U(g)(n, \lambda) := U_n(g_n)(\lambda)\,.$$

Then U is an unitary operator and satisfies

$$(UTU^{-1}(f))(n, \lambda) = \lambda f(n, \lambda)$$

for all $f \in L^2(M, d\mu)$, that is $m(n, \lambda) = \lambda$. □

Theorem 4.2.11 shows that every self-adjoint bounded operator is unitarily equivalent to a multiplication operator on a suitable L^2-space. In the following example, for a particular self-adjoint operator T we contruct "by hands" a multiplication operator to which T is unitarily equivalent.

Example 4.2.12. Let $H = \ell^2(\mathbb{Z})$ be the Hilbert space of all $\mathbb{C}$-valued sequences $x = (x_n)_{n\in\mathbb{Z}}$ such that $\|x\|^2 = \sum_{n\in\mathbb{Z}} |x_n|^2 < \infty$. We denote by $L\colon H \to H$ the left translation operator, defined by $(L(x))_n := x_{n+1}$ for all $n \in \mathbb{Z}$, and by $R\colon H \to H$ the right translation operator, defined by $(R(x))_n := x_{n-1}$ for all $n \in \mathbb{Z}$ (the operator R has been already introduced in Example 2.3.11-(1)). It is easy to verify that $L^* = R$ and $R^* = L$ and, hence, the operator $T := L + R$ is self-adjoint.

We now define the operator $U\colon \ell^2(\mathbb{Z}) \to L^2((0,1))$ by setting

$$U(x) := \sum_{n\in\mathbb{Z}} x_n e^{2\pi i n x}, \qquad x \in (0,1).$$

The sequence $(e^{2\pi i n x})_{n\in\mathbb{Z}}$ of functions forms a complete orthonormal system of $L^2((0,1))$. This implies the surjectivity of U and the fact that U preserves the norm. Therefore, U is an unitary operator.

Finally, we observe that ULU^{-1} and URU^{-1} are the multiplication operators by the functions $e^{-2\pi i x}$ and $e^{2\pi i x}$ respectively. So, it follows that UTU^{-1} is the multiplication operator by the function $2\cos(2\pi x)$.

Notes

We suggest the reading of the interesting classical paper by Halmos [5] for further information on the spectral theorem.

4.3 Exercises

Exercise 4.3.1. Prove Theorem 4.1.4.

Exercise 4.3.2. If T is a normal operator on a complex Hilbert space and P is a complex polynomial, prove that $P(T)$ is a normal operator too.

Exercise 4.3.3. Let P and Q be two complex polynomials, Prove that, if $P(T)$ is invertible, then $(P(T))^{-1}$ and $Q(T)$ commute. Further, prove that, if also $Q(T)$ is invertible then its inverse operator commutes with $(P(T))^{-1}$.

Exercise 4.3.4. Given a function $f \in C(K, \mathbb{C})$, where K is a compact set of $\mathbb{R}$, $\varepsilon > 0$ and $\lambda \in K$, prove that there exists a continuous function $g_\varepsilon : K \to \mathbb{C}$ such that $g_\varepsilon(\lambda) = 1$, $\|g_\varepsilon\|_\infty = 1$ and $|(f(\lambda) - f(\eta))g_\varepsilon(\eta)| \le \varepsilon$ for every $\eta \in K$.

Exercise 4.3.5. Prove that a self-adjoint operator on a finite dimensional Hilbert space has a cyclic vector if and only if all the eigenspaces have dimension one.

Exercise 4.3.6. Let S and T be two non trivial compact self-adjoint operators on a infinite dimensional Hilbert space H such that $ST = TS$. Prove that S and T have the same sequence of orthogonal eigenvectors.

Exercise 4.3.7. Let T be a compact self-adjoint operator on a infinite dimensional separable Hilbert space H. Prove that there exists an orthonormal basis $(x_n)_{n\in\mathbb{N}}$ of H such that each x_n is an eigenvector corresponding to some real eigenvalue λ_n and for every $\lambda \notin \sigma(T)$ we have

$$R(\lambda, T)x = \sum_{n=1}^{\infty} \frac{\langle x, x_n\rangle}{\lambda - \lambda_n} x_n, \qquad x \in H.$$

Exercise 4.3.8. Let $T \in \mathcal{L}(H)$ be a self-adjoint operator on the Hilbert space H, $f \in C(\sigma(T), \mathbb{R})$ and $g \in C(f(\sigma(T)), \mathbb{C})$. Prove that $g(f(T)) = (g \circ f)(T)$.

Appendix C

Stone Weierstrass Theorem and Riesz representation theorem

C.1 Stone-Weierstrass theorem

Definition C.1.1. Let X be a not empty set and let A be a subspace of the space of all $\mathbb{R}$-valued (resp. $\mathbb{C}$-valued) functions on X. The space A is said to be a real (resp. complex) functional algebra if for every $f, g \in A$ also fg belongs to A. A subspace B of A is said to be a subalgebra of A if B is an algebra.

Example C.1.2. Let X be a topological space and let $C(X, \mathbb{R})$ be the space of all $\mathbb{R}$-valued continuous functions on X. Then $C(X, \mathbb{R})$ is a real functional algebra. The space $C_b(X, \mathbb{R})$ of all real bounded and continuous functions on X is a subalgebra of $C(X, \mathbb{R})$.

As it is well-known, $C_b(X, \mathbb{R})$ endowed with the norm

$$\|f\|_\infty := \sup_{x \in X} |f(x)|, \qquad f \in C_b(X, \mathbb{R}),$$

is a Banach space.

Definition C.1.3. Let X be a not empty set and let $f, g : X \to \mathbb{R}$ be two functions on X. We denote by $f \vee g$ and $f \wedge g$ the functions defined by setting

$$(f \vee g)(x) := \max\{f(x),\ g(x)\}, \qquad (f \wedge g)(x) := \min\{f(x), g(x)\} \qquad x \in X\,.$$

Remark C.1.4. We point out that

$$f \vee g = \frac{f + g + |f - g|}{2}, \qquad f \wedge g = \frac{f + g - |f - g|}{2}\,,$$

as it is easy to prove. Thus, if f and g are two $\mathbb{R}$-valued continuous functions on a topological space X, then $f \vee g$ and $f \wedge g$ are also $\mathbb{R}$-valued continuous functions on X.

For what follows it is useful to recall the following notation

$$\binom{\alpha}{n} = \frac{\alpha(\alpha - 1)\ldots(\alpha - n + 1)}{n!}, \qquad \alpha \in \mathbb{R},\ n \in \mathbb{N},$$

and the following result.

Lemma C.1.5. *If $\alpha > 0$, then for every $x \in [-1,1]$,*

$$\sum_{n=0}^{\infty} \binom{\alpha}{n} x^n = (1+x)^\alpha$$

and the convergence of the series is total on $[-1,1]$.

Proof. As it is well-known, the binomial series absolutely converges in $]-1,1[$ and totally converges in every compact interval $[a,b] \subset]-1,1[$. The absolute convergence of the binomial series for $x=-1$ and for $x=1$ follows from Raabe criterion, after having observed that

$$\lim_{n\to+\infty} n \left(\frac{|\binom{\alpha}{n}|}{|\binom{\alpha}{n+1}|} - 1 \right) = 1+\alpha\,. \quad \square$$

Theorem C.1.6. *Let X be a topological space and let A be a subalgebra of $C_b(X,\mathbb{R})$. Then the closure $\overline{A}$ of A in $(C_b(X,\mathbb{R}), \|\cdot\|_\infty)$ is a subalgebra of $C_b(X,\mathbb{R})$. Moreover, if $f, g \in \overline{A}$, then $|f|$, $f \vee g$ and $f \wedge g \in \overline{A}$.*

Proof. Let $f, g \in \overline{A}$ be fixed. Then there exists two sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ in A such that $\|f_n - f\|_\infty \to 0$ and $\|g_n - g\|_\infty \to 0$ as n tends to ∞. Since A is an algebra, $(f_n+g_n)_{n\in\mathbb{N}}$, $(\alpha f_n)_{n\in\mathbb{N}}$, $(f_n g_n)_{n\in\mathbb{N}}$ are all contained in A for every $\alpha\in\mathbb{R}$. From the following inequalities

$$\begin{aligned}
&\|f_n + g_n - (f+g)\|_\infty \le \|f_n - f\|_\infty + \|g_n - g\|_\infty \\
&\|\alpha f_n - \alpha f\|_\infty = |\alpha| \|f_n - f\|_\infty \qquad (\alpha\in\mathbb{R}) \\
&\|f_n g_n - fg\|_\infty = \|f(g-g_n) + g_n(f-f_n)\|_\infty \\
&\qquad\qquad\qquad \le \|f\|\|g-g_n\| + \|g_n\|\|f-f_n\|
\end{aligned}$$

it follows that $f_n + g_n \to f+g$, $\alpha f_n \to \alpha f$ and $f_n g_n \to fg$ in $(C_b(X,\mathbb{R}), \|\cdot\|_\infty)$ as n tends to ∞, taking into account that $(\|g_n\|)_{n\in\mathbb{N}}$ is a bounded sequence. Thus, also $f+g$, αf and fg belongs to $\overline{A}$. Accordingly, $\overline{A}$ is a subalgebra of $C_b(X,\mathbb{R})$.

Let us now fix $0 \neq f \in \overline{A}$. In order to show that $|f| \in \overline{A}$, we proceed as follows.

We observe that by Lemma C.1.5 we have

$$\sqrt{t} = [1+(t-1)]^{\frac{1}{2}} = \sum_{k=0}^{\infty} \binom{\frac{1}{2}}{k} (t-1)^k, \qquad t\in[0,1]$$

and the convergence of the series is uniform on [0,1].

If we set $p_n(t) := \sum_{k=0}^n \binom{\frac{1}{2}}{k}(t-1)^k$, for $t\in[0,1]$, then p_n is a polynomial of degree n and $\lim_{n\to+\infty} p_n(0) = 0$. Thus, the polynomial $q_n := p_n - p_n(0)$ is homogeneous and $\lim_{n\to\infty} q_n(t) = \lim_{n\to\infty}(p_n(t) - p_n(0)) = \sqrt{t}$ uniformly on $[0,1]$. In particular,

$$|t| = \sqrt{t^2} = \lim_{n\to\infty} q_n(t^2)$$

uniformly on $[-1,1]$.

Since $\left|\frac{f}{\|f\|_\infty}\right| \le 1$, we can write

$$\text{(B.1)} \qquad \frac{|f|}{\|f\|_\infty} = \lim_{n\to\infty} q_n\left(\frac{f^2}{\|f\|_\infty^2}\right) \quad \text{in } (C_b(X,\mathbb{R}), \|\cdot\|_\infty)\,.$$

Since $f^2 \in \overline{A}$ and hence $\frac{f^2}{\|f\|^2} \in \overline{A}$, also $q_n\left(\frac{f^2}{\|f\|_\infty^2}\right) \in \overline{A}$. By (B.1) it follows that $\frac{|f|}{\|f\|} \in \overline{A}$ and hence, $|f| \in \overline{A}$. The remaining assertions follow now from Remark C.1.4. □

Definition C.1.7. Let X be a non empty set and let A be a set of $\mathbb{R}$-valued functions defined on X. We say that A separates points of X if

$$\forall x, y \in X,\ x \neq y,\ \ \exists f \in A \text{ such that } \ f(x) \neq f(y).$$

Lemma C.1.8. *Let X be a non empty set and let A be a vector space of $\mathbb{R}$-valued functions on X which separates points of X and contains the constant functions. Then*

$$\forall x, y \in X,\ x \neq y,\ \forall a, b \in \mathbb{R}\ \ \exists f \in A \textit{ such that } \ f(x) = a \ \text{ and } \ f(y) = b\,.$$

Proof. Fix $x, y \in X$ such that $x \neq y$. By assumption there exists $g \in A$ such that $g(x) \neq g(y)$. Now, if we set $f := a + \frac{b-a}{g(y)-g(x)}(g - g(x))$, then the function f belongs to A because it is a linear combination of the function g and a constant function. Moreover, the function f clearly verifies $f(x) = a$ and $f(y) = b$. □

Theorem C.1.9 (Stone-Weierstrass theorem – real case)**.** *Let X be a compact topological space and let A be a subalgebra of $C(X,\mathbb{R})$ which separates points of X and contains the constant functions. Then A is a dense subalgebra of $C(X,\mathbb{R})$.*

Proof. Let us fix $f \in C(X,\mathbb{R})$ and $\varepsilon > 0$. We will prove that there exists $g \in \overline{A}$ such that $\|g - f\|_\infty < \varepsilon$.

For each $x, y \in X$ with $x \neq y$ we choose a function $h_{x,y} \in A$ such that $h_{x,y}(x) = f(x)$ and $h_{x,y}(y) = f(y)$. In the case $x = y$, we set $h_{x,x} = f$. Next, we consider the set $V_{x,y}$ defined by

$$V_{x,y} := \{t \in X \mid h_{x,y}(t) < f(t) + \varepsilon\}\,.$$

Clearly, x and y belong to $V_{x,y}$. Moreover, $V_{x,y}$ is an open subset of X. Indeed, the function $\varphi := h_{x,y} - f$ is continuous in X and hence, $V_{x,y} = \varphi^{-1}((-\infty, \varepsilon))$ is an open subset of X.

Since for any given $x \in X$ the family $\{V_{x,y} \mid y \in X\}$ is an open covering of X and X is compact, there exist $y_1, \ldots, y_n \in X$ such that

$$X = V_{x,y_1} \cup \ldots \cup V_{x,y_n}\,.$$

Now, for any given $x \in X$, let us consider the function $g_x := h_{x,y_1} \wedge \cdots \wedge h_{x,y_m}$. Then, by Theorem C.1.6, $g_x \in \overline{A}$ and $g_x(x) = f(x)$. Moreover, if $t \in X$, then there exists $i \in \{1, \ldots, n\}$ such that $t \in V_{x,y_i}$. Thus, $g_x(t) \leq h_{x,y_i}(t) < f(t) + \varepsilon$.

For any given $x \in X$ we set

$$W_x := \{t \in X :\ g_x(t) > f(t) - \varepsilon\}\,.$$

Arguing as before, we can easily show that $x \in W_x$ and that W_x is an open subset of X. Accordingly, $\{W_x \mid x \in X\}$ is also an open covering of X. Therefore, there exist $x_1, \ldots, x_m \in X$ such that

$$X = W_{x_1} \cup \cdots \cup W_{x_m}\,.$$

Now, let us consider the function $g := g_{x_1} \vee \cdots \vee g_{x_m}$. Then, by Theorem C.1.6, g belongs to $\overline{A}$. Moreover, fix $t \in X$. We have already shown that $g_x(t) < f(t) + \varepsilon$ for every $x \in X$, so that $g(t) < f(t)+\varepsilon$. To prove that $g(t) > f(t)-\varepsilon$, we observe that there exists $j \in \{1, \dots, m\}$ such that $t \in W_{x_j}$, so that $g_{x_j}(t) > f(t) - \varepsilon$, thereby implying that $g(t) > f(t) - \varepsilon$. Summing up, we have proved that $|g(t) - f(t)| < \varepsilon$. Due to the arbitrariness of $t \in X$, we conclude that $\|g - f\|_\infty < \varepsilon$. □

Corollary C.1.10 (Weierstrass approximation theorem)**.** *Let K be a compact subset of $\mathbb{R}$ and $f : K \to \mathbb{R}$ be a continuous function on K. Then, there exists a polynomial P such that $\|f - P\|_\infty < \varepsilon$.*

Proof. Let A be the set of all the restrictions to K of polynomials with real coefficients. Then A is a subalgebra of $C(K)$ which contains the constant functions and separates points of K (it suffices to consider the polynomial $P(x) = x$). Then, the assertion follows, by applying Theorem C.1.9. □

We now present some remarks on validity of the Stone-Weierstrass theorem in the complex case.

Example C.1.11. Let us consider a compact set $K \subset \mathbb{C}$ with non empty interior set $\mathrm{int}(K)$ and the set defined by

$$A := \{f \in C(K, \mathbb{C}) : \ f \text{ is analytic in } \mathrm{int}(K)\}.$$

Then A is a subalgebra of $C(K, \mathbb{C})$, contains the constant functions and separates points of K (it suffices to consider $f(z) = z$). Moreover, A is a closed subalgebra of $C(K, \mathbb{C})$ because uniform limits of sequences of analytic functions are analytic functions. On the other hand, $A \neq C(K, \mathbb{C})$ because, for instance the function $g(z) = |z|$ does not belong to A. Therefore, Stone-Weierstrass theorem is not longer true in the complex case as formulated above.

A more sophisticated example shows also that the space of all complex polynomials is not dense in the space of all complex continuous functions on a compact subset of $\mathbb{C}$ (see [11, Example 3.120]).

Theorem C.1.12 (Stone-Weierstrass theorem – complex case)**.** *Let X be a compact topological space and let A be a subalgebra of $C(X, \mathbb{C})$ which separates points of X, contains the constant functions and is closed with respect to conjugation, i.e., $\overline{f} \in A$ for all $f \in A$. Then, A is a dense subalgebra of $C(X, \mathbb{C})$.*

Proof. Let $A_0 := \{\mathrm{Re}\, f \mid f \in A\}$. Then A_0 is a vector subspace of $C(X, \mathbb{R})$. In particular, if $f \in A$, then $\mathrm{Im}\, f = -\mathrm{Re}\, i f \in A_0$. Moreover, for every $f, g \in A$,

$$\mathrm{Re}\, f \cdot \mathrm{Re}\, g = \mathrm{Re}\, \frac{1}{2}(fg + f\overline{g}) \in A_0 .$$

Thus, A_0 is a subalgebra of $C(X, \mathbb{R})$ which contains the constant functions. Actually, $A_0 \subset A$ because $\mathrm{Re} f = \frac{f+\overline{f}}{2} \in A$ for all $f \in A$.

Moreover, if $x, y \in X$, with $x \neq y$, then there exists $h \in A$ such that $h(x) \neq h(y)$. This implies that $\mathrm{Re}\, h(x) \neq \mathrm{Re}\, h(y)$ or $\mathrm{Im}\, h(x) \neq \mathrm{Im}\, h(y)$. Thus, A_0 separates points of X. Then, by Stone–Weierstass theorem C.1.9 we can conclude that A_0 is a dense subalgebra of $C(X, \mathbb{R})$.

Now, let $f \in C(X, \mathbb{C})$ be fixed. As it has been proved in Theorem C.1.9, for every $\varepsilon > 0$ there exist $h_1,\ h_2 \in A_0$ such that

$$\|\operatorname{Re} f - h_1\|_\infty < \frac{\varepsilon}{2}, \qquad \|\operatorname{Im} f - h_2\|_\infty < \frac{\varepsilon}{2},$$

from which it follows that $\|f - g\|_\infty < \varepsilon$, with $g := (h_1 + ih_2) \in A$. This completes the proof. □

Corollary C.1.13. *Let K be a compact subset of $\mathbb{R}$ and $f : K \to \mathbb{C}$ be a continuous function. Then, there exists a polynomial P such that $\|f - P\|_\infty < \varepsilon$.*

Proof. Let A be the set of all the restrictions to K of polynomials with complex coefficients. Then, A is a subalgebra of $C(K, \mathbb{C})$ which contains the constant functions, separates points of K (it suffices to consider $P(x) = x$) and is closed with respect to conjugation. The assertion follows by Theorem C.1.12. □

Corollary C.1.14. *Let $X = \{z \in \mathbb{C} \mid |z| = 1\}$. Then, the vector space of all functions of type*

$$P(z) = \sum_{k=-n}^{n} a_k z^k, \qquad (a_k \in \mathbb{C}),$$

is dense in $C(X, \mathbb{C})$.

Proof. The proof is immediate. We just observe that the set of all the functions P, defined as above, is closed with respect to conjugation since $\overline{z} = z^{-1}$ for every $z \in X$. □

Corollary C.1.15. *Let K be a compact subset of $\mathbb{C}$ and $f : K \to \mathbb{C}$ be a continuous function on K. For every $\varepsilon > 0$ there exists a $\mathbb{C}$-valued polynomial $P(\cdot,\cdot)$ of two variables such that $\sup_{z\in K} |f(z) - P(z, \overline{z})| < \varepsilon$.*

Proof. It suffices to observe that the set of functions of type $P(z, \overline{z})$, with P a $\mathbb{C}$-valued polynomial of two variables, is a subalgebra of $C(K, \mathbb{C})$ which satisfies all the assumptions of Theorem C.1.12. □

C.2 Riesz Representation Theorem

We recall the Riesz Representation Theorem which will be one the main tools for the proof of the Spectral representation theorem. For more details and for its proof we refer the reader, e.g., to [10, Chapter 2].

Theorem C.2.1 (Riesz representation theorem)**.** *Let K be a compact space and Λ be a continuous linear functional on $C(K)$. Then, there exists a unique regular Borel measure μ on K such that*

$$\Lambda(f) = \int_K f d\mu, \qquad f \in C(K),$$

and $\|\Lambda\|' = |\mu|(K)$. Moreover, μ is positive if and only if Λ is positive.

Bibliography

[1] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[2] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[3] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[4] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[5] P.R. Halmos, What does the Spectral Theorem say?, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[6] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[7] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981

[8] M. Reed, B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[9] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[10] W. Rudin, Real and Complex Analysis, 3^{rd} ed., McGraw-Hill, 1986.

[11] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[12] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

November 18, 2021

Lecture 5

Spectral representation theorem for bounded operators II

After the previous two dense lectures, this lecture is intended at the same time to take a breath and satisfy a curiosity, namely how to obtain a spectral representation theorem for normal operators. The latter is a fundamental ingredient for the study of unbounded selfadjoint operators, but its proof is usually skipped.
To obtain a spectral representation theorem for normal operators, we need to define the functional calculus not just for real-valued functions, but more generally for functions of one complex variable, since, in general, the spectrum of a normal operator contains also complex numbers. Note that the polynomials are not dense in the space of continuous functions defined on a compact set of the complex plane, as it is observed in Appendix C. Thus, to apply the Stone-Weierstrass theorem we will need to consider not only the polynomials but also their conjugate to obtain a dense set. Once we have constructed a functional calculus, the proof of the spectral theorem for normal operators will run analogously to that for self-adjoint operators.

In this lecture, we will follow the approach of the paper by Whitley [13].

5.1 Spectral representation theorem for normal operators

Definition 5.1.1. Let $T \in \mathcal{L}(H)$ and let $P(\lambda, \overline{\lambda}) = \sum_{n+m\leq N} a_{nm}\lambda^n\overline{\lambda}^m$ for every $\lambda \in \mathbb{C}$, where the coefficients a_{nm} belong to $\mathbb{C}$. Then, we set

$$P(T, T^*) := \sum_{n+m\leq N} a_{nm}T^n(T^*)^m.$$

To go on in the construction of the functional calculus, a fundamental step consists in proving that, if T is a normal operator, then

$$\|P(T, T^*)\| = \sup\{|P(\lambda, \overline{\lambda})| \mid \lambda \in \sigma(T)\}.$$

For this purpose, let us prove the following preliminary result.

Lemma 5.1.2. *Let $T \in \mathcal{L}(H)$ be a normal operator such that $0 \in \sigma(T)$. Then, for every $\varepsilon > 0$ there exists a closed subspace $M \neq \{0\}$ of H with the following properties:*

(1) *for every operator $B \in \mathcal{L}(H)$, which commutes with TT^*, M is invariant for B and its adjoint operator B^*;*

(2) $T_{|M} \in \mathcal{L}(M)$ *and* $\|T_{|M}\| \le \varepsilon$.

Proof. (1) Let us set $A := TT^*$. Since $0 \in \sigma(T)$, we can apply Proposition 2.3.6 and infer that there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset H$, with $\|x_n\| = 1$ for every $n \in \mathbb{N}$, such that Tx_n converges to 0 as n tends to ∞. This means that Ax_n tends to 0, so that $0 \in \sigma(A)$.

Fix $\varepsilon > 0$ and consider the function $h : \mathbb{R} \to \mathbb{R}$, defined by

$$h(t) = \begin{cases} 1 & |t| \le \frac{\varepsilon^2}{2}, \\ 2\left(1 - \dfrac{|t|}{\varepsilon^2}\right) & \frac{\varepsilon^2}{2} \le |t| \le \varepsilon^2, \\ 0 & |t| \ge \varepsilon^2. \end{cases}$$

Clearly this function is continuous and $\sup_{t\in\mathbb{R}} |th(t)| \le \varepsilon^2$. Let f be the restriction of h to $\sigma(A)$. Since A is self-adjoint, the operator $f(A)$ is well defined by Theorem 4.2.4.

Observe that the set $M := \{x \in H : f(A)x = x\}$ is a closed subspace of H. Moreover, if $B \in \mathcal{L}(H)$ commutes with A, then, by Theorem 4.2.4, B commutes also with $f(A)$. Hence,

$$Bx = Bf(A)x = f(A)Bx,$$

for every $x \in M$, so that $Bx \in M$. This means that M is invariant for B. On the other hand

$$B^*A = B^*TT^* = (TT^*B)^* = (BTT^*)^* = TT^*B^* = AB^*,$$

so that B^* commutes with A and, hence, M is invariant also for B^*. Property (1) is now proved.

(2) Let us observe that T commutes with TT^*, so that $T(M) \subseteq M$. Moreover, for every $x \in M$, with $\|x\| = 1$, it holds that

$$\|Ax\| = \|Af(A)x\| \le \|Af(A)\| = \sup\{|\lambda f(\lambda)| \mid \lambda \in \sigma(A)\} \le \varepsilon^2.$$

We thus conclude that $\|Tx\|^2 = \langle Tx, Tx\rangle = \langle Ax, x\rangle \le \varepsilon^2$ for every $x \in M$ with $\|x\| = 1$. Therefore, $\|T_{|M}\| \le \varepsilon$. Property (2) is proved.

To prove that $M \ne \{0\}$, we observe that

$$\|(I - f(A))f(2A)\| = \sup\{|(1 - f(\lambda))f(2\lambda)| : \lambda \in \sigma(A)\} = 0,$$

since, if $f(2\lambda) \ne 0$, then $f(\lambda) = 1$. As a consequence, we deduce that the range of $f(2A)$ is contained in M and it contains some element $x \ne 0$ since

$$\|f(2A)\| = \sup\{|f(2\lambda)| : \lambda \in \sigma(A)\} \ge |f(0)| = 1. \quad \square$$

Lemma 5.1.3 (Spectral mapping theorem for normal operators). *Let $T \in \mathcal{L}(H)$ be a normal operator and let $P(\lambda, \overline{\lambda}) = \sum_{n+m\le N} a_{n,m}\lambda^n\overline{\lambda}^m$ for every $\lambda \in \mathbb{C}$, with coefficients $a_{nm} \in \mathbb{C}$. Then,*

$$\sigma(P(T, T^*)) = \{P(\lambda, \overline{\lambda}) \mid \lambda \in \sigma(T)\}. \tag{5.1}$$

Proof. Fix $\lambda \in \sigma(T)$ and consider a sequence $(x_k)_{k\in\mathbb{N}} \subset H$, with $\|x_k\| = 1$ for every $k \in \mathbb{N}$, such that $(\lambda - T)x_k$ converges to 0 as k tends to ∞ (cf. Proposition 2.3.6). Since T is a normal operator, we can apply Lemma 2.3.5 to infer that $\|(\overline{\lambda} - T^*)x_k\| = \|(\lambda - T)x_k\|$ for every $k \in \mathbb{N}$ and, then, conclude that $\|(\overline{\lambda} - T^*)x_k\|$ converges to 0 as k tends to ∞. Using the following chain of equalities

$$\begin{aligned}(P(T,T^*) - P(\lambda,\overline{\lambda}))x_k &= \sum_{n+m\le N} a_{nm}(T^n(T^*)^m - \lambda^n\overline{\lambda}^m)x_k \\ &= \sum_{n+m\le N} a_{nm}[T^n((T^*)^m - \overline{\lambda}^m)x_k + \overline{\lambda}^m(T^n - \lambda^n)x_k] \\ &= \sum_{n+m\le N} a_{nm}[T^n(T^{*(m-1)} + \cdots + \overline{\lambda}^{m-1})(T^* - \overline{\lambda})x_k \\ &\qquad\qquad + \overline{\lambda}^m(T^{n-1} + \cdots + \lambda^{n-1})(T-\lambda)x_k],\end{aligned}$$

we conclude that $(P(T,T^*) - P(\lambda,\overline{\lambda}))x_k$ converges to 0 as k tends to ∞. Hence, $P(\lambda,\overline{\lambda}) \in \sigma(P(T,T^*))$ and the inclusion $\{P(\lambda,\overline{\lambda}) : \lambda \in \sigma(T)\} \subset \sigma(P(T,T^*))$ follows.

To prove the other inclusion, we fix $\mu \in \sigma(P(T,T^*))$. Then, $B := P(T,T^*) - \mu$ is a normal operator and $0 \in \sigma(B)$. We can thus apply Lemma 5.1.2 to infer that:

(i) for every $n \in \mathbb{N}$, there exists a closed subspace $M_n \neq \{0\}$, which is invariant for B and B^*, with $\|B_{|M_n}\| \le 1/n$;

(ii) M_n is invariant also for T and T^* since T commutes with BB^*.

The operator $T_{|M_n}$ is clearly normal. In view of Proposition 2.1.1, $\sigma(T_{|M_n}) \neq \varnothing$.

For every $n \in \mathbb{N}$, fix $\lambda_n \in \sigma(T_{|M_n})$. Then, there exists $y_n \in M_n$, with $\|y_n\| = 1$, such that $\|(\lambda_n - T)y_n\| \le 1/n$, due to Proposition 2.3.6. The sequence $(\lambda_n)_{n\in\mathbb{N}}$ is bounded by $\|T\|$. Therefore, up to a subsequence, $(\lambda_n)_{n\in\mathbb{N}}$ converges to some λ, which belongs to $\sigma(T)$ since

$$\|(\lambda - T)y_n\| \le |\lambda - \lambda_n| + \|(\lambda_n - T)y_n\| \le |\lambda - \lambda_n| + \frac{1}{n}$$

for every $n \in \mathbb{N}$. This implies that $(\lambda - T)y_n$ converges to 0 as n tends to ∞. Arguing as in the first part of the proof, we can show that $(P(T,T^*) - P(\lambda,\overline{\lambda}))y_n$ tends to 0. On the other hand, $y_n \in M_n$ for every $n \in \mathbb{N}$ so that

$$\|(P(T,T^*) - \mu)y_n\| = \|By_n\| \le \|B_{|M_n}\| \cdot \|y_n\| \le \frac{1}{n}.$$

It follows that $(P(T,T^*) - \mu)y_n$ tends to 0 as n tends to ∞ so that $\mu = P(\lambda,\overline{\lambda})$. □

We can now construct a functional calculus for normal operators.

Theorem 5.1.4. *Let $T \in \mathcal{L}(H)$ be a normal operator. Then, there exists a unique linear operator*

$$\Phi\colon C(\sigma(T),\mathbb{C}) \to \mathcal{L}(H)$$

with the following properties:

(1) $\Phi(fg) = \Phi(f)\Phi(g)$, $\Phi(\lambda f) = \lambda\Phi(f)$ *for every* $f, g \in C(\sigma(T), \mathbb{C})$ *and* $\lambda \in \mathbb{C}$*;*

(2) $\Phi(1) = I$ *and* $\Phi(\overline{f}) = \Phi(f)^*$ *for every* $f \in C(\sigma(T), \mathbb{C})$*;*

(3) $\|\Phi(f)\| = \|f\|_\infty$ *for every* $f \in C(\sigma(T), \mathbb{C})$*;*

(4) *if* $f = id_{\sigma(T)}$*, then* $\Phi(f) = T$*;*

(5) $\sigma(\Phi(f)) = \{f(\lambda) : \lambda \in \sigma(T)\}$ *for every* $f \in C(\sigma(T), \mathbb{C})$*;*

(6) *if* $S \in \mathcal{L}(H)$ *commutes with* T *and* T^**, then* S *commutes with* $\Phi(f)$ *for every* $f \in C(\sigma(T), \mathbb{C})$*;*

(7) *if* $f \in C(\sigma(T), \mathbb{R})$ *is nonnegative, then* $\Phi(f) \geq 0$.

Proof. For every polynomial $P(\lambda, \overline{\lambda}) = \sum_{m+n \leq N} a_{n,m}\lambda^n(\overline{\lambda})^m$, we define $\Phi(P) = P(T, T^*)$. Due to Proposition 2.3.5-(4) and Lemma 5.1.3, $\|\Phi(P)\| = \sup\{|P(\lambda, \overline{\lambda})| : \lambda \in \sigma(T)\}$. Hence, Φ is a linear isometry from the space $\mathcal{A}$ of the functions of type $\lambda \mapsto P(\lambda, \overline{\lambda})$, with $\lambda \in \sigma(T)$, in $(\mathcal{L}(H), \|\cdot\|)$. We can uniquely extend Φ to a linear and continuous mapping $\tilde{\Phi}$ from the closure of $\mathcal{A}$ in $(C(\sigma(T), \mathbb{C}), \|\cdot\|_\infty)$, with values in $(\mathcal{L}(H), \|\cdot\|)$. On the other hand, the closure of $\mathcal{A}$ coincides with $(C(\sigma(T), \mathbb{C}), \|\cdot\|_\infty)$, due to Corollary C.1.15.

To ease the notation, we still denote by Φ the extension $\tilde{\Phi}$. By construction, the operator $\Phi \colon C(\sigma(T), \mathbb{C}) \to \mathcal{L}(H)$ is a linear isometry, so that property (3) follows. Properties (1) and (2) are clearly satisfied if f and g are functions in $\mathcal{A}$ and, hence, they can be easily extended to the case where f and g are continuous functions on $C(\sigma(T), \mathbb{C})$, using a density argument. Property (4) follows straightforwardly from the definition.

Let us prove property (5). Fix $\mu \in \sigma(T)$ and let $(P_n)_{n\in\mathbb{N}}$ be a sequence of polynomials such that $P_n(\lambda, \overline{\lambda}) \to f(\lambda)$ uniformly on $\sigma(T)$. Then, the sequence $(P_n(\mu, \overline{\mu}) - P_n(T, T^*))_{n\in\mathbb{N}}$ converges to $f(\mu) - \Phi(f)$ in $(\mathcal{L}(H), \|\cdot\|)$. On the other hand, for every $n \in \mathbb{N}$, $P_n(\mu, \overline{\mu}) \in \sigma(P_n(T, T^*))$ so that $P_n(\mu, \overline{\mu}) - P_n(T, T^*)$ is not invertible. This implies that $f(\mu) - \Phi(f)$ is not invertible (otherwise there would admit a neighborhood consisting of invertible operators) and, hence, $f(\mu) \in \sigma(\Phi(f))$. Vice versa, fix $\mu \in \mathbb{C} \setminus f(\sigma(T))$. Then, $\mu - f(\lambda) \neq 0$ for every $\lambda \in \sigma(T)$ so that the function $g = 1/(\mu - f)$ belongs to $C(\sigma(T), \mathbb{C})$. From properties (1) and (4) it follows that

$$\Phi(g)(\mu - \Phi(f)) = (\mu - \Phi(f))\Phi(g) = I,$$

i.e., $\mu - \Phi(f)$ is invertible and, hence, $\mu \notin \sigma(\Phi(f))$.

Property (6) is immediate to prove if $f \in \mathcal{A}$. The general case then follows by a density argument.

In order to prove property (7), we observe that if $f \geq 0$ on $\sigma(T)$, then $f = g^2$ for some continuous real-valued function g. By property (1) it follows that $\Phi(f) = \Phi(g^2) = \Phi(g)^2$, where $\Phi(g)$ is a self-adjoint operator due to property (2). Therefore, for every $x \in H$, it holds that

$$\langle \Phi(f)(x), x\rangle = \langle \Phi(g)^2(x), x\rangle = \langle \Phi(g)(x), \Phi(g)(x)\rangle = \|\Phi(g)(x)\|^2 \geq 0\,,$$

i.e., $\Phi(f) \geq 0$.

To prove the uniqueness of Φ it suffices to observe that if $\Psi \colon C(\sigma(T), \mathbb{C}) \to \mathcal{L}(H)$ is another mapping with properties (1) to (4), then $\Psi(f) = \Phi(f)$ for every $f \in \mathcal{A}$. Hence, by density it follows that $\Psi \equiv \Phi$. $\square$

In what follows, sometimes we will write $f(T)$ instead of $\Phi(f)$ to stress the dependence on T.

Remark 5.1.5. Property (6) in Theorem 5.1.4 can be improved thanks to Fuglede's Theorem which states that, if an operator S commutes with T, then S commutes also with T^*. For a proof of such a theorem, we refer the reader to [5].

Let $T \in \mathcal{L}(H)$ be a normal operator and fix $x \in H$. Denote by L the functional on $C(\sigma(T), \mathbb{C})$ defined by

$$C(\sigma(T), \mathbb{C}) \ni f \mapsto L(f) := \langle f(T)(x), x \rangle \, .$$

This functional is linear, continuous and positive due to Theorem 5.1.4-(7). Then, by the Riesz representation theorem, we know that there exists a unique positive and regular Borel measure μ_x, defined on the compact set $\sigma(T)$, such that

$$L(f) = \langle f(T)(x), x \rangle = \int_{\sigma(T)} f d\mu_x \, .$$

This measure μ_x is called *spectral measure* associated with x.

The following result is the counterpart for normal operators of Lemma 4.2.8. In such a lemma, we required the existence of a cyclic vector in H, i.e., the existence of $x \in H$ such that the span of the vectors $T^k x$, when $k \in \mathbb{N} \cup \{0\}$, is dense in H. Here, we need a slightly different assumption.

Lemma 5.1.6. *Let $T \in \mathcal{L}(H)$ be a normal operator. Suppose that there exists $x \in H$ such that* $\operatorname{span}\{T^n (T^*)^m x : n, m \in \mathbb{N} \cup \{0\}\}$ *is dense in H. Then, there exists an unitary operator $U \colon H \to L^2(\sigma(T), \mu_x)$ such that*

$$(UTU^{-1})(f)(\lambda) = \lambda f(\lambda) \quad \mu_x - a.e.$$

for every $f \in L^2(\sigma(T), \mu_x)$.

Proof. For every $f \in C(\sigma(T), \mathbb{C})$, we set

$$U\Phi(f)(x) := f \, .$$

The operator U is well defined in the space $\{\Phi(f)(x) : f \in C(\sigma(T), \mathbb{C})\}$. Indeed, if $f, g \in C(\sigma(T), \mathbb{C})$ are functions such that $\Phi(f)(x) = \Phi(g)(x)$, then

$$\begin{aligned} \Phi(f) T^n (T^*)^m (x) =& \Phi(f)\Phi(z^n \overline{z}^m)(x) = \Phi(z^n \overline{z}^m)\Phi(f)(x) \\ =& \Phi(z^n \overline{z}^m)\Phi(g)(\psi) = \Phi(g) T^n (T^*)^m (x) \end{aligned}$$

for every n, $m \in \mathbb{N} \cup \{0\}$, i.e., $\Phi(f) = \Phi(g)$ on a dense subspace of H. The continuity of $\Phi(f)$ and $\Phi(g)$ implies that $\Phi(f) = \Phi(g)$ on H so that

$$0 = \|\Phi(f - g)\| = \|f - g\|_\infty \, ,$$

i.e., $f \equiv g$. Moreover, for every $f \in C(\sigma(T), \mathbb{C})$ it holds that

$$\|\Phi(f)(x)\|^2 = \langle x, \Phi(f)^* \Phi(f)(x) \rangle = \langle x, \Phi(\overline{f} f)(x) \rangle$$

$$=\langle \Phi(\overline{f}f)(x), x\rangle = \int_{\sigma(T)} |f|^2 d\mu_x \, .$$

This means that U is an isometry from $(\{\Phi(f)(x) : f \in C(\sigma(T), \mathbb{C})\}, \|\cdot\|)$ into $L^2(\sigma(T), \mu_x)$. Since the space $\{\Phi(f)(x) : f \in C(\sigma(T), \mathbb{C})\}$ is dense in H, we can extend U as an isometry from H into $L^2(\sigma(T), \mu_x)$. On the other hand, since $C(\sigma(T), \mathbb{C})$ is a dense subspace of $L^2(\sigma(T), \mu_x)$, it follows that U is also surjective.

Now, we observe that

$$(UTU^{-1})(f)(\lambda) = (UT\Phi(f)(x))(\lambda) = (U\Phi(yf)(x))(\lambda) = \lambda f(\lambda)$$

for every $f \in C(\sigma(T), \mathbb{C})$. This identity can be extended to any $f \in L^2(\sigma(T), \mu_x)$ using the density of $C(\sigma(T), \mathbb{C})$ into $L^2(\sigma(T), \mu_x)$. □

To extend this result to a general normal and bounded operator, as we did in the case of self-adjoint operators, we prove the following lemma.

Lemma 5.1.7. *Let H be a separable complex Hilbert space and let $T \in \mathcal{L}(H)$ be a normal operator. Then, there exists a family $(H_n)_{n\in J}$ of closed subspaces of H, with $J \subseteq \mathbb{N}$ finite or infinite, such that*

(i) $H = \bigoplus_{n\in J} H_n$;

(ii) *H_n is invariant for T and T^*;*

(iii) *for each $n \in J$ there exists $x_n \in H_n$ such that* $\overline{\text{span}\{T^k(T^*)^m x_n \mid k, m \in \mathbb{N}\cup\{0\}\}} = H_n$.

Proof. Let $(e_n)_{n\in\mathbb{N}}$ be a complete orthonormal system of H. Set $x_1 := e_1$ and

$$H_1 := \overline{\text{span}\{T^k(T^*)^m x_1 : k, m \in \mathbb{N}\cup\{0\}\}}.$$

Then, H_1 is invariant with respect to T and T^*. If $e_n \in H_1$ for every $n \in \mathbb{N}$, then $H_1 = H$. In such a case, the proof is complete.

Suppose that e_n does not belong to H_1 for some $n \in \mathbb{N}$ and denote by n_1 the first index such that $e_{n_1} \notin H_1$. This means that $e_n \in H_1$ for every $n < n_1$. Denote by $P_{H_1^\perp}$ the orthogonal projection onto the closed subspace $H_1^\perp$ and set $x_2 := P_{H_1^\perp}(e_{n_1})$. Note that $x_2 \neq 0$ since $e_{n_1} \neq H_1$. Moreover, since T is a normal operator and both T and T^* leave H_1 invariant, it follows that T and T^* leave invariant also $H_1^\perp$ (see Exercise 5.2.1). Let us set $H_2 := \overline{\text{span}\{T^k(T^*)^m x_2 : k, m \in \mathbb{N}\cup\{0\}\}}$. Clearly, $H_2 \subset H_1^\perp$ and, in particular, H_2 is invariant with respect to T and T^*. If $H = H_1 \oplus H_2$, then the proof is complete.

Suppose that $H_1 \oplus H_2 \subsetneq H$ and let n_2 be the first index such that $e_{n_2} \notin H_1 \oplus H_2$. Denote by $P_{(H_1\oplus H_2)^\perp}$ the orthogonal projection onto the closed subspace $(H_1 \oplus H_2)^\perp$ and set $x_3 := P_{(H_1\oplus H_2)^\perp}(e_{n_2})$. Iterating this argument, we obtain that either $H = H_1 \oplus H_2 \oplus \ldots H_N$ for some $N \in \mathbb{N}$, where, for every $i = 1, \ldots, N$, H_i satisfies (iii) for some $x_i \in H$, and is invariant with respect to T and T^*, or there exists a family $(H_i)_{i\in\mathbb{N}}$ of closed and mutually orthogonal subspaces of H, which are invariant with respect to T and T^* and such that property (iii) holds true. In both cases, $e_n \in H_1$ for every $n < n_1$, $e_n \in H_1 \oplus H_2$ for every $n_1 \leq n < n_2$ and so on. This ensures that $(e_n)_{n\in\mathbb{N}} \subset \bigoplus_{i\in\mathbb{N}} H_i$ and, hence, $H = \bigoplus_{i\in\mathbb{N}} H_i$. □

Theorem 5.1.8. *Let H be a separable complex Hilbert space and let $T \in \mathcal{L}(H)$ be a normal operator. Then, there exist a family $(\mu_n)_{n\in J}$ of positive Borel measures defined on $\sigma(T)$, with $J \subseteq \mathbb{N}$ finite or infinite, and a unitary operator*

$$U\colon H \to \bigoplus_{n\in J} L^2(\sigma(T), d\mu_n)$$

such that

$$(UTU^{-1}(\psi))_n(\lambda) = \lambda\psi_n(\lambda) \quad \mu_n - a.e.$$

for every $\psi = (\psi_n)_{n\in J} \in \bigoplus_{n\in J} L^2(\sigma(T), d\mu_n)$ and every $n \in J$.

Proof. The assertion follows applying first Lemma 5.1.7, to determine the decomposition, and then Lemma 5.1.6 to each component, observing that $T_{|H_n}$ is a normal operator, since H_n is invariant with respect to T and T^*. We thus obtain that the measure μ_n is the spectral measure associated to some vector $x_n \in H_n$ and it is defined on $\sigma(T_{|H_n})$. Extending such a measure to $\sigma(T)$, setting $\mu_n = 0$ on $\sigma(T) \setminus \sigma(T_{|H_n})$, we conclude the proof. □

Mimicking the proof of Theorem 4.2.11, we can prove the spectral mapping theorem for normal operators in its classical formulation.

Theorem 5.1.9 (Spectral theorem for normal operators)**.** *Let H be a separable complex Hilbert space and let $T \in \mathcal{L}(H)$ be a normal operator. Then, there exist a measurable space (M, μ) with finite measure, a bounded and measurable function $m\colon M \to \mathbb{C}$ and a unitary operator $U\colon H \to L^2(M, d\mu)$ such that*

$$(UTU^{-1}(f))(\lambda) = m(\lambda)f(\lambda) \quad \mu - a.e.$$

for every $f \in L^2(M, d\mu)$.

5.2 Exercises

Exercise 5.2.1. Let T be a normal operator on a complex Hilbert space H and let M be a subspace of H. Prove that if M is invariant for T and T^*, then $M^\perp$ is invariant for T and T^*.

Exercise 5.2.2. Prove Theorem 5.1.9.

Exercise 5.2.3. Let T be a normal operator defined on a separable complex Hilbert space such that $T^4 = T^3$. Prove that T is selfadjoint.

Exercise 5.2.4. Let T be a normal operator on a separable complex Hilbert space H. Prove that

1. if T is positive, then there exists a unique positive operator S such that $S^2 = T$ (S is called the square root of T)
2. if there exists $\theta \in (0, \pi)$ such that $\sigma(T) \subseteq \{z \in \mathbb{C} \mid |Arg(z)| < \theta\}$, then for every $n \in \mathbb{N}$ there exists a normal operator S on H such that $S^n = T$. Is S unique?;

Exercise 5.2.5. Let S and T be two normal operators on a complex Hilbert space H. Denote by $\mathcal{A}_S$ (resp., by $\mathcal{A}_T$) the range of the operator $\Phi_S \colon C(\sigma(S), \mathbb{C}) \to \mathcal{L}(H)$ (resp. $\Phi_T \colon C(\sigma(T), \mathbb{C}) \to \mathcal{L}(H)$) constructed in Theorem 5.1.4. Prove that $\sigma(S) = \sigma(T)$ if and only if there exists an isometric isomorphism Λ from $\mathcal{A}_S$ onto $\mathcal{A}_T$ such that $\Lambda(S) = T$.

Exercise 5.2.6. Let T be a normal operator defined on a complex Hilbert space H and $\Phi \colon C(\sigma(T), \mathbb{C}) \to \mathcal{L}(H)$ be the operator constructed in Theorem 5.1.4. Prove that if $Tx = \lambda x$ for some $x \in H$, $x \neq 0$, and $\lambda \in \mathbb{C}$, then $\Phi(f)x = f(\lambda)x$ for every $f \in C(\sigma(T), \mathbb{C})$.

Exercise 5.2.7. Let T be a normal operator on a complex separable Hilbert space H. Prove that T is compact if and only if the following two conditions hold:

(i) $\sigma(T)$ consists of at most a sequence of eigenvalues, including 0, which possibly accumulates at 0;

(ii) $\dim \ker(\lambda - T) < \infty$ for all $\lambda \neq 0$.

Bibliography

[1] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[2] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[3] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[4] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[5] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[6] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[7] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981

[8] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[9] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[10] W. Rudin, Real and Complex Analysis, 3^{rd} ed., McGraw-Hill, 1986.

[11] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[12] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958

[13] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

November 24, 2021

Lecture 6

Self-adjoint and dissipative operators on Hilbert spaces

In this lecture we define the adjoint of an unbounded linear operator on a Hilbert space and study its properties. We introduce symmetric and self-adjoint operators on a Hilbert space and characterize the latter operators by a range condition, see Theorem 6.1.16. We also discuss about dissipative operators on Hilbert spaces. It is simple to see that the dissipativity of $\pm iT$ completely characterizes the symmetry of the densely defined operator T, see Theorem 6.2.4. This permits us to show that the spectrum of a self-adjoint operator is real.

Along the lecture, $(H, \|\cdot\|)$ will denote a Hilbert space over $\mathbb{C}$ with inner product $\langle \cdot, \cdot \rangle$.

6.1 Adjoint of an unbounded operator, self-adjoint and symmetric operators

Definition 6.1.1. Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator. We set

$$D(T^*) := \{y \in H : \exists y^* \in H \text{ such that } \langle Tx, y\rangle = \langle x, y^*\rangle \ \ \forall x \in D(T)\}. \tag{6.1}$$

Notice that, for every $y \in D(T^*)$, the element y^* in (6.1) is unique. Indeed, suppose that $y_1^* \in H$ and $y_2^* \in H$ are such that

$$\langle x, y_1^*\rangle = \langle Tx, y\rangle = \langle x, y_2^*\rangle, \qquad x \in D(T).$$

Hence,

$$\langle x, y_1^* - y_2^*\rangle = 0, \qquad x \in D(T).$$

Then, the density of $D(T)$ in H implies that $y_1^* = y_2^*$. Therefore, the following definition makes sense.

Definition 6.1.2. Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator, The operator T^*, defined by $T^*y := y^*$ for every $y \in D(T^*)$, where y^* is the unique element in H such that $\langle Tx, y\rangle = \langle x, y^*\rangle$ for every $x \in D(T)$, is called the adjoint operator of T.

It is easy to see that $T^* : D(T^*) \to H$ is a linear operator. The following result characterizes the elements of $D(T^*)$.

Proposition 6.1.3. *If* $T : D(T) \subseteq H \to H$ *is a densely defined linear operator, then*

$$y \in D(T^*) \iff \exists c > 0 \text{ such that } |\langle Tx, y\rangle| \le c\|x\| \ \ \forall x \in D(T).$$

Proof. "$\Longrightarrow$": Fix $y \in D(T^*)$. Then, from (6.1) it follows that there exists $y^* \in H$ such that

$$|\langle Tx, y\rangle| = |\langle x, y^*\rangle| \le \|y^*\|\,\|x\| \quad \forall x \in D(T).$$

"$\Longleftarrow$": Consider the linear functional $S : D(T) \to \mathbb{C}$, defined by $Sx = \langle Tx, y\rangle$ for every $x \in D(T)$. By assumptions, $\|Sx\| \le c\|x\|$ for every $x \in D(T)$. Since $D(T)$ is dense in H, the operator S can be uniquely extended to a bounded operator on H, still denoted by S. Therefore, the Riesz-Fréchet theorem implies that there exists $y^* \in H$ such that $Sx = \langle x, y^*\rangle$ for every $x \in H$. In particular, we can infer that

$$\langle Tx, y\rangle = \langle x, y^*\rangle, \qquad x \in D(T).$$

This formula clearly shows that $y \in D(T^*)$. □

Example 6.1.4. Let $H = L^2(\Omega, \mu)$, where (Ω, μ) is a σ-finite measure space. Let us consider the multiplication operator M_m, associated to the measurable function $m : \Omega \to \mathbb{C}$, whose domain is $D(M_m) = \{f \in L^2(\Omega, \mu) : mf \in L^2(\Omega, \mu)\}$, see Section 2.2. Then, $M_m^* = M_{\overline{m}}$.

Indeed, if $f \in D(M_{\overline{m}})$ then

$$\langle mh, f\rangle = \int_\Omega mh\overline{f}d\mu = \int_\Omega h\overline{(\overline{m}f)}d\mu = \langle h, \overline{m}f\rangle, \qquad h \in D(M_m).$$

Thus, we have proved that $D(M_{\overline{m}}) \subseteq D(M_m^*)$ and $M_m^* f = M_{\overline{m}} f$ for every $f \in D(M_{\overline{m}})$.

Vice versa, suppose that $f \in D(M_m^*)$. Then, there exists $g \in H$ such that

$$\langle mh, f\rangle = \langle h, g\rangle, \qquad h \in D(M_m),$$

which implies that

$$\int_\Omega mh\overline{f}d\mu = \int_\Omega h\overline{g}d\mu, \qquad h \in D(M_m)$$

or, equivalently,

$$\int_\Omega h\psi d\mu = 0, \qquad h \in D(M_m),$$

where $\psi = m\overline{f} - \overline{g}$. From this formula, we can infer that $\psi = 0$ μ-almost everywhere on Ω. For this purpose, we consider an increasing sequence $(\Omega_n)_{n\in\mathbb{N}}$ of measurable subsets of Ω, with finite measure, whose union is Ω. Moreover, for every $n \in \mathbb{N}$ we introduce the function $h_n = |\psi|^{-1}\overline{\psi}\chi_{A_n}$, where $A_n = \{x \in \Omega_n : \psi(x) \neq 0 \text{ and } |m(x)| \le n\}$. Clearly, each function h_n belongs to $D(M_m)$. Moreover,

$$0 = \int_\Omega h_n\psi d\mu = \int_\Omega |\psi|\chi_{\{x\in\Omega_n : |m(x)|\le n\}}d\mu$$

and this formula implies that $\psi = 0$ μ-almost everywhere on the set $\{x \in \Omega_n : |m(x)| \le n\}$. Letting n tend to ∞, we conclude that $\psi = 0$ almost everywhere in Ω or, equivalently, $\overline{m}f = g \in H$. This means that $f \in D(M_{\overline{m}})$.

As the following example shows, in general T^* is not densely defined.

Example 6.1.5. Fix $f_0 \in H := L^2(\mathbb{R};\mathbb{C})$, with $f_0 \neq 0$, and a locally bounded measurable function $f \notin L^2(\mathbb{R})$. Let us consider the linear operator $(T, D(T))$, defined by

$$D(T) := \{ g \in L^2(\mathbb{R};\mathbb{C}) \mid fg \in L^1(\mathbb{R};\mathbb{C}) \}, \qquad Tg := \langle g, f \rangle \cdot f_0 .$$

Since $D(T)$ contains the space $C_c(\mathbb{R};\mathbb{C})$ of all the continuous functions on $\mathbb{R}$ with compact support, one deduces that $D(T)$ is dense in $L^2(\mathbb{R};\mathbb{C})$. On the other hand, $D(T^*)$ is not a dense subspace of H. Indeed, if $h \in D(T^*)$ then

$$\begin{aligned}\langle g, T^*h\rangle =& \langle Tg, h\rangle = \langle \langle g, f\rangle \cdot f_0, h\rangle = \langle g, f\rangle \cdot \langle f_0, h\rangle \\ =& \langle g, \overline{\langle f_0, h\rangle} \cdot f\rangle = \langle g, \langle h, f_0\rangle \cdot f\rangle\end{aligned}$$

for every $g \in D(T)$. Since $C_c(\mathbb{R};\mathbb{C})$ is dense in $L^2(\mathbb{R};\mathbb{C})$, we conclude that $T^*h = \langle f_0, h\rangle \cdot f$. Next, we observe that, since $f \notin L^2(\mathbb{R};\mathbb{C})$, it follows that $\langle f_0, h\rangle = 0$ for every $h \in D(T^*)$. Therefore, $D(T^*)$ is not dense in H, otherwise $f_0 = 0$, which is not the case.

In the following proposition, we state and prove some basic properties of adjoint operators of unbounded linear operators. Given two linear operators S and T, we use the notation $S \subset T$ when T is an extension of the operator S, i.e., when $D(S) \subset D(T)$ and $Tx = Sx$ for every $x \in D(S)$.

Proposition 6.1.6. *Let $S : D(S) \subset H \to H$ and $T : D(T) \subset H \to H$ be densely defined linear operators. The following properties hold true.*

(1) *If $S \subset T$, then $T^* \subset S^*$.*

(2) *If T^* is densely defined, then $T \subset T^{**} := (T^*)^*$.*

Proof. (1) Fix $y \in D(T^*)$. Then, for every $x \in D(S) \subset D(T)$ it holds that

$$\langle Sx, y\rangle = \langle Tx, y\rangle = \langle x, T^*y\rangle .$$

It follows that $y \in D(S^*)$ and $S^*y = T^*y$.

(2) Fix $x \in D(T)$. Then,

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

for every $y \in D(T^*)$. It follows that $x \in D(T^{**})$ and $T^{**}x = Tx$. □

Proposition 6.1.7. *Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator, with $\overline{\mathrm{rg}(T)} = H$. Then, the following properties are satisfied.*

(1) *T^* is injective.*

(2) *If T is injective, then $(T^*)^{-1} = (T^{-1})^*$.*

Proof. (1) Fix $y \in D(T^*)$ such that $T^*y = 0$. Then,

$$\langle Tx, y\rangle = \langle x, T^*y\rangle = 0$$

for every $x \in D(T)$. This implies that $\langle z, y\rangle = 0$ for every $z \in \mathrm{rg}(T)$. Since $\mathrm{rg}(T)$ is dense in H, we conclude that $y = 0$. Hence, T^* is injective.

(2) Since $D(T^{-1}) = \mathrm{rg}(T)$, it follows that $D(T^{-1})$ is a dense subspace of H. Hence, the adjoint operator $(T^{-1})^*$ is well defined.

Let us prove that $(T^*)^{-1} \subset (T^{-1})^*$ and $(T^{-1})^* \subset (T^*)^{-1}$. We begin by proving the first property. For this purpose, we fix $y^* \in D((T^*)^{-1}) = \mathrm{rg}(T^*)$. Then, there exists $y \in D(T^*)$ such that $T^*y = y^*$. If $z \in D(T^{-1})$, then $T^{-1}z \in D(T)$ and therefore

$$\langle T^{-1}z, y^*\rangle = \langle T^{-1}z, T^*y\rangle = \langle TT^{-1}z, y\rangle = \langle z, y\rangle = \langle z, (T^*)^{-1}y^*\rangle .$$

Hence, $y^* \in D((T^{-1})^*)$ and $(T^{-1})^*y^* = (T^*)^{-1}y^*$. The first property is proved.

Next, fix $y^* \in D((T^{-1})^*)$. Then,

$$\langle T^{-1}x, y^*\rangle = \langle x, (T^{-1})^*y^*\rangle$$

for every $x \in D(T^{-1}) = \mathrm{rg}(T)$. Taking, $x = Tz$ in the previous formula, it thus follows that

$$\langle z, y^*\rangle = \langle Tz, (T^{-1})^*y^*\rangle$$

for every $z \in D(T)$. Hence, $(T^{-1})^*y^* \in D(T^*)$, $T^*((T^{-1})^*y^*) = y^*$ and, as a byproduct, $y^* \in \mathrm{rg}(T^*) = D((T^*)^{-1})$. This completes the proof. □

Definition 6.1.8. Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator. T is a symmetric operator if $T \subset T^*$, i.e., if

$$\langle Tx, y\rangle = \langle x, Ty\rangle, \qquad x, y \in D(T).$$

T is called self-adjoint if $T = T^*$.

Example 6.1.9. Let us consider the multiplication operator M_m introduced in 6.1.4. It is easy to check that M_m is symmetric if and only if $\mathrm{Im}(m) = 0$. In such a case, $D(M_m) = D(M_m^*)$ so that the operator M_m is also self-adjoint.

Remark 6.1.10. Every self-adjoint operator is clearly symmetric. On the other hand, there exist symmetric operators which are not self-adjoint. Consider for instance the operator $T : W_0^{1,2}((0,1);\mathbb{C}) \subset L^2((0,1);\mathbb{C}) \to L^2((0,1);\mathbb{C})$ defined by[1]

$$Tf = if', \qquad f \in W_0^{1,2}((0,1);\mathbb{C}),$$

where $W_0^{1,2}((0,1);\mathbb{C}) = \{f \in W^{1,2}((0,1);\mathbb{C}) : f(0) = f(1) = 0\}$. Then, $D(T) = W_0^{1,2}((0,1);\mathbb{C})$ is a dense subspace of $L^2((0,1);\mathbb{C})$. Moreover, for every pair of functions $f \in W_0^{1,2}((0,1);\mathbb{C})$ and $g \in W^{1,2}((0,1);\mathbb{C})$, an integration by parts shows that

$$\langle Tf, g\rangle = \int_0^1 if'\overline{g}dx = -\int_0^1 if\overline{g'}dx = \int_0^1 f\overline{(ig')}dx.$$

The previous formula implies that $W^{1,2}((0,1);\mathbb{C})$ is contained in $D(T^*)$ and $T^*g = ig'$ for every $g \in W_0^{1,2}((0,1);\mathbb{C})$. In particular, this shows that T is symmetric. On the other hand, if $g \in D(T^*)$ then, for every $f \in D(T) = W_0^{1,2}((0,1);\mathbb{C})$ it holds that

$$\langle Tf, g\rangle = \langle f, T^*g\rangle ,$$

[1]For the main properties of the space $W^{1,2}((0,1);\mathbb{C})$ we refer the reader to [1, Chapter VIII]

i.e.,

$$\int_0^1 if'\overline{g}dx = \int_0^1 f\overline{T^*g}dx\,.$$

Therefore,

$$\int_0^1 f'(i\overline{g})dx = \int_0^1 f\overline{T^*g}dx$$

for every $f \in C_c^\infty([0,1];\mathbb{C})$. Recalling the definition of the space $W^{1,2}((0,1);\mathbb{C})$, we deduce that $ig \in W^{1,2}((0,1);\mathbb{C})$. Therefore, $D(T^*) \subseteq W^{1,2}((0,1);\mathbb{C})$. We have so proved that $D(T^*) = W^{1,2}((0,1);\mathbb{C})$. As a byproduct, it follows that T cannot be self-adjoint. Finally, we observe that T^* is not symmetric. Indeed, if $f, g \in W^{1,2}((0,1);\mathbb{C})$ it holds that

$$\langle T^*f, g\rangle = \int_0^1 if'\overline{g}dx = if(1)g(1) - if(0)g(0) + \int_0^1 f\overline{ig'}dx,$$

so that $\langle T^*f, g, \rangle \neq \langle f, T^*g\rangle$ unless $f(1)g(1) - f(0)g(0) = 0$.

Symmetric linear operators on Hilbert spaces are bounded as the following result shows.

Theorem 6.1.11 (Hellinger-Toeplitz theorem). *If $T : H \to H$ is a symmetric linear operator, then $T \in \mathcal{L}(H)$. In particular, T is self-adjoint.*

Proof. In view of Theorem 1.2.3, it suffices to prove that T is closed. Fix a sequence $(x_n)_{n\in\mathbb{N}} \subseteq H$ such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} Tx_n = y$ for some $x, y \in H$. Note that, for every $z \in H$, it holds that

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

As a consequence, $Tx = y$. □

Proposition 6.1.12. *Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator. Then, the following properties hold true.*

(1) *T^* is closed.*

(2) *T is closable if and only if $D(T^*)$ is dense in H. In such a case, $\overline{T} = T^{**}$.*

(3) *If T is closable, then $(\overline{T})^* = T^*$.*

Proof. We first observe that on the product space $H \times H$, we can define in a natural way an inner product by setting

$$\langle (x_1, y_1), (x_2, y_2)\rangle_{H\times H} := \langle x_1, x_2\rangle_H + \langle y_1, y_2\rangle_H$$

for every $(x_1, y_1), (x_2, y_2) \in H \times H$. It is easy to check that the space $H \times H$, endowed with the inner product defined above, is a Hilbert space. Moreover, the linear operator $V : H \times H \to H \times H$, defined by $V(x, y) := (-y, x)$ for every $(x, y) \in H \times H$ preserves the inner product (i.e., it is a unitary operator), is surjective and $V^2 = -I$. In particular, for every subspace $E \subseteq H \times H$ the following formula holds true:

$$V(E^\perp) = V(E)^\perp\,. \tag{6.2}$$

If $S\colon D(S) \subseteq H \to H$ is a densely defined linear operator, then it holds that

$$\begin{aligned}(x,y) \in [V(\mathcal{G}(S))]^{\perp} &\iff \langle (x,y), V(x_1,y_1)\rangle_{H\times H} = 0 \ \ \forall (x_1,y_1) \in \mathcal{G}(S)\\ &\iff \langle (x,y), (-Sz, z)\rangle_{H\times H} = 0 \ \ \forall z \in D(S)\\ &\iff -\langle x, Sz\rangle + \langle y, z\rangle = 0 \ \ \forall z \in D(S)\\ &\iff \langle x, Sz\rangle = \langle y, z\rangle \ \ \forall z \in D(S)\\ &\iff x \in D(S^*) \text{ and } S^*x = y\\ &\iff (x,y) \in \mathcal{G}(S^*).\end{aligned}$$

We have thus proved that

$$\mathcal{G}(S^*) = [V(\mathcal{G}(S))]^{\perp}. \tag{6.3}$$

We can now prove the three properties in the statement of the proposition.

(1) Since the orthogonal space of a subspace of a Hilbert space is closed, from (6.3) it follows that $\mathcal{G}(T^*)$ is a closed subspace of $H \times H$ and, hence, T^* is a closed linear operator.

(2) As it is easy to check,

$$\overline{\mathcal{G}(T)} = [\mathcal{G}(T)^{\perp}]^{\perp}.$$

Since the operator V defined above satisfies properties (6.2), (6.3) and $V^2 = -I$, it follows that

$$\overline{\mathcal{G}(T)} = V^2\left[[\mathcal{G}(T)^{\perp}]^{\perp}\right] = \left[V[V(\mathcal{G}(T))]^{\perp}\right]^{\perp} = [V(\mathcal{G}(T^*))]^{\perp}\ . \tag{6.4}$$

Let us suppose that $\overline{D(T^*)} = H$. Then, applying first (6.4) and then (6.3), we obtain that

$$\overline{\mathcal{G}(T)} = [V(\mathcal{G}(T^*))]^{\perp} = \mathcal{G}(T^{**})\,.$$

This ensures that $(T^{**}, D(T^{**}))$ is a closed linear operator. Hence, taking Proposition 6.1.6(2) into account, we conclude that T is closable and $\overline{T} = T^{**}$.

Vice versa, let us suppose that T is a closable operator. By contradiction, we suppose that $D(T^*)$ is not dense in H. Then, $D(T^*)^{\perp}$ is a proper subspace of H and, hence, there exists $0 \neq x \in D(T^*)^{\perp}$. So for every $y \in D(T^*)$ it holds that

$$\langle (x,0), (y, T^*y)\rangle_{H\times H} = \langle x, y\rangle + \langle 0, T^*y\rangle = 0\,,$$

i.e., $(x,0) \in [\mathcal{G}(T^*)]^{\perp}$. Therefore, $(0,x) \in V[\mathcal{G}(T^*)^{\perp}] = [V(\mathcal{G}(T^*))]^{\perp} = \overline{\mathcal{G}(T)}$, by (6.4). Now, since $x \neq 0$ and $(0,x) \in \overline{\mathcal{G}(T)}$, the space $\overline{\mathcal{G}(T)}$ cannot be the graph of a linear operator and, therefore, T is not closable, which is a contradiction.

(3) Due to property (1), the linear operator T^* is closed. This allows us to apply property (2) to the operator T^* to infer that

$$T^* = \overline{T^*} = (T^*)^{**} = (T^{**})^* = (\overline{T})^*. \quad \square$$

Corollary 6.1.13. *If $T : D(T) \subseteq H \to H$ is a symmetric densely defined linear operator, then T is closable and $\overline{T} = T^{**}$.*

Proof. It suffices to observe that $T \subset T^*$ and apply properties (1), (2) in Proposition 6.1.12. $\square$

Remark 6.1.14. If $T : D(T) \subseteq H \to H$ is a densely defined symmetric operator, then $T \subset T^*$ and, hence, $T^{**} = \overline{T} \subset T^*$. If T is also closed, then

$$T = \overline{T} = T^{**} \subset T^*. \tag{6.5}$$

Consequently, if T is a symmetric closed operator, then T is self-adjoint if and only if T^* is symmetric.

Proposition 6.1.15. *Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator. Then, T is symmetric if and only if $\langle Tx, x\rangle \in \mathbb{R}$ for every $x \in D(T)$.*

Proof. Let us suppose that T is symmetric. Then,

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

for every $x \in D(T)$, from which it follows that $\langle Tx, x\rangle \in \mathbb{R}$.

Vice versa, let us assume that $\langle Tz, z\rangle \in \mathbb{R}$ for every $z \in D(T)$. Then, for every $x, y \in D(T)$ it holds that

$$\langle Ty, x\rangle + \langle Tx, y\rangle = -\langle T(x-y), x-y)\rangle + \langle Tx, x\rangle + \langle Ty, y\rangle \in \mathbb{R}\,.$$

Therefore, $\mathrm{Im}\langle Tx, y\rangle = -\mathrm{Im}\langle Ty, x\rangle = \mathrm{Im}\langle x, Ty\rangle$. Analogously,

$$i\langle Ty, x\rangle - i\langle Tx, y\rangle = -\langle T(x-iy), x-iy)\rangle + \langle Tx, x\rangle + \langle Ty, y\rangle \in \mathbb{R}$$

for every $x, y \in D(T)$. Therefore, $\mathrm{Re}\langle Tx, y\rangle = \mathrm{Im}\, i\langle Tx, y\rangle = \mathrm{Im}\, i\langle Ty, x\rangle = \mathrm{Re}\langle Ty, x\rangle = \mathrm{Re}\langle x, Ty\rangle$. We have thus proved that $\langle Tx, y\rangle = \langle x, Ty\rangle$ for every $x, y \in D(T)$, i.e., T is symmetric. □

Theorem 6.1.16. *Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator. If T is symmetric, then the following properties are equivalent.*

(i) *T is self-adjoint.*

(ii) *T is closed and* $\ker(T^* \pm i) = \{0\}$.

(iii) $\mathrm{rg}(T \pm i) = H$.

Proof. "(i) ⇒ (ii)": Since $T = T^*$, due to Proposition 6.1.12(i) we can conclude that T is closed. Next, we fix $x \in D(T^*) = D(T)$ such that $T^*x = Tx = ix$. Then, recalling that T is symmetric, it follows that

$$i\langle x, x\rangle = \langle ix, x\rangle = \langle Tx, x\rangle$$

which implies that $x = 0$ since $\langle Tx, x\rangle \in \mathbb{R}$ due to Proposition 6.1.15. In a completely similar way, we can show that $\ker(T^* + i) = \{0\}$.

"(ii) ⇒ (iii)": Let us first prove that $\mathrm{rg}(T - i)$ is a dense subspace of H. For this purpose, we fix $y \in H$ such that $\langle Tz - iz, y\rangle = 0$ for every $z \in D(T)$. Then,

$$\langle Tz, y\rangle = \langle z, -iy\rangle$$

for every $z \in D(T)$. As a byproduct, we deduce that $y \in D(T^*)$ and $T^*y = -iy$. Therefore, from property (ii) it follows that $y = 0$. We have so proved that $[\mathrm{rg}(T - i)]^{\perp} = \{0\}$, which implies that $\mathrm{rg}(T - i)$ is a dense subspace of H.

Let us now prove that $\mathrm{rg}(T-i)$ is a closed subspace of H. For this purpose, we observe that, due to Proposition 6.1.15,

$$\begin{aligned}\|(T-i)x\|^2 =&\|Tx\|^2+\|x\|^2+2\mathrm{Re}\langle Tx,-ix\rangle\\ =&\|Tx\|^2+\|x\|^2+2\mathrm{Re}[i\langle Tx,x\rangle]\\ =&\|Tx\|^2+\|x\|^2 \end{aligned}\tag{6.6}$$

for every $x\in D(T)$. If $(x_n)_{n\in\mathbb{N}}\subset D(T)$ is a sequence such that $\lim_{n\to\infty}(Tx_n-ix_n)=y_0\in H$, then $(Tx_n-ix_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in H. On the other hand, formula (6.6) implies that

$$\|x_n-x_m\|^2+\|Tx_n-Tx_m\|^2=\|Tx_n-ix_n-Tx_m+ix_m\|^2$$

for every $m,n\in\mathbb{N}$. Consequently, also $(x_n)_{n\in\mathbb{N}}$ and $(Tx_n)_{n\in\mathbb{N}}$ are Cauchy sequences in H. In particular, they converge, respectively, to some $x_0\in H$ and $z_0\in H$. Since T is a closed operator, it follows that $x_0\in D(T)$ and $z_0=Tx_0$, i.e., $y_0=(T-i)x_0\in\mathrm{rg}(T-i)$.

Arguing in the same way, we can prove that $\mathrm{rg}(T+i)$ is closed.

"(iii) $\Rightarrow$ (i)": To begin with, we show that $\ker(T^*-i)=\{0\}$. For this purpose, we fix $z\in D(T^*)$ such that $T^*z=iz$. Then,

$$\langle Tx+ix,z\rangle=\langle x,T^*z-iz\rangle=0,\qquad x\in D(T).$$

Since $\mathrm{rg}(T+i)=H$ by assumptions, it follows that $z=0$.

Now, fix $y\in D(T^*)$. By assumptions, there exists $x\in D(T)$ such that $(T-i)x=(T^*-i)y$. Since T is symmetric, it follows that $D(T)\subset D(T^*)$ and, consequently, $y-x\in D(T^*)$ and $(T^*-i)(y-x)=0$. Hence, $y-x=0$ and so $y\in D(T)$. We have so proved that $D(T)=D(T^*)$, or, equivalently, that T is self-adjoint. □

Remark 6.1.17. The assumption $\ker(T\pm i)=\{0\}$ in (ii) cannot be removed. Indeed, the linear operator

$$T:W_0^{1,2}((0,1);\mathbb{C})\subset L^2((0,1);\mathbb{C})\to L^2((0,1);\mathbb{C}),\qquad Tf=if',$$

is symmetric and $D(T^*)=W^{1,2}((0,1);\mathbb{C})$ (cf. Remark 6.1.10). Thus, T is not self-adjoint. Moreover, $(T,D(T))$ is also a closed operator. Indeed, let $(\psi_n)_{n\in\mathbb{N}}\subset W_0^{1,2}((0,1);\mathbb{C})$ be a sequence such that $\lim_{n\to\infty}\psi_n=\psi$ and $\lim_{n\to\infty}T\psi_n=\varphi$ in $L^2((0,1);\mathbb{C})$. Then, for every $\xi\in C_c^\infty((0,1);\mathbb{C})$, it holds that

$$\int_0^1\psi\xi'dx=\lim_{n\to\infty}\int_0^1\psi_n\xi'dx=-\lim_{n\to\infty}\int_0^1\psi_n'\xi dx=i\int_0^1\varphi\xi dx.$$

This ensures that $\psi'\in L^2((0,1);\mathbb{C})$ and $\psi'=-i\varphi$ or, equivalently, $\varphi=i\psi'$. We can thus conclude that ψ_n converges to ψ in $W^{1,2}((0,1);\mathbb{C})$ as n tends to ∞. Since $(\psi_n)_{n\in\mathbb{N}}\subset W_0^{1,2}((0,1))$ and $W_0^{1,2}((0,1);\mathbb{C})$ is a closed subspace of $W^{1,2}((0,1);\mathbb{C})$, it also follows that $\psi\in W_0^{1,2}((0,1);\mathbb{C})$. This completes the proof of the closedness of T. On the other hand, the functions $f_\pm\in D(T^*)$ defined by $f_\pm(x):=e^{\pm x}$ are such that $T^*f_\pm=\pm if_\pm$. Hence, $\ker(T^*\pm i)\neq\{0\}$.

6.2 Dissipative operators

Definition 6.2.1. A linear operator $T : D(T) \subseteq H \to H$ is called dissipative if

$$\mathrm{Re}\langle Tx, x\rangle \leq 0 \qquad x \in D(T).$$

Example 6.2.2. Let us consider the linear operator T on $L^2((0,1),\mathbb{C})$, defined by

$$D(T) := \{f \in C^1([0,1],\mathbb{C}) : f(0) = f(1) = 0\}, \qquad Tf := f'\,.$$

If $f \in D(T)$, then

$$\langle Tf, f\rangle = \int_0^1 f'\overline{f}dx = -\int_0^1 f\overline{f'}dx = -\overline{\int_0^1 f'\overline{f}dx}\,.$$

This implies that $\mathrm{Re}\langle Tf, f\rangle = 0$. The arbitrariness of $f \in D(T)$ yields the dissipativity of T.

Example 6.2.3. The multiplication operator M_m introduced in Example 6.1.4 is dissipative if and only if for every $f \in D(M_m)$

$$\mathrm{Re}\int_\Omega m|f|^2 dx \leq 0,$$

namely if and only if $\mathrm{Re}(m) \leq 0$ μ-almost everywhere in Ω. Indeed, consider the same sequence $(\Omega_n)_{n\in\mathbb{N}}$ of subsets of Ω as in Example 6.1.4 and, for every $k, n \in \mathbb{N}$ let $f_{k,n} = \chi_{A_{k,n}}$, where $A_{k,n} = \{x \in \Omega_n : \mathrm{Re}(m) \geq 1/k \text{ and } |m(x)| \leq n\}$. Clearly, each function $f_{k,n}$ belongs to $D(M_m)$. Therefore,

$$0 \geq \int_\Omega \mathrm{Re}(m)|f_{k,n}|^2 d\mu \geq \int_\Omega k^{-1}\chi_{A_{k,n}} d\mu = k^{-1}\mu(A_{k,n}),$$

so that $\mu(A_{k,n}) = 0$. Letting first n and, then, k tend to ∞, we conclude that $\mathrm{Re}(m) \leq 0$ μ-almost everywhere on Ω.

Theorem 6.2.4. *Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator. Then, the following properties are equivalent.*

(i) *T is symmetric.*

(ii) *$\pm iT$ is dissipative.*

Proof. "(i) ⇒ (ii)": Since T is symmetric, $\langle Tx, x\rangle \in \mathbb{R}$ due to Proposition 6.1.15. Therefore, for every $x \in D(T)$ it holds that

$$\mathrm{Re}\langle \pm iTx, x\rangle = \pm\mathrm{Re}(i\langle Tx, x\rangle) = 0.$$

This implies that $\pm iT$ is dissipative.

"(ii) ⇒ (i)": Due to Proposition 6.1.15 it suffices to show that $\langle Tx, x\rangle \in \mathbb{R}$ for every $x \in D(T)$. For this purpose, we observe for every $x \in D(T)$

$$\mathrm{Im}\langle Tx, x\rangle = -\mathrm{Re}\langle iTx, x\rangle = \mathrm{Re}\langle -iTx, x\rangle.$$

Since $\pm iT$ is dissipative, it follows that $\mathrm{Re}\langle iTx, x\rangle = 0$. So, $\mathrm{Im}\langle Tx, x\rangle = 0$ for every $x \in D(T)$. Hence, $\langle Tx, x\rangle = \mathrm{Re}\langle Tx, x\rangle \in \mathbb{R}$ for every $x \in D(T)$ and this implies that T is symmetric. □

Proposition 6.2.5. *A densely defined linear operator $T : D(T) \subseteq H \to H$ is dissipative if and only if $\|\lambda x - Tx\| \geq \lambda\|x\|$ for every $x \in D(T)$ and $\lambda > 0$.*

Proof. Let us suppose that T is dissipative. Then, for every $x \in D(T)$ and $\lambda > 0$ it holds that

$$\|\lambda x - Tx\|\|x\| \geq |\langle \lambda x - Tx, x\rangle| \geq \mathrm{Re}\langle \lambda x - Tx, x\rangle = \lambda\|x\|^2 - \mathrm{Re}\langle Tx, x\rangle \geq \lambda\|x\|^2.$$

Vice versa, if $x \in D(T)$, then

$$\lambda^2\|x\|^2 \leq \|\lambda x - Tx\|^2 = \lambda^2\|x\|^2 + \|Tx\|^2 - 2\lambda\mathrm{Re}\langle Tx, x\rangle\,.$$

for every $\lambda > 0$. This implies that $\|Tx\|^2 - 2\lambda\mathrm{Re}\langle Tx, x\rangle \geq 0$ for every $\lambda > 0$ so that $\mathrm{Re}(Tx, x) \leq 0$. □

Corollary 6.2.6. *If $T : D(T) \subseteq H \to H$ is a dissipative operator, then $\lambda - T$ is injective for every $\lambda > 0$. If, in addition, T is also closed, then $\mathrm{Rg}(\lambda - T)$ is a closed in H for every $\lambda > 0$.*

Proof. Fix $\lambda > 0$. The injectivity of $\lambda - T$ follows immediately from Proposition 6.2.5.

Let us assume now that T is also closed and that $(x_n)_{n\in\mathbb{N}} \subset D(T)$ is a sequence such that $\lambda x_n - Tx_n$ converges to some $y \in H$ as n tends to ∞. Due to Proposition 6.2.5, we know that

$$\lambda\|x_n - x_m\| \leq \big\|\lambda x_n - Tx_n - \lambda x_m + Tx_m\big\|, \qquad n, m \in \mathbb{N}\,.$$

It turns out that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in H and, therefore, it converges to some $x \in H$. As a byproduct, $\lim_{n\to\infty} Tx_n = \lambda x - y$. Since T is closed, we can conclude that $x \in D(T)$ and $y = \lambda x - Tx$. This shows that $\mathrm{Rg}(\lambda - T)$ is a closed subspace of H. □

Theorem 6.2.7. *Let $T : D(T) \subseteq H \to H$ be a dissipative operator. If there exists $\lambda_0 \in \mathbb{C}$, with positive real part, such that $(\lambda_0 - T)(D(T)) = H$, then $\lambda_0 \in \rho(T)$. In particular, $\mathbb{C}_+ = \{\lambda \in \mathbb{C} \mid \mathrm{Re}\,\lambda > 0\} \subseteq \rho(T)$ and the following inequality holds true*

$$\|R(\lambda, T)\| \leq \frac{1}{\mathrm{Re}\,\lambda}, \qquad \forall \lambda \in \mathbb{C}_+\,. \tag{6.7}$$

Proof. We just need to prove that the operator $\lambda_0 - T$ is injective since $(\lambda_0 - T)(D(T)) = H$ by assumptions. For this purpose, we fix $x \in D(T)$, set $y := \lambda_0 x - Tx$ and observe that

$$\begin{aligned}\mathrm{Re}\lambda_0\,\|x\|^2 =&\mathrm{Re}\lambda_0\,\langle x, x\rangle = \mathrm{Re}\langle \lambda_0 x, x\rangle\\ =&\mathrm{Re}\langle y + Tx, x\rangle = \mathrm{Re}\langle y, x\rangle + \mathrm{Re}\langle Tx, x\rangle\\ \leq&\mathrm{Re}\langle y, x\rangle \leq \|x\| \cdot \|y\|\,.\end{aligned} \tag{6.8}$$

If $y = 0$, then from (6.8) it follows that $x = 0$ as well. Hence, the operator $\lambda_0 - T$ is injective. Moreover, inequality (6.8) also implies that

$$\|(\lambda_0 - T)^{-1}y\| = \|x\| \leq \frac{1}{\mathrm{Re}\lambda_0}\|y\|.$$

Thus, we have proved that $\lambda_0 \in \rho(T)$ and $\|R(\lambda_0, T)\| \leq (\mathrm{Re}\,\lambda_0)^{-1}$.

If $\lambda \in \mathbb{C}_+ \cap \rho(T)$ with $\lambda \neq \lambda_0$, then, using the previous argument, we can show that $\|R(\lambda, T)\| \leq (\mathrm{Re}\,\lambda)^{-1}$.

To complete the proof, it suffices to prove that $\mathbb{C}_+ \cap \rho(T) = \mathbb{C}_+$. For this purpose, we argue as follows. We observe that $\mathbb{C}_+ \cap \rho(T)$ is a not empty open subset of $\mathbb{C}_+$. Let us prove that it is also a closed subset of $\mathbb{C}_+$. Let $(\lambda_n)_{n\in\mathbb{N}} \subset \rho(T) \cap \mathbb{C}_+$ be a sequence converging to some $\lambda \in \mathbb{C}_+$. Without loss of generality, we can assume that $\mathrm{Re}\lambda_n \geq c > 0$ for every $n \in \mathbb{N}$. Therefore, $\|R(\lambda_n, T)\| \leq c^{-1}$, in view of (6.7). Since $(\lambda_n)_{n\in\mathbb{N}}$ converges to λ, there exists $\overline{n}$ such that

$$|\lambda - \lambda_{\overline{n}}| < c \leq \frac{1}{\|R(\lambda_{\overline{n}}, T)\|}\,.$$

Therefore, $\lambda \in \rho(T)$ due to Proposition 1.3.5(1). We have so proved that $\mathbb{C}_+ \cap \rho(T)$ is a closed subset of $\mathbb{C}_+$. Since $\mathbb{C}_+ \cap \rho(T) \neq \varnothing$ is an open and closed subset of $\mathbb{C}_+$, which is a connected set, we can infer that $\mathbb{C}_+ \cap \rho(T) = \mathbb{C}_+$. □

Proposition 6.2.8. *Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator. If T is also symmetric, then the following properties are satisfied.*

(1) *If there exists $\lambda \in \mathbb{C}$, with $\mathrm{Im}\,\lambda > 0$, such that $(\lambda - T)(D(T)) = H$, then $\{\mu \in \mathbb{C} : \mathrm{Im}\,\mu > 0\} \subseteq \rho(T)$.*

(2) *If there exists $\lambda \in \mathbb{C}$, with $\mathrm{Im}\,\lambda < 0$, such that $(\lambda - T)(D(T)) = H$, then $\{\mu \in \mathbb{C} : \mathrm{Im}\,\mu < 0\} \subseteq \rho(T)$.*

Proof. Let us observe that, by virtue of Theorem 6.2.4, the operator $\pm iT$ is dissipative since T is symmetric.

(1) Fix $\lambda \in \mathbb{C}$ with positive imaginary part and set $\mu := -i\lambda$. Note that $\mathrm{Re}\,\mu = \mathrm{Im}\,\lambda > 0$ and the operator $\mu + iT$ is surjective by assumptions. We can thus apply the previous theorem to $-iT$ to infer that $\mathbb{C}_+ \subseteq \rho(-iT)$, or, equivalently, $i\mathbb{C}_+ \subseteq \rho(T)$. Since $i\mathbb{C}_+ = \{z \in \mathbb{C} \mid \mathrm{Im}\,z > 0\}$, the assertion follows.

(2) The property can be obtained applying the argument in the proof of (1) to the operator iT. □

Remark 6.2.9. From Proposition 6.2.8 it follows that, if T is a densely defined, symmetric operator, then the spectrum $\sigma(T)$ satisfies one and only one of the following conditions:

- $\sigma(T) \subseteq \{\mu \in \mathbb{C} \mid \mathrm{Im}\,\mu \geq 0\}$.
- $\sigma(T) \subseteq \{\mu \in \mathbb{C} \mid \mathrm{Im}\,\mu \leq 0\}$.
- $\sigma(T) = \mathbb{C}$.
- $\sigma(T) \subseteq \mathbb{R}$ (if there exist $\lambda_1, \lambda_2 \in \rho(T)$, with $\mathrm{Im}\,\lambda_1 > 0$ and $\mathrm{Im}\,\lambda_2 < 0$).

Corollary 6.2.10. *Let $T : D(T) \subseteq H \to H$ be a densely defined linear operator. If T is self-adjoint, then $\sigma(T) \subseteq \mathbb{R}$.*

Proof. Since T is self-adjoint, it follows that $\mathrm{Rg}(T \pm i) = H$ in view of Theorem 6.1.16. Therefore, Proposition 6.2.8 implies that $\{z \in \mathbb{C} : \mathrm{Im} z \neq 0\} \subseteq \rho(T)$ so that $\sigma(T) \subseteq \mathbb{R}$. □

Proposition 6.2.11. *Let $T : D(T) \subseteq H \to H$ be a symmetric and dissipative densely defined linear operator. The following properties are equivalent.*

(i) *T is self-adjoint*

(ii) $\sigma(T) \subseteq (-\infty, 0]$.

(iii) $(I - T)(D(T)) = H$.

Proof. "(i)⇒ (ii)": Due to Corollary 6.2.10, we just need to show that, if $\lambda > 0$, then $\lambda \in \rho(T)$.

If $\lambda > 0$, then the operator $\lambda - T$ is injective and its range is closed, due to Corollary 6.2.6 and Theorem 6.1.16. On the other hand,

$$\{0\} = \ker(\lambda - T) = [\mathrm{Rg}(\lambda - T)]^{\perp},$$

see Exercise 6.3.1, since T is self-adjoint, so that $(\lambda - T)(D(T))$ is a dense subspace of H. Therefore, $(\lambda - T)(D(T)) = H$, i.e., $\lambda - T$ is also surjective. By Proposition 1.3.2, $(\lambda - T)^{-1}$ belongs to $\mathcal{L}(H)$. We have so proved that $\lambda \in \rho(T)$.

"(ii) ⇒ (i)": From property (ii), it follows that $\pm i \in \rho(T)$ and, therefore, T is self-adjoint due to Theorem 6.1.16.

"(ii) ⇒ (iii)": From property (ii) it follows that $1 \in \rho(T)$.

"(iii) ⇒ (ii)": Theorem 6.2.7 implies that $\mathbb{C}_+ \subseteq \rho(T)$. Therefore, there exist $\lambda_1, \lambda_2 \in \rho(T)$, with $\mathrm{Im}\lambda_1 > 0$ and $\mathrm{Im}\lambda_2 < 0$. Remark 6.2.9 guarantees that $\pm i \in \rho(T)$ and, thus, T is self-adjoint due to Theorem 6.1.16. □

Example 6.2.12. Let $H = L^2((0,1);\mathbb{C})$ and let $(A, D(A))$ be the operator, defined by

$$D(A) = W_0^{1,2}((0,1);\mathbb{C}) \cap W^{2,2}((0,1);\mathbb{C}), \qquad Af = f''.$$

The operator $(A, D(A))$ is the Laplacian operator with homogeneous Dirichlet boundary conditions. Then, A is a self-adjoint dissipative operator on $L^2((0,1);\mathbb{C})$, see Exercise 6.3.6.

6.3 Exercises

Exercise 6.3.1. Let $T\colon D(T) \subseteq H \to H$ be a densely defined operator on a complex Hilbert space H. Prove that

(i) $(\mathrm{rg}\, T)^{\perp} = \ker T^*$.

(ii) If T is closed, then $(\mathrm{rg}\, T^*)^{\perp} = \ker T$.

Exercise 6.3.2. Let $T\colon D(T) \subseteq H \to H$ be a symmetric densely defined operator on a complex Hilbert space H. Prove that

(i) If $\overline{\mathrm{rg}\, T} = H$, then T is one-to-one.

(ii) If $T = T^*$ and T is one-to-one, then $\overline{\mathrm{rg}\, T} = H$ and T^{-1} is self-adjoint.

Exercise 6.3.3. Let $T\colon D(T) \subseteq H \to H$ be a symmetric densely defined operator on a complex Hilbert space H. Prove that if $\lambda \in \mathbb{C}$ is an approximate eigenvalue of T, then $\lambda \in \mathbb{R}$.

Exercise 6.3.4. Let $T : D(T) \subseteq H \to H$ and $S : D(S) \subseteq H \to H$ be self-adjoint operators on a complex Hilbert space H. Is $S + T$ a self-adjoint operator?

Exercise 6.3.5. Prove that a densely defined linear dissipative operator on a Hilbert space is closable and its closure is dissipative.

Exercise 6.3.6. On $L^2((0,1);\mathbb{C})$ consider the operator, defined by

$$D(A) = W_0^{1,2}((0,1);\mathbb{C}) \cap W^{2,2}((0,1);\mathbb{C}), \qquad Af = f''.$$

Prove that A is a self-adjoint and dissipative operator on $L^2((0,1);\mathbb{C})$.

Exercise 6.3.7. Let $p, q : [a,b] \to \mathbb{R}$ be respectively a strictly positive smooth function and a continuous function. Set $Lu = -(pu')' + qu$ for every $u \in D(L)$, where

$$D(L) = \{y \in C^1([a,b]) \cap C^2((a,b)) \mid \alpha_1 u(a) + \alpha_2 u'(a) = \beta_1 u(b) + \beta_2 u'(b) = 0\}$$

with $\alpha_i, \beta_i \in \mathbb{R}$, $\alpha_1^2 + \alpha_2^2 > 0, \beta_1^2 + \beta_2^2 > 0$. Prove that $(L, D(L))$ is a symmetric operator on $L^2([a,b])$ (L is a called a regular Sturm-Lioville operator).
Is it self-adjoint?
[**Hint:** prove that $(Lu)v - (Lu)u = ((-pu')v + u(pv'))'$ for every $u, v \in D(L)$.)]

Bibliography

[1] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[2] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[3] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[4] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[5] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[6] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[7] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981

[8] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[9] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[10] W. Rudin, Real and Complex Analysis, 3^{rd} ed., McGraw-Hill, 1986.

[11] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[12] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958

[13] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

December 1, 2021

Lecture 7

Spectral representation theorem for unbounded operators II

In this lecture, we conclude our journey through spectral theory. The last steps are the spectral representation theorem for self-adjoint (possibly) unbounded operators, and its version for self-adjoint operators with compact resolvent. In the latter situation, adding the assumption of positivity, we also prove the celebrated Rayleigh-Ritz variational formula, which gives an approximation of the eigenvalues of the operator. As an application of this formula, we compare the sequences of eigenvalues associated with two positive self-adjoint operators with compact resolvent.

7.1 Representation theorem for unbounded self-adjoint operators

In this section, we prove the following important result.

Theorem 7.1.1. *Let H be a separable complex Hilbert space and let $T : D(T) \subseteq H \to H$ be a self-adjoint operator. Then, there exist a measurable space (Y, μ), with finite measure μ, a unitary operator $U : H \to L^2(Y, d\mu)$ and a μ-measurable function $q : Y \to \mathbb{R}$ such that*

(1) $x \in D(T) \iff Ux \in D(M_q)$;

(2) $Tx = U^{-1}M_qUx$ *for every* $x \in D(T)$.

Proof. By virtue of Theorem 6.1.16, the operators $T + i$ and $T - i$, defined on $D(T)$, are one to one and closed. Moreover, $\mathrm{rg}(T \pm i) = H$. Therefore, the operators $(T + i)^{-1}$ and $(T - i)^{-1}$ are well defined and, in particular, they are bounded on H. Moreover, they commute each other thanks to the resolvent identity.

Next, we observe that, for every $x, y \in D(T)$, it holds that

$$\begin{aligned}\langle (T-i)x, (T+i)^{-1}(T+i)y\rangle =& \langle (T-i)x, y\rangle \\ =& \langle x, (T+i)y\rangle \\ =& \langle (T-i)^{-1}(T-i)x, (T+i)y\rangle .\end{aligned}$$

Since $\mathrm{rg}(T \pm i) = H$, it follows that

$$\langle z_1, (T+i)^{-1}z_2\rangle = \langle (T-i)^{-1}z_1, z_2\rangle$$

for every $z_1, z_2 \in H$. This ensures that $((T+i)^{-1})^* = (T-i)^{-1}$, so that $(T+i)^{-1}$ is a normal operator. We can thus apply Theorem 5.1.9 and infer that there exist a measurable space (Y, μ), with finite measure μ, a unitary operator $U : H \to L^2(Y, d\mu)$ and a μ-measurable and bounded function $m : Y \to \mathbb{C}$ such that $U(T+i)^{-1}U^{-1} = M_m$. Since $\ker(T+i)^{-1} = \{0\}$, it follows that $m \neq 0$ μ-a.e. so that we can define the function $q := m^{-1} - i$. Clearly, q is a μ-measurable function.

We can now prove properties (1) and (2).

Fix $x \in D(T)$ and set $y = (T+i)x$. Then, $x = (T+i)^{-1}y = U^{-1}M_mUy$, so that $Ux = M_mUy$, which shows that $(U^{-1}M_{\frac{1}{m}}U)x = y = Tx + ix$. It thus follows that $Tx = (U^{-1}M_{\frac{1}{m}}U)x - iU^{-1}Ux = U^{-1}M_qUx$ and property (2) follows.

From the previous formula, we also deduce that $Ux \in D(M_q)$, so that the implication "$\Longrightarrow$" in (1) follows. To prove the other implication, we fix $x \in H$ such that $Ux \in D(M_q)$ and observe that

$$U^{-1}M_qUx = (U^{-1}M_{\frac{1}{m}}U)x - ix\,.$$

Set $z := U^{-1}M_{\frac{1}{m}}Ux$. It is easy to check that $x = (T+i)^{-1}z$, so that $x \in D(T)$.

Finally, recalling that T is a self-adjoint operator, we can apply Corollary 6.2.10 to claim that $\sigma(T) \subseteq \mathbb{R}$. Moreover, since the operator M_q is self-adjoint, we can apply Exercise 2.3.4 to conclude that q is a real-valued function. □

7.2 Spectral representation theorem for self-adjoint operators

In this section, we recall that the spectrum of an operator with compact resolvent consists only of eigenvalues. This permits us to diagonalize every self-adjoint operator with compact resolvent on complex separable Hilbert spaces.

Definition 7.2.1. Let H be a Hilbert space. We say that a linear operator $T\colon D(T) \subseteq H \to H$ has compact resolvent if $\rho(T) \neq \varnothing$ and $R(\lambda, T)$ is a compact operator for all $\lambda \in \rho(T)$.

The following result gives a useful characterization of operators with compact resolvent.

Proposition 7.2.2. *Let H be a Hilbert space. Let $T : D(T) \subseteq H \to H$ be a linear operator with $\rho(T) \neq \varnothing$. Then, T has compact resolvent if and only if the canonical embedding ι : $(D(T), \|\cdot\|_{D(T)}) \hookrightarrow H$ is compact, where $\|\cdot\|_{D(T)}$ denotes the graph norm.*

Proof. We set $H_1 = (D(T), \|\cdot\|_{D(T)})$ and observe that, for every $\lambda \in \rho(T)$, the graph norm $\|\cdot\|_{D(T)}$ is equivalent to the norm

$$\|x\|_\lambda := \|(\lambda - T)x\|, \qquad x \in D(T).$$

Indeed, for every $x \in D(T)$, it holds that

$$\begin{aligned}
\|x\|_{D(T)} =& \|x\| + \|Tx\| \\
=& \|R(\lambda, T)(\lambda - T)x\| + \|\lambda x - (\lambda - T)x\| \\
\le& \|R(\lambda, T)(\lambda - T)x\| + \|\lambda R(\lambda, T)(\lambda x - Tx)\| + \|(\lambda - T)x\| \\
=& [(1 + |\lambda|)\|R(\lambda, T)\| + 1]\|(\lambda - T)x\|
\end{aligned}$$

$$=[(1+|\lambda|)\|R(\lambda,T)\|+1]\|x\|_\lambda$$

and

$$\|x\|_\lambda \le |\lambda|\|x\| + \|Tx\| \le \max\{|\lambda|,1\}\|x\|_{D(T)}.$$

Therefore, for every $\lambda \in \rho(T)$, the operator $R(\lambda,T)\colon H \to H_1$ is an isomorphism with continuous inverse $\lambda - T$.

Now, let us suppose that T has compact resolvent. Then, we can write $\iota = R(\lambda,T)(\lambda-T)$, for any $\lambda \in \rho(T)$, where $\lambda - T\colon H_1 \to H$ is a continuous linear operator and $R(\lambda,T)\colon H \to H$ is a compact operator. Therefore, ι is clearly a compact operator.

Vice versa, let us suppose that $\iota\colon H_1 \to H$ is a compact operator. Since we can write $R(\lambda,T) = \iota R(\lambda,T)$ for all $\lambda \in \rho(T)$, with $R(\lambda,T)$ on the right acting continuously from H into H_1, it follows that the operator $R(\lambda,T)$ is compact. □

The following result shows that the spectrum of an operator with compact resolvent consists only of eigenvalues.

Theorem 7.2.3. *Let H be a complex Hilbert space and let $T : D(T) \subseteq H \to H$ be a linear operator on H with compact resolvent. Then, the following properties are satisfied.*

(1) $\sigma(T) = \sigma_p(T)$.

(2) $\sigma(T)$ *is finite or* $\sigma(T) = \{\lambda_n \mid n \in \mathbb{N}\} \subseteq \mathbb{C}$ *with* $|\lambda_n| \to \infty$.

(3) $\dim\ker(\lambda - T)$ *is finite for all* $\lambda \in \sigma(T)$.

Proof. The result follows from Proposition 1.3.9 and Proposition 3.2.14. □

As in the finite dimensional case, one can show that every self-adjoint operator with compact resolvent on separable complex Hilbert spaces can be diagonalized.

Theorem 7.2.4. *Let H be a separable complex Hilbert space and let $T : D(T) \subseteq H \to H$ be a self-adjoint operator with compact resolvent. Then, there exist a sequence $(\lambda_n))n \in \mathbb{N} \subseteq \mathbb{R}$ and a complete orthonormal system $(e_n)_{n\in\mathbb{N}}$ of H, with $e_n \in D(T)$ for all $n \in \mathbb{N}$, such that*

(1) $Te_n = \lambda_n e_n$ *for all* $n \in \mathbb{N}$,

(2) $D(T) = \{x \in H \mid (\lambda_n\langle x, e_n\rangle)_{n\in\mathbb{N}} \in \ell^2\}$,

(3) $Tx = \sum_{n=1}^\infty \lambda_n\langle x, e_n\rangle e_n$ *for all* $x \in D(T)$.

Proof. By Theorem 7.2.3 there exists $\mu > 0$ such that $\mu \in \rho(T)$. The operator $R(\mu,T)$ is compact because T is an operator with compact resolvent. Moreover, $R(\mu,T)$ is a self-adjoint operator. Indeed, for fixed $y_1, y_2 \in H$, there exist $x_1, x_2 \in D(T)$ such that $y_i = (\mu - T)x_i$ for $i = 1, 2$ and, since T is self-adjoint, it follows that

$$\langle R(\mu,T)y_1, y_2\rangle = \langle x_1, (\mu-T)x_2\rangle = \langle(\mu-T)x_1, x_2\rangle = \langle y_1, R(\mu,T)y_2\rangle\,.$$

We can then apply Theorem 4.1.4 to conclude that there exist a complete orthonormal system $(e_n)_{n\in\mathbb{N}}$ of H and a sequence $(\alpha_n)_{n\in\mathbb{N}}$ of real numbers such that $R(\mu,T)e_n = \alpha_n e_n$ for all $n \in \mathbb{N}$ and

$$R(\mu,T)x = \sum_{n=1}^\infty \alpha_n\langle x, e_n\rangle e_n, \qquad x \in H.$$

Since $R(\mu, T)$ is an injective operator, each eigenvalue α_n differs from 0. Thus, $e_n \in D(T)$ and $Te_n = \lambda_n e_n$ with $\lambda_n := \mu - \alpha_n^{-1} \in \mathbb{R}$, for all $n \in \mathbb{N}$. So, the proof of property (1) is complete.

Next, we fix $x \in D(T)$. Then, thanks to the fact that $(e_n)_{n\in\mathbb{N}}$ is a complete orthonormal system of H, we have

$$(\lambda_n \langle x, e_n\rangle)_{n\in\mathbb{N}} = (\langle x, Te_n\rangle)_{n\in\mathbb{N}} = (\langle Tx, e_n\rangle)_{n\in\mathbb{N}} \in \ell^2$$

and

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n \, .$$

From this, property (3) and the inclusion "$\subseteq$" in property (2) follow.

To prove the other inclusion, we proceed as follows. We fix $x \in H$ such that $(\lambda_n \langle x, e_n\rangle)_{n\in\mathbb{N}} \in \ell^2$. Then, for each $k \in \mathbb{N}$ we set

$$x_k := \sum_{n=1}^{k} \langle x, e_n\rangle e_n \quad \text{and} \quad y_k := \sum_{n=1}^{k} \lambda_n \langle x, e_n\rangle e_n \, .$$

It is clear that $x_k \in D(T)$ and $Tx_k = y_k$ for all $k \in \mathbb{N}$. Moreover, x_k converges to x and Tx_k converges to $\sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n$ in H as k tends to ∞. Since T is a closed operator (see Theorem 6.1.16), we deduce that $x \in D(T)$ and $Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n$. This completes the proof. □

Example 7.2.5. The Laplace operator A with homogeneous Dirichlet boundary conditions considered in Example 6.2.12 has compact resolvent. To show this, we first observe that the embedding $\iota : (D(A), \|\cdot\|_{D(A)}) \hookrightarrow W^{1,2}((0,1))$ is continuous. Indeed, if $(f_n)_{n\in\mathbb{N}} \subseteq D(A)$ converges to f with respect to the graph norm and f_n tends to g in $W^{1,2}((0,1))$, then $\lim_{n\to\infty} f_n = f$ and $\lim_{n\to\infty} f_n = g$ in $L^2((0,1))$. Therefore, $f = g$ and this shows that the embedding operator $\iota : (D(A), \|\cdot\|_{D(A)}) \to W^{1,2}((0,1))$ is closed. Since $(D(A), \|\cdot\|_A)$ is a Banach space and $(A, D(A))$ is a closed operator, we can apply the closed graph theorem to conclude that ι is a continuous operator and, consequently, the embedding is continuous. Moreover, the embedding $W^{1,2}((0,1)) \hookrightarrow L^2((0,1))$ is compact, cf. [1, Theorem 8.8]. Thus, $(D(A), \|\cdot\|_{D(A)})$ is compactly embededded into $L^2((0,1))$. By Proposition 7.2.2, it follows that $(A, D(A))$ has compact resolvent. Finally, one can show that

$$Af = \sum_{n=1}^{\infty} n^2\pi^2 \left(\int_0^1 f(x) e_n(x) dx \right) e_n$$

for every $f \in L^2((0,1))$, where $e_n(x) = \sqrt{2}\sin(n\pi x)$ for every $x \in (0,1)$ (see Exercise 7.4.7).

7.3 Positive operators and minimax theorems for their eigenvalues

In this section, we apply Theorems 7.1.1 and 7.2.4 to study some important properties of positive self-adjoint operators on separable complex Hilbert spaces.

Definition 7.3.1. Let H be a complex Hilbert space and let $T : D(T) \subseteq H \to H$ be a symmetric linear operator. The operator T is said to be positive (for short $T \geq 0$) if

$$\langle Tx, x\rangle \geq 0, \qquad x \in D(T).$$

If S and T are two symmetric linear operators on H and $D(S) = D(T)$, then we write $S \leq T$ if $T - S \geq 0$.

Remark 7.3.2. If $c \in \mathbb{R}$, then

$$T \geq cI \iff \langle Tx, x\rangle \geq c\|x\|^2 \ \ \forall x \in D(T).$$

In particular, if T is a positive symmetric operator, then $-T$ is a dissipative operator.

Thanks to the Spectral representation theorem 7.1.1 we can state and prove the following characterization.

Theorem 7.3.3. *Let H be a separable complex Hilbert space, let $T : D(T) \subseteq H \to H$ be a self-adjoint linear operator and let $c \in \mathbb{R}$. Then, the following properties are equivalent.*

(i) $\langle Tx, x\rangle \geq c\|x\|^2$ *for all* $x \in D(T)$.

(ii) $\sigma(T) \subseteq [c, \infty)$.

In particular, $T \geq 0$ *if and only if* $\sigma(T) \subseteq [0, \infty)$.

Proof. By Theorem 7.1.1 there exist a measurable space (Y, μ) with finite measure, a μ-measurable function $q : Y \to \mathbb{R}$ and an unitary operator $U : H \to L^2(Y, d\mu)$ such that $T = U^{-1}M_qU$. Then, we can write

$$\begin{aligned}
\langle Tx, x\rangle \geq c\|x\|^2, \quad \forall x \in D(T) &\iff \langle U^{-1}M_qUx, x\rangle \geq c\|x\|^2, \quad \forall x \in D(T)\\
&\iff \langle U^{-1}M_qf, U^{-1}f\rangle \geq c\|U^{-1}f\|^2, \quad \forall f \in D(M_q)\\
&\iff \langle M_qf, f\rangle \geq c\|f\|^2, \quad \forall f \in D(M_q)\\
&\iff \int_Y q|f|^2 d\mu \geq c\int_Y |f|^2 d\mu, \quad \forall f \in D(M_q)\\
&\iff q \geq c \ \ \mu - \text{a.e.}\\
&\iff q_{\text{ess}}(\Omega) \subseteq [c, \infty)
\end{aligned}$$

and Proposition 2.2.1(4) allows us to complete the proof. □

Theorem 7.3.4 (Rayleigh-Ritz variational formula)**.** *Let H be a complex separable Hilbert space and let $T : D(T) \subseteq H \to H$ be a positive self-adjoint linear operator with compact resolvent. Let $(\lambda_n)_{n\in\mathbb{N}}$ be the sequence of eigenvalues of T sorted in ascending order and repeated according to their multiplicity. Then, for all $n \in \mathbb{N}$,*

$$\lambda_n = \inf\{\lambda(L) \mid L \textit{ is a subspace of } D(T),\ \dim L = n\} \tag{7.1}$$

where

$$\lambda(L) := \sup\{\langle Tx, x\rangle \mid x \in L \text{ and } \|x\| = 1\}\,. \tag{7.2}$$

Proof. We first observe that, if L is a finite dimensional subspace of H such that $L \subset D(T)$, then $T_{|L}$ is clearly a bounded operator and, hence, there exists a positive constant c such that $0 \le \langle Tx, x\rangle \le c\|x\|^2$ for all $x \in L$. Thus, $0 \le \lambda(L) < \infty$.

For each $n \in \mathbb{N}$, we set $\mu_n := \inf\{\lambda(L) \mid L \text{ is a subspace of } D(T),\ \dim L = n\}$ and prove $\mu_n = \lambda_n$.

By Theorem 7.2.4, there exists a complete orthonormal system $(e_n)_{n\in\mathbb{N}} \subset D(T)$ of H such that $Te_n = \lambda_n e_n$ for all $n \in \mathbb{N}$, and $Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n$ for all $x \in D(T)$. For a fixed $n \in \mathbb{N}$, let $L := \operatorname{span}\{e_1, \dots, e_n\}$. If $x \in L$ with $\|x\| = 1$, then

$$x = \sum_{i=1}^{n} \langle x, e_i\rangle e_i, \qquad Tx = \sum_{i=1}^{n} \lambda_i \langle x, e_i\rangle e_i\,,$$

and, hence,

$$\langle Tx, x\rangle = \sum_{i=1}^{n} \lambda_i |\langle x, e_i\rangle|^2 \le \lambda_n \sum_{i=1}^{n} |\langle x, e_i\rangle|^2 = \lambda_n \|x\|^2 = \lambda_n\,.$$

It follows that $\lambda(L) \le \lambda_n$. This implies that $\mu_n \le \lambda_n$, by taking into account the definition of μ_n.

Vice versa, let us fix a subspace L of $D(T)$ with dimension equal to n and consider the orthogonal projection P from H onto $G = \operatorname{span}\{e_1, \dots, e_{n-1}\}$, defined by setting

$$Px = \sum_{i=1}^{n-1} \langle x, e_i\rangle e_i, \qquad x \in H.$$

Then, there exists $x \in L$ with $\|x\| = 1$ such that $Px = 0$ because $\dim G = n - 1 < \dim L$. Thus,

$$x = \sum_{i=n}^{\infty} \langle x, e_i\rangle e_i, \qquad Tx = \sum_{i=n}^{\infty} \lambda_i \langle x, e_i\rangle e_i\,.$$

It follows that

$$\langle Tx, x\rangle = \sum_{i=n}^{\infty} \lambda_i |\langle x, e_i\rangle|^2 \ge \lambda_n \sum_{i=n}^{\infty} |\langle x, e_i\rangle|^2 = \lambda_n \|x\|^2 = \lambda_n.$$

Therefore, $\lambda(L) \ge \lambda_n$. Since L is arbitrary we conclude that $\mu_n \ge \lambda_n$ for all $n \in \mathbb{N}$. This completes the proof. □

The following result is an immediate consequence of Theorem 7.3.4.

Corollary 7.3.5. *Let H be a separable complex Hilbert space and let $T_1\colon D(T_1) \subseteq H \to H$ and $T_2\colon D(T_2) \subseteq H \to H$ be two positive self-adjoint operators with compact resolvent such that $T_1 \le T_2$. Let $(\lambda_n^{(1)})_{n\in\mathbb{N}}$ and $(\lambda_n^{(2)})_{n\in\mathbb{N}}$ be the sequences of the eigenvalues of T_1 and T_2, respectively, sorted in ascending order and repeated according to their multiplicity. Then, for every $n \in \mathbb{N}$, it holds that*

$$\lambda_n^{(1)} \le \lambda_n^{(2)}\,. \tag{7.3}$$

Proof. By assumptions, $T_1 \le T_2$ and, hence, $D := D(T_1) = D(T_2)$ and $\langle T_1 x, x\rangle \le \langle T_2 x, x\rangle$ for all $x \in D$. Then,

$$\begin{aligned}\lambda^{(1)}(L) =& \sup\{\langle T_1 x, x\rangle \mid x \in L \text{ and } \|x\| = 1\}\\ \le& \lambda^{(2)}(L) = \sup\{\langle T_2 x, x\rangle \mid x \in L \text{ and } \|x\| = 1\}\end{aligned}$$

for every subspace $L \subset D$ with $\dim L = n$ and for every $n \in \mathbb{N}$. Passing to the lower extremes the assertion is proved, thanks to the equality (7.1). □

7.4 Exercises

Exercise 7.4.1. Let $T : D(T) \to H$ be a positive self-adjoint linear operator on a separable Hilbert space H.

(1) Prove that there exists a positive self-adjoint linear operator $S : D(S) \to H$ such that $S^2 = T$.

(2) Let A be a closed operator which commutes with T. Prove that A commutes with S.

(3) **Uniqueness:** Use 2. to show that if there exists a positive self-adjoint operator B such that $B^2 = T$, then $B = S$.

Exercise 7.4.2. Let $T : D(T) \subseteq H \to H$ be a densely defined, positive and symmetric linear operator on a Hilbert space H. Prove that T is essentially self-adjoint (i.e. $\overline{T}$ is self-adjoint) if and only if $\mathrm{rg}(1 + T)$ is dense in H.

Exercise 7.4.3. Let $T : D(T) \subseteq H \to H$ be a self-adjoint operator on a separable Hilbert space H. Prove that the measure space (Y, μ) and the μ-measurable function $q : Y \to \mathbb{R}$ in Theorem 7.1.1 can be chosen so that $q \in L^p(Y, d\mu)$ for all $p \in [1, \infty)$.

Exercise 7.4.4. Let $T : D(T) \subseteq H \to H$ be a self-adjoint linear operator on a separable Hilbert space H and let α, β be real numbers with $\alpha < \beta$. Show that the following conditions are equivalent:

(1) $(\alpha, \beta) \subseteq \rho(T)$;

(2) $\|2Tx - (\alpha + \beta)x\| \ge (\beta - \alpha)\|x\|$ for all $x \in D(T)$;

(3) $(T - \beta)(T - \alpha)$ is a positive operator.

Exercise 7.4.5. Let $T : D(T) \subseteq H \to H$ be a self-adjoint linear operator on a separable Hilbert space H and let $x \in D(T)$ be such that $\|x\| = 1$. Suppose that $\sigma(T) \cap (\alpha, \beta) = \{\lambda\}$ and $\eta := \langle x, Tx\rangle \in (\alpha, \beta)$ for some real numbers α, β, with $\alpha < \beta$. Show that

$$\eta - \frac{\varepsilon^2}{\beta - \eta} \le \lambda \le \eta + \frac{\varepsilon^2}{\eta - \alpha}, \quad \text{where } \varepsilon^2 := \|(T - \eta)x\|^2 \text{ (Temple's inequality).}$$

Exercise 7.4.6. Let $T : D(T) \subseteq H \to H$ be a self-adjoint operator on a separable Hilbert space H. Prove that each isolated point of the spectrum $\sigma(T)$ is an eigenvalue of T.

Exercise 7.4.7. Let A be the Laplace operator with Dirichlet boundary conditions considered in Example 7.2.5. Prove that

$$Af = \sum_{n=1}^{\infty} n^2\pi^2 \left(\int_0^1 f(x)e_n(x)dx \right) e_n$$

for every $f \in L^2((0,1))$, where $e_n(x) = \sqrt{2}\sin(n\pi x)$ for every $x \in (0,1)$.

Bibliography

[1] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[2] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[3] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[4] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[5] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[6] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[7] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981

[8] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[9] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[10] W. Rudin, Real and Complex Analysis, 3^{rd} ed., McGraw-Hill, 1986.

[11] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[12] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958

[13] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

December 9, 2021

Lecture 8

Strongly continuous semigroups

In this and in the next lecture, we introduce the concept of semigroup of bounded linear operators and study two important classes of semigroups: the *strongly continuous* and the *analytic semigroups*.

The concept of semigroup of bounded operators generalizes what is known since the first courses of Calculus: the solutions of the system $u' = Au$ of d ordinary differential equations with constant coefficients are given by $u(t) = e^{tA}c$ for every $t \in \mathbb{R}$, where $c \in \mathbb{R}^d$ is an arbitrary vector and

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n, \qquad t \in \mathbb{R}. \tag{8.1}$$

The previous formula can be straightforwardly extended to the case when the matrix A is replaced by a bounded operator in a Banach space X. Indeed,

$$\left\| \sum_{n=m}^{m+p} \frac{t^n}{n!} A^n \right\| \le \sum_{n=m}^{m+p} \frac{t^n}{n!} \|A\|^n$$

for every $m, p \in \mathbb{N}$ and the real valued series $\sum_{n=0}^{\infty} \frac{t^n}{n!} \|A\|^n$ converges locally uniformly in $\mathbb{R}$ (to $e^{t\|A\|}$). Hence, the series $\sum_{n=0}^{\infty} \frac{t^n}{n!} A^n$ converges in $\mathcal{L}(X)$, locally uniformly with respect to $t \in \mathbb{R}$. Set $a_k = t^k A^k / k!$ and $b_k = s^k A^k / k!$ for any $k \in \mathbb{N} \cup \{0\}$. Then,

$$\sum_{k=0}^{n} a_k b_{n-k} = A^n \sum_{k=0}^{n} \frac{t^k s^{n-k}}{k!(n-k)!} = \frac{(t+s)^n}{n!} A^n, \qquad n \in \mathbb{N} \cup \{0\}.$$

Hence, as for the Cauchy product of scalar series, one can see that

$$\begin{aligned} e^{tA} e^{sA} &= \sum_{i=0}^{\infty} \frac{t^i}{i!} A^i \cdot \sum_{j=0}^{\infty} \frac{s^j}{j!} A^j = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k} \\ &= \sum_{n=0}^{\infty} \frac{(t+s)^n}{n!} A^n = e^{(t+s)A} = e^{sA} e^{tA}. \end{aligned}$$

Based on the above remarks, we can now give the following definition.

Definition 8.0.1. A family $\{T(t) : t \geq 0\}$ (in the sequel simply denoted by $\{T(t)\}$) of bounded linear operators on a Banach space X is called a semigroup of bounded operators if it satisfies the semigroup property, i.e., $T(0) = I$ and $T(t+s) = T(t)T(s)$ for every $s, t > 0$.

It follows that, for each $A \in \mathcal{L}(X)$, $\{e^{tA} : t \geq 0\}$ is a semigroup of bounded operators on the Banach space X.

As a matter of fact, this class of semigroups, usually referred as *uniformly continuous semigroups*, is too small. For this reason, we need to go further in the study of semigroups.

Throughout this lecture, X will denote a complex Banach space and $\|\cdot\|$ its norm. We introduce the strongly continuous semigroups, their infinitesimal generators and state and prove their main properties. We collect important properties a semigroup generator must have and provide necessary and sufficient conditions on an operator to be the infinitesimal generator of a strongly continuous semigroup.

8.1 Definitions and basic properties

In this section we introduce the strongly continuous semigroups and prove some basic properties. In particular, we show that for every strongly continuous semigroup there exist $M \geq 1$ and $\omega \in \mathbb{R}$ such that $\|T(t)\| \leq Me^{\omega t}$ for every $t > 0$.

Definition 8.1.1. A family $\{T(t) : t \geq 0\}$ of bounded operators on X, which satisfies the semigroup property, is a strongly continuous semigroup (or C_0-semigroup[1]) if the function $t \mapsto T(t)x$ is continuous in $[0, \infty)$ with values in X, for each $x \in X$.

The following example is crucial and we will use it in the next lecture to show the main differences between analytic and C_0-semigroups.

Example 8.1.2. On $X = BUC(\mathbb{R})$, the set of all bounded and uniformly continuous functions $f : \mathbb{R} \to \mathbb{R}$, consider the family $\{T(t)\}$ of linear operators defined by $(T(t)f)(x) = f(x+t)$ for each $x \in \mathbb{R}$, $t \geq 0$ and $f \in BUC(\mathbb{R})$. This is a C_0-semigroup on X. Indeed, $T(t)f$ tends to f uniformly in $\mathbb{R}$ as $t \to 0^+$ if and only if $\sup_{x\in\mathbb{R}} |f(x+t) - f(x)|$ vanishes as $t \to 0^+$. But this condition is a rewriting of the definition of uniform continuity. The semigroup property is straightforward to prove.

We stress that this semigroup, called the semigroup of the left translations, cannot be written in the form (8.1) for some bounded operator A. Indeed, if this were the case, then

$$\|e^{tA} - I\| \leq \sum_{n=1}^{\infty} \frac{t^n}{n!}\|A\|^n = e^{t\|A\|} - 1, \qquad t \geq 0,$$

and, consequently, $\lim_{t\to 0^+} \|e^{tA} - I\| = 0$. The semigroup of left translations does not satisfy this property since $\|T(t) - I\| = 2$ for every $t > 0$. To check this claim, fix $t > 0$ and consider the function $f_t \in BUC(\mathbb{R})$, defined by $f_t(x) = \sin(\pi x/t)$ for each $x \in \mathbb{R}$. As it is immediately seen, $\|f_t\|_\infty = 1$ and $f_t(x+t) = -f_t(x)$ for every $x \in \mathbb{R}$. Hence,

$$\|T(t) - I\| \geq \sup_{x\in\mathbb{R}} |f_t(x+t) - f_t(x)| = 2 \sup_{x\in\mathbb{R}} |f_t(x)| = 2.$$

[1] C_0 or $(C, 0)$ abbreviates Cesàro summable of order zero, which means the continuity property $\lim_{t\to 0} T(t)x = x$ for every $x \in X$.

On the other hand, $\|T(t)f\|_\infty = \|f\|_\infty$ for every $f \in BUC(\mathbb{R})$ and, consequently, $\|T(t)-I\| \leq \|T(t)\| + 1 = 2$ for every $t \geq 0$.

Actually, $\{T(t)\}$ is a group of bounded operators, since $T(t)$ can be defined, using the same rule, also for $t < 0$.

Now, we prove that the function $t \mapsto \|T(t)\|$ grows at most exponentially at infinity.

Proposition 8.1.3. *There exist $M \geq 1$ and $\omega \in \mathbb{R}$ such that*

$$\|T(t)\| \leq Me^{\omega t}, \qquad t \geq 0. \tag{8.2}$$

Proof. The core of the proof consists in showing that there exists $\delta > 0$ such that

$$\sup_{t\in[0,\delta]} \|T(t)\| < \infty. \tag{8.3}$$

Once this property is established, the semigroup property allows us to complete the proof. Indeed, if $t > \delta$, then there exist $n \in \mathbb{N}$ and $r \in [0, \delta)$ such that $t = n\delta + r$. By applying the semigroup property, we conclude that $T(t) = T(r)(T(\delta))^n$. Hence, denoting by M the supremum in (8.3) and observing that $M \geq 1$, we get

$$\begin{aligned}\|T(t)\| \leq& \|T(r)\|\|T(\delta)\|^n\\ \leq& M^{n+1} = M\exp(n\log(M)) \leq Me^{\frac{\log(M)}{\delta}t}\end{aligned}$$

and the assertion follows with $\omega = \delta^{-1}\log(M)$.

To prove (8.3), we argue by contradiction. We suppose that (8.3) does not hold true. Then, we can determine a positive sequence $(t_n)_{n\in\mathbb{N}}$ converging to zero such that $\|T(t_n)\|$ diverges to ∞ as n tends to ∞. Since $T(t_n)x$ converges to x as n tends to ∞, for every $x \in X$, the uniform boundedness principle leads us to a contradiction. □

Remark 8.1.4. A semigroup of bounded operators $\{T(t)\}$ on a Banach space X is called *uniformly continuous* if the map $t \mapsto T(t) \in \mathcal{L}(X)$ is continuous (with respect to the operator norm). One can prove that a semigroup $\{T(t)\}$ on X is uniformly continuous if and only if there is $A \in \mathcal{L}(X)$ such that $T(t) = e^{tA}$ for all $t \geq 0$ (see Exercise 8.5.3). So such semigroups satisfy (8.2) with $M = 1$. In the case of strongly continuous semigroups it is quite possible to have $M > 1$ (see Exercise 8.5.5).

Corollary 8.1.5. *A semigroup $\{T(t)\}$ of bounded operators on X is strongly continuous if and only if the function $t \mapsto T(t)x$ is continuous at $t = 0$ for each $x \in X$.*

Proof. Clearly, if the semigroup $\{T(t)\}$ is strongly continuous, then, in particular, the mapping $t \mapsto T(t)x$ is continuous at $t = 0$ for every $x \in X$.

Vice versa, if $\{T(t)\}$ is a semigroup such that the mapping $t \mapsto T(t)x$ is continuous at $t = 0$, then from the proof of Proposition 8.1.3 we deduce that there exist $M \geq 1$ and $\omega \geq 0$ such that $\|T(t)\| \leq Me^{\omega t}$ for every $t > 0$. Fix $t_0 > 0$ and observe that, if $t \geq t_0$, then

$$\|T(t)x - T(t_0)x\| = \|T(t_0)[T(t-t_0)x - x]\| \leq Me^{\omega t_0}\|T(t-t_0)x - x\|$$

and the last side of the previous chain of inequalities tends to 0 as t tends to t_0^+. Similarly, if $t < t_0$, then

$$\|T(t)x - T(t_0)x\| = \|T(t)[x - T(t_0-t)x]\| \leq Me^{\omega t}\|T(t_0-t)x - x\|$$

and also in this case the last side vanishes as t tends to t_0^-. Hence, the function $t \mapsto T(t)x$ is continuous at t_0. The arbitrariness of $t_0 > 0$ shows that the semigroup $\{T(t)\}$ is strongly continuous. □

Definition 8.1.6. The growth bound $\omega_0(T(\cdot))$ of a C_0-semigroup $\{T(t)\}$ is defined as the infimum of the set

$$\{\omega \in \mathbb{R} \,:\, \exists\, M = M_\omega \geq 1 \text{ such that } \|T(t)\| \leq M\mathrm{e}^{\omega t} \text{ for every } t \geq 0\}.$$

Remark 8.1.7. (i) We stress that $\omega_0(T(\cdot))$ could be also equal to $-\infty$ (see Exercise 8.5.6).

(ii) In general, $\omega_0(T(\cdot))$ is just an infimum and not a minimum. Consider for instance, the semigroup $\{T(t)\}$ in $\mathbb{R}^2$, defined by

$$T(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad t \geq 0.$$

Here, $\omega_0(T(\cdot)) = 0$, but clearly the function $t \mapsto T(t)$ is not bounded in $[0, \infty)$ with values in $\mathcal{L}(\mathbb{R}^2)$.

Example 8.1.8. For $f \in L^p(\mathbb{R})$ we define

$$(S(t)f)(s) := f(t+s) \quad \text{for } s \in \mathbb{R},\, t \geq 0.$$

Then $S(t)$ is a linear isometry on $L^p(\mathbb{R})$. Moreover, $S(\cdot)$ has the semigroup property. We call $\{S(t)\}$ the *semigroup of left translations* on $L^p(\mathbb{R})$. Furthermore, for $p \in [1, \infty)$ the semigroup $\{S(t)\}$ is strongly continuous on $L^p(\mathbb{R})$ endowed with its usual norm.

In fact, recall that the semigroup of left translations is strongly continuous on the space of bounded uniformly continuous functions and that the set of continuous functions with compact support is dense in $L^p(\mathbb{R})$. Taking $f \in C_c(\mathbb{R})$ and $\alpha, \beta \in \mathbb{R}$ such that supp $f \subset [\alpha, \beta]$, we see that

$$\|f - f(\cdot + t)\|_p^p = \int_{\mathbb{R}} |f(s) - f(s+t)|^p ds \leq (\beta - \alpha) \sup_{s \in [\alpha, \beta]} |f(s) - f(s+t)|^p,$$

which tends to zero as $t \to 0$ by the uniform continuity of f. Since $\|S(t)\| \leq 1$, the statement follows by Exercise 8.5.1 and Corollary 8.1.5.

Example 8.1.9. Let (Ω, μ) be a σ-finite measure space and let $m : \Omega \to \mathbb{C}$ be a measurable function such that

$$\sup\{\mathrm{Re}\,\lambda \mid \lambda \in m_{\mathrm{ess}}(\Omega)\} < \infty.$$

Fix $1 \leq p < \infty$ and set

$$T_m(t)f = e^{tm}f, \qquad f \in L^p(\Omega, \mu), \;\; t \geq 0. \tag{8.4}$$

Then, it is immediate to prove that $\{T_m(t)\}$ is a C_0-semigroup on $L^p(\Omega, \mu)$. Moreover it is a uniformly continuous semigroup if and only if $m \in L^\infty(\Omega, \mu)$ (see Exercise 8.5.4).

8.2 The infinitesimal generator

One message what we would like to transmit is that if we have a *semigroup*, then there exists a *differential equation* governed by an operator whose solutions are explicitly written in terms of the semigroup.

In this section, we begin to make the claim more clear, showing that to any C_0-semigroup it is possible to associate a linear operator, called the *infinitesimal generator* of the semigroup, and studying the main properties of this operator.

Definition 8.2.1. The infinitesimal generator, or simply the generator, A of a C_0-semigroup $\{T(t)\}$ is the operator defined as follows:

$$\begin{cases} D(A) = \left\{ x \in X : \exists \lim\limits_{t \to 0^+} \dfrac{T(t)x - x}{t} \in X \right\}, \\ Ax = \lim_{t \to 0^+} \dfrac{T(t)x - x}{t}, \quad x \in D(A). \end{cases}$$

Proposition 8.2.2. *The generator A of a C_0-semigroup $\{T(t)\}$ satisfies the following properties.*

(i) *A is a linear operator satisfying $AT(t) = T(t)A$ on $D(A)$ for every $t \geq 0$. Moreover, for each $x \in D(A)$, the function $u = T(\cdot)x$ belongs to $C^1([0,\infty); X) \cap C([0,\infty); D(A))$ and solves the Cauchy problem*

$$\begin{cases} u'(t) = Au(t), \quad t \geq 0, \\ u(0) = x. \end{cases} \tag{8.5}$$

(ii) *For each $t > 0$ and $x \in X$, $\int_0^t T(s)x\,ds \in D(A)$, where the integral has to be understood as the Riemann integral of the continuous function $s \mapsto T(s)x$, see Appendix B, and*

$$T(t)x - x = A \int_0^t T(s)x ds. \tag{8.6}$$

In particular, if $x \in D(A)$, then

$$A \int_0^t T(s)x\,ds = \int_0^t T(s)Ax\,ds. \tag{8.7}$$

(iii) *A is closed and $D(A)$ is dense in X.*

(iv) *The operator A completely characterizes the semigroup $\{T(t)\}$ in the sense that there exists no other semigroup which admits A as generator.*

Proof. (i) It is straightforward to check that $D(A)$ is a linear subspace of X (in particular, $D(A) \neq \varnothing$ since $0 \in D(A)$) and that A is linear. To complete the proof of property (i), we fix $x \in D(A)$ and $t > 0$. Using the semigroup property and the continuity of the operator $T(t)$, we get

$$\lim_{h \to 0^+} \frac{T(h)T(t)x - T(t)x}{h} = \lim_{h \to 0^+} T(t)\frac{T(h)x - x}{h}$$

$$=T(t)\lim_{h\to 0^+}\frac{T(h)x-x}{h}=T(t)Ax.$$

Hence, $T(t)x\in D(A)$ and $AT(t)x=T(t)Ax$. Since $T(h)T(t)=T(t+h)$, the above computation shows that the function $T(\cdot)x$ is differentiable from the right in $[0,\infty)$ and the right derivative coincides with the function $AT(\cdot)x$. To prove that the previous function is also differentiable from the left in $(0,\infty)$, we observe that

$$\begin{aligned}\frac{T(t+h)x-T(t)x}{h}-T(t)Ax=&T(t+h)\frac{T(-h)x-x}{-h}-T(t)Ax\\ =&T(t+h)\left(\frac{T(-h)x-x}{-h}-Ax\right)\\ &+T(t+h)Ax-T(t)Ax.\end{aligned}$$

Taking Proposition 8.1.3 into account, we obtain

$$\begin{aligned}&\left\|\frac{T(t+h)x-T(t)x}{h}-T(t)Ax\right\|\\ \le&\|T(t+h)\|\left\|\frac{T(-h)x-x}{-h}-Ax\right\|+\|T(t+h)Ax-T(t)Ax\|\\ \le&Me^{\omega(t+h)}\left\|\frac{T(-h)x-x}{-h}-Ax\right\|+\|T(t+h)Ax-T(t)Ax\|,\end{aligned}$$

for some $M\ge 1$ and $\omega\in\mathbb{R}$. Letting h tend to 0^- we conclude that the function $T(\cdot)x$ is differentiable from the left at t. Hence, the function $T(\cdot)x$ is differentiable is $[0,\infty)$ and since its derivative is the function $T(\cdot)Ax$, which is continuous in $[0,\infty)$, it follows that $T(\cdot)x$ belongs to $C^1([0,\infty);X)\cap C([0,\infty);D(A))$ and solves problem (8.5).

(ii) Fix $x\in X$, $t>0$, set $y=\displaystyle\int_0^t T(s)x\,ds$ and observe that

$$\begin{aligned}\frac{T(h)y-y}{h}&=\frac{1}{h}\left(T(h)\int_0^t T(s)x\,ds-\int_0^t T(s)x\,ds\right)\\ &=\frac{1}{h}\left(\int_0^t T(h+s)x\,ds-\int_0^t T(s)x\,ds\right)\\ &=\frac{1}{h}\left(\int_h^{t+h} T(s)x\,ds-\int_0^t T(s)x\,ds\right)\\ &=\frac{1}{h}\left(\int_t^{t+h} T(s)x\,ds-\int_0^h T(s)x\,ds\right).\end{aligned}$$

Since $s\mapsto T(s)x$ is continuous, taking the limit as h tend to 0^+ gives

$$\lim_{h\to 0^+}\frac{T(h)y-y}{h}=T(t)x-x.$$

Hence, $y\in D(A)$ and (8.6) follows. In the particular case when $x\in D(A)$, by property (i), we know that $T(s)Ax=AT(s)x=\frac{d}{ds}T(s)x$ for every $s\ge 0$. Formula (8.7) follows.

(iii) Let $(x_n)_{n\in\mathbb{N}} \subset D(A)$ be such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} Ax_n = y$ for some x, $y \in X$. By (ii), we know that

$$\frac{T(h)x_n - x_n}{h} = \frac{1}{h}\int_0^h T(s)Ax_n\, ds, \qquad n \in \mathbb{N},\ h > 0. \tag{8.8}$$

Since $\|T(s)Ax_n - T(s)y\| \le Me^{\omega s}\|Ax_n - y\| \le Me^{\omega^+ h}\|Ax_n - y\|$ for every $s \in [0, h]$, $T(\cdot)Ax_n$ converges to $T(\cdot)y$, uniformly in $[0, h]$. Hence, letting n tend to ∞ in both the sides of (8.8), we conclude that

$$\frac{T(h)x - x}{h} = \frac{1}{h}\int_0^h T(s)y\, ds, \qquad h > 0.$$

Since the function $T(\cdot)y$ is continuous in $[0, \infty)$, letting h tend to 0^+ it follows that $x \in D(A)$ and $Ax = y$. Hence, A is a closed.

To prove that $D(A)$ is a dense subspace of X, we fix $x \in X$ and, for each $n \in \mathbb{N}$, we set $x_n = n\displaystyle\int_0^{1/n} T(s)xds$. By property (ii), $x_n \in D(A)$ for every $n \in \mathbb{N}$. Moreover, $\lim_{n\to\infty} x_n = x$.

(iv) Suppose that $\{S(t)\}$ is another C_0-semigroup having A as generator. Fix $x \in D(A)$, $t > 0$ and consider the function $u : [0, t] \to X$, defined by $u(s) := T(t - s)S(s)x$ for $s \in [0, t]$. As it is easily seen, u is differentiable in $[0, t]$ with identically vanishing derivative. This implies that $u(t) = u(0)$, i.e., $S(t)x = T(t)x$. Since $D(A)$ is dense in X and t is arbitrarily fixed in $(0, \infty)$, we conclude that $T(t) = S(t)$ for every $t > 0$. □

The following result shows that the resolvent set of the generator A is not empty and, for each $\lambda \in \rho(A)$, the operator $R(\lambda, A)$ can be written in terms of the semigroup.

Proposition 8.2.3. *Let $\{T(t)\}$ be a C_0-semigroup with generator A and let $M \ge 1$ and $\omega \in \mathbb{R}$ be such that $\|T(t)\| \le Me^{\omega t}$ for every $t \ge 0$. Then, $\rho(A) \supset \{\lambda \in \mathbb{C} : \mathrm{Re}\lambda > \omega\}$ and*

$$R(\lambda, A)x = \int_0^\infty e^{-\lambda t}T(t)x\, dt \tag{8.9}$$

for every $x \in X$ and $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > \omega$. Moreover,

$$\|(R(\lambda, A))^n\| \le \frac{M}{(\mathrm{Re}\lambda - \omega)^n}. \tag{8.10}$$

Proof. Fix $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > \omega$. Then, the operator defined by the right-hand side of (8.9) is well defined, linear and continuous in X, since

$$\|e^{-\lambda t}T(t)x\| \le Me^{-(\mathrm{Re}\lambda - \omega)t}\|x\|$$

for $t \ge 0$. In particular,

$$\left\|\int_0^\infty e^{-\lambda t}T(t)x\, dt\right\| \le M\|x\|\int_0^\infty e^{-(\mathrm{Re}\lambda - \omega)t}dt = \frac{M}{\mathrm{Re}\lambda - \omega}\|x\|, \qquad x \in X. \tag{8.11}$$

To prove that this operator, which we denote by R_λ, is the inverse of $\lambda I - A$, we observe that, since A is a closed operator and the function $t \mapsto e^{-\mathrm{Re}\lambda t}\|T(t)Ax\|$ is integrable in $[0, \infty)$

for every $x \in D(A)$, $R_\lambda x$ belongs to $D(A)$ for every $x \in D(A)$ and, integrating by parts, we get

$$\begin{aligned}R_\lambda Ax = &\int_0^\infty e^{-\lambda t}T(t)Ax\,dt = \lim_{n\to\infty}\int_0^n e^{-\lambda t}\frac{d}{dt}T(t)x\,dt\\ =&\lim_{n\to\infty}\left(e^{-\lambda n}T(n)x - x + \lambda\int_0^n e^{-\lambda t}T(t)x\,dt\right) = -x + \lambda R_\lambda x.\end{aligned}$$

Hence, $R_\lambda(\lambda I - A)x = x$ for every $x \in D(A)$. This shows that the operator $\lambda I - A$ is injective. To prove that it is also surjective, we fix $x \in X$ and prove that $R_\lambda x \in D(A)$ and $AR_\lambda x = \lambda R_\lambda - x$. For this purpose, we fix $h > 0$ and observe that

$$\begin{aligned}\frac{T(h)R_\lambda x - R_\lambda x}{h} =&\frac{1}{h}\bigg(\int_0^\infty e^{-\lambda t}T(t+h)x\,dt - \int_0^\infty e^{-\lambda t}T(t)x\,dt\bigg)\\ =&\frac{e^{\lambda h}}{h}\int_h^\infty e^{-\lambda s}T(s)x\,ds - \frac{1}{h}\int_0^\infty e^{-\lambda t}T(t)x\,dt\\ =&\frac{e^{\lambda h}-1}{h}\int_0^\infty e^{-\lambda s}T(s)x\,ds - \frac{e^{\lambda h}}{h}\int_0^h e^{-\lambda s}T(s)x\,ds.\end{aligned}$$

Letting h tend to 0^+, we conclude that

$$\lim_{h\to 0^+}\frac{T(h)R_\lambda x - R_\lambda x}{h} = \lambda R_\lambda x - x,$$

so that $R_\lambda x \in D(A)$ and the claim follows.

To complete the proof, let us prove (8.10) with $n > 1$, since the case $n = 1$ follows from (8.11). Fix $x \in X$. By Proposition 1.3.5, we know that the function $\lambda \mapsto R(\lambda, A)x$ is holomorphic in $\rho(A)$ and, applying the dominated convergence theorem to the formula (8.9), it can be easily checked that

$$\frac{d^n}{d\lambda^n}R(\lambda, A)x = \int_0^\infty (-1)^n t^n e^{-\lambda t}T(t)x\,dt$$

for $n \in \mathbb{N}$ and $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > \omega$. Let us compute the derivatives of the function $R(\cdot, A)x$ in a different way, using the resolvent identity. This identity, see Proposition 1.3.5.(4), shows that

$$\frac{d^n}{d\lambda^n}R(\lambda, A) = (-1)^n n!(R(\lambda, A))^{n+1}, \qquad \lambda \in \mathbb{C},\ \mathrm{Re}\lambda > \omega.$$

From these last two formulas it follows that

$$(R(\lambda, A))^n x = \frac{1}{(n-1)!}\int_0^\infty t^{n-1}e^{-\lambda t}T(t)x\,dt.$$

Hence,

$$\|(R(\lambda, A))^n x\| \le \frac{M}{(n-1)!}\|x\|\int_0^\infty t^{n-1}e^{-\mathrm{Re}\lambda t}e^{\omega t}dt = \frac{M}{(\mathrm{Re}\lambda - \omega)^n}\|x\|,$$

which completes the proof. □

Let us go back to the semigroup of left translations in Example 8.1.2 and characterize its generator.

Example 8.2.4. Let us show that the generator A of the left translations semigroup in Example 8.1.2 is the first order derivative with $BUC^1(\mathbb{R})$ as domain, where $BUC^1(\mathbb{R})$ denotes the set of all continuously differentiable functions $f:\mathbb{R}\to\mathbb{R}$ with bounded uniformly continuous first-order derivative. Fix $f\in D(A)$. Then, the function $t^{-1}(f(\cdot+t)-f)$ converges to Af in $BUC(\mathbb{R})$ as t tends to 0^+. In particular, for each $x\in\mathbb{R}$, the ratio $t^{-1}(f(x+t)-f(x))$ converges to $(Af)(x)$. It thus follows that f is differentiable from the right in $\mathbb{R}$ and $f'=Af\in BUC(\mathbb{R})$. Thanks to Exercise 8.5.2, we conclude that $f\in BUC^1(\mathbb{R})$.

Conversely, suppose that $f\in BUC^1(\mathbb{R})$. Then, by the fundamental theorem of calculus, we can write

$$\frac{f(x+t)-f(x)}{t}=\frac{1}{t}\int_0^t f'(x+s)\,ds,\qquad x\in\mathbb{R},\ \ t>0.$$

Fix $\varepsilon>0$ and let $\delta>0$ be such that $|f'(x_2)-f'(x_1)|\le\varepsilon$ for every $x_1,x_2\in\mathbb{R}$ with $|x_2-x_1|\le\delta$. Then,

$$\left|\frac{f(x+t)-f(x)}{t}-f'(x)\right|\le\frac{1}{t}\int_0^t|f'(x+s)-f'(x)|\,ds\le\varepsilon$$

for every $x\in\mathbb{R}$ if $t\in(0,\delta]$. Hence, $t^{-1}(T(t)f-f)$ converges to f' uniformly in $\mathbb{R}$ as t tends to 0^+. This shows that $BUC^1(\mathbb{R})\subset D(A)$.

Example 8.2.5. The generator of the multiplication semigroup $\{\mathrm{T}_m(t)\}$ on $L^p(\Omega)$ considered in Example 8.1.9 is the multiplication operator M_m introduced in Section 2.2 (see Exercise 8.5.4).

8.3 Hille-Yosida theorem

By the results of the previous section we know that to any C_0-semigroup $\{T(t)\}$ on X, we can associate an operator $A:D(A)\subset X\to X$, the infinitesimal generator, which satisfies the following properties:

(i) A is closed and densely defined;

(ii) $\rho(A)$ contains the right-halfplane $\{\lambda\in\mathbb{C}:\mathrm{Re}\lambda>\omega_0\}$ for some $\omega_0\in\mathbb{R}$;

(iii) for each $\omega>\omega_0$ there exists a positive constant M such that

$$\|(R(\lambda,A))^n\|\le M(\mathrm{Re}\lambda-\omega)^{-n}$$

for every $n\in\mathbb{N}$ and $\lambda\in\mathbb{C}$ with $\mathrm{Re}\lambda>\omega$.

A natural question arises: is any linear operator $A:D(A)\subset X\to X$, which satisfies the above three properties, the infinitesimal generator of a C_0-semigroup on X? The answer is positive and given by the famous Hille-Yosida theorem

Theorem 8.3.1 (Hille-Yosida)**.** *Let $A:D(A)\subset X\to X$ be a linear operator on a Banach space X. Then the following properties are equivalent.*

(a) *A generates a C_0-semigroup $\{T(t)\}$ on X.*

(b) *A satisfies the above properties (i)-(iii).*

Proof. In view of the remarks at the beginning of this section, we just need to show that (b)⇒(a). Without loss of generality we can assume that $\omega = 0$. Indeed, the operator $\widetilde{A} = A - \omega I$, satisfies properties (i)-(iii), with $\omega = 0$. If we prove that $\widetilde{A}$ generates a C_0-semigroup $\{S(t)\}$, then the operator A generates the C_0-semigroup $\{e^{\omega t}S(t)\}$.

Since it is rather long, we split the proof into steps.

Step 1. Here, we define the Yosida approximation of the operator A, i.e., the operators $A_n : X \to X$ defined by $A_n = nAR(n, A)$ for $n \in \mathbb{N}$. Since

$$AR(n, A) = (A - nI)R(n, A) + nR(n, A) = nR(n, A) - I,$$

each operator A_n is bounded in X. We claim that $\lim_{n\to\infty} A_n x = Ax$ for each $x \in D(A)$. To prove the claim, it suffices to show that $nR(n, A)y$ tends to y, as n tends to ∞, for each $y \in X$. Indeed, $A_n x = nR(n, A)Ax$ for $x \in D(A)$. First, we suppose that $y \in D(A)$. Then,

$$nR(n, A)y = R(n, A)(ny - Ay + Ay) = y + R(n, A)Ay, \qquad n \in \mathbb{N},$$

and $\|R(n, A)Ay\| \le Mn^{-1}\|Ay\|$ vanishes as $n \to \infty$. Hence, $nR(n, A)y$ tends to y as $n \to \infty$. If $y \in X$, then we consider a sequence $(y_m)_{m\in\mathbb{N}} \subset D(A)$, converging to y as $m \to \infty$. Since $\|nR(n, A)\|_{L(X)} \le M$ for every $n \in \mathbb{N}$, we can estimate

$$\begin{aligned}\|nR(n, A)y - y\| \le& \|nR(n, A)(y - y_m)\| + \|nR(n, A)y_m - y_m\| + \|y_m - y\|\\ \le&(M + 1)\|y_m - y\| + \|nR(n, A)y_m - y_m\|\end{aligned}$$

for $m, n \in \mathbb{N}$. Hence,

$$\begin{aligned}\limsup_{n\to\infty} \|nR(n, A)y - y\| \le&(M + 1)\|y - y_m\| + \lim_{n\to\infty} \|nR(n, A)y_m - y_m\|\\ =&(M + 1)\|y_m - y\|\end{aligned}$$

for every $m \in \mathbb{N}$. Letting m tend to ∞, we conclude that

$$\limsup_{n\to\infty} \|nR(n, A)y - y\| = 0.$$

Hence, $nR(n, A)y$ converges to y as $n \to \infty$.

Step 2. For any $n \in \mathbb{N}$, we introduce the uniform continuous semigroup $\{T_n(t)\}$ on X, defined by

$$T_n(t) = e^{tA_n} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A_n^k, \qquad t \ge 0,$$

and prove that, for each $x \in X$, $(T_n(\cdot)x)_{n\in\mathbb{N}}$ is a Cauchy sequence in $C([0, T]; X)$ for every $T > 0$.

To begin with, we observe that, since $A_n = n^2R(n, A) - nI$, $T_n(t) = e^{-nt}e^{tn^2R(n,A)}$ for each $t \ge 0$. Hence, recalling that $\|(nR(n, A))^k\|_{L(X)} \le M$ for $k \in \mathbb{N}$, we deduce that

$$\|T_n(t)\| \le e^{-nt}\left\| \sum_{k=0}^{\infty} \frac{t^k}{k!} (n^2R(n, A))^k \right\|_{\mathcal{L}(X)}$$

$$\begin{aligned} &\le e^{-nt}\sum_{k=0}^{\infty}\frac{(tn)^k}{k!}\|(nR(n,A))^k\| \\ &\le Me^{-nt}e^{nt} = M \end{aligned} \tag{8.12}$$

for every $t \ge 0$ and $n \in \mathbb{N}$.

Next, we fix $T > 0$, $t \in (0,T]$, $x \in D(A)$ and introduce the function $u_{m,n} : [0,t] \to X$, defined by

$$u_{m,n}(s) = T_m(t-s)T_n(s)x, \qquad s \in [0,t].$$

As it is easily checked $u_{m,n}(t) = T_n(t)x$ and $u_{m,n}(0) = T_m(t)x$. Moreover, $u_{m,n}$ is differentiable in $[0,t]$ and

$$u'_{m,n}(s) = -T_m(t-s)A_mT_n(s)x + T_m(t-s)T_n(s)A_nx$$

for every $s \in [0,t]$. Note that A_m commutes with A_n since $R(n,A)$ commutes with $R(m,A)$ (see Proposition 1.3.5). As a byproduct, A_m commutes with the semigroup $\{T_n(t)\}$ and, consequently,

$$u'_{m,n}(s) = T_m(t-s)T_n(s)(A_mx - A_nx), \qquad s \in [0,t].$$

Thus,

$$\begin{aligned} \|T_n(t)x - T_m(t)x\| \le &\int_0^t \|T_m(t-s)T_n(s)(A_mx - A_nx)\|\,ds \\ \le &M^2t\|A_mx - A_nx\| \le M^2T\|A_mx - A_nx\|. \end{aligned}$$

Since T has been arbitrarily fixed and $x \in D(A)$, by Step 1 $(T_n(\cdot)x)_{n\in\mathbb{N}}$ is a Cauchy sequence in $C([0,T];X)$ for every $T > 0$.

If $x \in X$, then there exists a sequence $(x_k)_{k\in\mathbb{N}} \subset D(A)$ such that $\lim_{k\to\infty}\|x_k - x\| = 0$. Then, taking (8.12) into account, for each $k,m,n \in \mathbb{N}$ and $T > 0$, we can estimate

$$\begin{aligned} \|T_n(\cdot)x - T_m(\cdot)x\|_{C([0,T];X)} \le &\|T_n(\cdot)(x-x_k)\|_{C([0,T];X)} \\ &+ \|T_n(\cdot)x_k - T_m(\cdot)x_k\|_{C([0,T];X)} \\ &+ \|T_m(\cdot)(x-x_k)\|_{C([0,T];X)} \\ \le &2M\|x - x_k\| + M^2T\|A_mx_k - A_nx_k\|. \end{aligned}$$

Hence, for every fixed $\varepsilon > 0$ we can fix $k_0, n_0 \in \mathbb{N}$ such that $2M\|x - x_{k_0}\| \le \varepsilon/2$ and $M^2T\|A_mx_{k_0} - A_nx_{k_0}\| \le \varepsilon/2$ for each $m,n \ge n_0$. It follows that $\|T_n(\cdot)x - T_m(\cdot)x\|_{C([0,T];X)} \le \varepsilon$ for each $m,n \ge n_0$. Hence, $(T_n(\cdot)x)_{n\in\mathbb{N}}$ is a Cauchy sequence in $C([0,T];X)$ for every $x \in X$.

Step 3. Here, we define a C_0-semigroup $\{T(t)\}$. Since each operator $T_n(t)$ is linear and $(T_n(\cdot)x)_{n\in\mathbb{N}}$ is a Cauchy sequence in $[0,T]$ for each $T > 0$ and $x \in X$, there exists a family of linear operators $\{T(t)\}$ such that $\lim_{n\to\infty} T_n(t)x = T(t)x$ for every $x \in X$ and the convergence is uniform on each interval $[0,T]$. So, the function $t \mapsto T(t)x$ is continuous on $[0,\infty)$ for each $x \in X$. As is immediately seen, $T(0) = I$, since $T_n(0) = I$ for every $n \in \mathbb{N}$. Moreover, from the estimate $\|T_n(t)x\| \le M\|x\|$, which holds true for every $t \ge 0$, $x \in X$ and $n \in \mathbb{N}$, it follows that $T(t)$ is a bounded linear operator and $\|T(t)\| \le M$ for any $t \ge 0$. To conclude that $\{T(t)\}$ is a C_0-semigroup we need to check the semigroup property. This follows from letting

n tend to ∞ in the formula $T_n(t+s)x = T_n(t)T_n(s)x$ which holds true for each $t, s > 0$, $x \in X$ and $n \in \mathbb{N}$.

Step 4. Here, we complete the proof. Since $\{T(t)\}$ is a C_0-semigroup on X which satisfies the estimate $\|T(t)\| \leq M$ for every $t \geq 0$, by Propositions 8.2.2 and 8.2.3 it admits an infinitesimal generator B, whose resolvent set contains the line $(0, \infty)$.

To show that $B = A$, we begin by showing that $D(A) \subset D(B)$ and $Bx = Ax$ for every $x \in D(A)$. For this purpose, we fix $x \in D(A)$, $n \in \mathbb{N}$ and observe that the function $T_n(\cdot)x$ is differentiable in $[0, \infty)$ with $\frac{d}{dt}T_n(t)x = T_n(t)A_nx$ for each $t \geq 0$. We claim that $D_tT_n(\cdot)x$ converges to the function $T(\cdot)Ax$ locally uniformly in $[0, \infty)$. To check the claim, we observe that

$$\begin{aligned}&\|D_tT_n(\cdot)x - T(\cdot)Ax\|_{C([0,T];X)}\\ \leq &M\|A_nx - Ax\| + \|T_n(\cdot)Ax - T(\cdot)Ax\|_{C([0,T];X)}\end{aligned}$$

for every $T > 0$ and the right-hand side of the previous inequality converges to zero as n tends to ∞ by Steps 1 and 3. The claim is proved.

Since $T_n(\cdot)x$ converges to $T(\cdot)x$, locally uniformly in $[0, \infty)$, the function $T(\cdot)x$ is differentiable in $[0, \infty)$ and $\frac{d}{dt}T(t)x = T(t)Ax$ for every $t \geq 0$. This, in particular, shows that $x \in D(B)$ and $Bx = Ax$.

To show that $D(B) = D(A)$, we observe that $1 \in \rho(A) \cap \rho(B)$ and

$$X = (I - A)(D(A)) = (I - B)(D(A)) \subset (I - B)(D(B)) = X.$$

Hence, $(I - B)(D(A)) = (I - B)(D(B))$. Since the operator $I - B$ is injective, we obtain that $D(A) = D(B)$ and we are done. □

Remark 8.3.2. Given a closed operator A, such that $\Pi = \{\lambda \in \mathbb{C} : \text{Re}\lambda > \omega\} \subseteq \rho(A)$ for some $\omega \in \mathbb{R}$, we need to check the infinitely many conditions $\|(R(\lambda, A))^n\| \leq M(\text{Re}\lambda - \omega)^{-n}$ for every $\lambda \in \Pi$ to establish whether A generates a C_0-semigroup or not. In the particular case when $M = 1$, things are easier since once the previous condition is proved with $n = 1$, it can be easily extended to every $n > 1$ just observing that $\|(R(\lambda, A))^n\| \leq \|R(\lambda, A)\|^n$ for each $n \in \mathbb{N}$.

Remark 8.3.3. At the very beginning of this chapter, we have introduced uniform continuous semigroups, which are actually groups since the operator e^{tA} is defined for every real value of t and the semigroup property is satisfied for every $s, t \in \mathbb{R}$. On the other hand, C_0-semigroups are in general defined only on the line $[0, \infty)$. Indeed, suppose that the C_0-semigroup is actually a group. Then, $\{T(t)\}$ and $\{T(-t)\}$ are C_0-semigroups. Clearly, if A is the infinitesimal generator of $\{T(t)\}$, then, $-A$ generates the semigroup $\{T(-t)\}$. Hence, the Hille-Yosida theorem implies that the resolvent set of A should contain the halfplanes $\{\lambda \in \mathbb{C} : \text{Re}\lambda > \omega\}$ and $\{\lambda \in \mathbb{C} : \text{Re}\lambda < -\omega\}$ for some $\omega \geq 0$. Moreover, $\|(R(\lambda, A))^n\|_{L(X)} \leq M(|\text{Re}\lambda| - \omega)^{-n}$ for every $\lambda \in \mathbb{C}$, such that $|\text{Re}\lambda| > \omega$, every $n \in \mathbb{N}$ and some positive constant M. These conditions are also sufficient for an operator A to be the generator of a C_0-group.

An useful criterium to guarantee that a closed operator generates a C_0-semigroup is the well celebrated Lumer-Phillips theorem. To state it, we introduce the concept of dissipative operator in the general settting of operators on Banach spaces.

Definition 8.3.4. An operator $A : D(A) \subset X \to X$ is called dissipative if $\|\lambda x - Ax\| \geq \lambda\|x\|$ for each $\lambda > 0$ and $x \in X$.

Remark 8.3.5. The property of an operator to be dissipative can be stated also in a different way. In fact, a linear operator $A : D(A) \subset X \to X$ is dissipative if and only if for every $x \in D(A)$ there exists $x' \in F(x)$, where $F(x) := \{x' \in X' : x'(x) = \|x\|^2 = \|x'\|_{X'}^2\}$, such that $\mathrm{Re}(x'(Ax)) \leq 0$. In particular, if X is a Hilbert space, then A is dissipative if and only if $\mathrm{Re}\langle Ax, x\rangle \leq 0$ for every $x \in D(A)$ (see Proposition 6.2.5).

Theorem 8.3.6 (Lumer-Phillips theorem)**.** *Let $A : D(A) \subset X \to X$ be a dissipative operator with dense domain and such that $\rho(A) \cap (0, +\infty) \neq \varnothing$. Then, A generates a C_0-semigroup of contractions on X (i.e., $\|T(t)\| \leq 1$ for every $t \geq 0$).*

Proof. Fix $\lambda_0 \in \rho(A) \cap (0, \infty)$. Then, the dissipativity of A implies that $\|R(\lambda_0, A)\| \leq \lambda_0^{-1}$. The first part of the proof of Proposition 1.3.5, shows that the open ball $B(\lambda_0, r)$ is contained in $\rho(A)$ when $r = \|R(\lambda_0, A)\|^{-1}$. Since $r \geq \lambda_0$, the interval $(0, 2\lambda_0)$ is contained in $\rho(A)$. Now, replacing λ_0 with $3\lambda_0/2$ we conclude that the interval $(0, 5\lambda_0/2)$ is contained in $\rho(A)$. Starting from $2\lambda_0$ instead of $3\lambda_0/2$, we obtain that $(0, 3\lambda_0) \subset \rho(A)$. Iterating this procedure, we can show that $(0, \infty) \subset \rho(A)$ as claimed. □

As in the case of densely defined operators in Hilbert spaces, we can introduce the notion of adjoint of densely defined operators in Banach spaces.

Definition 8.3.7. Let $A \colon D(A) \subseteq X \to X$ be a densely defined linear operator. We set

$$D(A') := \{x' \in X' \colon \exists f' \in X' \text{ such that } \langle Ax, x'\rangle = \langle x, f'\rangle \ \forall x \in D(A)\},$$

where $\langle x, g'\rangle = g'(x)$ for every $x \in X$ and $g' \in X'$.

Since $D(A)$ is dense in X, such an element f' can be proved to be unique and therefore we can define $A'y' = f'$. Moreover, it is easy to deduce from the above definition that A' is closed.

Corollary 8.3.8. *If $A : D(A) \subset X \to X$ is densely defined and dissipative and its adjoint operator A' is dissipative as well, then the closure $\overline{A}$ of A generates a C_0-semigroup of contractions.*

Proof. First of all, we prove that A is closable. This is clear if H is a Hilbert space (see Exercise 6.3.5). The same proof works also in the case of Banach spaces as we show here below. For this purpose, we fix a sequence $(x_n)_{n\in\mathbb{N}} \subset D(A)$ converging to zero and such that Ax_n converges to some $y \in X$ as n tends to ∞. Since $D(A)$ is dense in X, we can determine a sequence $(y_n)_{n\in\mathbb{N}} \subset D(A)$ converging to y as n tends to ∞. From the dissipativity of A we can infer that

$$\begin{aligned}\|k^2x_n - kAx_n + ky_m - Ay_m\| =&\|k(kx_n + y_m) - A(kx_n + y_m)\|\\ \geq& k\|kx_n + y_m\|\end{aligned}$$

for every $k, m, n \in \mathbb{N}$. Letting n tend to ∞ in the first and last side of the previous chain of inequalities gives

$$\| - ky + ky_m - Ay_m\| \geq k\|y_m\|, \qquad k, m \in \mathbb{N},$$

or, equivalently, dividing by k,

$$\| - y + y_m - k^{-1}Ay_m\| \geq \|y_m\|, \qquad k, m \in \mathbb{N},$$

so that, letting k tend to ∞, we obtain that $\|y - y_m\| \geq \|y_m\|$. Finally, letting $m \to \infty$, we can easily infer that $y = 0$. This shows that A is closable.

Next, we show that the range of the operator $I - A$ is dense in X. For this purpose, we fix a functional $x' \in X'$ such that

$$x'(x - Ax) = 0, \qquad x \in D(A). \tag{8.13}$$

From (8.13) we deduce that $x' \in D(A')$ and $x' - A'x' = 0$. Since A' is dissipative, in particular it is a one-to-one operator. Hence, $x' = 0$ and we conclude that $(I - A)(D(A))$ is dense in X.

Since $\overline{A}$ extends the operator A, also the range of the operator $I - \overline{A}$ is dense in X for every $x \in D(A)$. Actually, $(I - \overline{A})(D(\overline{A})) = X$. Indeed, fix $y \in Y$ and consider a sequence $(x_n)_{n\in\mathbb{N}} \subset D(A)$ such that $y_n = x_n - Ax_n$ converges to y as n tends to ∞. From the dissipativity of A we can infer that $\|x_n - x_m\| \leq \|y_n - y_m\|$ for every $m, n \in \mathbb{N}$ so that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X. Therefore it converges to some $x \in X$. As a byproduct also $\overline{A}x_n = Ax_n$ converges in X to $x - y$. Since $\overline{A}$ is a closed operator, x belongs to $D(A)$ and $\overline{A}x = x - y$ or, equivalently, $x - \overline{A}x = y$. Thus, y belongs to the range of the operator $I - \overline{A}$, which coincides with X. Applying Lumer-Phillips theorem, we conclude that $\overline{A}$ is the generator of a strongly continuous semigroup of contractions. □

8.4 Notes and Remarks

Operator semigroups has been widely studied during the last decades and there are many monographs dealing with them. We mention here the excellent graduate texts by Engel and Nagel [5,6]. The first milestone in the theory was the opus of Hille and Phillips [10]. Important later references are the books by Belleni-Morante [1], Goldstein [8] and Pazy [12].

8.5 Exercises

Exercise 8.5.1. Prove that a mapping $T : [0, \infty) \to \mathcal{L}(X)$ with the semigroup property is strongly continuous on X if and only if it is locally bounded and there is a dense subset $D \subseteq X$ on which T is strongly continuous.

Exercise 8.5.2. Prove that if $f : [a, b) \to X\,(a, b \in \mathbb{R}, a < b)$, where X is a Banach space, is continuous and admits right-derivative at each point $[a, b)$ and the right-derivative is continuous in $[a, b)$, then f is Fréchet differentiable in $[a, b)$.

Exercise 8.5.3. Prove that a semigroup $\{T(t)\}$ on a Banach space X is uniformly continuous if and only if there exists an operator $A \in \mathcal{L}(X)$ such that $T(t) = e^{tA}$ for all $t \geq 0$.

Exercise 8.5.4. Prove that the generator of the multiplication semigroup $\{T_m(t)\}$, introduced in Example 8.1.9, is the multiplication operator M_m. Moreover, prove that $\{T_m(t)\}$ is uniformly continuous if and only if $m \in L^\infty(\Omega, \mu)$.

Exercise 8.5.5. Consider the Hilbert space $L^2((0, 1), \mu)$, where μ is the measure defined by

$$\mu(\Omega) := 2\lambda(\Omega \cap (0, 1/2)) + \lambda(\Omega \cap (1/2, 1))$$

for each Lebesgue measurable set $\Omega \subseteq (0, 1)$, and λ is the Lebesgue measure. Define the family of operators on $L^2((0, 1), \mu)$

$$T(t)f(s) := \begin{cases} f(s + t) & \text{for } s + t \leq 1, \\ 0 & \text{for } s + t > 1. \end{cases}$$

(i) Prove that $\{T(t)\}$ defines a C_0-semigroup on $L^2((0,1),\mu)$.

(ii) Prove that the estimate $\|T(t)\| \leq Me^{t\omega}$ cannot hold for every $t > 0$, with $M < 2$, whichever $\omega \in \mathbb{R}$ you choose.

Exercise 8.5.6. On the Banach space $X = \{f \in C([0,1]) : f(1) = 0\}$, endowed with the sup-norm, consider the family $\{T(t)\}$ of operators, defined by

$$(T(t)f)(x) = \begin{cases} f(x+t), & \text{if } x+t \leq 1, \\ 0, & \text{otherwise,} \end{cases}$$

for every $t \geq 0$. Prove that $\{T(t)\}$ is a C_0-semigroup on $C([0,1])$, and show that $\omega_0(T(\cdot)) = -\infty$.

Exercise 8.5.7. Let $X = C([0,1])$ and consider the operator $Af = f$" with domain

$$D(A) = \{f \in C^2([0,1]) : f'(0) + \alpha f(0) = f'(1) + \beta f(1) = 0\}$$

for some α, $\beta \in \mathbb{R}$. Show that A generates C_0-semigroup on X.

Exercise 8.5.8. Let $X := L^p((1,\infty))$, $1 \leq p < \infty$ and $(T(t)f)(s) := f(se^t)$. Show that $\{T(t)\}$ is a C_0-semigroup and that $\omega_0(T(\cdot)) = -\frac{1}{p}$. Can you identify its generator?

Exercise 8.5.9.

1. Let $\{T(t)\}$ be a C_0-semigroup with generator $A : D(A) \to H$ on a Hilbert space H. Prove that the adjoint $\{(T(t))^*\}$ is the C_0-semigroup generated by A'.

2. **Stone's Theorem:** Prove that a densely defined linear operator $A : D(A) \to H$ on a Hilbert space H generates a unitary group $\{T(t)\}_{t\in\mathbb{R}}$ if and only if A is skew-adjoint, i.e. $A^* = -A$.

Bibliography

[1] A. Belleni-Morante, Applied Semigroups and Evolution Equations, Oxford University Press, 1979.

[2] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[3] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[4] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[5] K.-J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Graduate Text in Mathematics **194**, Springer-Verlag, 2000.

[6] K.-J. Engel and R. Nagel, Short Course on Operator Semigroups, Universitext, Springer-Verlag, 2006.

[7] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[8] J.A. Goldstein, Semigroups of Operators and Applications, Oxford University Press, 1985.

[9] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[10] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[11] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981.

[12] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Appl. Math.Sci. **44**, Springer-Verlag,1983.

[13] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[14] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[15] W. Rudin, Real and Complex Analysis, 3^{rd} ed., McGraw-Hill, 1986.

[16] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[17] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958.

[18] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861.

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

December 15, 2021

Lecture 9

Analytic semigroups

In this lecture, we keep on the study of semigroups of bounded operators introducing analytic semigroups and studying their main properties. By X we still denote a complex Banach space with norm $\|\cdot\|$. Moreover, given a curve $\gamma : I \to \mathbb{C}$ ($I \subset \mathbb{R}$ being an interval) and a smooth enough function f, defined at least on the support of γ, we set $\int_{-\gamma} f(\lambda)d\lambda := -\int_\gamma f(\lambda)d\lambda$.

9.1 Prelude

In the previous lecture, we have seen that, to every bounded operator on X, we can associate a uniformly continuous semigroup by setting

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n, \qquad t \in \mathbb{R}. \tag{9.1}$$

This formula cannot be extended to the case when A is not defined in the whole X since the domain of the powers A^n becomes smaller and smaller. On the other hand, when $\rho(A)$ is not empty, the resolvent operator is defined and bounded in X. Hence, the idea is to look for a formula for e^{tA} which involves the operators $R(\lambda, A)$.

We start by proving an integral representation formula for uniformly continuous semigroups.

Proposition 9.1.1. *Let $A \in \mathcal{L}(X)$ and let γ_r ($r > \|A\|$) be the curve, defined by $\gamma_r(t) = re^{it}$ for $t \in [0, 2\pi]$. Then,*

$$e^{tA} = \frac{1}{2\pi i} \int_{\gamma_r} e^{t\lambda} R(\lambda, A) d\lambda, \qquad t \in \mathbb{R}. \tag{9.2}$$

Proof. By Proposition 2.1.1, we can write

$$R(\lambda, A) = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}}$$

for every $|\lambda| > \|A\|$. Hence,

$$e^{\lambda t} R(\lambda, A) = \sum_{n=0}^{\infty} \frac{t^n \lambda^n}{n!} \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}}.$$

Integrating both sides of the previous formula along the curve γ_r and observing that the series and the integral commute, we get

$$\frac{1}{2\pi i}\int_{\gamma_r} e^{t\lambda}R(\lambda,A)d\lambda = \frac{1}{2\pi i}\sum_{n=0}^{\infty}\frac{t^n}{n!}\sum_{k=0}^{\infty}A^k\int_{\gamma_r}\lambda^{n-k-1}d\lambda = e^{tA}, \tag{9.3}$$

since

$$\int_{\gamma_r}\lambda^{n-k-1}d\lambda = \begin{cases} 2\pi i, & n=k, \\ 0, & \text{otherwise.}\end{cases}$$

□

Note that the right-hand side of formula (9.2) makes sense for each closed operator $A : D(A) \subset X \to X$ whose spectrum is bounded, and is independent of $r > 0$. Moreover, if we set $T(0) = I$, then the family $\{T(t)\}$ defines a semigroup on X. Indeed, take $s, t > 0$ and let $r > 0$ be such that $\sigma(A) \subset B(0,r)$. Then,

$$\begin{aligned}
T(t)T(s) &= -\frac{1}{4\pi^2}\int_{\gamma_{2r}} e^{t\lambda}R(\lambda,A)d\lambda\int_{\gamma_r} e^{s\mu}R(\mu,A)d\mu \\
&= -\frac{1}{4\pi^2}\int_{\gamma_{2r}} e^{t\lambda}d\lambda\int_{\gamma_r} e^{s\mu}R(\lambda,A)R(\mu,A)d\mu \\
&= -\frac{1}{4\pi^2}\int_{\gamma_{2r}} e^{t\lambda}d\lambda\int_{\gamma_r} \frac{e^{s\mu}}{\mu-\lambda}(R(\lambda,A)-R(\mu,A))d\mu \\
&= -\frac{1}{4\pi^2}\int_{\gamma_{2r}} e^{t\lambda}R(\lambda,A)d\lambda\int_{\gamma_r}\frac{e^{s\mu}}{\mu-\lambda}d\mu - \frac{1}{4\pi^2}\int_{\gamma_r} e^{s\mu}R(\mu,A)d\mu\int_{\gamma_{2r}}\frac{e^{t\lambda}}{\lambda-\mu}d\lambda,
\end{aligned}$$

where in the last integral term we have changed the order of integration. Since $\lambda \notin \overline{B(0,r)}$, the function $\mu \mapsto (\mu-\lambda)^{-1}e^{s\mu}$ is holomorphic in $\overline{B(0,r)}$ and, consequently, by the Cauchy integral theorem, $\int_{\gamma_r}(\mu-\lambda)^{-1}e^{s\mu}d\mu = 0$. On the contrary, the function $\lambda \mapsto (\lambda-\mu)^{-1}e^{t\lambda}$ has a simple pole at $\lambda = \mu \in B(0,2r)$. Hence, by the residue theorem, $\int_{\gamma_{2r}}(\lambda-\mu)^{-1}e^{t\lambda}d\lambda = (2\pi i)e^{t\mu}$. It thus follows that

$$T(t)T(s) = \frac{1}{2\pi i}\int_{\gamma_{2r}} e^{(s+t)\mu}R(\mu,A)d\mu = T(t+s).$$

As we know, in general, the spectrum of a closed operator is not bounded (see e.g. Example 1.3.6). Hence, formula (9.2) cannot be used to define a semigroup associated to an unbounded operator A. But, as we will show in the next section, we can overcome this difficulty by changing the path of integration, provided that the operator A satisfies nice spectral properties.

9.2 Sectorial operators and analytical semigroups

As announced, here we introduce an important class of closed operators, the so-called, *sectorial operators*, to which we can associate a semigroup through a variant of formula (9.2).

Definition 9.2.1. A linear operator $A : D(A) \subset X \to X$ is called *sectorial* in X if there exist $\omega \in \mathbb{R}$, $\theta_0 \in (\pi/2, \pi)$ and $M > 0$ such that the resolvent set of A contains the sector $\Sigma_{\omega,\theta_0} = \{\lambda \in \mathbb{C} : \lambda \neq \omega, |\arg(\lambda-\omega)| < \theta_0\}$ and $\|R(\lambda,A)\| \le M|\lambda-\omega|^{-1}$ for every $\lambda \in \Sigma_{\omega,\theta_0}$.

Throughout this lecture, we summarize the properties listed in Definition 9.2.1, by saying that $A \in S(\omega, \theta_0, M)$.

Given $A \in S(\omega, \theta_0, M)$, we can define an operator $T(t)$, for each $t > 0$, by using formula (9.2), where we replace the curve γ_r by the "union" of three curves $-\gamma_{1,r,\eta,\omega}$, $\gamma_{2,r,\eta,\omega}$ and $\gamma_{3,r,\eta,\omega}$, where $\gamma_{2k+1,r,\eta,\omega} : [r, \infty) \to \mathbb{C}$ $(k = 0, 1)$ is defined by $\gamma_{2k+1,r,\eta,\omega}(\rho) = \omega + \rho e^{(-1)^{k+1} i\eta}$, for each $\rho \geq r$, and $\gamma_{2,r,\eta,\omega} : [-\eta, \eta] \to \mathbb{C}$ is defined by $\gamma_{2,r,\eta,\omega}(\theta) = \omega + re^{i\theta}$ for each $\theta \in [-\eta, \eta]$. Here, $r > 0$ and $\eta \in (\pi/2, \theta_0)$ are arbitrarily fixed. More precisely, we set for $t > 0$

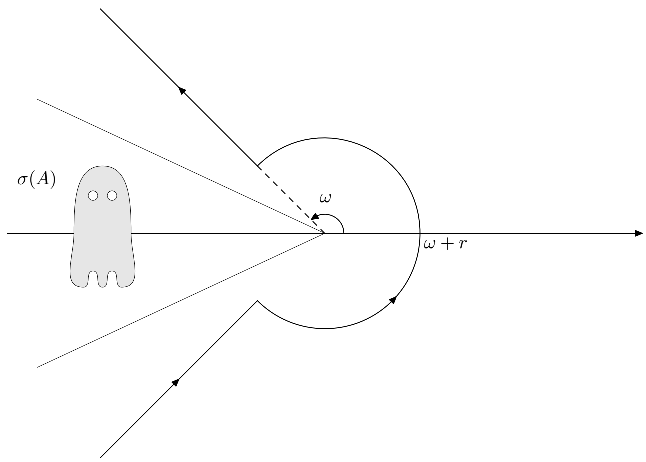

Figure 9.1: the support of the union of the three curves $\gamma_{1,r,\eta,\omega}$, $\gamma_{2,r,\eta,\omega}$ and $\gamma_{3,r,\eta,\omega}$.

$$
\begin{aligned}
T(t) = & -\frac{1}{2\pi i}\int_{\gamma_{1,r,\eta,\omega}} e^{\lambda t} R(\lambda, A)d\lambda + \frac{1}{2\pi i}\int_{\gamma_{2,r,\eta,\omega}} e^{\lambda t} R(\lambda, A)d\lambda \\
& + \frac{1}{2\pi i}\int_{\gamma_{3,r,\eta,\omega}} e^{\lambda t} R(\lambda, A)d\lambda \\
= & \frac{e^{\omega t}}{2\pi i}\bigg(\int_r^\infty e^{\rho\cos(\eta)t}[e^{i(\eta+\rho\sin(\eta)t)} R(\omega + \rho e^{i\eta}, A) - e^{-i(\eta+\rho\sin(\eta)t)} R(\omega + \rho e^{-i\eta}, A)]d\rho \\
& + \int_{-\eta}^{\eta} e^{(r\cos(\theta)+ir\sin(\theta))t} R(\omega + re^{i\theta}, A) ire^{i\theta} d\theta\bigg). \qquad (9.4)
\end{aligned}
$$

To ease the notation, we denote by $\int_{\gamma_{r,\eta,\omega}} e^{t\lambda} R(\lambda, A)d\lambda$ the last two lines of formula (9.4).

Note that $T(t)$ is well defined since the definition is independent of r and η. Indeed, by Proposition 1.3.5, the function $\lambda \mapsto v(\lambda) = e^{t\lambda} R(\lambda, A)$ is an holomorphic function in the sector Σ_{ω,θ_0}, with values in $\mathcal{L}(X)$. This in particular shows that the integral of v over $\gamma_{2,r,\eta,\omega}$ is well defined. Similarly, since

$$
\|e^{t\rho\cos(\eta)\pm it\rho\sin(\eta)} R(\omega + \rho\cos(\eta) \pm i\rho\sin(\eta), A)\| \leq Mr^{-1} e^{t\rho\cos(\eta)} \qquad (9.5)
$$

for every $\rho \geq r$ and $\cos(\eta) < 0$, it follows immediately that also the integrals of v over the curves $\gamma_{1,r,\eta,\omega}$ and $\gamma_{3,r,\eta,\omega}$ are well defined.

Now, we fix $r' > 0$ and $\eta' \in (\pi/2, \theta_0)$ and denote by D the region lying between the supports of the curves $\gamma_{j,r,\eta,\omega}$ and $\gamma_{j,r',\eta',\omega}$ $(j = 1, 2, 3)$ and by D_n $(n \in \mathbb{N})$ the intersection of D with the closed ball centered at zero with radius n.

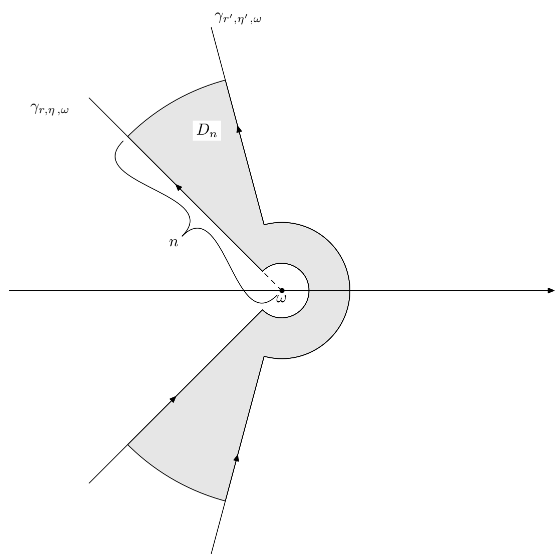

Figure 9.2: the region D_n.

Denote by γ a curve which parameterizes D_n, obtained from the curves $\gamma_{j,r,\eta,\omega}$, $\gamma_{j,r',\eta',\omega}$ and the canonical parametrization of the arc of $\partial B(0,n)$. The Cauchy integral theorem implies that

$$\int_\gamma e^{t\lambda} R(\lambda, A) d\lambda = 0.$$

By estimate (9.5) the integrals on the two arcs contained in $\partial B(0,n)$ vanish as n tends to ∞. From these remarks, we deduce that

$$\int_{\gamma_{r,\eta,\omega}} e^{t\lambda} R(\lambda, A) d\lambda = \int_{\gamma_{r',\eta',\omega}} e^{t\lambda} R(\lambda, A) d\lambda$$

as claimed.

We can now study the main properties of the operators $T(t)$, $t > 0$.

Theorem 9.2.2. *Let $A \in S(\omega, \theta_0, M)$ and let $T(t)$ be given by (9.4) for $t > 0$. Then, the following properties hold true.*

(i) *For each $x \in X$, $k \in \mathbb{N}$ and $t > 0$, $T(t)x$ belomgs to $D(A^k)$. Further, if $x \in D(A^k)$, then $A^k T(t)x = T(t)A^k x$ for every $t \geq 0$.*

(ii) *If we set $T(0) = I$, then the family $\{T(t)\}$ is a semigroup of bounded linear operators.*

(iii) *There exist positive constants M_k ($k \in \mathbb{N} \cup \{0\}$) such that*

$$\|t^k (A - \omega I)^k T(t)\| \leq M_k e^{\omega t}, \qquad t > 0, \; k \in \mathbb{N} \cup \{0\}. \tag{9.6}$$

(iv) *The function $t \mapsto T(t)$ belongs to $C^\infty((0,\infty); \mathcal{L}(X))$ and $D_t^k T(t) = A^k T(t)$ for every $t > 0$. Moreover, the function $t \mapsto T(t)$ admits an analytic extension to the sector $\Sigma_{0,\theta_0 - \pi/2}$, given by*

$$T(z) = \frac{1}{2\pi i} \int_{\gamma_{r,\theta'_z,\omega}} e^{\lambda z} R(\lambda, A)\, d\lambda, \qquad z \in \Sigma_{0,\theta_0-\pi/2},$$

where θ'_z is arbitrarily fixed in $(\pi/2, \theta_0 - \arg(z))$.

Proof. As in the proof of Hille-Yosida theorem, we replace A with the operator $A-\omega I$, which is sectorial in Σ_{0,θ_0}.

(i) The core of the proof is the case $k = 1$. Fix $t > 0$ and $x \in X$. Since $e^{t\lambda}AR(\lambda, A) = -e^{\lambda t}I + \lambda e^{\lambda t}R(\lambda, A)$, it follows that

$$\|e^{\lambda t}I - \lambda e^{\lambda t}R(\lambda, A)\| \le e^{\mathrm{Re}\,\lambda t}(1 + \|\lambda R(\lambda, A)\| \le e^{\mathrm{Re}\,\lambda t}(1 + M)$$

and $\cos(\eta) < 0$, the function $\lambda \mapsto e^{t\lambda}AR(\lambda, A)x$ is integrable along the curves $\gamma_{1,r,\eta,0}$ and $\gamma_{3,r,\eta,0}$ for each $x \in X$. Of course, being a continuous function it is also integrable along the curve $\gamma_{2,r,\eta,0}$. Therefore, we can invoke Proposition B.0.3, which guarantees that $T(t)x \in D(A)$ and

$$\begin{aligned} AT(t)x =&\frac{1}{2\pi i}\int_{\gamma_{r,\eta,0}} e^{t\lambda}AR(\lambda, A)x\,d\lambda = -\frac{x}{2\pi i}\int_{\gamma_{r,\eta,0}} e^{t\lambda}\,d\lambda + \frac{1}{2\pi i}\int_{\gamma_{r,\eta,0}} \lambda e^{t\lambda}R(\lambda, A)x\,d\lambda \\ =&\frac{1}{2\pi i}\int_{\gamma_{r,\eta,0}} \lambda e^{t\lambda}R(\lambda, A)\,\lambda, \end{aligned} \tag{9.7}$$

since

$$\int_{\gamma_{r,\eta,0}} e^{t\lambda}\,d\lambda = 0. \tag{9.8}$$

If $x \in D(A)$, then $AR(\lambda, A)x = R(\lambda, A)Ax$ and, hence, $AT(t)x = T(t)Ax$. Iterating this argument, we can show that $T(t)x \in D(A^k)$ for every $k \in \mathbb{N}$ and

$$A^kT(t)x = \frac{1}{2\pi i}\int_{\gamma_{r,\eta,0}} \lambda^k e^{\lambda t}R(\lambda, A)x\,d\lambda \tag{9.9}$$

and, if $x \in D(A^k)$, then $A^kT(t)x = T(t)A^kx$.

(ii) The semigroup property can be proved arguing as in the last part of Section 9.1. Fix $s, t > 0$. Writing

$$T(t) = \frac{1}{2\pi}\int_{\gamma_{r,\eta,0}} e^{\lambda t}R(\lambda, A)\,d\lambda, \qquad T(s) = \frac{1}{2\pi}\int_{\gamma_{2r,\eta',0}} e^{\lambda s}R(\lambda, A)\,d\lambda$$

for each $r > 0$, $\pi/2 < \eta' < \eta < \theta_0$ and using the resolvent identity it follows that

$$\begin{aligned} T(t)T(s) =& -\frac{1}{4\pi^2}\int_{\gamma_{r,\eta,0}} e^{\lambda t}R(\lambda, A)\,d\lambda \int_{\gamma_{2r,\eta',0}} \frac{e^{\mu s}}{\mu-\lambda}\,d\mu \\ &+ \frac{1}{4\pi^2}\int_{\gamma_{2r,\eta',0}} e^{\mu s}R(\mu, A)\,d\mu \int_{\gamma_{r,\eta,0}} \frac{e^{\lambda t}}{\mu-\lambda}\,d\lambda. \end{aligned}$$

Since

$$\int_{\gamma_{2r,\eta',0}} e^{\mu s}\frac{d\mu}{\mu-\lambda} = 2\pi i e^{s\lambda}, \quad \text{for } \lambda \in \gamma_{r,\eta,0} \text{ and } \int_{\gamma_{r,\eta,0}} e^{\lambda t}\frac{d\lambda}{\mu-\lambda} = 0 \text{ for } \mu \in \gamma_{2r,\eta',0}, \tag{9.10}$$

we conclude that $T(t)T(s) = T(t + s)$.

(iii) We fix $t > 0$ and take the norm of the three integrals which define $T(t)$ (see (9.4)). Since we are assuming that $\omega = 0$, from the estimate in Definition 9.2.1 we get

$$\|T(t)\| \le \frac{M}{\pi}\int_r^\infty e^{t\rho\cos(\eta)}\rho^{-1}d\rho + \frac{M}{2\pi}\int_{-\eta}^{\eta} e^{tr\cos(\theta)}d\theta \le \frac{M}{\pi}\left(\frac{e^{tr\cos(\eta)}}{tr|\cos(\eta)|} + e^{tr}\right)$$

for each $r > 0$ and $\eta \in (\pi/2, \theta_0)$. Note that, if we take the same r for every $t > 0$, we end up with an estimate for $\|T(t)\|$ which is singular as t tends to 0. To overcome this difficulty, it suffices to replace r by $1/t$ and we get (9.6) with $M_0 = M\pi^{-1}(|\cos(\eta)|^{-1}e^{\cos(\eta)} + e)$.

Estimating the norm of $AT(t)$ is easier. We do not need to take a radius which depends on t. It is enough to use formula (9.7) observing that $\|\lambda R(\lambda, A)\| \le M$. We thus get

$$\|AT(t)\| \le \frac{M}{\pi}\left(\frac{e^{tr\cos(\eta)}}{t|\cos(\eta)|} + \frac{Mr}{2\pi}\int_{-\eta}^{\eta} e^{tr\cos\theta}d\theta\right)$$

for each $r > 0$. Letting r tend to 0^+, by dominated convergence we get

$$\|AT(t)\| \le \frac{M}{\pi t|\cos(\eta)|}, \qquad t > 0. \tag{9.11}$$

To estimate the operator norm of $A^kT(t)$ for $k > 1$ it suffices to use (9.11), property (i) and the semigroup property to write $A^kT(t) = (AT(t/k))^k$ and, hence, estimate $\|A^kT(t)\| \le \|AT(t/k)\|^k$.

(iv) By applying repeatedly the dominated convergence theorem, it can be easily shown that the function $t \mapsto T(t)$ belongs to $C^\infty((0,\infty); \mathcal{L}(X))$ and

$$D_t^kT(t) = \frac{1}{2\pi i}\int_{\gamma_{r,\eta,\omega}} \lambda^k e^{\lambda t}R(\lambda, A)\, d\lambda, \qquad t > 0.$$

From this formula and (9.9) it follows that $D_t^kT(t) = A^kT(t)$ for every $t > 0$ and $k \in \mathbb{N}$.

To complete the proof, we fix $\alpha \in (0, \theta_0 - \pi/2)$. Then, the function

$$z \mapsto T(z) = \frac{1}{2\pi i}\int_{\gamma_{r,\theta_0-\alpha,\omega}} e^{z\lambda}R(\lambda, A)\, d\lambda,$$

is well defined and holomorphic in the sector $\Sigma_{0,\theta_0-\pi/2-\alpha}$. Indeed, if $\lambda = \rho e^{i(\theta_0-\alpha)}$ and $z = |z|e^{i\varphi}$ belongs to $\Sigma_{0,\theta_0-\pi/2-\alpha}$, then $\mathrm{Re}(z\lambda) = \rho|z|\cos(\theta_0 - \alpha + \varphi)$ and $\cos(\theta_0 - \alpha + \varphi)$ is negative since $\theta_0 - \alpha + \varphi \in (\pi/2, 3\pi/2)$. Hence, we can differentiate under the integral sign, taking the dominated convergence theorem into account, and conclude that the map $z \mapsto T(z)$ is holomorphic in $\Sigma_{0,\theta_0-\pi/2-\alpha}$. Since $\Sigma_{0,\theta_0-\pi/2} = \bigcup_{\alpha\in(0,\theta_0-\pi/2)}\Sigma_{0,\theta_0-\pi/2-\alpha}$, the conclusion follows. □

Remark 9.2.3. From property (iii) in Theorem 9.2.2 it follows that there exists a positive constant C_k such that $\|t^kA^kT(t)\| \le C_ke^{\omega t}$ for $t > 0$. Indeed,

$$A^k = (A - \omega I + \omega I)^k = \sum_{n=0}^{k}\binom{k}{n}\omega^{k-n}(A - \omega I)^n$$

and, therefore,

$$\|A^kT(t)\| \le \sum_{n=0}^{k}\binom{k}{n}\omega^{k-n}\|(A - \omega I)^nT(t)\| \le \sum_{n=0}^{k}\binom{k}{n}\omega^{k-n}M_nt^{-n}, \qquad t > 0.$$

If $t \le 1$, then from the previous estimate we immediately get

$$\|A^k T(t)\| \le t^{-k} \sum_{n=0}^{k} \binom{k}{n} \omega^{k-n} M_n.$$

On the other, if $t > 1$, then there exists a positive constant C_ω such that $t^{-n}\omega^{k-n} \le C_\omega t^{-k} e^{\omega t}$ for every $n < k$. We thus conclude that

$$\|A^k T(t)\| \le C_\omega t^{-k} e^{\omega t} \sum_{n=0}^{k} \binom{k}{n} M_n$$

and we are done.

In view of Theorem 9.2.2 we can now give the following definition.

Definition 9.2.4. Let A be a sectorial operator. The family of operators $\{T(t)\}$, defined by (9.4) for $t > 0$ and such that $T(0) = I$, is called *analytic semigroup generated by* A (in X).

Is each analytic semigroup strongly continuous? The answer is in general negative since $D(A)$ may be not dense in X. In any case, for each $x \in X$, $T(t)x$ converges to x as $t \to 0^+$ in a "weak" sense as the following proposition shows.

Proposition 9.2.5. *Let $\{T(t)\}$ be an analytic semigroup associated with an operator $A \in S(\omega, \theta_0, M)$. Then, $\lim_{t\to 0^+} T(t)x = x$ if and only if $x \in \overline{D(A)}$. As a byproduct, for each $\lambda \in \rho(A)$ and $x \in X$, $\lim_{t\to 0^+} R(\lambda, A)T(t)x = R(\lambda, A)x$.*

Proof. By replacing A with $A - \alpha I$ for some constant α, we can assume without loss of generality that $\omega = 0$.
Using the Cauchy integral theorem, it follows that

$$\begin{aligned} T(t)x - x &= \frac{1}{2\pi i} \int_{\gamma_{r,\eta,\omega}} e^{\lambda t} \left[R(\lambda, A)x - \frac{1}{\lambda} x \right] d\lambda \\ &= \frac{1}{2\pi i} \int_{\gamma_{r,\eta,\omega}} \frac{e^{\lambda t}}{\lambda} R(\lambda, A)Ax \, d\lambda \end{aligned}$$

for $x \in D(A)$ and $t > 0$. Since $\|\lambda^{-1} R(\lambda, A)Ax\| \le C\|Ax\||\lambda|^{-2}$, it follows from the dominated convergence theorem and Cauchy's theorem that

$$\lim_{t\to 0^+} (T(t)x - x) = \frac{1}{2\pi i} \int_{\gamma_{r,\eta,\omega}} \frac{1}{\lambda} R(\lambda, A)Ax \, d\lambda = 0$$

for $x \in D(A)$ and hence for $x \in \overline{D(A)}$, since $\|T(t)\| \le M$ for every $t \in [0, 1]$. The other implication is obtained from the inclusion $T(t)X \subset D(A)$ for every $t > 0$, see Theorem 9.2.2. Now, the last statement follows from the fact that $R(\lambda, A)$ commute with $T(t)$ for every $t > 0$ and $\lambda \in \rho(A)$. □

Remark 9.2.6. By Theorem 9.2.2, $T(t)$ maps X into $D(A)$ for every $t > 0$. Hence, it leaves $\overline{D(A)}$ invariant. Moreover, by Proposition 9.2.5, $T(t)x$ converges to x as $t \to 0^+$ for every $x \in \overline{D(A)}$. It follows that the restriction of $\{T(t)\}$ to $\overline{D(A)}$ is a C_0-semigroup. Note that $\overline{D(A)}$ is the largest subspace of X, where the restriction of $\{T(t)\}$ is a C_0-semigroup. Indeed, as already remarked, $T(t)x \in D(A)$ for each $x \in X$ and $t > 0$. Hence, if $T(t)x$ converges to x as $t \to 0^+$, then, necessarily, $x \in \overline{D(A)}$.

In the following proposition, we show other interesting properties of analytic semigroups. In particular, we show that the infinitesimal generator of the restriction of $\{T(t)\}$ to $\overline{D(A)}$ is the part of A in $\overline{D(A)}$, i.e., the operator defined in $D := \{x \in D(A) : Ax \in \overline{D(A)}\}$ by $A_{|\overline{D(A)}}x = Ax$ for each $x \in D$.

Proposition 9.2.7. *The following properties hold true.*

(i) *For each $x \in X$ and $t > 0$, $\int_0^t T(s)x\,ds \in D(A)$ and*

$$A\int_0^t T(s)x\,ds = T(t)x - x. \tag{9.12}$$

If, in addition, $x \in D(A)$, then

$$T(t)x - x = \int_0^t T(s)Ax\,ds, \qquad t \geq 0. \tag{9.13}$$

(ii) *If $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega$, then*

$$R(\lambda, A) = \int_0^\infty e^{-\lambda t}T(t)\,dt. \tag{9.14}$$

(iii) *If $x \in D(A)$ and $Ax \in \overline{D(A)}$, then $\lim_{t\to 0^+}(T(t)x - x)/t = Ax$. Conversely, if $z := \lim_{t\to 0^+}(T(t)x - x)/t$ exists, then $x \in D(A)$ and $Ax = z \in \overline{D(A)}$.*

Proof. (i) Fix $t > 0$ and $x \in X$. Since the function $T(\cdot)x$ is differentiable in $(0, \infty)$ and $D_tT(\cdot)x = AT(\cdot)x$ (see Theorem 9.2.2) then

$$\begin{aligned}\int_\varepsilon^t T(s)x\,ds =& \int_\varepsilon^t ((\omega+1)I - A)R(\omega+1, A)T(s)x\,ds\\ =&(\omega+1)\int_\varepsilon^t R(\omega+1, A)T(s)x\,ds - \int_\varepsilon^t \frac{d}{ds}(R(\omega+1, A)T(s)x)\,ds\\ =&(\omega+1)R(\omega+1, A)\int_\varepsilon^t T(s)x\,ds - T(t)R(\omega+1, A)x + T(\varepsilon)R(\omega+1, A)x\end{aligned}$$

for each $\varepsilon \in (0, t)$. Recalling that the function $T(\cdot)x$ is bounded in $[0, t]$ and continuous in $(0, +\infty)$ and taking Proposition 9.2.5 into account, we can let ε tend to 0 and obtain

$$\int_0^t T(s)x\,ds = (\omega+1)R(\omega+1, A)\int_0^t T(s)x\,ds - R(\omega+1, A)(T(t)x - x). \tag{9.15}$$

Thus, $\int_0^t T(s)x\,ds$ belongs to $D(A)$ and applying the operator $((\omega+1)I - A)$ to both the sides of (9.15), formula (9.12) follows.

If $x \in D(A)$, then the function $AT(\cdot)x$ is continuous in $(0, t]$ and bounded in $[0, t]$. Hence,

$$A\int_\varepsilon^t T(s)x\,ds = \int_\varepsilon^t AT(s)x\,ds$$

and, letting ε tend to 0^+, we conclude that

$$\lim_{\varepsilon\to 0^+} A\int_\varepsilon^t T(s)x\,ds = \int_0^t AT(s)x\,ds$$

and formula (9.13) follows from Proposition B.0.3(ii), since A is a closed operator.

(ii) Fix $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega$ and consider the operator $A - \lambda I$. This operator is sectorial and the associated semigroup $\{S(t)\}$ is defined by $S(t) = e^{-\lambda t}T(t)$ for $t \geq 0$. From formula (9.12) we know that

$$(A-\lambda)\int_0^t S(r)x\,dr = S(t)x - x, \qquad t>0,\ \ x\in X.$$

Since $\|T(t)\| \leq Me^{\omega t}$ for $t \geq 0$ (see Theorem 9.2.2), the function $S(t)x$ vanishes as t tends to ∞. Hence, $\lim_{t\to\infty}(A-\lambda)\int_0^t S(r)x\,dr = -x$ for every $x \in X$. Since $\int_0^t S(r)x\,dr$ converges to $\int_0^\infty S(r)x\,dr$ as t tends to ∞, by the closedness of A, $\int_0^t S(r)x\,dr$ belongs to $D(A)$ and

$$(\lambda - A)\int_0^\infty e^{-\lambda t}T(t)x\,dt = x, \qquad x\in X.$$

Taking into account that $\lambda \in \rho(A)$, formula (9.14) follows at once.

(iii) Fix $x \in D(A)$ such that $Ax \in \overline{D(A)}$. Then, by property (i), we can write

$$\frac{T(t)x-x}{t} = \frac{1}{t}A\int_0^t T(s)x\,ds = \frac{1}{t}\int_0^t T(s)Ax\,ds. \tag{9.16}$$

Since $Ax \in \overline{D(A)}$, by Proposition 9.2.5 the function $T(\cdot)Ax$ is continuous in $[0,\infty)$. Hence, letting t tend to 0^+ in (9.16), we conclude that $\lim_{t\to 0^+}(T(t)x-x)/t = Ax$.

Vice versa, suppose that the limit $z := \lim_{t\to 0^+}(T(t)x-x)/t$ exists. Then, clearly, $T(t)x$ converges to x as $t \to 0^+$, so that x belong to $\overline{D(A)}$. Since $T(t)x \in D(A)$ for every $t > 0$, also z belongs to $\overline{D(A)}$. Moreover, using property (i) and recalling that $R(\omega+1, A)$ commutes with $T(s)$, we get

$$\begin{aligned} R(\omega+1,A)z &= \lim_{t\to 0} R(\omega+1,A)\frac{T(t)x-x}{t} = \lim_{t\to 0} t^{-1}R(\omega+1,A)A\int_0^t T(s)x\,ds \\ &= -\lim_{t\to 0}\frac{1}{t}\int_0^t T(s)x\,ds + (\omega+1)\frac{1}{t}\int_0^t R(\omega+1,A)T(s)x\,ds. \end{aligned} \tag{9.17}$$

Since $x \in \overline{D(A)}$, the function $T(\cdot)x$ is continuous in $[0,\infty)$. Hence $\lim_{t\to 0^+} t^{-1}\int_0^t T(s)x\,ds = x$. Similarly, $R(\omega+1, A)T(t)x$ converges to $R(\omega+1, A)x$ as $t \to 0^+$, by Proposition 9.2.5. Hence, from (9.17) it follows that $x = R(\omega+1, A)((\omega+1)x - z) \in D(A)$ and $Ax = z$. □

We now provide a sufficient condition for a closed operator to be sectorial. This criterium is particularly used in the applications.

Proposition 9.2.8. *Let $A : D(A) \subset X \mapsto X$ be a closed operator such that $\Pi = \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > \omega\} \subset \rho(A)$ for some $\omega \in \mathbb{R}$, and*

$$\|\lambda R(\lambda, A)\| \leq M, \qquad \lambda \in \Pi, \tag{9.18}$$

for some constant $M \geq 1$. Then, A is a sectorial operator.

Proof. The same argument as in the proof of Proposition 1.3.5 (which uses Lemma 1.3.4) shows that the open ball $B(\omega \pm ir, |\omega + ir|/M)$ is contained in $\rho(A)$ for each $r > 0$. Since $|\omega + ir| \geq r$, the union of such balls and the halfplane Π properly contains the sector $\Sigma_{\omega,\theta_{2M}}$, where $\theta_{2M} = \pi - \arctan(2M)$. Observe that each $\lambda \in \Sigma_{\omega,\theta_{2M}}$ with $\mathrm{Re}\,\lambda \le \omega$ can be written in the form $\lambda = \omega \pm ir - (2M)^{-1}(\theta r)$ for some $\theta \in [0,1)$. Moreover, since

$$R(\lambda, A) = R(\omega \pm ir, A)(I - (2M)^{-1}\theta r R(\omega \pm ir, A))^{-1}$$

and $\|(I - (2M)^{-1}\theta r R(\omega \pm ir, A))^{-1}\| \le 2$, it follows that

$$\|R(\lambda, A)\| \le \frac{2M}{|\omega \pm ir|} \le \frac{2M}{r} \le \frac{\sqrt{4M^2+1}}{|\lambda - \omega|}.$$

On the other hand, if $\lambda \in \Sigma_{\omega,\theta_{2M}}$ with $\mathrm{Re}\,\lambda > \omega$, then from (9.18) it follows easily that $\|R(\lambda, A)\| \le M|\lambda - \omega|^{-1}$ and this completes the proof. □

We conclude this lecture with two important remarks.

Remark 9.2.9. (i) In this and in the previous lecture we have dealt with *complex* Banach spaces. In particular, the sectorial operators have been defined through an integral on an unbounded curve in the complex plane.

As a matter of fact, in many applications one has to deal with closed operators A on *real* Banach spaces X. This (apparent) difficult can be overcome by complexifying both the Banach space X and the operator A. The complexification $X_{\mathbb{C}}$ of X is the set $X_{\mathbb{C}} = \{x + iy : x, y \in X\}$, which is a Banach space when endowed with the norm $\|x + iy\|_{\widetilde{X}} = \sup_{-\pi \le \theta \le \pi} \|x\cos\theta + y\sin\theta\|$ for every $x + iy \in X_{\mathbb{C}}$ and the operations $(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$ and $(\lambda_1 + i\lambda_2)(x_1 + iy_1) = \lambda_1 x_1 - \lambda_2 y_1 + i(\lambda_1 y_1 + \lambda_2 x_1)$ for every $x_1 + iy_1, x_2 + iy_2 \in X_{\mathbb{C}}$ and every $\lambda_1 + i\lambda_2 \in \mathbb{C}$. (Note that, in general, the map $(x + iy) \mapsto \sqrt{\|x\|^2 + \|y\|^2}$ is not a norm, see Exercise 9.4.5). The complexification of the operator A is the operator $A_{\mathbb{C}} : D(A_{\mathbb{C}}) \subset X_{\mathbb{C}} \to X_{\mathbb{C}}$, defined as follows; $D(A_{\mathbb{C}}) = \{x + iy : x, y \in D(A)\}$ and $A_{\mathbb{C}}(x + iy) = Ax + iAy$ for every $x + iy \in D(A_{\mathbb{C}})$. Clearly, if we identify X with the set $\{x + i0 : x \in X\}$, then the restriction of the operator $A_{\mathbb{C}}$ to X is the operator A.

Suppose that the operator $A_{\mathbb{C}}$ is sectorial in $X_{\mathbb{C}}$ and denote by $\{T_{\mathbb{C}}(t)\}$ the associated analytic semigroup. We claim that it leaves X invariant. To prove the claim, it turns out useful to replace the three curves used to define each operator $T(t)$ by the union of the two curves γ_- and γ_+, where $\gamma_\pm : [0,\infty) \to \mathbb{C}$ are defined by $\gamma_\pm(\rho) = \omega + 1 + \rho e^{\pm i\theta}$ for every $\rho \ge 0$, where θ is fixed in $(\pi/2, \theta_0)$. For each $t > 0$ and $x \in X$ it holds that

$$\begin{aligned} T_{\mathbb{C}}(t)x = \frac{1}{2\pi i}\int_0^\infty e^{(\omega+1)t}[&e^{\rho t e^{i\theta}} R(\omega + 1 + \rho e^{i\theta}, A_{\mathbb{C}})x \\ &- \overline{e^{\rho t e^{i\theta}}} R(\omega + 1 + \rho e^{-i\theta}, A_{\mathbb{C}})x]d\rho. \end{aligned}$$

Note that $e^{\rho t e^{i\theta}} R(\omega + 1 + \rho e^{i\theta}, A_{\mathbb{C}})x - \overline{e^{\rho t e^{i\theta}}} R(\omega + 1 + \rho e^{-i\theta}, A_{\mathbb{C}})x$ belongs to iX, so that $T_{\mathbb{C}}(t)x$ belongs to X. Indeed, it is easy to check that, if we set $\overline{z + iw} = z - iw$ for every $z + iw \in X_{\mathbb{C}}$, then $R(\overline{\lambda}, A_{\mathbb{C}})x = \overline{R(\lambda, A_{\mathbb{C}})x}$ for every $\lambda \in \rho(A_{\mathbb{C}})$ and $x \in X$. Thus,

$$\begin{aligned} &e^{\rho t e^{i\theta}} R(\omega + 1 + \rho e^{i\theta}, A_{\mathbb{C}})x - \overline{e^{\rho t e^{i\theta}}} R(\omega + 1 + \rho e^{-i\theta}, A_{\mathbb{C}})x \\ =&e^{\rho t e^{i\theta}} R(\omega + 1 + \rho e^{i\theta}, A_{\mathbb{C}})x - \overline{e^{\rho t e^{i\theta}}}\,\overline{R(\omega + 1 + \rho e^{i\theta}, A_{\mathbb{C}})x} \\ =&2i\mathrm{Im}(e^{\rho t e^{i\theta}} R(\omega + 1 + \rho e^{i\theta}, A_{\mathbb{C}})x), \end{aligned}$$

where $\mathrm{Im}(x+iy) = y$ for every $x + iy \in X_{\mathbb{C}}$.
If we set $T(t) = T_{\mathbb{C}}(t)_{|X}$ for $t \geq 0$, then we define a semigroup of bounded operators in X, which satisfies the properties that we have established so far.

(ii) From this and the previous lecture, a very crucial difference between strongly continuous semigroups and analytic semigroup arises: the analytic semigroups take a datum x in the space X and make it smoother (indeed, $T(t)x$ belongs to $\bigcap_{k\in\mathbb{N}} D(A^k)$ for every $t > 0$). Strongly continuous semigroups do not enjoy this property. Think for instance to the semigroup of left translations on $X = BUC(\mathbb{R})$. Since $T(t)f = f(\cdot + t)$ for every $t > 0$, $T(t)$ has no smoothing effects on f.

9.3 Notes

We mention here that the characterization of the sectoriality of a given operator A involves only a single resolvent estimate, see Proposition 9.2.8. This should be compared with the Hille-Yosida theorem, which characterizes the infinitesimal generators of C_0-semigroups. We also mention that the single resolvent estimate, which a sectorial operator A should satisfy, yields regularity properties for the solution to the associated Cauchy problem

$$\begin{cases} u'(t) = Au(t), & t \in (0, \infty), \\ u(0) = x \end{cases}$$

which is $u = T(\cdot)x$, see Theorem 9.2.2.(i).
For more details on sectorial operators and analytic semigroups, we refer the reader, e.g., to [5, 8, 13, 14] and the recent monograph [12].

9.4 Exercises

Exercise 9.4.1. Prove formulas (9.8) and (9.10).

Exercise 9.4.2. Let $A : D(A) \subset X \to X$ be sectorial, let $\alpha \in \mathbb{C}$, and consider the operators $B : D(B) = D(A) \to X$ and $C : D(C) = D(A) \to X$, defined, respectivey, by $Bx = Ax - \alpha x$ and $Cx = \alpha Ax$ for every $x \in D(A)$. Prove that the operator B is sectorial, and that the associated semigroup $\{S(t)\}$ is defined by $S(t) = e^{-\alpha t}T(t)$ for every $t \geq 0$. For which α the operator C is sectorial?

Exercise 9.4.3. Let $A : D(A) \subset X \to X$ be a linear operator, Prove that if A and $-A$ are sectorial operators in X, then A is bounded.

Exercise 9.4.4. Let X_k, $k = 1, \ldots, n$ be Banach spaces, and let $A_k : D(A_k) \to X_k$ be sectorial operators. Set

$$X = \prod_{k=1}^{n} X_k, \qquad D(A) = \prod_{k=1}^{n} D(A_k),$$

and $A(x_1, \ldots, x_n) = (A_1x_1, \ldots, A_nx_n)$, and show that A is a sectorial operator in X, endowed with the product norm $\|(x_1, \ldots, x_n)\| = \left(\sum_{k=1}^{n} \|x_k\|^2\right)^{1/2}$.

Exercise 9.4.5. Let X be a real Banach space. Prove that the function $f : X \times X \to \mathbb{R}$ defined by $f(x,y) = \sqrt{\|x\|^2 + \|y\|^2}$ for every $x, y \in X$, does not satisfy, in general, the homogeneity property.

Exercise 9.4.6. (i) Prove that if A with domain $D(A)$ generates a C_0-group on a Banach space X, then A^2 generates an analytic semigroup on X.

(ii) Prove that the operator $Af = f'$ for $f \in D(A) = W^{1,p}(\mathbb{R})$ generates a C_0-group on $L^p(\mathbb{R})$, $1 < p < \infty$.

(iii) Deduce that the operator $Bf = f''$ for $f \in D(B) = W^{2,p}(\mathbb{R})$ generates an analytic semigroup on $L^p(\mathbb{R})$, $1 < p < \infty$.

Exercise 9.4.7. (i) Prove that the operator $Af = f''$ for any $f \in D(A) = C_b^2(\mathbb{R})$ is sectorial in $C_b(\mathbb{R})$ and, hence, generates an analytic semigroups.

(ii) Prove that the operator $Af = f''$ for any $f \in D(A) = \{u \in C_b^2([0,1]) : u(0) = u(1) = 0\}$ generates an analytic semigroup in $C([0,1])$. Is this semigroup strongly continuous?

(iii) Prove that the operator $Af = f''$ for any $f \in D(A) = \{u \in C_b^2([0,1]) : u'(0) = u'(1) = 0\}$ generates an analytic semigroup in $C([0,1])$. Is this semigroup strongly continuous?

Exercise 9.4.8. Prove that, given a normal operator A on a separable Hilbert space, then the operators e^{tA}, defined according to the functional calculus introduced in Lecture 5, and according to formula (9.2), coincide.

Bibliography

[1] A. Belleni-Morante, Applied Semigroups and Evolution Equations, Oxford University Press, 1979.

[2] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[3] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[4] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[5] K.-J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Graduate Text in Mathematics **194**, Springer-Verlag, 2000.

[6] K.-J. Engel and R. Nagel, Short Course on Operator Semigroups, Universitext, Springer-Verlag, 2006.

[7] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[8] J.A. Goldstein, Semigroups of Operators and Applications, Oxford University Press, 1985.

[9] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[10] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[11] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981.

[12] L. Lorenzi and A. Rhandi, Semigroups of Bounded Operators and Second-Order Elliptic and Parabolic Partial Differential Equations, CRC Press, Taylor & Francis Group, 2021.

[13] A. Lunardi, Analytic Semigroups and Optimal Regularity in Parabolic Problems, Birkhäuser Verlag, 1995.

[14] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Appl. Math.Sci. **44**, Springer-Verlag,1983.

[15] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[16] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[17] W. Rudin, Real and Complex Analysis, 3$^{\text{rd}}$ ed., McGraw-Hill, 1986.

[18] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[19] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958.

[20] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861.

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

January 13, 2022

Lecture 10

Non homogeneous abstract Cauchy problems

In this lecture, we analyze the abstract Cauchy problem

$$\begin{cases} u'(t) = Au(t) + f(t), & t \in (0, T], \\ u(0) = x, \end{cases} \tag{10.1}$$

on a Banach space X, when $f : [0, T] \to X$ is a continuous function, $x \in X$ and $A : D(A) \subset X \to X$ is the infinitesimal generators of a strongly continuous semigroups or a sectorial operator. As we will see, when A is not sectorial quite strong regularity assumptions on x and f should be assumed to guarantee the existence of a solution to (10.1). Such conditions can be sensibly weakened when A is a sectorial operator, due to the smoothing properties of the associated semigroup. All the results of this lecture will be "tested" on a relevant model, namely the Gauss-Weierstrass semigroup, in Lecture 11.

We introduce the following definition.

Definition 10.0.1. Let $f : [0, T] \to X$ be a continuous function and let $x \in X$. Then, a strict (resp. classical) solution to problem (10.1) is a function $u \in C^1([0, T]; X) \cap C([0, T]; D(A))$ (resp. $u \in C^1((0, T]; X) \cap C((0, T]; D(A)) \cap C([0, T]; X)$) which satisfies the differential equation $u' = Au + f$ in $(0, T]$ and satisfies the condition $u(0) = x$.

At it is readily seen, the main difference between strict and classical solutions is the regularity at $t = 0$.

Throughout this lecture, we write $Y \hookrightarrow X$ when Y is a Banach space, continuously embedded into X.

10.1 The case when A generates a C_0-semigroup

From the definition of strict solution it follows immediately that, if problem (10.1) admits a strict solution, then $x \in D(A)$.

The following proposition shows that problem (10.1) admits at most a unique solution and provides us with a representation formula for the solutions in terms of the semigroup $\{T(t)\}$ generated by the operator A.

Proposition 10.1.1. *Fix $f \in C([0,T],X)$ and $x \in D(A)$. If u is a strict solution to the Cauchy problem* (10.1)*, then*

$$u(t) = T(t)x + \int_0^t T(t-s)f(s)\,ds, \qquad t \in [0,T]. \tag{10.2}$$

Proof. Let u be a strict solution to (10.1) and fix $t \in (0,T]$. Since u belongs to $C^1([0,T];X) \cap C([0,T];D(A))$, the function $w : [0,t] \to X$, defined by $w(s) = T(t-s)u(s)$, $s \in [0,t]$, belongs to $C^1([0,t],X)$ and

$$w'(s) = -\,AT(t-s)u(s) + T(t-s)(Au(s) + f(s)) = T(t-s)f(s), \qquad s \in [0,t].$$

Integrating the previous formula in $[0,t]$ and observing that $w(0) = T(t)x$ and $w(t) = u(t)$, the assertion follows. □

Formula (10.2) is the so-called variation-of-constants formula. Whenever the integral in (10.2) makes sense (for instance, when the function $t \mapsto \|f(t)\|$ belongs to $L^1((0,T))$, the function u defined by (10.2) is said to be a *mild solution* of (10.1).

Remark 10.1.2. As it is easily seen, if $x \in X$ and $f \in C([0,T];X)$, then the mild solution to the Cauchy problem (10.1) is continuous in $[0,T]$, with values in X. Moreover, there exists a positive constant C, independent of x and f, such that

$$\|u\|_{C([0,T];X)} \le C(\|x\| + \|f\|_{C([0,T];X)}).$$

The following theorem, provides us with two existence and uniqueness results for problem (10.1). For this purpose we recall that the graph norm is defined by $\|x\|_{D(A)} := \|x\| + \|Ax\|$ for $x \in D(A)$.

Theorem 10.1.3. *Let $x \in D(A)$ and $f \in C([0,T];D(A))$ (resp. $f \in C^1([0,T];X)$). Then, the Cauchy problem* (10.1) *admits a unique strict solution u. Moreover, there exists a positive constant C, independent of x and f, such that*

$$\|u\|_{C^1([0,T];X)} + \|u\|_{C([0,T];D(A))} \le C(\|x\|_{D(A)} + \|f\|_{C([0,T];D(A))}), \tag{10.3}$$

if $f \in C([0,T];D(A))$, and

$$\|u\|_{C^1([0,T];X)} + \|u\|_{C([0,T];D(A))} \le C(\|x\|_{D(A)} + \|f\|_{C^1([0,T];X)}) \tag{10.4}$$

if $f \in C^1([0,T];X)$.

In the proof of the previous theorem as well as in the proof of the forthcoming Theorems 10.2.14 and 10.2.16, we will take advantage of the following proposition.

Proposition 10.1.4. *Let u be the mild solution to the Cauchy problem* (10.1)*, corresponding to $x \in X$ and $f \in C([0,T];X)$. Then, u is a strict solution to the Cauchy problem* (10.1) *if and only if $u \in C([0,T];D(A))$ or, equivalently, if and only if $u \in C^1([0,T];X)$.*

In the proof of Proposition 10.1.4, we will use the following result, which shows that the mild solution satisfies an integrated version of (10.1).

Lemma 10.1.5. *For every $f \in C([0,T];X)$ and $x \in X$, let u be the function defined by* (10.2). *Then, for every $t \in [0,T]$, $\int_0^t u(s)ds$ belongs to $D(A)$, and*

$$u(t) = x + A\int_0^t u(s)ds + \int_0^t f(s)ds, \qquad t \in [0,T]. \tag{10.5}$$

Proof. We fix $t \in (0,T]$ and integrate function u over the interval $[0,T]$. Interchanging the order of integration, we obtain that

$$\int_0^t u(s)ds = \int_0^t T(s)xds + \int_0^t d\sigma \int_0^{t-\sigma} T(s)f(\sigma)ds. \tag{10.6}$$

The first term in the right-hand side of (10.6) belongs to $D(A)$ due to Proposition 8.2.2(ii). The same proposition shows that the integral $\int_0^{t-\sigma} T(\tau)f(\sigma)d\tau$ belongs to $D(A)$ and $A\int_\sigma^t T(s-\sigma)f(\sigma)ds = (T(t-\sigma)-I)f(\sigma)$, so that, applying Proposition B.0.3, we conclude that also the second integral in the right-hand side of (10.6) belongs to $D(A)$. Hence, $\int_0^t u(s)ds$ belongs to $D(A)$ and, again, Proposition B.0.3 shows that

$$A\int_0^t u(s)ds = T(t)x - x + \int_0^t \big(T(t-\sigma)-I\big)f(\sigma)d\sigma.$$

The assertion follows. □

Proof of Proposition 10.1.4. Of course, it suffices to prove that if $u \in C([0,T];D(A))$ or $u \in C^1([0,T];X)$, then u is a strict solution to problem (10.1). In view of Remark 10.1.2, we already know that $u \in C([0,T];X)$ and $u(0) = x$. Using the integral representation formula (10.5), we can write

$$\frac{u(t+h)-u(t)}{h} = A\bigg(\frac{1}{h}\int_t^{t+h} u(s)ds\bigg) + \frac{1}{h}\int_t^{t+h} f(s)ds \tag{10.7}$$

for every $t \in [0,T]$ and h such that $t+h \in [0,T]$.

Since f is continuous in $[0,T]$, the second integral in the right-hand side of (10.7) converges to $f(t)$ as h tends to 0, for every $t \in [0,T]$. Similarly, the integral mean of u over the interval with endpoints t and $t+h$ converges to $u(t)$ as h tends to 0 for every $t \in [0,T]$. Assume that $u \in C^1([0,T];X)$. Then, the left-hand side of (10.7) converges to $u'(t)$ as h tends to 0 for every $t \in [0,T]$. By difference, the first integral term in the right-hand side of (10.7) converges to $u'(t) - f(t)$ for every $t \in [0,T]$. The closedness of A implies that $u(t) \in D(A)$ and $Au(t) = u'(t) - f(t)$ for every $t \in [0,T]$.

On the other hand, if $u \in C([0,T];D(A))$ then

$$A\bigg(\frac{1}{h}\int_t^{t+h} u(s)ds\bigg) = \frac{1}{h}\int_t^{t+h} Au(s)ds$$

for every $t \in [0,T]$. Hence, the above term converges to $Au(t)$ as h tends to 0 for every $t \in [0,T]$. By difference, u is differentiable in $[0,T]$ and $u'(t) = Au(t) + f(t)$ for such values of t. The proof is complete. □

We can now prove Theorem 10.1.3.

Proof of Theorem 10.1.3. By Proposition 8.2.2(i), we already know that the function $t \mapsto T(t)x$ belongs to $C^1([0,T];X) \cap C([0,T];D(A))$ and solves the Cauchy problem (10.1), with $f = 0$. Hence, to complete the proof, we just need to deal with the integral term in the definition of the mild solution. We denote by v such a term.

We first assume that $f \in C([0,T];D(A))$. Then, for every $t \in [0,T]$ the function $s \mapsto AT(t-s)f(s) = T(t-s)Af(s)$ is continuous in $[0,T]$. This implies that $v(t)$ belongs to $D(A)$ for every $t \in [0,T]$ and

$$Av(t) = \int_0^t T(t-s)Af(s)ds \quad \text{for every } t \in [0,T]. \tag{10.8}$$

Hence, Av is continuous in $[0,T]$ with values in X. Taking Remark 10.1.2 into account, we obtain that $v \in C([0,T];D(A))$. Since v is the mild solution of the problem (10.1), using Proposition 10.1.4, we conclude that the mild solution u is actually a strict solution. Moreover, applying (10.8), one deduces (10.3).

Suppose now that $f \in C^1([0,T];X)$ and fix $t \in [0,T]$ and $h \in \mathbb{R}$ such that $t+h \in [0,T]$. Then

$$\begin{aligned}\frac{v(t+h)-v(t)}{h} =& \frac{1}{h}\bigg(\int_0^{t+h} T(s)f(t+h-s)ds - \int_0^t T(s)f(t-s)ds\bigg)\\ =& \int_0^t T(s)\frac{f(t+h-s)-f(t-s)}{h}ds + \frac{1}{h}\int_t^{t+h} T(s)f(t+h-s)ds\\ =& I_{1,h} + I_{2,h}.\end{aligned}$$

As it is easily seen,

$$\bigg\|I_{1,h} - \int_0^t T(s)f'(t-s)ds\bigg\| \le M_{0,T}\int_0^t \bigg\|\frac{f(t+h-s)-f(t-s)}{h} - f'(t-s)\bigg\|ds$$

and the right-hand side of the previous inequality vanishes as h tends to 0. As far as $I_{2,h}$ is concerned, estimating $\|T(s)(f(t+h-s)-f(0))\| \le M_{0,T}\|f(t+h-s)-f(0)\|$, where $M_{0,T} = \sup_{t\in[0,T]}\|T(t)\|_{\mathcal{L}(X)}$, we can write

$$\|I_{2,h} - T(t)f(0)\| \le M_{0,T}\frac{1}{h}\int_t^{t+h}\|f(t+h-s)-f(0)\|ds + \frac{1}{h}\int_t^{t+h}\|T(s)f(0)\|ds.$$

Clearly, the term $\dfrac{1}{h}\displaystyle\int_t^{t+h}\|T(s)f(0)\|ds$ tends to $\|f(0)\|$ as h tends to 0, since the function $s \mapsto \|T(s)f(0)\|$ is continuous in $[0,\infty)$. Similarly, the continuity of the function f shows that $\dfrac{1}{h}\displaystyle\int_t^{t+h}\|T(s)(f(t+h-s)-f(0))\|ds$ vanishes as h tends to 0.

We have so proved that v is differentiable at t and

$$v'(t) = \int_0^t T(s)f'(t-s)ds + T(t)f(0), \qquad t \in [0,T]. \tag{10.9}$$

From this formula, we easily deduce that $v' \in C([0,T];X)$ so that, Proposition 10.1.4 allows us again to conclude that the mild solution to (10.1) is actually a strict solution. Furthermore, taking (10.9) into account, one can easily verify that (10.4) holds. □

Remark 10.1.6. Without any additional information on the operator A, the assumptions of Theorem 10.1.3 cannot be weakened. Indeed, consider the Cauchy problem (10.1), associated to the realization A of the first-order derivatives in $BUC(\mathbb{R})$, having $BUC^1(\mathbb{R})$ as domain, with $x = 0$ and $f(t) = T(t)\psi$ for every $t \in [0,T]$, where $\{T(t)\}$ is the semigroup of left-translations, $\psi \in BUC(\mathbb{R})$ is not differentiable in $\mathbb{R}$. The mild solution to such Cauchy problem is the function u, defined by

$$u(t) = \int_0^t T(t-s)T(s)\psi\, ds = tT(t)\psi = t\psi(t+\cdot), \qquad t \geq 0,$$

which is not differentiable in $[0,\infty)$.

We remark that, differently from analytic semigroups, strongly continuous semigroups do not enjoy, in general, any smoothing property and this prevents us from weakening the regularity assumptions on the data to guarantee the existence of a solution of the problem (10.1).

10.2 The case when A is a sectorial operator

In this section, we assume that A is a sectorial operator and denote by $\{T(t)\}$ the associated semigroup. Moreover, for every $T > 0$ and $k = 0, 1, 2$ we set $M_{k,T} = \sup\limits_{0<t\leq T} \|t^k A^k T(t)\|_{\mathcal{L}(X)}$.

10.2.1 The interpolation spaces

In the analysis of the Cauchy problem (10.1) a crucial role is played by the so-called interpolation space $D_A(\alpha,\infty)$. Roughly speaking, for $\alpha \in (0,1)$, $D_A(\alpha,\infty)$ is characterized as the subset of X of all the elements x for which the function $t \mapsto AT(t)x$, exhibits, as t tends to 0, an intermediate behaviour between the behaviour it has when $x \in X$ and when $x \in D(A)$. More, rigorously,

Definition 10.2.1. For every $\alpha \in (0,1)$, $D_A(\alpha,\infty)$ is the set of all $x \in X$ such that $[x]_{D_A(\alpha,\infty)} := \sup_{0<t\leq 1} \|t^{1-\alpha}AT(t)x\| < \infty$.

$D_A(\alpha,\infty)$ is a Banach space when endowed with the norm $x \mapsto \|x\|_{D_A(\alpha,\infty)} = \|x\| + [x]_{D_A(\alpha,\infty)}$. Moreover, $x \in D_A(\alpha,\infty)$ if and only $[x]'_\alpha = \sup_{0<t\leq\tau} \|t^{1-\alpha}AT(t)x\| < \infty$ for some $\tau > 0$. The norm of $D_A(\alpha,\infty)$ is equivalent to the norm $x \mapsto \|x\| + [x]'_\alpha$ (see Esercises 10.4.1 and 10.4.2(a)).

Remark 10.2.2. (i) Using the semigroup law and Remark 9.2.3, it is easy to obtain an estimate for the blow up rate of $\|A^k T(t)\|_{\mathcal{L}(D_A(\alpha,\infty);X)}$ as t tends to 0. Indeed, for each $x \in D_A(\alpha,\infty)$ we can estimate

$$\sup_{0<t\leq 1} \|t^{k-\alpha}A^kT(t)x\| \leq \sup_{0<t\leq 1} \|t^{k-1}A^{k-1}T(t/2)\|_{\mathcal{L}(X)} \|t^{1-\alpha}AT(t/2)x\| \leq C\|x\|_{D_A(\alpha,\infty)}.$$

(ii) It is clear that, if $x \in D_A(\alpha,\infty)$ and $T > 0$, then the function $s \mapsto \|AT(s)x\|$ belongs to $L^1((0,T))$, so that, by Proposition 9.2.7(ii), $T(t)x - x = \int_0^t AT(s)x ds$ for every $t \in [0,1]$. From this formula we deduce easily that

$$\|T(t)x - x\| \leq \alpha^{-1}[x]_{D_A(\alpha,\infty)}t^\alpha, \qquad t \in [0,1], \tag{10.10}$$

and Proposition 9.2.7(iii) shows that $D_A(\alpha,\infty) \hookrightarrow \overline{D(A)}$. Moreover, from the definition of the space $D_A(\alpha,\infty)$, it follows also that $D(A) \hookrightarrow D_A(\alpha,\infty) \hookrightarrow D_A(\beta,\infty)$ for every $0<\beta<\alpha<1$.

Actually, condition (10.10) provides us with an equivalent characterization of the interpolation space $D_A(\alpha,\infty)$, as the following proposition shows.

Proposition 10.2.3. *For every $\alpha\in(0,1)$ the equality*

$$D_A(\alpha,\infty)=\Big\{x\in X:\ [[x]]_{D_A(\alpha,\infty)}=\sup_{0<t\le 1} t^{-\alpha}\|T(t)x-x\|<\infty\Big\}$$

holds, and the norm $x\mapsto \|x\|+[[x]]_{D_A(\alpha,\infty)}$ is equivalent to the norm of $D_A(\alpha,\infty)$.

Proof. In view of the remark above, we just need to show that, if $[[x]]_{D_A(\alpha,\infty)}<\infty$, then $x\in D_A(\alpha,\infty)$ and $[x]_{D_A(\alpha,\infty)}\le C[[x]]_{D_A(\alpha,\infty)}$ for some positive constant C, independent of x. For this purpose, we fix $x\in X$ such that $[[x]]_{D_A(\alpha,\infty)}<\infty$ and write

$$AT(t)x=AT(t)\frac{1}{t}\int_0^t(x-T(s)x)ds+T(t)\frac{1}{t}A\int_0^t T(s)xds$$

for $t\in(0,1]$. Then,

$$\begin{aligned}\|t^{1-\alpha}AT(t)x\| \le & t^{1-\alpha}\frac{M_{1,1}}{t^2}\int_0^t s^\alpha\frac{\|x-T(s)x\|}{s^\alpha}ds+M_{0,1}t^{-\alpha}\|T(t)x-x\|\\ \le & t^{1-\alpha}\frac{M_{1,1}}{t^2}[[x]]_{D_A(\alpha,\infty)}\int_0^t s^\alpha ds+M_{0,1}[[x]]_{D_A(\alpha,\infty)}\\ = & \left(\frac{M_{1,1}}{1+\alpha}+M_{0,1}\right)[[x]]_{D_A(\alpha,\infty)},\end{aligned}$$

and the assertion follows. □

As a corollary of the previous proposition, we obtain another characterization of the interpolation space $D_A(\alpha,\infty)$.

Corollary 10.2.4. *The function $t\mapsto T(t)x$ belongs to $C^\alpha([0,1];X)$ if and only if $x\in D_A(\alpha,\infty)$. In this case, it belongs to $C^\alpha([0,T];X)$ for every $T>0$.*

Proof. Since $T(0)x=x$, it is clear that, if the function $t\mapsto T(t)x$ is α-Hölder continuous in $[0,1]$, then $x\in D_A(\alpha,\infty)$.

Let us suppose that $x\in D_A(\alpha,\infty)$ and fix $s,t\in[0,1]$ with $t>s$. Then

$$\begin{aligned}\|T(t)x-T(s)x\| = & \|T(s)(T(t-s)x-x)\|\le\|T(s)\|_{\mathcal{L}(X)}\|T(t-s)x-x\|\\ \le & M_{0,1}[[x]]_{D_A(\alpha,\infty)}(t-s)^\alpha,\end{aligned}$$

so that the function $t\mapsto T(t)x$ belongs to $C^\alpha([0,1];X)$.

The last part of the proof is left to the reader as an exercise (see Exercise 10.4.3). □

Next, we prove the following remarkable property of the interpolation space $D_A(\alpha,\infty)$.

Proposition 10.2.5. *For every $x \in D(A)$ it holds that $[x]_{D_A(\alpha,\infty)} \le M_{0,1}^\alpha M_{1,1}^{1-\alpha}\|Ax\|^\alpha \|x\|^{1-\alpha}$. As a consequence, for every $T>0$ and $n \in \mathbb{N}$, it follows that*

$$\sup_{0<t\le T} \|t^{n+\alpha}A^nT(t)\|_{\mathcal{L}(X,D_A(\alpha,\infty))} < \infty. \tag{10.11}$$

Proof. Fix $x \in D(A)$ and $t \in (0,1]$. Then, we can estimate

$$\|t^{1-\alpha}AT(t)x\| \le M_{0,1}t^{1-\alpha}\|Ax\|, \qquad \|t^{1-\alpha}AT(t)x\| \le M_{1,1}t^{-\alpha}\|x\|.$$

Combining these two estimates, we deduce that

$$\|t^{1-\alpha}AT(t)x\| \le (M_{0,1}t^{1-\alpha}\|Ax\|)^\alpha (M_{1,1}t^{-\alpha}\|x\|)^{1-\alpha} = M_{0,1}^\alpha M_{1,1}^{1-\alpha}\|Ax\|^\alpha\|x\|^{1-\alpha}.$$

Taking the supremum over $(0,1]$, the first part of the assertion follows.

Applying the estimate $[x]_{D_A(\alpha,\infty)} \le M_{0,1}^\alpha M_{1,1}^{1-\alpha}\|Ax\|^\alpha\|x\|^{1-\alpha}$ with x being replaced with $T(t)x$, which belongs to $D(A)$ for every $t>0$ and $x \in X$, we obtain

$$\|T(t)x\|_{D_A(\alpha,\infty)} \le M_{0,1}^\alpha M_{1,1}^{1-\alpha}\|AT(t)x\|^\alpha\|T(t)x\|^{1-\alpha} + M_{0,T}\|x\| \le \frac{C_T}{t^\alpha}\|x\|, \qquad t \in (0,T]. \tag{10.12}$$

Finally, writing $\|A^nT(t)x\|_{D_A(\alpha,\infty)} \le \|A^nT(t/2)\|_{\mathcal{L}(X)}\|T(t/2)x\|_{D_A(\alpha,\infty)}$ for $t>0$ and $n \in \mathbb{N}$ and using Remark 9.2.3, we complete the proof. □

Proposition 10.2.5, shows that the norm of $D_A(\alpha,\infty)$ can be interpolated between the norms of X and $D(A)$. More precisely, there exists a positive constant C, independent of x, such that $\|x\|_{D_A(\alpha,\infty)} \le C\|x\|_{D(A)}^\alpha\|x\|^{1-\alpha}$ for every $x \in D(A)$. This amounts to say that $D_A(\alpha,\infty)$ is a space of class J_α between X and $D(A)$, according to the following definition.

Definition 10.2.6. Let X, Y, Z be three Banach spaces with $Z \hookrightarrow Y \hookrightarrow X$ and let $\alpha \in (0,1)$. The Banach space Y is said to be of class J_α between X and Z if there exists $C>0$ such that $\|z\|_Y \le C\|z\|_Z^\alpha\|z\|_X^{1-\alpha}$ for every $z \in Z$.

As the following proposition shows, Formula (10.11) can be extended to the case when $D_A(\alpha,\infty)$ is replaced by a space of class J_α between X and $D(A)$.

Proposition 10.2.7. *Let X_α be a space of class J_α between X and $D(A)$, for some $\alpha \in (0,1)$. Then, for every $k \in \mathbb{N}\cup\{0\}$ there exist constants $M_{k,\alpha}>0$ such that*

$$\|A^kT(t)\|_{\mathcal{L}(X,X_\alpha)} \le \frac{M_{k,\alpha}}{t^{k+\alpha}}, \qquad t \in (0,1]. \tag{10.13}$$

Moreover, $D_A(\beta,\infty) \hookrightarrow X_\alpha$, for every $\beta \in (\alpha,1)$.

Proof. Since $\|A^kT(t)x\|_{X_\alpha} \le C(\|A^kT(t)x\|_{D(A)})^\alpha(\|A^kT(t)x\|_X)^{1-\alpha}$ for every $x \in X$, using Remark 9.2.3, formula (10.13) follows easily.

A similar argument shows that $\|AT(s)x\|_{X_\alpha} \le C_\beta s^{-1+\beta}\|x\|_{D_A(\beta,\infty)}$ for every $s \in (0,1]$, $x \in D_A(\beta,\infty)$, $\beta \in (\alpha,1)$ and some positive constant C_β, independent of x. Hence, writing

$$x = T(1)x - \int_0^1 AT(s)xds.$$

and observing that $T(1)x \in D(A) \hookrightarrow X_\alpha$ and

$$\left\|\int_0^1 AT(s)xds\right\|_{X_\alpha} \le C_\beta\|x\|_{D_A(\beta,\infty)}\int_0^1 t^{\beta-1}dt,$$

we conclude that $x \in X_\alpha$ and the embedding $D_A(\beta,\infty) \hookrightarrow X_\alpha$ follows. □

Remark 10.2.8. In general a space of class J_α between X and $D(A)$ may not be contained in any interpolation space $D_A(\beta,\infty)$. Indeed, if $X = C([0,1])$, A is the realization of the second-order derivative with homogeneous Dirichlet boundary conditions in X, i.e. $D(A) = \{u \in C^2([0,1]) : \ u(0) = u(1) = 0\}$ and $Au = u''$ for every $u \in D(A)$, then $C^1([0,1])$ is of class $J_{1/2}$ between X and $\overline{D(A)}$ but there exists no $\beta \in (0,1)$ such that $C^1([0,1]) \subset D_A(\beta,\infty)$ because the functions in $\overline{D(A)}$ vanish at $x = 0$ and at $x = 1$ and $D_A(\beta,\infty) \subset \overline{D(A)}$ (see Remark 10.2.2). Similarly, the part A_α of A in X_α could not be sectorial. Since $D(A) \hookrightarrow X_\alpha \hookrightarrow X$, it follows that the function $t \mapsto T(t)$ is analytic in $(0,\infty)$ with values in $\mathcal{L}(X_\alpha)$. In particular, the function $t \mapsto \|T(t)\|_{\mathcal{L}(X_\alpha)}$ is bounded in bounded closed intervals contained in $(0,\infty)$, but it could blow up as t tends to 0.

To conclude this subsection, we introduce the interpolation spaces of order greater than one.

Definition 10.2.9. For every $k \in \mathbb{N}$ and $\alpha \in (0,1)$ we set $D_A(k+\alpha,\infty) = \{x \in D(A^k) : A^k x \in D_A(\alpha,\infty)\}$, endowed with the norm $x \mapsto \|x\|_{D_A(k+\alpha,\infty)} = \|x\|_{D(A^k)} + [A^k x]_{D_A(\alpha,\infty)}$.

It is easy to check that $D_A(k+\alpha,\infty)$ are Banach spaces (see Exercise 10.4.1). Moreover, from Corollary 10.2.4 it follows that the function $t \mapsto u(t) := T(t)x$ belongs to the space $C^\alpha([0,T];D(A))$ for every $T > 0$ if and only if x belongs to $D_A(\alpha+1,\infty)$. Similarly, since $\frac{d}{dt}T(t)x = T(t)Ax$ for $x \in D(A)$, u belongs to $C^{1+\alpha}([0,1];X)$ (and then to $C^{1+\alpha}([0,T];X)$ for all $T > 0$) if and only if x belongs to $D_A(\alpha+1,\infty)$.

Let us denote by A_α the part of A in $D_A(\alpha,\infty)$ ($\alpha \in (0,1)$), i.e., the operator $A_\alpha : D_A(\alpha+1,\infty) \to D_A(\alpha,\infty)$, defined by $A_\alpha x = Ax$ for every $x \in D_A(\alpha+1,\infty)$, is a sectorial operator as the following proposition shows.

Proposition 10.2.10. *For $\alpha \in (0,1)$ the resolvent set of A_α contains $\rho(A)$, $R(\lambda,A_\alpha)$ is the restriction of $R(\lambda,A)$ to $D_A(\alpha,\infty)$, and the inequality $\|R(\lambda,A_\alpha)\|_{\mathcal{L}(D_A(\alpha,\infty))} \le \|R(\lambda,A)\|_{\mathcal{L}(X)}$ holds for every $\lambda \in \rho(A)$. In particular, A_α is a sectorial operator and the associated semigroup is the restriction of $T(t)$ to $D_A(\alpha,\infty)$.*

Proof. Fix $\lambda \in \rho(A)$ and $x \in D_A(\alpha,\infty)$. The resolvent equation $\lambda y - Ay = x$ has a unique solution $y \in D(A)$. Since $D(A) \subset D_A(\alpha,\infty)$, Ay belongs to $D_A(\alpha,\infty)$, so that $y = R(\lambda,A)x \in D_A(\alpha+1,\infty)$. Moreover for $t \in (0,1]$, we can estimate

$$\|t^{1-\alpha}AT(t)R(\lambda,A)x\| = \|R(\lambda,A)t^{1-\alpha}AT(t)x\| \le \|R(\lambda,A)\|_{\mathcal{L}(X)}\|t^{1-\alpha}AT(t)x\|$$

and conclude that $[R(\lambda,A)x]_{D_A(\alpha,\infty)} \le \|R(\lambda,A)\|_{\mathcal{L}(X)}[x]_{D_A(\alpha,\infty)}$. Hence, $\|R(\lambda,A_\alpha)\|_{\mathcal{L}(D_A(\alpha,\infty))} \le \|R(\lambda,A)\|_{\mathcal{L}(X)}$ and this is enough to conclude that A_α is a sectorial operator. □

10.2.2 Existence of a solution of the Cauchy problem (10.1)

To begin with, we observe that, if problem (10.1) admits a strict (resp. classical) solution, then $x \in D(A)$ and $Ax + f(0) \in \overline{D(A)}$ (resp. $x \in \overline{D(A)}$).

As in the case when A is the generator of a strongly continuous semigroup, we can prove that the classical solution to problem (10.1) is unique, if it exists.

Proposition 10.2.11. *Fix $f \in C([0,T];X)$ and $x \in \overline{D(A)}$. If u is a classical solution to the Cauchy problem* (10.1)*, then*

$$u(t) = T(t)x + \int_0^t T(t-s)f(s)ds, \qquad t \in [0,T]. \tag{10.14}$$

Proof. The assertion can be obtained arguing as in the proof of Proposition 10.1.1 with a slight change that we illustrate here. As in the proof of the quoted proposition, for every $t \in (0,T]$ we introduce the same function $w : [0,t] \to X$, which, now, is continuously differentiable only in $(0,t)$ and $w'(s) = T(t-s)f(s)$ for every $s \in (0,t)$. Integrating the previous formula in $[\varepsilon, t-\varepsilon]$ and, then, letting ε tend to 0^+, we get the assertion. □

In view of Proposition 10.2.11 the problem of the existence of a classical or strict solution to problem (10.1) is reduced to the problem of the regularity of the mild solution. In general, even for $x = 0$ the continuity of f is not sufficient to guarantee that the mild solution is classical. Trying to show that $u(t) \in D(A)$ by estimating $\|AT(t-s)f(s)\|$ is useless, because $\|AT(t-s)f(s)\| \le C\|f\|_\infty (t-s)^{-1}$ and this is not sufficient to make the integral convergent. More sophisticated arguments, such as in the proof of Proposition 9.2.7(ii), do not work (see Exercise 10.4.7 for a rigorous counterexample).

Without any further regularity assumption on f besides its continuity in $[0,T]$ with values in X, the integral term in the mild solution u to problem (10.1) may not be differentiable in $(0,T]$. Nevertheless, it is much smoother than in the case $\{T(t)\}$ is merely a C_0-semigroup: in fact, for every $\alpha \in (0,1)$, u belongs to the space $C^\alpha([0,T];X)$ of all α-Hölder continuous functions $v : [0,T] \to X$, endowed with the norm $\|v\|_{C^\alpha([0,T];X)} = \|v\|_\infty + [v]_{C^\alpha([0,T];X)}$, where

$$[v]_{C^\alpha([0,T];X)} = \sup_{\substack{t,s\in[0,T]\\ t\neq s}} \frac{\|v(t)-v(s)\|}{|t-s|^\alpha}.$$

Proposition 10.2.12. *Let $f \in C([0,T];X)$. Then, the function*

$$v(t) = (T(\cdot) * f)(t) := \int_0^t T(t-s)f(s)ds, \qquad t \in [0,T],$$

belongs to $C^\alpha([0,T];X)$ for every $\alpha \in (0,1)$, and there exists a positive constant $C = C(\alpha,T)$ such that

$$\|v\|_{C^\alpha([0,T];X)} \le C\|f\|_{C([0,T];X)}. \tag{10.15}$$

Proof. A straightforward computation shows that

$$\|v(t)\| \le M_{0,T}t\|f\|_\infty, \qquad t \in [0,T], \tag{10.16}$$

so that v is bounded in $[0,T]$ with values in X. Moreover, using Proposition 9.2.7, for $0 \le s \le t \le T$ we can split

$$\begin{aligned} v(t)-v(s) &= \int_0^s \left(T(t-\sigma)-T(s-\sigma)\right) f(\sigma)d\sigma + \int_s^t T(t-\sigma)f(\sigma)d\sigma \\ &= \int_0^s d\sigma \int_{s-\sigma}^{t-\sigma} AT(\tau)f(\sigma)d\tau + \int_s^t T(t-\sigma)f(\sigma)d\sigma \end{aligned}$$

and estimate

$$\begin{aligned} \|v(t)-v(s)\| \le& M_{1,T}\|f\|_\infty \int_0^s d\sigma \int_{s-\sigma}^{t-\sigma} \frac{d\tau}{\tau} + M_{0,T}\|f\|_\infty (t-s) \\ \le& M_{1,T}\|f\|_\infty \int_0^s \frac{d\sigma}{(s-\sigma)^\alpha} \int_{s-\sigma}^{t-\sigma} \frac{1}{\tau^{1-\alpha}}d\tau + M_{0,T}\|f\|_\infty (t-s) \end{aligned}$$

$$\leq\left[\frac{M_{1,T}T^{1-\alpha}}{\alpha(1-\alpha)}(t-s)^{\alpha}+M_{0,T}(t-s)\right]\|f\|_{\infty}, \tag{10.17}$$

so that v is α-Hölder continuous. Combining (10.16) and (10.17), estimate (10.15) follows. □

Proposition 10.1.4 can be written, with no differences in the proof, also for classical solutions. More precisely,

Proposition 10.2.13. *Let u be the mild solution to* (10.1) *corresponding to $x \in \overline{D(A)}$ and $f \in C([0,T];X)$. Then, u is the classical solution of the Cauchy problem* (10.1) *if and only if $u \in C((0,T];D(A))$ or, equivalently, if and only if $u \in C^1((0,T];X)$.*

We can now give two sufficient conditions for the existence of a classical or strict solution to problem (10.1). For this purpose, for every Banach space Y we denote by $B([0,T];Y)$ the space of all bounded functions $f:[0,T]\to Y$.

Theorem 10.2.14 (Time regularity)**.** *Let $x \in X$, $f \in C^{\alpha}([0,T];X)$ for some $\alpha \in (0,1)$ and let u be the mild solution to problem* (10.1)*. Then, u belongs to $C^{\alpha}([\varepsilon,T];D(A)) \cap C^{1+\alpha}([\varepsilon,T];X)$, for every $\varepsilon \in (0,T)$, and the following statements hold:*

(i) *if $x \in \overline{D(A)}$, then u is the classical solution to* (10.1)*;*

(ii) *if $x \in D(A)$ and $Ax + f(0) \in \overline{D(A)}$, then u is a strict solution to* (10.1) *and there exists a positive constant C, independent of x and f, such that*

$$\|u\|_{C^1([0,T];X)}+\|u\|_{C([0,T];D(A))}\leq C(\|x\|_{D(A)}+\|f\|_{C^{\alpha}([0,T];X)}). \tag{10.18}$$

(iii) *if $x \in D(A)$ and $Ax + f(0) \in D_A(\alpha,\infty)$, then u' and Au belong to $C^{\alpha}([0,T];X)$, u' belongs to $B([0,T];D_A(\alpha,\infty))$, and there exists a positive constant C, independent of x and f, such that*

$$\begin{aligned}&\|u\|_{C^{1+\alpha}([0,T];X)}+\|Au\|_{C^{\alpha}([0,T];X)}+\|u'\|_{B([0,T];D_A(\alpha,\infty))}\\ \leq\, &C(\|f\|_{C^{\alpha}([0,T];X)}+\|x\|_{D(A)}+\|Ax+f(0)\|_{D_A(\alpha,\infty)}).\end{aligned} \tag{10.19}$$

Proof. To begin with, for every $t \in [0,T]$ we split $u(t) = u_1(t) + u_2(t)$, where

$$u_1(t)=\int_0^t T(t-s)(f(s)-f(t))ds, \qquad u_2(t)=T(t)x+\int_0^t T(t-s)f(t)ds.$$

Both $u_1(t)$ and $u_2(t)$ belong to $D(A)$ for every $t \in (0,T]$. Indeed, $u_2(t)$ belongs to $D(A)$ in view of Proposition 9.2.7. On the other hand, since f is α-Hölder continuous in $[0,T]$, we can estimate $\|AT(t-s)(f(s)-f(t))\| \leq M_{1,T}(t-s)^{\alpha-1}[f]_{C^{\alpha}([0,T];X)}$ for every $s \in [0,t)$, so that the function $s \mapsto T(t-s)(f(s)-f(t))$ is integrable with values in $D(A)$. This is enough to infer that $u_1(t) \in D(A)$ for every $t \in [0,T]$.

Further, we observe that

$$Au_1(t)=\int_0^t AT(t-s)(f(s)-f(t))ds, \qquad Au_2(t)=AT(t)x+(T(t)-I)f(t)$$

for every $t \in (0,T]$ (actually, the first formula holds true also with $t=0$).

Analysis of the function Au_1. Clearly, we can estimate

$$\|Au_1(t)\| \le \alpha^{-1} M_{1,T} T^{\alpha} [f]_{C^\alpha([0,T];X)}, \qquad t \in [0,T]. \tag{10.20}$$

Moreover, for $s, t \in [0,T]$, with $s < t$, we can write

$$\begin{aligned}
Au_1(t) - Au_1(s) &= \int_0^s (AT(t-\sigma) - AT(s-\sigma))\,(f(\sigma) - f(s))d\sigma \\
&\quad + \int_0^s AT(t-\sigma)(f(s) - f(t))d\sigma + \int_s^t AT(t-\sigma)(f(\sigma) - f(t))d\sigma \\
&= \int_0^s d\sigma \int_{s-\sigma}^{t-\sigma} A^2 T(\tau)(f(\sigma) - f(s))d\tau \\
&\quad + (T(t) - T(t-s))(f(s) - f(t)) + \int_s^t AT(t-\sigma)(f(\sigma) - f(t))d\sigma,
\end{aligned}$$

so that

$$\begin{aligned}
\|Au_1(t) - Au_1(s)\| \le & M_{2,T}[f]_{C^\alpha([0,T];X)} \int_0^s (s-\sigma)^\alpha d\sigma \int_{s-\sigma}^{t-\sigma} \tau^{-2} d\tau \\
&+ 2M_{0,T}[f]_{C^\alpha([0,T];X)}(t-s)^\alpha + M_{1,T}[f]_{C^\alpha([0,T];X)} \int_s^t (t-\sigma)^{\alpha-1} d\sigma \\
\le & M_{2,T}[f]_{C^\alpha([0,T];X)} \int_0^s d\sigma \int_{s-\sigma}^{t-\sigma} \tau^{\alpha-2} d\tau \\
&+ (2M_{0,T} + M_{1,T}\alpha^{-1})[f]_{C^\alpha([0,T];X)}(t-s)^\alpha \\
\le & \left(\frac{M_{2,T}}{\alpha(1-\alpha)} + 2M_{0,T} + \frac{M_{1,T}}{\alpha} \right) [f]_{C^\alpha([0,T];X)}(t-s)^\alpha. \qquad (10.21)
\end{aligned}$$

Then, Au_1 is α-Hölder continuous in $[0,T]$. From estimates (10.20) and (10.21) it follows that

$$\|Au_1\|_{C^\alpha([0,T];X)} \le C_1 \|f\|_{C^\alpha([0,T];X)} \tag{10.22}$$

for some positive constant C_1, independent of f.

Finally, we prove that Au_1 is bounded in $[0,T]$ with values in $D_A(\alpha,\infty)$. For this purpose, we fix $t \in (0,T]$ and observe that

$$\begin{aligned}
\sup_{\xi\in(0,1]} \|\xi^{1-\alpha} AT(\xi) Au_1(t)\| \le & \sup_{\xi\in(0,1]} \xi^{1-\alpha} \left\| \int_0^t A^2 T(t+\xi-s)(f(s) - f(t))ds \right\| \\
\le & M_{2,T+1}[f]_{C^\alpha([0,T];X)} \sup_{\xi\in(0,1]} \xi^{1-\alpha} \int_0^t (t-s)^\alpha (t+\xi-s)^{-2} ds \\
\le & M_{2,T+1}[f]_{C^\alpha([0,T];X)} \int_0^\infty \sigma^\alpha (\sigma+1)^{-2} d\sigma.
\end{aligned}$$

Hence, $Au_1(t)$ belongs to $D_A(\alpha,\infty)$ and there exists a positive constant C_2, independent of t and f, such that

$$\|Au_1(t)\|_{D_A(\alpha,\infty)} \le C_2 \|f\|_{C^\alpha([0,T];X)}, \qquad t \in [0,T]. \tag{10.23}$$

Analysis of the function Au_2. To begin with, we prove that Au_2 is α-Hölder continuous in $[\varepsilon,T]$ with values in X for every $\varepsilon\in(0,T)$. For this purpose, we fix such an $\varepsilon\in(0,T)$. Clearly, the function $t\mapsto AT(t)x$ is α-Hölder continuous in $[\varepsilon,T]$ since it is differentiable in $(0,\infty)$, due to Theorem 9.2.2. Further,

$$
\begin{aligned}
\|(T(t)-I)f(t)-(T(s)-I)f(s)\| \le& \|T(t)f(t)-T(t)f(s)\|+\|T(t)f(s)-T(s)f(s)\|\\
&+\|f(s)-f(t)\|\\
\le& M_{0,T}[f]_{C^\alpha([0,T];X)}|t-s|^\alpha+M_{1,T}\|f\|_\infty\varepsilon^{-\alpha}\int_s^t\sigma^{\alpha-1}d\sigma\\
&+[f]_{C^\alpha([0,T];X)}|t-s|^\alpha\\
\le& \big(M_{0,T}[f]_{C^\alpha([0,T];X)}+\alpha^{-1}M_{1,T}\|f\|_\infty\varepsilon^{-\alpha}+1\big)(t-s)^\alpha
\end{aligned}
$$

for every $s,t\in[\varepsilon,T]$, with $s<t$. Hence, Au_2 is α-Hölder continuous in $[\varepsilon,T]$.

Let us now suppose that $x\in D(A)$ with $Ax+f(0)\in\overline{D(A)}$. Then, we can split $Au_2(t)=T(t)(Ax+f(0))+T(t)(f(t)-f(0))-f(t)$ for every $t\in[0,T]$. From the previous formula, it follows immediately that the function Au_2 is continuous also at $t=0$. Moreover,

$$
\|Au_2\|_{C([0,T];X)}\le M_{0,T}(\|Ax+f(0)\|+2\|f\|_{C([0,T];X)}). \tag{10.24}
$$

Next, we observe that the function $t\mapsto T(t)(f(t)-f(0))$ is α-Hölder continuous in $[0,T]$ with values in X. Indeed,

$$
\begin{aligned}
&\|T(t)(f(t)-f(0))-T(s)(f(s)-f(0))\|\\
\le&\|(T(t)-T(s))(f(s)-f(0))\|+\|T(t)(f(t)-f(s))\|\\
\le&M_{1,T}[f]_{C^\alpha([0,T];X)}s^\alpha\int_s^t\frac{d\sigma}{\sigma}+M_{0,T}[f]_{C^\alpha([0,T];X)}(t-s)^\alpha\\
\le&M_{1,T}[f]_{C^\alpha([0,T];X)}\int_s^t\sigma^{\alpha-1}d\sigma+M_{0,T}[f]_{C^\alpha([0,T];X)}(t-s)^\alpha\\
\le&\bigg(\frac{M_{1,T}}{\alpha}+M_{0,T}\bigg)(t-s)^\alpha[f]_{C^\alpha([0,T];X)}
\end{aligned}
$$

for every $s,t\in[0,T]$ such that $s\le t$. Hence, Au_2 is α-Hölder continuous in $[0,T]$ if and only if $Ax+f(0)\in D_A(\alpha,\infty)$. In such a case,

$$
\|Au_2\|_{C^\alpha([0,T];X)}\le C_3(\|Ax+f(0)\|_{D_A(\alpha,\infty)}+\|f\|_{C^\alpha([0,T];X)}) \tag{10.25}
$$

for some positive constant C_3, independent of f and x, as all the other forthcoming constants.

Finally, we observe that the function Au_2+f is bounded in $[0,T]$ with values in $D_A(\alpha,\infty)$. For this purpose, we fix $t\in[0,T]$ and observe that

$$
\begin{aligned}
&\|\xi^{1-\alpha}AT(\xi)(Au_2(t)+f(t))\|\\
\le&\|\xi^{1-\alpha}AT(t+\xi)(Ax+f(0))+\|\xi^{1-\alpha}AT(t+\xi)(f(t)-f(0))\|\\
\le&M_{0,T}[Ax+f(0)]_{D_A(\alpha,\infty)}+M_{1,T+1}[f]_{C^\alpha([0,T];X)}\xi^{1-\alpha}(t+\xi)^{-1}t^\alpha\\
\le&M_{0,T}[Ax+f(0)]_{D_A(\alpha,\infty)}+M_{1,T+1}[f]_{C^\alpha([0,T];X)}
\end{aligned}
$$

for every $\xi \in (0,1]$. This shows that $Au_2(t) + f(t)$ belongs to $D_A(\alpha,\infty)$. The previous estimate, combined with (10.25) shows that

$$\sup_{t\in[0,T]} \|Au_2(t) + f(t)\|_{D_A(\alpha,\infty)} \le C_4(\|Ax + f(0)\|_{D_A(\alpha,\infty)} + \|f\|_{C^\alpha([0,T];X)}). \tag{10.26}$$

We can now complete the proof. First of all, we observe that since the function $t \mapsto T(t)x$ is differentiable in $(0,\infty)$, then it belongs to $C^\alpha([\varepsilon,T];X)$. This, combined with Proposition 10.2.12, shows that $u \in C^\alpha([\varepsilon,T];X)$ for every $\varepsilon \in (0,T)$. Moreover, the results of the first part of the proof show that $Au \in C^\alpha([\varepsilon,T];X)$ so that $u \in C^\alpha([\varepsilon,T];D(A))$. Proposition 10.2.13 allows us to conclude that u is differentiable in $(0,T]$ with values in X and $u' = Au+f$ in $(0,T]$. It thus follows that $u' \in C^\alpha([\varepsilon,T];X)$ as well.

(i) Since $x \in \overline{D(A)}$, Propositions 9.2.5 and 10.2.12 show that $u \in C([0,T];X)$ and $u(0) = x$. By this and the above results, we conclude that u is a classical solution to the Cauchy problem (10.1).

(ii) If $x \in D(A)$ and $Ax + f(0) \in \overline{D(A)}$, then the functions Au_1 and Au_2 are continuous in $[0,T]$, so that $Au \in C([0,T];X)$. Moreover, from (10.22) and (10.24) it follows that

$$\|Au\|_{C([0,T];X)} \le C_5(\|x\|_{D(A)} + \|f\|_{C^\alpha([0,T];X)}) \tag{10.27}$$

Moreover, since $D(A) \subset D_A(\alpha,\infty)$, combining Corollary 10.2.4 and Proposition 10.2.12 we conclude that $u \in C^\alpha([0,T];X)$ and

$$\|u\|_{C^\alpha([0,T];D(A))} \le C_6(\|x\|_{D(A)} + \|f\|_{C([0,T];X)}). \tag{10.28}$$

Proposition 10.1.4 implies that u is a strict solution to the Cauchy problem (10.1). Moreover, estimates (10.27) and (10.28) yield (10.18).

(iii) If $x \in D(A)$ and $Ax+f(0) \in D_A(\alpha,\infty)$, then clearly, the function u is a strict solution to problem (10.1). Moreover, the functions Au_1 and Au_2 belong to $C^\alpha([0,T];X)$ so that Au and, by difference, u' belong to $C^\alpha([0,T];X)$. Finally, since Au_1 and $Au_2 + f$ are bounded in $[0,T]$ with values in $D_A(\alpha,\infty)$, the function $u' = Au_1 + Au_2 + f$ is bounded in $[0,T]$ with values in $D_A(\alpha,\infty)$. Estimate (10.19) follows from (10.22), (10.23), (10.25), (10.26) and (10.28). □

Remark 10.2.15. The first part of the proof of Theorem 10.2.14 shows that the condition $Ax + f(0) \in D_A(\alpha,\infty)$ is also necessary for problem (10.1) to have a solution u with $Au \in C^\alpha([0,T];X)$. If this condition is satisfied, then $Au(t) + f(t)$ belongs to $D_A(\alpha,\infty)$ for every $t \in [0,T]$.

Theorem 10.2.16 (Spatial regularity)**.** *For every $f \in C([0,T];X) \cap B([0,T];D_A(\alpha,\infty))$ (for some $\alpha \in (0,1)$) and $x \in X$, the mild solution to (10.1) belongs to $C^1((0,T];X) \cap C((0,T];D(A))$ and is locally bounded in $(0,T]$ with values in $D_A(\alpha+1,\infty)$. Moreover, the following properties hold true.*

(i) *If $x \in \overline{D(A)}$, then u is the classical solution of (10.1);*

(ii) *If $x \in D(A)$, $Ax \in \overline{D(A)}$, then u is the strict solution of (10.1);*

(iii) *If $x \in D_A(\alpha+1,\infty)$, then u' and Au belong to $B([0,T];D_A(\alpha,\infty)) \cap C([0,T];X)$, Au belongs to $C^\alpha([0,T];X)$, and there exists a positive constant C, independent of x and f, such that*

$$\|u'\|_{B([0,T];D_A(\alpha,\infty))} + \|Au\|_{B([0,T];D_A(\alpha,\infty))} + \|Au\|_{C^\alpha([0,T];X)}$$

$$\leq C(\|f\|_{B([0,T];D_A(\alpha,\infty))} + \|x\|_{D_A(\alpha,\infty)}). \tag{10.29}$$

Proof. We split $u(t) = T(t)x + v(t)$ for every $t \in [0,T]$ and analyze in details the function v.

Analysis of v. For notational convenience, for every $\tau > 0$, we set

$$M_{k,\alpha,\tau} = \sup_{0<t\leq\tau} \|t^{k-\alpha}A^kT(t)\|_{\mathcal{L}(D_A(\alpha,\infty),X)}, \qquad k = 1,2.$$

Let us prove that $v \in C^\alpha([0,T];D(A))$. By Proposition 10.2.12, we already know that $v \in C^\alpha([0,T];X)$. Moreover, since $\|AT(t-s)f(s)\| \leq M_{1,\alpha,T}(t-s)^{\alpha-1}\|f\|_{B([0,T];D_A(\alpha,\infty)}$ for every $s \in (0,t)$ and every $t \in (0,T]$, it follows immediately that $v(t)$ belongs to $D(A)$ for every $t \in [0,T]$ and

$$\|Av(t)\| \leq \alpha^{-1}T^\alpha M_{1,\alpha,T}\|f\|_{B([0,T];D_A(\alpha,\infty)}. \tag{10.30}$$

Next, we observe that

$$\begin{aligned}
\|Av(t) - Av(s)\| \leq & \left\|A\int_0^s (T(t-\sigma) - T(s-\sigma))\, f(\sigma)d\sigma\right\| + \left\|A\int_s^t T(t-\sigma)f(\sigma)d\sigma\right\| \\
\leq & \|f\|_{B([0,T];D_A(\alpha,\infty))}\left(M_{2,\alpha,T}\int_0^s d\sigma\int_{s-\sigma}^{t-\sigma}\tau^{\alpha-2}d\tau + M_{1,\alpha,T}\int_s^t (t-\sigma)^{\alpha-1}d\sigma\right) \\
\leq & \left(\frac{M_{2,\alpha,T}}{\alpha(1-\alpha)} + \frac{M_{1,\alpha,T}}{\alpha}\right)(t-s)^\alpha\|f\|_{B([0,T];D_A(\alpha,\infty))}
\end{aligned} \tag{10.31}$$

for every $s,t \in [0,T]$, with $s \leq t$. Hence, Av is α-Hölder continuous in $[0,T]$. From (10.15), (10.30) and (10.31) we deduce that

$$\|v\|_{C^\alpha([0,T];D(A))} \leq C_1\|f\|_{B([0,T];D_A(\alpha,\infty))} \tag{10.32}$$

for some positive constant C_1, independent of f.

Finally, to prove that Av is bounded in $[0,T]$ with values in $D_A(\alpha,\infty)$, we fix $\xi \in (0,1]$ and estimate

$$\begin{aligned}
\|\xi^{1-\alpha}AT(\xi)Av(t)\| = & \xi^{1-\alpha}\left\|\int_0^t A^2T(t+\xi-s)f(s)ds\right\| \\
\leq & M_{2,\alpha,T+1}\|f\|_{B([0,T];D_A(\alpha,\infty))}\xi^{1-\alpha}\int_0^t (t+\xi-s)^{\alpha-2}ds \\
\leq & \frac{M_{2,\alpha,T+1}}{1-\alpha}\|f\|_{B([0,T];D_A(\alpha,\infty))}.
\end{aligned} \tag{10.33}$$

Taking the supremm with respect to $\xi \in (0,1]$, we conclude that $Av \in B([0,T];D_A(\alpha,\infty))$. Moreover, estimates (10.30) and (10.33) show that

$$\|Av\|_{B([0,T];D_A(\alpha,\infty))} \leq C_2\|f\|_{B([0,T];D_A(\alpha,\infty))} \tag{10.34}$$

for some positive constant C_2, independent of f.

We can now complete the proof. If $x \in X$, then we know from Theorem 9.2.2 that the function $t \mapsto T(t)x$ belongs to $C((0,T];D(A^2))$. Hence, the function u itself belongs to $C((0,T];D(A))$ and Proposition 10.2.13 shows that it belongs also to $C^1((0,T];X)$, Since

$D(A^2) \hookrightarrow D_A(\alpha+1,\infty)$, it also follows that u is locally bounded in $(0,T]$ with values in $D_A(\alpha+1,\infty)$.

(i) Since the function v is continuous at 0, where it vanishes, and the function $t \mapsto T(t)x$ is continuous at $t=0$, since $x \in \overline{D(A)}$ (see Proposition 9.2.5), u is continuous at $t=0$, and $u(0)=x$. Hence, u is a classical solution to problem (10.1).

(ii) If $x \in D(A)$ and $Ax \in \overline{D(A)}$, then the function $t \mapsto T(t)x$ belongs to $C([0,T];D(A))$, so that u belongs to $C([0,T];D(A))$ as well. By Proposition 10.2.13, u belongs also to $C^1([0,T];X)$ and is a strict solution to the Cauchy problem (10.1).

(iii) If $x \in D_A(\alpha+1,\infty)$, then, clearly, u is the strict solution to the equation (10.1). Moreover, the function $t \mapsto AT(t)x = T(t)Ax$ is α-Hölder continuous in $[0,T]$ due to Corollary 10.2.4. It is also bounded in $[0,T]$ with values in $D_A(\alpha,\infty)$. As a byproduct, we infer that the function $t \mapsto T(t)x$ is α-Hölder continuous in $[0,T]$ with values in $D(A)$ and bounded in $[0,T]$ with values in $D_A(\alpha+1,\infty)$. The assertion follows immediately, taking also (10.32) and (10.34) into account. □

Remark 10.2.17. The results proved in Theorems 10.2.14 and 10.2.16 can be applied also to Cauchy problems set in intervals $[a,b] \neq [0,T]$. It suffices to shift the time variable to transform the problem in an equivalent problem set in the interval $[0,b-a]$. In particular, the mild solution to the initial value problem $u(a)=x$ for the equation $u' = Au + f$ in (a,b) is given, when $f \in C([a,b];X)$, by

$$u(t) = T(t-a)x + \int_a^t T(t-s)f(s)ds, \qquad t \in [a,b].$$

10.3 Notes

It is worth observing that, by adopting the definition in Proposition 10.2.3, the interpolation spaces $D_A(\alpha,\infty)$ can be defined also in the case of strongly continuous semigroups. In this case, they are also known as Favard spaces (see [6, Chapter II, 5.b]).

10.4 Exercises

Exercise 10.4.1. Show that, for every $\alpha \in (0,\infty) \setminus \mathbb{N}$, $D_A(\alpha,\infty)$ is a Banach space.

Exercise 10.4.2. Fix $\alpha \in (0,1)$.

(a) Prove that $x \in D_A(\alpha,\infty)$ if and only if $[x]'_\alpha = \sup_{t\in(0,\tau]} \|t^{1-\alpha}AT(t)x\| < \infty$ for some $\tau > 0$, and that the norm of $D_A(\alpha,\infty)$ is equivalent to the norm $x \mapsto \|x\| + [x]'_{D_A(\alpha,\infty)}$.

(b) Prove that, if $\omega < 0$ in Definition 9.2.1 then $D_A(\alpha,\infty)$ is the set of $x \in X$ such that $\|x\|_\alpha = \sup_{t>0} \|t^{1-\alpha}AT(t)x\| < \infty\}$, and that $x \mapsto \|x\|_\alpha$ is an equivalent norm in $D_A(\alpha,\infty)$. What about $\omega = 0$?

Exercise 10.4.3. Fix $T > 1$ and $x \in D_A(\alpha,\infty)$ for some $\alpha \in (0,1)$. Prove that the function $t \mapsto T(t)x$ belongs to $C^\alpha([0,T];X)$ and there exists a positive constant C, independent of x, such that $\|T(\cdot)x\|_{C^\alpha([0,T])} \le C\|x\|_{D_A(\alpha,\infty)}$.

Exercise 10.4.4. Show that the closure of $D(A)$ in $D_A(\alpha,\infty)$ is the subspace of all $x \in X$ such that $\lim_{t\to 0} t^{1-\alpha}AT(t)x = 0$. This implies that, even if $D(A)$ is dense in X, it is not necessarily dense in $D_A(\alpha,\infty)$.
[**Hint:** to prove that $T(t)x - x$ tends to zero in $D_A(\alpha,\infty)$ provided $t^{1-\alpha}AT(t)x$ vanishes as t tends to 0, split the sup over $(0,1]$ in the definition of $[\,\cdot\,]_{D_A(\alpha,\infty)}$ into the sup over $(0,\varepsilon]$ and over $[\varepsilon,1]$, for ε small].

Exercise 10.4.5. Prove that $D(A)$ is of class $J_{1/2}$ between X and $D(A^2)$.
[**Hint:** If $\omega = 0$, then use formula (9.12) to get $\|Ax\| \le M_1\|x\|/t + Mt\|A^2x\|$ for each $t > 0$ and then take the minimum for $t \in (0,\infty)$. If $\omega > 0$, then replace A by $A - \omega I$...].

Exercise 10.4.6. (a) Following the proof of Proposition 10.2.5, show that $D_A(\alpha,\infty)$ is of class $J_{\alpha/\theta}$ between X and $D_A(\theta,\infty)$, for every $0 < \alpha < \theta < 1$.

(b) Show that any space of class J_α between X and $D(A)$ is of class $J_{\alpha/\theta}$ between X and $D_A(\theta,\infty)$, for every $\theta \in (\alpha, 1)$.

(c) Using (a), prove that every function which is continuous with values in X and bounded with values in $D_A(\theta,\infty)$ in an interval $[a,b]$, is also continuous with values in $D_A(\alpha,\infty)$ in $[a,b]$, for $\alpha < \theta$.

Exercise 10.4.7. Let $f \in C_b((0,T);X)$ (the set of bounded and continuous functions $f : (0,T) \to X$, endowed with the sup-norm), set $v = (T(\cdot) * f)$ and let X_α be a space of class J_α between X and $D(A)$ ($\alpha \in (0,1)$). Using estimates (10.11) and (10.12), and the technique of Proposition 10.2.12 prove that

(a) $v \in B([0,T];X_\alpha)$ and $\|u\|_{B([0,T];X_\alpha)} \le C_1(\alpha)\|f\|_\infty$;

(b) $v \in C^{1-\alpha}([0,T];X_\alpha)$ and $\|v\|_{C^\alpha([0,T];X_\alpha)} \le C_2(\alpha)\|f\|_\infty$.

Exercise 10.4.8. Let $\alpha \in (0,1)$ and $a < b \in \mathbb{R}$. Prove that if a function u belongs to $C^{1+\alpha}([a,b];X) \cap C^\alpha([a,b];D(A))$ then u' is bounded in $[a,b]$ with values in $D_A(\alpha,\infty)$.
[**Hint:** set $u_0 = u(a)$, $f(t) = u'(t) - Au(t)$, and use Theorem 10.2.14(iii) and Remark 10.2.15].

Exercise 10.4.9. For $p \in [1,\infty)$, consider the sectorial operator $A_p : D(A_p) = \{(x_n) \in \ell^p : (nx_n) \in \ell^p\} \to \ell^p$, defined by $A_p(x_n) = -(nx_n)$ for every $(x_n) \in D(A_p)$. Prove that there exists at least a function $f \in C([0,T];\ell^p)$ such that the mild solution v of (10.1) corresponding to the initial value $x = 0$ is not a strict solution.
[**Hint**: By contradiction, suppose that for every $f \in C([0,1];\ell^p)$ the mild solution to the Cauchy problem (10.1) is a strict solution.

(i) Use the closed graph theorem to show that the linear operator $S : C([0,1];\ell^p) \to C([0,1];D(A_p))$, defined by $S(t)f = \displaystyle\int_0^t T(t-s)f(s)ds$ for every $t \in [0,1]$ and $f \in C([0,1];\ell^p)$, where $\{T(t)\}$ is the analytic semigroup associated with operator A_p, is bounded.

(ii) Let (e_n) be the canonical basis of ℓ^p and consider a nonzero continuous function $g : [0,\infty) \to [0,1]$ with support contained in $[1/2,1]$. For every $n \in \mathbb{N}$ and $t \in [0,1]$, set $f_n(t) = g(2^n(1-t))e_{2^n}$. Prove that $f_n \in C([0,1];\ell^p)$ and $\|f_n\|_\infty \le 1$. Further, set $h_N = f_1 + \cdots + f_N$ and prove that $h_N \in C([0,1];\ell_p)$ with $\|h_N\|_\infty \le 1$.

(iii) Show that $S(1)f_n = c2^{-n}e_{2^n}$ where $c = \int_0^\infty e^{-s}g(s)ds$ and deduce that $\|S(1)h_N\|_{D(A_p)} \ge cN^{1/p}$ (this implies that $S(1)$ is unbounded, contradicting (i)).]

Bibliography

[1] A. Belleni-Morante, Applied Semigroups and Evolution Equations, Oxford University Press, 1979.

[2] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[3] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[4] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[5] K.-J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Graduate Text in Mathematics **194**, Springer-Verlag, 2000.

[6] K.-J. Engel and R. Nagel, Short Course on Operator Semigroups, Universitext, Springer-Verlag, 2006.

[7] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[8] J.A. Goldstein, Semigroups of Operators and Applications, Oxford University Press, 1985.

[9] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[10] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[11] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981.

[12] L. Lorenzi and A. Rhandi, Semigroups of Bounded Operators and Second-Order Elliptic and Parabolic Partial Differential Equations, CRC Press, Taylor & Francis Group, 2021.

[13] A. Lunardi, Analytic Semigroups and Optimal Regularity in Parabolic Problems, Birkhäuser Verlag, 1995.

[14] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Appl. Math.Sci. **44**, Springer-Verlag,1983.

[15] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[16] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[17] W. Rudin, Real and Complex Analysis, 3rd ed., McGraw-Hill, 1986.

[18] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[19] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958.

[20] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861.

[21] A. Zygmund, Trigonometric Series, Cambridge Univ. Press., 2nd Edition Reprinted (1968).

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

January 19, 2022

Lecture 11

A remarkable example: the Gauss-Weierstrass semigroup

In this lecture, we discuss a notable example of analytic semigroup: the so-called Gauss-Weierstrass semigroup, which is naturally associated to the heat equation. We confine our analysis to the space $C_b(\mathbb{R}^d;\mathbb{C})$ of all bounded and continuous complex-valued functions over $\mathbb{R}^d$. We stress that the Gauss-Weierstrass semigroup is analytic also in the usual L^p-spaces over $\mathbb{R}^d$ for $p\in[1,\infty)$.

11.1 Heat equation

In this section, we consider the homogeneous Cauchy problem

$$\begin{cases} D_t u(t,x)=\Delta u(t,x), & t>0, \quad x\in\mathbb{R}^d,\\ u(0,x)=f(x), & \qquad\qquad x\in\mathbb{R}^d,\end{cases} \tag{11.1}$$

where Δu denotes the trace of the Hessian matrix of u, i.e.,

$$\Delta u=\sum_{j=1}^d \frac{\partial^2 u}{\partial x_j^2}.$$

The Laplacian Δ is the prototype of a uniformly elliptic operator with bounded coefficients in $\mathbb{R}^d$ and it has the peculiarity that, for every $f\in C_b(\mathbb{R}^d;\mathbb{C})$, there exists an explicit formula for the classical solution (see the forthcoming Definition 11.1.1).

In this lecture we will use also the space $C_b^\alpha(\mathbb{R}^d;\mathbb{C})$ ($\alpha>0$) of all the functions $f:\mathbb{R}^d\to\mathbb{C}$ which are bounded and admit bounded derivatives up to the order $[\alpha]$ and their derivatives of order $[\alpha]$ are $(\alpha-[\alpha])$-Hölder continuous in $\mathbb{R}^d$ (if $\alpha\notin\mathbb{N}$). This is a Banach space when endowed with the norm

$$\|f\|_{C_b^\alpha(\mathbb{R}^d;\mathbb{C})}=\sum_{|\beta|\le[\alpha]}\left\|\frac{\partial^\beta f}{\partial x^\beta}\right\|_\infty+\sum_{|\beta|=[\alpha]}\left[\frac{\partial^\beta f}{\partial x^\beta}\right]_{C^{\alpha-[\alpha]}(\mathbb{R}^d;\mathbb{C})},$$

where $[g]_{C^{\alpha-[\alpha]}(\mathbb{R}^d;\mathbb{C})}=\sup\left\{\frac{|g(x)-g(y)|}{|x-y|^{\alpha-[\alpha]}}:x,y\in\mathbb{R}^d,\ x\neq y\right\}$.

Following the definition of classical solution of abstract Cauchy problems introduced in Lecture 10, we give the following definition.

Definition 11.1.1. A function $u : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ is a classical solution to problem (11.1) if (i) $u \in C([0,\infty) \times \mathbb{R}^d; \mathbb{C})$, (ii) it is continuously differentiable in $(0,\infty) \times \mathbb{R}^d$, once with respect to the time variable and twice with respect to the spatial variables, (iii) it satisfies the differential equation and the initial condition in (11.1).

We now introduce the so-called *fundamental solution* to the differential equation in (11.1), i.e., the function $K : (0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ defined by

$$K(t,x) = (4\pi t)^{-\frac{d}{2}} e^{-\frac{|x|^2}{4t}}, \qquad t > 0,\ \ x \in \mathbb{R}^d, \tag{11.2}$$

and study its main properties.

Lemma 11.1.2. *The following properties are satisfied.*

(i) $K \in C^\infty((0,\infty) \times \mathbb{R}^d; \mathbb{R})$;

(ii) $\displaystyle\int_{\mathbb{R}^d} K(t,x)dx = 1$ *for all* $t > 0$;

(iii) $D_j K(t,x) = -\dfrac{x_j}{2t} K(t,x)$ *for all* $(t,x) \in (0,\infty) \times \mathbb{R}^d$;

(iv) $D_{ij} K(t,x) = \left(\dfrac{x_i x_j}{4t^2} - \dfrac{\delta_{ij}}{2t}\right) K(t,x)$ *for all* $(t,x) \in (0,\infty) \times \mathbb{R}^d$;

(v) $D_t K(t,x) = \Delta K(t,x)$ *for all* $(t,x) \in (0,\infty) \times \mathbb{R}^d$.

Proof. We limit ourselves to proving property (v), since the remaining ones are straightforward to prove.

From property (iv) it follows that

$$\Delta K(t,x) = \left(\frac{|x|^2}{4t^2} - \frac{d}{2t}\right) K(t,x), \qquad (t,x) \in (0,\infty) \times \mathbb{R}^d.$$

On the other hand,

$$D_t K(t,x) = -d(4\pi t)^{-\frac{d}{2}-1} 2\pi e^{-\frac{|x|^2}{4t}} + (4\pi t)^{-\frac{d}{2}} \frac{|x|^2}{4t^2} e^{-\frac{|x|^2}{4t}} = \left(\frac{|x|^2}{4t^2} - \frac{d}{2t}\right) K(t,x)$$

for all $(t,x) \in (0,\infty) \times \mathbb{R}^d$. Property (v) follows at once. □

Now, we can prove the existence and uniqueness of a classical solution to the Cauchy problem (11.1).

Theorem 11.1.3. *For each* $f \in C_b(\mathbb{R}^d; \mathbb{C})$, *the Cauchy problem* (11.1) *has a unique classical solution* u *given by the following formula:*

$$u(t,x) = \begin{cases} f(x), & t = 0, \quad x \in \mathbb{R}^d, \\ \displaystyle\int_{\mathbb{R}^d} K(t, x-y) f(y) dy, & t > 0, \quad x \in \mathbb{R}^d. \end{cases} \tag{11.3}$$

Moreover, $\|u(t,\cdot)\|_\infty \le \|f\|_\infty$ *for all* $t \ge 0$.

Proof. To begin with, we observe that, if u is given by (11.3), then

$$|u(t,x)| \leq \|f\|_\infty \int_{\mathbb{R}^d} K(t,x-y)dy = \|f\|_\infty \int_{\mathbb{R}^d} K(t,y)dy = \|f\|_\infty, \qquad (t,x) \in (0,\infty)\times\mathbb{R}^d,$$

by property (ii) in Lemma 11.1.2. It thus follows that $\|u(t,\cdot)\|_\infty \leq \|f\|_\infty$ for all $t \geq 0$ as claimed.

Next, we observe that, since $K \in C^\infty((0,\infty)\times\mathbb{R}^d;\mathbb{R})$ (see Lemma 11.1.2(i)), by the dominated convergence theorem we conclude that $u \in C^\infty((0,\infty)\times\mathbb{R}^d;\mathbb{C})$, see Exercise 11.4.1, and

$$D_t u(t,x) - \Delta u(t,x) = \int_{\mathbb{R}^d} (D_t K(t,x-y) - \Delta K(t,x-y))f(y)dy = 0$$

for every $(t,x) \in (0,\infty)\times\mathbb{R}^d$, thanks to Lemma 11.1.2(v).

To prove that u is a classical solution to problem (11.1) we need to show that u is continuous on $\{0\}\times\mathbb{R}^d$, where it equals the function f. For this purpose, we fix $t>0$ and, using the fact that $\int_{\mathbb{R}^d} K(t,x)\,dx = 1$ for all $t>0$, we split

$$\begin{aligned}
|u(t,x) - f(x_0)| =& \left| \int_{\mathbb{R}^d} K(t,y)f(x-y)\,dy - \int_{\mathbb{R}^d} K(t,y)f(x_0)\,dy \right| \\
\leq& \int_{\mathbb{R}^d} K(t,y)|f(x-y) - f(x_0)|\,dy \\
\leq& \int_{\mathbb{R}^d} K(t,y)|f(x-y) - f(x)|\,dy + |f(x) - f(x_0)|
\end{aligned}$$

for $x, x_0 \in \mathbb{R}^d$. By the change of variable $z = y/\sqrt{t}$ we can rewrite

$$\begin{aligned}
\int_{\mathbb{R}^d} K(t,y)|f(x-y) - f(x)|dy =& (4\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|z|^2}{4}} |f(x-\sqrt{t}z) - f(x)|dz \\
=& (4\pi)^{-\frac{d}{2}} \int_{B(0,r)} e^{-\frac{|z|^2}{4}} |f(x-\sqrt{t}z) - f(x)|dz \\
&+ (4\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d\setminus B(0,r)} e^{-\frac{|z|^2}{4}} |f(x-\sqrt{t}z) - f(x)|dz \\
\leq& (4\pi)^{-\frac{d}{2}} \int_{B(0,r)} e^{-\frac{|z|^2}{4}} |f(x-\sqrt{t}z) - f(x)|dz \\
&+ 2(4\pi)^{-\frac{d}{2}} \|f\|_\infty \int_{\mathbb{R}^d\setminus B(0,r)} e^{-\frac{|z|^2}{4}} dz
\end{aligned} \tag{11.4}$$

for each $r>0$. Fix $\varepsilon>0$. By the absolutely continuity of the measure $e^{-|z|^2/4}dz$, we can find out $r_0>0$ such that

$$2(4\pi)^{-\frac{d}{2}} \|f\|_\infty \int_{\mathbb{R}^d\setminus B(0,r_0)} e^{-\frac{|z|^2}{4}} dz \leq \frac{\varepsilon}{2}.$$

As far as the first term in the last side of (11.4) is concerned, we notice that since f is continuous in $\mathbb{R}^d$, it is uniformly continuous in each compact set K of $\mathbb{R}^d$. Hence, if $x \in B(x_0,1)$, $z \in B(0,r_0)$ and $t<1$, then both $x-\sqrt{t}z$ and x belong to $K = \overline{B(0,M)}$

where $M = |x_0| + 1 + r_0$. Let $\delta > 0$ be such that $|f(z_2) - f(z_1)| \le \varepsilon/2$ if $z_1, z_2 \in K$ satisfy $|z_2 - z_1| \le \delta$. If $t \le t_0 := \delta^2 r_0^{-2}$, then $|f(x - \sqrt{t}z) - f(x)| \le \varepsilon/2$ for all $x \in B(x_0, 1)$, so that

$$(4\pi)^{-\frac{d}{2}} \int_{B(0,r_0)} e^{-\frac{|z|^2}{4}} |f(x - \sqrt{t}z) - f(x)|\, dz \le \frac{\varepsilon}{2}(4\pi)^{-\frac{d}{2}} \int_{B(0,r_0)} e^{-\frac{|z|^2}{4}} dz$$
$$\le \frac{\varepsilon}{2}(4\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|z|^2}{4}} dz = \frac{\varepsilon}{2}$$

and we conclude that $|u(t,x) - f(x_0)| \le \varepsilon + |f(x) - f(x_0)|$ for every $(t,x) \in (0, t_0] \times B(x_0, 1)$. It thus follows that

$$\limsup_{(t,x)\to(0,x_0)} |u(t,x) - f(x_0)| \le \varepsilon$$

and the arbitrariness of $\varepsilon > 0$ implies that $u(t,x)$ tends to $f(x_0)$ as (t,x) tends to $(0, x_0)$.

To complete the proof, we show the uniqueness of the classical solution to the Cauchy problem (11.1). For this purpose, we introduce the "*Lyapunov*" function $\varphi : \mathbb{R}^d \to \mathbb{R}$, defined by $\varphi(x) = 1 + |x|^2$ for every $x \in \mathbb{R}^d$. As it is easily seen, $\Delta\varphi(x) = 2d \le 2d\varphi(x)$ for every $x \in \mathbb{R}^d$. Suppose that v is another classical solution to the Cauchy problem (11.1). Then, the function $w = u - v$ satisfies the heat equation $D_t w = \Delta w$ in $(0, \infty) \times \mathbb{R}^d$ and it identically vanishes at $t = 0$. Let us prove that w identically vanishes on $[0, \infty) \times \mathbb{R}^d$. Separating real and imaginary parts of w, we can assume that w is a real-valued function. For every $n \in \mathbb{N}$, we consider the function $\zeta_n : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ defined by $\zeta_n(t,x) = e^{-2dt} w(t,x) - n^{-1}\varphi(x)$ for every $(t,x) \in [0, \infty) \times \mathbb{R}^d$. This function is negative on $\{0\} \times \mathbb{R}^d$. Moreover, it satisfies the differential inequality $D_t\zeta_n \le \Delta\zeta_n - 2d\zeta_n$ in $(0, \infty) \times \mathbb{R}^d$. Fix $T > 0$ and observe that, since $\varphi(x)$ tends to ∞ as $|x|$ tends to ∞, it follows that $\sup_{t\in[0,T]} \zeta_n(t,x)$ tends to $-\infty$ as $|x|$ tends to ∞. Hence, ζ_n admits maximum at some point $(t,x) \in [0,T] \times \mathbb{R}^d$. Denote by (t_n, x_n) a point where ζ_n achieves it maximum value. If $t_n = 0$, then ζ_n is negative on $[0,T] \times \mathbb{R}^d$. Suppose that $t_n > 0$. Then, the Hessian matrix of ζ_n is negative definite at (t_n, x_n), so that $\Delta\zeta_n(t_n, x_n) \le 0$. Since t_n is a maximum point of the function $\zeta_n(\cdot, x_n)$, it follows that $D_t\zeta_n(t_n, x_n) \ge 0$. Hence, $0 \le D_t\zeta_n(t_n, x_n) \le \Delta\zeta_n(t_n, x_n) - 2d\zeta_n(t_n, x_n) \le -2d\zeta_n(t_n, x_n)$ and, again, we conclude that $\zeta_n(t_n, x_n) \le 0$ so that ζ_n is non positive in $[0,T] \times \mathbb{R}^d$. Letting n tend to ∞, we deduce that $w \le 0$ in $[0,T] \times \mathbb{R}^d$. The same arguments applied to the function $-w$ yields that $-w \le 0$ in $[0,T] \times \mathbb{R}^d$. It follows that w identically vanishes in $[0,T] \times \mathbb{R}^d$. □

Remark 11.1.4. Up to now we have just verified that formula (11.3) defines the (unique) classical solution to the Cauchy problem (11.1). A natural question is how formula (11.3) can be derived. The (formal) answer is based on the use of the Fourier transform. For each $f \in L^1(\mathbb{R}^d; \mathbb{C}) \cap L^2(\mathbb{R}^d; \mathbb{C})$, the Fourier transform of f is the function $\mathcal{F}(f)$ defined by

$$\mathcal{F}(f)(\xi) := (2\pi)^{-\frac{d}{2}} \int_{-\infty}^{\infty} \mathrm{e}^{-i\langle \xi, x\rangle} f(x) dx, \qquad \xi \in \mathbb{R}^d,$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product of $\mathbb{R}^d$. By Plancherel's theorem, the Fourier transform can be extended to a bounded linear operator on $L^2(\mathbb{R}^d; \mathbb{C})$. We recall that

$$\mathcal{F}(D_j f)(\xi) = i\xi_j \mathcal{F}(f)(\xi), \qquad \xi \in \mathbb{R}^d, \ j = 1, \ldots, d,$$

for every smooth enough function f. If we take the Fourier transform of both the sides of the equation $D_t u = \Delta u$ with respect to x and interchange the actions of $\mathcal{F}$ and the time derivative, then we deduce that the function $\widehat{u}$, defined by $\widehat{u}(t,\xi)=(\mathcal{F}(u(t,\cdot)))(\xi)$ for every $t\geq 0$ and $\xi\in\mathbb{R}^d$, solves the Cauchy problem

$$\begin{cases} D_t\widehat{u}(t,\xi)=-|\xi|^2\widehat{u}(t,\xi), & (t,\xi)\in(0,\infty)\times\mathbb{R}^d,\\ \widehat{u}(0,\xi)=\widehat{f}(\xi), & \xi\in\mathbb{R}^d.\end{cases}$$

This is a Cauchy problem for an ordinary differential equation, since ξ plays the role of a parameter, and it is easy to see that

$$\widehat{u}(t,\xi)=\mathrm{e}^{-t|\xi|^2}\widehat{f}(\xi),\qquad t\geq 0,\ \ \xi\in\mathbb{R}^d. \tag{11.5}$$

To get back to u, we take the inverse Fourier transform of the right-hand side of (11.5), recalling that the inverse Fourier transform of the product of two functions is the convolution of the inverse Fourier transforms of the two factors, and the inverse Fourier transform of the function $\xi\mapsto \mathrm{e}^{-t|\xi|^2}$ is the function $x\mapsto(4\pi t)^{-d/2}\mathrm{e}^{-|x|^2/(4t)}$ for $t>0$.

The results in Theorem 11.1.3 can be rephrased in the semigroup language. More precisely, if for each $f\in C_b(\mathbb{R}^d;\mathbb{C})$ and $t\geq 0$ we set $T(t)f=u(t,\cdot)$, where u is the unique classical solution to problem (11.1), we define a semigroup of bounded linear operators in $C_b(\mathbb{R}^d;\mathbb{C})$, usually called the *Gauss-Weierstrass* semigroup. Indeed, the uniqueness of the classical solution to problem (11.1) for each $f\in C_b(\mathbb{R}^d;\mathbb{C})$ shows that each operator $T(t)$ is linear on $C_b(\mathbb{R}^d;\mathbb{C})$. It also implies the semigroup rule, i.e.

$$T(t+s)f=T(t)T(s)f,\qquad f\in C_b(\mathbb{R}^d;\mathbb{C}),\ \ t,s\geq 0.$$

In other terms the value at time t of the classical solution of problem (11.1) with initial condition $T(s)f$ at time zero coincides with the value at time $t+s$ of the classical solution to problem (11.1) with f as initial condition at time zero. Moreover, for $t>0$ and $x\in\mathbb{R}^d$,

$$(T(t)f)(x)=(4\pi t)^{-\frac{d}{2}}\int_{\mathbb{R}^d}K(t,x-y)f(y)dy,\qquad t>0,\ \ x\in\mathbb{R}^d. \tag{11.6}$$

Remark 11.1.5. The Gauss-Weierstrass semigroup is not strongly continuous. It turns out that $T(t)f$ converges to f in $C_b(\mathbb{R}^d;\mathbb{C})$ as $t\to 0^+$ if and only if $f\in BUC(\mathbb{R}^d;\mathbb{C})$. Indeed, adapting the arguments in the last part of the proof of Theorem 11.1.3, we can easily show that, if f is uniformly continuous in $\mathbb{R}^d$, then $T(t)f$ converges uniformly in $\mathbb{R}^d$ to f as t tends to 0^+.

To prove the converse, we observe that, for every $t>0$, the function $T(t)f$ is continuously differentiable in $\mathbb{R}^d$ with bounded partial derivatives (see Exercise 11.4.1) Hence, if $T(t)f$ converges to f in $C_b(\mathbb{R}^d;\mathbb{C})$ as $t\to 0^+$, then f belongs to the closure of $C^1_b(\mathbb{R}^d;\mathbb{C})$ in $C_b(\mathbb{R}^d;\mathbb{C})$, which is $BUC(\mathbb{R}^d;\mathbb{C})$. See Exercise 11.4.2.

On the other hand, as the following theorem shows $\{T(t)\}$ is an analytic semigroup.

Theorem 11.1.6. *The Gauss-Weierstrass semigroup is analytic in* $C_b(\mathbb{R}^d;\mathbb{C})$.

In the proof of Theorem 11.1.6 we will take advantage of the following lemmata.

Lemma 11.1.7. *Let $\Omega\subset\mathbb{C}$ be an open set and let X be a complex Banach space. Further, let $\{F(\lambda):\lambda\in\Omega\}\subset\mathcal{L}(X)$ be a family of linear operators satisfying the resolvent identity*

$$F(\lambda)-F(\mu)=(\mu-\lambda)F(\lambda)F(\mu),\qquad \lambda,\mu\in\Omega.$$

If the operator $F(\lambda_0)$ is injective for some $\lambda_0\in\Omega$, then there exists a closed linear operator $A:D(A)\subset X\to X$ such that $\rho(A)$ contains Ω, and $R(\lambda,A)=F(\lambda)$ for each $\lambda\in\Omega$.

Proof. Fix $\lambda_0 \in \Omega$, and consider the operator $A : D(A) = \text{Range } F(\lambda_0) \to X$ defined by $Ax = \lambda_0 x - F(\lambda_0)^{-1}x$ for every $x \in D(A)$. For $\lambda \in \Omega$ and $y \in X$ the resolvent equation $\lambda x - Ax = y$ is equivalent to $(\lambda - \lambda_0)x + F(\lambda_0)^{-1}x = y$. Applying $F(\lambda)$ to both the sides of the previous equation, we obtain $(\lambda - \lambda_0)F(\lambda)x + F(\lambda)F(\lambda_0)^{-1}x = F(\lambda)y$, and using the resolvent identity it is easily seen that

$$F(\lambda)F(\lambda_0)^{-1} = F(\lambda_0)^{-1}F(\lambda) = (\lambda_0 - \lambda)F(\lambda) + I.$$

Hence, if x is solution of the resolvent equation, then $x = F(\lambda)y$. Let us check that $x = F(\lambda)y$ is actually a solution. In fact, $(\lambda - \lambda_0)F(\lambda)y + F(\lambda_0)^{-1}F(\lambda)y = y$, and therefore λ belongs to $\rho(A)$ and the equality $R(\lambda, A) = F(\lambda)$ holds. Furthermore, since $\{F(\lambda) : \lambda \in \Omega\} \subset \mathcal{L}(X)$, it follows that $\Omega \subseteq \rho(A)$, and the closedness of A follows from Remark 1.3.3. □

Lemma 11.1.8. *Let $I \subset \mathbb{R}$ be an interval and let $\psi : I \times \mathbb{R}^d \to \mathbb{C}$ be a continuous function such that $\|\psi(t, \cdot)\|_\infty \leq g(t)$ for each $t \in I$ and some function $g \in L^1(I; \mathbb{R})$. Then, the function $\Psi : \mathbb{R}^d \to \mathbb{C}$, defined by*

$$\Psi(x) = \int_I \psi(t, x)dt, \qquad x \in \mathbb{R}^d,$$

is bounded and continuous and

$$(T(t)\Psi)(x) = \int_I (T(t)\psi(s, \cdot))(x)ds, \qquad t > 0, \ x \in \mathbb{R}^d. \tag{11.7}$$

Proof. The dominated convergence theorem shows that the function Ψ is well defined and belongs to $C_b(\mathbb{R}^d; \mathbb{C})$ and formula (11.7) follows from Fubini theorem. Indeed, the function $(s, y) \mapsto e^{-|x-y|^2/(4t)}\psi(s, y)$ belongs to $L^1(I \times \mathbb{R}^d; \mathbb{C})$ for each $x \in \mathbb{R}^d$. Hence,

$$\begin{aligned}(T(t)\Psi)(x) =&(4\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{4t}} \left(\int_I \psi(s, y)\, ds \right) dy \\ =&(4\pi t)^{-\frac{d}{2}} \int_{I \times \mathbb{R}^d} e^{-\frac{|x-y|^2}{4t}} \psi(s, y)\, ds\, dy = (4\pi t)^{-\frac{d}{2}} \int_I ds \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{4t}} \psi(s, y)\, dy \\ =& \int_I (T(t)\psi(s, \cdot))(x)\, ds\end{aligned}$$

for $t > 0$ and $x \in \mathbb{R}^d$, and we are done. □

Proof of Theorem 11.1.6. We denote by Π the right-halfplane, i.e. the set of all $\lambda \in \mathbb{C}$ with positive real part. For each $\lambda \in \Pi$, we introduce the operator R_λ defined by

$$(R_\lambda f)(x) = \int_0^\infty e^{-\lambda t}(T(t)f)(x)dt, \qquad x \in \mathbb{R}^d, \ f \in C_b(\mathbb{R}^d; \mathbb{C}).$$

Observe that, by the dominated convergence theorem, $R_\lambda f$ is a bounded and continuous function in $\mathbb{R}^d$ for every $f \in C_b(\mathbb{R}^d; \mathbb{C})$ and $\lambda \in \Pi$. Moreover, R_λ is a bounded operator for every λ as above. Finally, if f is real valued and $\lambda > 0$, then R_λ is real valued too.

It is easily seen that $\{R_\lambda : \lambda \in \Pi\}$ is a resolvent family: indeed, taking Lemma 11.1.8 into account, we can show that, for every $\lambda, \mu \in \Pi$, it holds that

$$(R_\lambda R_\mu f)(x) = \int_0^\infty e^{-\lambda t} \left(T(t) \int_0^\infty e^{-\mu s}(T(s)f)(\cdot)\, ds \right)(x)\, dt$$

$$
\begin{aligned}
&=\int_0^\infty dt\int_0^\infty e^{-\lambda t-\mu s}(T(t+s)f)(x)\,ds\\
&=\int_0^\infty e^{-\mu\sigma}(T(\sigma)f)(x)\,d\sigma\int_0^\sigma e^{(\mu-\lambda)t}dt=\int_0^\infty e^{-\mu\sigma}(T(\sigma)f)(x)\frac{e^{(\mu-\lambda)\sigma}-1}{\mu-\lambda}d\sigma\\
&=\frac{1}{\mu-\lambda}[(R_\lambda f)(x)-(R_\mu f)(x)]
\end{aligned}
$$

for all $x\in\mathbb{R}^d$. Let us prove that R_λ is injective for each $\lambda\in\Pi$. For this purpose, we fix $\lambda_0\in\Pi$ and $f\in C_b(\mathbb{R}^d;\mathbb{C})$ such that $R_{\lambda_0}f\equiv 0$. The resolvent identity proved above shows that $R_\lambda f\equiv 0$ for each $\lambda\in\Pi$. Since, for every $x\in\mathbb{R}^d$, the function $\lambda\mapsto(R_\lambda f)(x)$ is the Laplace transform of the bounded and continuous function $t\mapsto(T(t)f)(x)$, by the uniqueness of the Laplace transform, we conclude that $(T(t)f)(x)=0$ for all $t\ge 0$. Taking $t=0$ we conclude that $f(x)=0$. The arbitrariness of $x\in\mathbb{R}^d$ implies that $f\equiv 0$. Hence, R_λ is injective.

By Lemma 11.1.7, it follows that there exists a closed operator $A:D(A)\subset C_b(\mathbb{R}^d;\mathbb{C})\to C_b(\mathbb{R}^d;\mathbb{C})$, whose resolvent set contains the right-halfplane, and such that $R(\lambda,A)=R_\lambda$ for each $\lambda\in\Pi$.

We claim that A is a sectorial operator. To check the claim, we will show that there exists a positive constant C such that

$$
\|R(\lambda,A)\|\le C|\lambda|^{-1},\qquad \lambda\in\mathbb{C},\ \ \mathrm{Re}\lambda\ge 1. \tag{11.8}
$$

Indeed, from (11.8) and Proposition 9.2.8, the sectoriality of A follows at once.

We begin by observing that, for each $f\in C_b(\mathbb{R}^d;\mathbb{C})$ and $x\in\mathbb{R}^d$, the function $t\mapsto(T(t)f)(x)$ can be extended to the right-halfplane with a holomorphic function. Indeed, for $x\in\mathbb{R}^d$, the function $K(\cdot,x):\Pi\to\mathbb{C}$, defined by $K(z,x)=(4\pi z)^{-d/2}e^{-|x|^2/(4z)}$ is holomorphic in Π. Moreover, $K(z,\cdot)\in L^1(\mathbb{R}^d;\mathbb{C})$ for every $z\in\Pi$ and

$$
\int_{\mathbb{R}^d}|K(z,x)|dx=\left(\frac{\mathrm{Re}z}{|z|}\right)^{-\frac{d}{2}},\qquad z\in\Pi.
$$

Hence, the function

$$
x\mapsto(T(z)f)(x)=\int_{\mathbb{R}^d}K(z,x-y)f(y)\,dz \tag{11.9}
$$

is well defined, bounded and continuous in $\mathbb{R}^d$, for every $z\in\Pi$, as it can be easily seen again applying the dominated convergence theorem. Moreover,

$$
\|T(z)f\|_\infty\le\left(\frac{\mathrm{Re}z}{|z|}\right)^{-\frac{d}{2}}\|f\|_\infty,\qquad z\in\Pi.
$$

In particular, if for each $\vartheta_0\in(0,\pi/2)$, we denote by Σ_{ϑ_0} the sector of all $\lambda\in\mathbb{C}\setminus\{0\}$ such that $|\arg(\lambda)|\le\vartheta_0$, then, from the previous estimate it follows immediately that

$$
\|T(z)f\|_\infty\le[1+(\tan(\vartheta_0))^2]^{\frac{d}{4}}\|f\|_\infty,\qquad z\in\Sigma_{\vartheta_0}.
$$

Now, we can prove (11.8). If $\lambda=a+iy$ with $a\ge 1$ and $y\ge 0$, then by Cauchy integral theorem we have

$$
(R(\lambda,A)f)(x)=\int_0^\infty e^{-\lambda t}(T(t)f)(x)\,dt=\int_\gamma e^{-\lambda z}(T(z)f)(x)\,dz,\qquad x\in\mathbb{R}^d,
$$

where $\gamma(s) = s - is$ for every $s \geq 0$. Therefore

$$\|R(\lambda, A)f\|_\infty \leq 2^{\frac{d}{4}}\|f\|_\infty \int_0^\infty e^{-(a+y)s} ds \leq 2^{\frac{d}{4}}|\lambda|^{-1}\|f\|_\infty.$$

If $y \leq 0$ then one gets the same estimate replacing the curve γ with the curve $\widetilde{\gamma}$, defined by $\widetilde{\gamma}(s) = s + is$ for every $s \geq 0$. Estimate (11.8) follows.

Denote by $\{S(t)\}$ the analytic semigroup generated in $C_b(\mathbb{R}^d;\mathbb{C})$ by operator A. By Proposition 8.2.3 it follows that, if $\lambda \in \mathbb{C}$ has sufficiently large real part, then

$$\int_0^\infty e^{-\lambda t}(S(t)f)(x)dt = (R(\lambda, A)f)(x) = (R_\lambda f)(x) = \int_0^\infty e^{-\lambda t}(T(t)f)(x)dt$$

for $x \in \mathbb{R}^d$, i.e.

$$\int_0^\infty e^{-\lambda t}[(S(t)f)(x) - (T(t)f)(x)]dt = 0, \qquad x \in \mathbb{R}^d.$$

Again, the uniqueness of the Laplace transform implies that $S(t)f \equiv T(t)f$ in $(0,\infty)$. We have so proved that the Gauss-Weierstrass semigroup is analytic in $C_b(\mathbb{R}^d;\mathbb{C})$. □

Remark 11.1.9. Without much effort, we can show that the sectorial operator A associated with the Gauss-Weierstrass semigroup is an extension of the operator $(\Delta, C_b^2(\mathbb{R}^d;\mathbb{C}))$. Indeed, fix $f \in C_b^2(\mathbb{R}^d;\mathbb{C})$. Recalling that each operator $T(t)$ commutes with the Laplacian and integrating by parts, we get

$$\begin{aligned}(R(1,A)(f-\Delta f))(x) =& \int_0^\infty e^{-t}(T(t)(f-\Delta f))(x)\,dt \\ =& \int_0^\infty e^{-t}((T(t)f)(x) - (\Delta T(t)f)(x))\,dt \\ =& \int_0^\infty e^{-t}(T(t)f)(x)\,dt - \int_0^\infty e^{-t}(D_tT(t)f)(x)\,dt \\ =&(R(1,A)f)(x) + f(x) - (R(1,A)f)(x) = f(x)\end{aligned}$$

for $x \in \mathbb{R}^d$. Hence, $R(1,A)(f-\Delta f) = f$. Applying the operator $I - A$ to both the sides of the previous equality we obtain $f - \Delta f = f - Af$, i.e., $Af = \Delta f$. Therefore, A is an extension of the operator $(\Delta, C_b^2(\mathbb{R}^d;\mathbb{C}))$.

If $d = 1$, then A actually coincides with the second order derivative with $C_b^2(\mathbb{R};\mathbb{C})$ as domain. On the other hand, if $d \geq 2$, then A is a proper extension of $(\Delta, C_b^2(\mathbb{R}^d;\mathbb{C}))$. In fact, $D(A) = \{u \in C_b(\mathbb{R}^d;\mathbb{C}) \cap W^{2,p}_{\rm loc}(\mathbb{R}^d;\mathbb{C})$, for all $p \in [1,\infty)$, and $\Delta u \in C_b(\mathbb{R}^d;\mathbb{C})\}$ and $Au = \Delta u$ for each $u \in D(A)$. The proof of the above characterization of $(A, D(A))$ is beyond the purposes of this lecture. We refer the interested reader to [12, Chapter 14].

11.2 The interpolation spaces $D_A(\alpha,\infty)$ and $D_A(\alpha+1,\infty)$

In this section we characterize the interpolation spaces $D_A(\alpha,\infty)$ and $D_A(\alpha+1,\infty)$.

Theorem 11.2.1. *For each $\alpha \in (0,1) \setminus \{1/2\}$ it holds that $D_A(\alpha,\infty) = C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})$, with equivalence of the respective norms.*

Proof. Throughout the proof, we denote by C a positive constant, independent of f and t, which may vary from line to line.

We begin by proving the embedding $C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C}) \hookrightarrow D_A(\alpha,\infty)$. For this purpose, we observe that, by a change of variable in the formula defining $T(t)f$, we can write

$$(T(t)f)(x) - f(x) = (4\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|y|^2}{4}} \big[f(x-\sqrt{t}y) - f(x)\big]dy, \qquad t>0,\ x\in\mathbb{R}^d, \tag{11.10}$$

for every $f \in C_b(\mathbb{R}^d;\mathbb{C})$. Consequently, if $f \in C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})$, with $\alpha \in (0,1/2)$, we can estimate

$$\|T(t)f - f\|_\infty \le (4\pi)^{-\frac{d}{2}} \|f\|_{C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})} t^\alpha \int_{\mathbb{R}^d} e^{-\frac{1}{4}|y|^2}|y|^{2\alpha}dy, \qquad t>0. \tag{11.11}$$

We claim that (11.11) holds true also when $\alpha \in (1/2,1)$. To prove the claim, we observe that, replacing y by $-y$ in (11.10), we can write

$$(T(t)f)(x) - f(x) = (4\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|y|^2}{4}} \big[f(x+\sqrt{t}y) - f(x)\big]dy, \qquad t>0,\ x\in\mathbb{R}^d. \tag{11.12}$$

Combining (11.10) and (11.12), we conclude that

$$(T(t)f)(x) - f(x) = \frac{1}{2}(4\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|y|^2}{4}} \big[f(x-\sqrt{t}y) - 2f(x) + f(x+\sqrt{t}y)\big]dy$$

for $(t,x) \in (0,\infty)\times\mathbb{R}^d$. Since $f \in C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})$, the first-order derivatives of f belong to $C_b^{2\alpha-1}(\mathbb{R}^d;\mathbb{C})$ and we can estimate

$$\begin{aligned}|f(x-\sqrt{t}y) - 2f(x) + f(x+\sqrt{t}y)| =& \sqrt{t}\left|\int_0^1 \langle \nabla f(x+\sigma\sqrt{t}y) - \nabla f(x-\sigma\sqrt{t}y), y\rangle d\sigma\right| \\ \le& \sqrt{t}|y| \int_0^1 |\nabla f(x+\sigma\sqrt{t}y) - \nabla f(x-\sigma\sqrt{t}y)|d\sigma \\ \le& C\|f\|_{C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})} t^\alpha |y|^{2\alpha}\end{aligned}$$

for every $t>0$ and $x,y\in\mathbb{R}^d$. From this estimate, (11.11) easily follows.

Since the integral term in the right-hand side of the previous formula converges, by Proposition 10.2.3 we deduce that $f \in D_A(\alpha,\infty)$ and $\|f\|_{D_A(\alpha,\infty)} \le C\|f\|_{C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})}$. The embedding $C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C}) \hookrightarrow D_A(\alpha,\infty)$ follows.

Conversely, assume that $f \in D_A(\alpha,\infty)$. We first consider the case $\alpha \in (0,1/2)$ and observe that for $t \in (0,1]$ and $x,y \in \mathbb{R}^d$ we can estimate

$$\begin{aligned}|f(x) - f(y)| \le& |(T(t)f)(x) - f(x)| + |(T(t)f)(x) - (T(t)f)(y)| + |(T(t)f)(y) - f(y)| \\ \le& 2\|f\|_{D_A(\alpha,\infty)} t^\alpha + \||\nabla T(t)f|\|_\infty |x-y|.\end{aligned} \tag{11.13}$$

This estimate can be straightforwardly extended to all $t>0$, observing that $|(T(t)f)(\xi) - f(\xi)| \le 2\|f\|_\infty$ for every $\xi\in\mathbb{R}^d$.

We would like to take $t = |x-y|^2$ in the previous formula. Unfortunately, the estimate $\||\nabla T(t)f|\|_\infty \le Ct^{-1/2}\|f\|_\infty$ in Exercise 11.4.1 is not enough sharp for our purposes. Indeed, using that estimate we would get $|f(x)-f(y)| \le 2\|f\|_{D_A(\alpha,\infty)}|x-y|^{2\alpha} + C$. To overcome this difficulty, we differentiate the formula

$$(T(n)f)(x) - (T(t)f)(x) = \int_t^n (D_tT(s)f)(x)ds = \int_t^n (\Delta T(s)f)(x)ds,$$

to get

$$(D_iT(n)f)(x) - (D_iT(t)f)(x) = \int_t^n (D_i\Delta T(s)f)(x)ds, \qquad t \in (0,n),\ x \in \mathbb{R}^d, \tag{11.14}$$

for every $i = 1, \ldots, d$. Since $f \in D_A(\alpha, \infty)$, we can estimate $\|\Delta T(t)f\|_\infty = \|AT(t)f\|_\infty \le t^{\alpha-1}[f]_{D_A(\alpha,\infty)}$ for every $t \in (0,1]$. Moreover, Exercise 11.4.1 shows that $\|\Delta T(t)f\|_\infty \le Ct^{-1}\|f\|_\infty$ for every $t \ge 1$. Combining these two estimates and replacing C with a larger constant, if needed, we conclude that $\|\Delta T(t)f\|_\infty \le Ct^{\alpha-1}\|f\|_{D_A(\alpha,\infty)}$ for every $t > 0$. Taking again Exercise 11.4.1 and the semigroup property into account, and observing that $T(t)$ and Δ commute on $C_b^2(\mathbb{R}^d;\mathbb{C})$, we can estimate

$$\begin{aligned}\|D_i\Delta T(s)f\|_\infty =&\|D_iT(s/2)\Delta T(s/2)f\|_\infty \le \|D_iT(s/2)\|_{L(C_b(\mathbb{R}^d))}\|\Delta T(s/2)f\|_\infty\\ \le& Cs^{\alpha-\frac{3}{2}}\|f\|_{D_A(\alpha,\infty)},\end{aligned} \tag{11.15}$$

so that we may let n tend to ∞ in (11.14) to get

$$(D_iT(t)f)(x) = -\int_t^\infty (D_i\Delta T(s)f)(x)ds, \qquad t > 0,\ x \in \mathbb{R}^d,$$

and

$$\|D_iT(t)f\|_\infty \le C\|f\|_{D_A(\alpha,\infty)}\int_t^\infty s^{\alpha-\frac{3}{2}}ds = C_\alpha t^{\alpha-\frac{1}{2}}\|f\|_{D_A(\alpha,\infty)} \tag{11.16}$$

for every $i = 1, \ldots, d$. This estimate is what we need to prove that f is 2α-Hölder continuous in $\mathbb{R}^d$. Indeed, replacing (11.16) in (11.13) and taking $t = |x-y|^2$, we obtain $|f(x) - f(y)| \le C\|f\|_{D_A(\alpha,\infty)}|x-y|^{2\alpha}$, so that $f \in C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})$ and $[f]_{C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})} \le C\|f\|_{D_A(\alpha,\infty)}$. We have so proved that $D_A(\alpha,\infty) \hookrightarrow C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})$. This completes the proof in the case $\alpha < 1/2$.

Let us now suppose that $\alpha \in (1/2, 1)$. As a first step, we prove that $f \in C_b^1(\mathbb{R}^d;\mathbb{C})$. For this purpose, we fix $0 < s < t$ and observe that

$$(D_iT(t)f)(x) - (D_iT(s)f)(x) = \int_s^t (D_i\Delta T(r)f)(x)dr, \qquad x \in \mathbb{R}^d,\ i = 1, \ldots, d.$$

Using estimate (11.15) we deduce that

$$|(D_iT(t)f)(x) - (D_iT(s)f)(x)| \le C\|f\|_{D_A(\alpha,\infty)}\Big(t^{\alpha-\frac{1}{2}} - s^{\alpha-\frac{1}{2}}\Big), \qquad x \in \mathbb{R}^d. \tag{11.17}$$

This estimate shows that $(\nabla_x T(1/n)f)_{n\in\mathbb{N}}$ is a Cauchy sequence in $(C_b(\mathbb{R}^d;\mathbb{C}))^d$. On the other hand, Remarks 10.2.2(ii) and 11.1.5 show that $D_A(\alpha,\infty) \subset BUC(\mathbb{R}^d;\mathbb{C})$. So, $(T(1/n)f)_{n\in\mathbb{N}}$ converges to f uniformly in $\mathbb{R}^d$ as n tends to ∞. As a byproduct, we conclude that f is continuously differentiable in $\mathbb{R}^d$. Moreover, from (11.17) it follows that

$$\|D_iT(t)f - D_if\|_\infty \le C\|f\|_{D_A(\alpha,\infty)}t^{\alpha-\frac{1}{2}}, \qquad t > 0,\ i = 1, \ldots, d. \tag{11.18}$$

Hence, taking Exercise 11.4.1 into accout, we deduce that

$$\|D_if\|_\infty \le \|D_iT(1)f - D_if\|_\infty + \|D_iT(1)f\|_\infty \le C\|f\|_{D_A(\alpha,\infty)}, \qquad i = 1, \ldots, d, \tag{11.19}$$

so that $f \in C_b^1(\mathbb{R}^d;\mathbb{C})$.

Fix $i \in \{1,\dots,d\}$. Since f is differentiable in $\mathbb{R}^d$, taking (11.18) into account we can estimate

$$\begin{aligned}|D_if(x)-D_if(y)| \le& |(D_iT(t)f)(x)-D_if(x)|+|(D_iT(t)f)(x)-(D_iT(t)f)(y)|\\ &+|(D_iT(t)f)(y)-D_if(y)|\\ \le& C\|f\|_{D_A(\alpha,\infty)}t^{\alpha-\frac{1}{2}}+\||\nabla_xD_iT(t)f|\|_\infty|x-y|. \end{aligned} \tag{11.20}$$

With the same arguments used to prove (11.15), we can show that

$$\|D_{ij}\Delta T(s)f\|_\infty \le Cs^{\alpha-2}\|f\|_{D_A(\alpha,\infty)}, \qquad s>0. \tag{11.21}$$

Therefore, since

$$(D_{ij}T(t)f)(x) = -\int_t^\infty (D_{ij}\Delta T(s)f)(x)ds, \qquad t>0,\ x\in\mathbb{R}^d,$$

using (11.21) we obtain that $\||\nabla_xD_iT(t)f|\|_\infty \le Ct^{\alpha-1}\|f\|_{D_A(\alpha,\infty)}$ for $t>0$. Replacing this estimate in (11.20), we get

$$|D_if(x)-D_if(y)| \le C\|f\|_{D_A(\alpha,\infty)}t^{\alpha-\frac{1}{2}}+Ct^{\alpha-1}\|f\|_{D_A(\alpha,\infty)}|x-y|$$

for $x,y\in\mathbb{R}^d$ and $t>0$. Taking $t=|x-y|^2$ we can infer that

$$[D_if]_{C_b^{2\alpha-1}(\mathbb{R}^d;\mathbb{C})} \le C\|f\|_{D_A(\alpha,\infty)}. \tag{11.22}$$

From (11.19) and (11.22) we conclude that $D_if\in C_b^{2\alpha-1}(\mathbb{R}^d;\mathbb{C})$ and $\|D_if\|_{C_b^{2\alpha-1}(\mathbb{R}^d;\mathbb{C})} \le C\|f\|_{D_A(\alpha,\infty)}$ for every $i=1,\dots,d$. The inclusion $D_A(\alpha,\infty)\hookrightarrow C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})$ follows at once. □

Remark 11.2.2. The previous theorem does not cover the case $\alpha=1/2$. One may think that $D_A(1/2,\infty)=C_b^1(\mathbb{R}^d;\mathbb{C})$ or $D_A(1/2,\infty)=\mathrm{Lip}_b(\mathbb{R}^d;\mathbb{C})$ (the space of bounded Lipschitz continuous functions $f:\mathbb{R}\to\mathbb{C}$). But this is not the case. Indeed,

$$D_A(1/2,\infty)=\left\{f\in C_b(\mathbb{R}^d;\mathbb{C}): \sup_{x\neq y}\frac{|f(x)+f(y)-2f((x+y)/2)|}{|x-y|}<\infty\right\},$$

and $\mathrm{Lip}_b(\mathbb{R}^d;\mathbb{C})$ is one of its proper subspaces (see e.g., [21]).

We are going to prove that the semigroup $T(t)$ leaves $C_b^\alpha(\mathbb{R}^d;\mathbb{C})$ invariant for any $t>0$ and $\alpha\in(0,1)$.

Proposition 11.2.3. *For each $f\in C_b^\alpha(\mathbb{R}^d;\mathbb{C})$ and $t>0$, the function $T(t)f$ belongs to $C_b^\alpha(\mathbb{R}^d;\mathbb{C})$. More precisely,*

$$\|T(t)f\|_{C_b^\alpha(\mathbb{R}^d;\mathbb{C})} \le C\|f\|_{C_b^\alpha(\mathbb{R}^d;\mathbb{C})}, \qquad t>0,$$

for some positive constant C, independent of f.

Proof. To prove the assertion, we observe that since $T(s)$ and Δ commute on $C_b^2(\mathbb{R}^d;\mathbb{C})$ and each operator $T(t)$ is a contraction in $C_b(\mathbb{R}^d;\mathbb{C})$, we can estimate

$$\|\Delta T(s+t)f\|_\infty = \|T(s)\Delta T(t)f\|_\infty \le \|\Delta T(t)f\|_\infty \le t^{\frac{\alpha}{2}-1}[f]_{D_A(\alpha/2,\infty)}$$

for all $t, s > 0$. Hence, $[T(s)f]_{D_A(\alpha/2,\infty)} \le [f]_{D_A(\alpha/2,\infty)}$ for any $s > 0$. Theorem 11.2.1 allows us to complete the proof. □

In terms of the bounded classical solution to problem (11.1), the previous result says that if $f \in C_b^\alpha(\mathbb{R}^d;\mathbb{C})$ for some $\alpha \in (0,1)$, then $u(t,\cdot) \in C_b^\alpha(\mathbb{R}^d;\mathbb{C})$ for any $t > 0$ and

$$\sup_{t\ge 0} \|u(t,\cdot)\|_{C_b^\alpha(\mathbb{R}^d;\mathbb{C})} \le C\|f\|_{C_b^\alpha(\mathbb{R}^d;\mathbb{C})}$$

for some constant C, independent of f.

Based on Theorem 11.2.1, we can now characterize the interpolation spaces $D_A(\alpha+1,\infty)$ for $\alpha \ne 1/2$.

Theorem 11.2.4. *For every $\alpha \in (0,1)\setminus\{1/2\}$, it holds that $D_A(\alpha+1,\infty) = C_b^{2+2\alpha}(\mathbb{R}^d;\mathbb{C})$, with equivalence of the respective norms.*

Proof. Recall that $D_A(\alpha+1,\infty) = \{f \in D(A) : Af \in D_A(\alpha,\infty)\}$ is equipped with the norm $\|f\|_{D_A(\alpha+1,\infty)} = \|f\|_{D(A)} + [Af]_{D_A(\alpha,\infty)}$, see Definition 10.2.9. The embedding $C_b^{2+2\alpha}(\mathbb{R}^d;\mathbb{C}) \subset D_A(\alpha+1,\infty)$ is an obvious consequence of the fact that $C_b^2(\mathbb{R}^d;\mathbb{C})$ is contained in $D(A)$ and $A = \Delta$ on $C_b(\mathbb{R}^d;\mathbb{C})$. Indeed, if $f \in C_b^{2+2\alpha}(\mathbb{R}^d;\mathbb{C})$, then $Af \in C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C}) \hookrightarrow D_A(\alpha,\infty)$. Hence, $\|f\|_\infty + \|Af\|_{D_A(\alpha,\infty)} \le C\|f\|_{C_b^{2+2\alpha}(\mathbb{R}^d;\mathbb{C})}$ for a constant C, which as all the other constants appearing in the proof (and still denoted by C) is independent of f.

To prove the other embedding, we fix $f \in D_A(\alpha+1,\infty)$. Then, $f = R(1,A)g$, where $g = f - Af \in D_A(\alpha,\infty) = C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})$, and by (9.14) we can write

$$f(x) = \int_0^\infty e^{-t}(T(t)g)(x)dt, \qquad x \in \mathbb{R}^d. \tag{11.23}$$

Since $g \in D_A(\alpha,\infty)$, the proof of Theorem 11.2.1 shows that $\|D_iT(t)g\|_\infty \le Ct^{\alpha-1/2}\|g\|_{D_A(\alpha,\infty)}$ for every $t > 0$ and $i = 1,\dots,d$, so that, using once again Exercise 11.4.1, we get

$$\begin{aligned}\|D_{ij}T(t)g\|_\infty =& \|D_jT(t/2)D_iT(t/2)g\|_\infty \\ \le& \|D_jT(t/2)\|\|D_iT(t/2)g\|_\infty \le Ct^{\alpha-1}\|g\|_{D_A(\alpha,\infty)}\end{aligned} \tag{11.24}$$

for $i,j = 1,\dots,d$. It follows that the functions $t \mapsto \|e^{-t}D_iT(t)g\|_\infty$ and $t \mapsto \|e^{-t}D_{ij}T(t)g\|_\infty$ $(i,j = 1,\dots,d)$ are integrable in $(0,\infty)$. Therefore, we can differentiate the right-hand side of (11.23) twice with respect to x and conclude that the function f belongs to $C_b^2(\mathbb{R}^d;\mathbb{C})$ and

$$\|u\|_{C_b^2(\mathbb{R}^d;\mathbb{C})} \le C\|g\|_{D_A(\alpha,\infty)}. \tag{11.25}$$

To prove that the second-order derivatives of f belong to $C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C}) = D_A(\alpha,\infty)$, it suffices to show that

$$\sup_{0<\xi\le 1} \|\xi^{1-\alpha}AT(\xi)D_{ij}f\|_\infty < \infty, \qquad i,j = 1,\dots,d.$$

For this purpose, we fix $\xi \in (0,1]$ and observe that

$$\begin{aligned}\|\xi^{1-\alpha} AT(\xi)D_{ij}f\|_\infty =&\left\| \int_0^{+\infty} \xi^{1-\alpha} e^{-t} AT(\xi + t/2) D_{ij} T(t/2) g dt \right\|_\infty \\ \le& Ct^{\alpha-1} \|g\|_{D_A(\alpha,\infty)} \int_0^\infty \frac{\xi^{1-\alpha}}{(2\xi + t)} dt \\ \le& C\|g\|_{D_A(\alpha,\infty)} \int_0^\infty \frac{1}{(1+s)s^{1-\alpha}} ds. \end{aligned} \tag{11.26}$$

From (11.25) and (11.26) we deduce that $\|f\|_{C_b^{2+2\alpha}(\mathbb{R}^d;\mathbb{C})} \le C\|g\|_{D_A(\alpha,\infty)}$. Since $\|g\|_{D_A(\alpha,\infty)} \le C\|f\|_{D_A(\alpha+1,\infty)}$, the embedding $D_A(\alpha+1,\infty) \hookrightarrow C_b^{2+2\alpha}(\mathbb{R}^d;\mathbb{C})$ follows. □

Corollary 11.2.5 (Schauder theorem)**.** *If $u \in C_b^2(\mathbb{R}^d;\mathbb{C})$ and $\Delta u \in C_b^\alpha(\mathbb{R}^d;\mathbb{C})$ for some $\alpha \in (0,1)$, then $u \in C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})$.*

Proof. It suffices to observe that function u belongs to $D_A(\alpha/2+1,\infty) = C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})$. □

Theorem 11.2.1 and Corollary 10.2.4 imply that the solution $u = T(\cdot)u_0$ of the Cauchy problem for the heat equation in $\mathbb{R}^d$,

$$\begin{cases} u_t(t,x) = \Delta u(t,x), & t>0, \quad x \in \mathbb{R}^d, \\ u(0,x) = u_0(x), & \qquad\quad\; x \in \mathbb{R}^d, \end{cases}$$

is α-Hölder continuous with respect to t on $[0,T]\times\mathbb{R}^d$, with Hölder constant independent of x, if and only if the initial datum u_0 belongs to $C_b^{2\alpha}(\mathbb{R}^d;\mathbb{C})$, where $\alpha \neq 1/2$. In this case, Proposition 10.2.10 implies that $\|u(t,\cdot)\|_{D_A(\alpha,\infty)} \le C\|u_0\|_{D_A(\alpha,\infty)}$ for $0 \le t \le T$, so that u is 2α-Hölder continuous with respect to x as well, with Hölder constant independent of t. We say that u belongs to the parabolic Hölder space $C^{\alpha,2\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})$, for all $T>0$.

This is a first example of a typical feature of second order parabolic partial differential equations: time regularity implies space regularity, and the degree of regularity with respect to time is one half of the regularity with respect to the space variables.

11.3 Optimal Schauder estimates

As a consequence of Theorems 10.2.14 and 10.2.16, we get a classical theorem of the theory of PDE's. We need some notation.

We recall that for $0<\alpha<1$ the parabolic Hölder space $C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})$ is the space of the continuous functions f such that

$$\|f\|_{C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})} := \|f\|_\infty + \sup_{x\in\mathbb{R}^d} [f(\cdot,x)]_{C^{\alpha/2}([0,T];\mathbb{C})} + \sup_{t\in[0,T]} [f(t,\cdot)]_{C_b^\alpha(\mathbb{R}^d;\mathbb{C})} < \infty$$

and $C^{1+\alpha/2,2+\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})$ is the space of the bounded functions u such that u_t, $D_{ij}u$ exist for all i, $j=1,\dots,N$ and belong to $C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})$. It is a Banach space when endowed with the norm

$$\|u\|_{C^{1+\alpha/2,2+\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})} := \|u\|_\infty + \sum_{i=1}^d \|D_i u\|_\infty$$

$$+\|D_t u\|_{C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})}+\sum_{i,j=1}^d \|D_{ij}u\|_{C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})}.$$

Note that $f\in C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})$ if and only if $t\mapsto f(t,\cdot)$ belongs to $C^{\alpha/2}([0,T];C_b(\mathbb{R}^d;\mathbb{C}))$ $\cap\, B([0,T];C_b^{\alpha}(\mathbb{R}^d;\mathbb{C}))$ (see Exercise 11.4.3).

Theorem 11.3.1 (Ladyzhenskaja-Solonnikov-Ural'ceva)**.** *For every* $u_0\in C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})$ $(\alpha\in(0,1))$*,* $f\in C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})$*, the Cauchy problem*

$$\begin{cases} D_t u(t,x)=\Delta u(t,x)+f(t,x), & t\in[0,T], \quad x\in\mathbb{R}^d,\\ u(0,x)=u_0(x), & x\in\mathbb{R}^d,\end{cases}\tag{11.27}$$

admits a unique solution $u\in C^{1+\alpha/2,2+\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})$*, and there exists a positive constant* C*, independent of* u_0 *and* f*, such that*

$$\|u\|_{C^{1+\alpha/2,2+\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})}\le C(\|u_0\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})}+\|f\|_{C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})}).$$

Proof. The function $t\mapsto f(t,\cdot)$ belongs to $C^{\alpha/2}([0,T];C_b(\mathbb{R}^d;\mathbb{C}))\cap B([0,T];D_A(\alpha/2,\infty))$, thanks to the characterization in Theorem 11.2.1. Similarly, the initial datum u_0 belongs to $D(A)$ and both Au_0 and $f(0,\cdot)$ belong to $D_A(\alpha/2,\infty)$. Thus, we may apply Theorems 10.2.14 and 10.2.16, with α being replaced by $\alpha/2$. They imply that the function u given by the variation of constants formula (10.14) is the unique strict solution to problem (10.1), with initial datum u_0 and with $f(t)=f(t,\cdot)$. Therefore, the function $u:[0,T]\times\mathbb{R}^d\to\mathbb{C}$, defined by

$$u(t,x):=u(t)(x)=(T(t)u_0)(x)+\int_0^t (T(t-s)f(s,\cdot))(x)ds, \qquad t\in[0,T],\ \ x\in\mathbb{R}^d,$$

is the unique bounded classical solution to (11.27) with bounded u_t. Moreover, Theorem 10.2.14 implies that $D_t u\in C^{\alpha/2}([0,T];C_b(\mathbb{R}^d;\mathbb{C}))\cap B([0,T];C_b^{\alpha}(\mathbb{R}^d;\mathbb{C}))$, so that $u_t\in C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})$, with norm bounded by $C(\|u_0\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})}+\|f\|_{C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})})$ for some $C>0$. Corollary 10.2.16 implies that u is bounded with values in $D_A(\alpha/2+1,\infty)$, so that $u(t,\cdot)\in C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})$ for each $t\in[0,T]$, and

$$\sup_{0\le t\le T}\|u(t,\cdot)\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})}\le C(\|u_0\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})}+\|f\|_{C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})}),$$

for some $C>0$, by estimate (10.29).

To finish the proof it remains to show that each second order space derivative $D_{ij}u$ is $\alpha/2$-Hölder continuous with respect to t. To this aim we use the interpolatory inequality

$$\|D_{ij}\varphi\|_\infty\le C(\|\varphi\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})})^{1-\alpha/2}(\|\varphi\|_{C_b^{\alpha}(\mathbb{R}^d;\mathbb{C})})^{\alpha/2},$$

that holds for every $\varphi\in C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})$, i, $j=1,\dots,d$ (see Exercise 11.4.4). Applying it to the function $\varphi=u(t,\cdot)-u(s,\cdot)$ we get

$$\begin{aligned}&\|D_{ij}u(t,\cdot)-D_{ij}u(s,\cdot)\|_\infty\\ \le& C(\|u(t,\cdot)-u(s,\cdot)\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})})^{1-\alpha/2}(\|u(t,\cdot)-u(s,\cdot)\|_{C_b^{\alpha}(\mathbb{R}^d;\mathbb{C})})^{\alpha/2}\\ \le& C(2\sup_{0\le t\le T}\|u(t,\cdot)\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})})^{1-\alpha/2}(|t-s|\sup_{0\le t\le T}\|u_t(t,\cdot)\|_{C_b^{\alpha}(\mathbb{R}^d;\mathbb{C})})^{\alpha/2}\\ \le& C'|t-s|^{\alpha/2}(\|u_0\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})}+\|f\|_{C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d;\mathbb{C})}),\end{aligned}$$

and the statement follows. □

11.4 Exercises

Exercise 11.4.1. Using Lemma 11.1.2, prove that, for every $t > 0$ and $f \in C_b(\mathbb{R}^d;\mathbb{C})$ the function $T(t)f$ belongs to $C_b^2(\mathbb{R}^d;\mathbb{C})$ and there exists a positive constant C, independent of t and f, such that

$$\sqrt{t}\||\nabla T(t)f|\|_\infty + t\|D^2T(t)f\|_\infty \le C\|f\|_\infty, \qquad t > 0.$$

Further, prove that the function $(t,x) \mapsto (T(t)f)(x)$ belongs to $C^\infty((0,\infty)\times\mathbb{R}^d;\mathbb{C})$.

Exercise 11.4.2. Prove that $BUC(\mathbb{R}^d;\mathbb{C})$ is the closure of $C^1(\mathbb{R}^d;\mathbb{C})$ with respect to the sup-norm.

Exercise 11.4.3. Prove that $f \in C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d)$ if and only if the function $t \mapsto f(t,\cdot)$ belongs to $C^{\alpha/2}([0,T];C_b(\mathbb{R}^d;\mathbb{C})) \cap B([0,T];C_b^\alpha(\mathbb{R}^d;\mathbb{C}))$.

Exercise 11.4.4. Prove that for every $\alpha \in (0,1)$ there exists a positive constant $C = C(\alpha)$ such that

$$\|D_i\varphi\|_\infty \le C(\|\varphi\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})})^{(1-\alpha)/2}(\|\varphi\|_{C_b^\alpha(\mathbb{R}^d;\mathbb{C})})^{(1+\alpha)/2},$$

$$\|D_{ij}\varphi\|_\infty \le C(\|\varphi\|_{C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})})^{1-\alpha/2}(\|\varphi\|_{C_b^\alpha(\mathbb{R}^d;\mathbb{C})})^{\alpha/2},$$

for every $\varphi \in C_b^{2+\alpha}(\mathbb{R}^d;\mathbb{C})$ and $i,j = 1,\ldots,d$.
[**Hint:** write $\varphi = \varphi - T(t)\varphi + T(t)\varphi = -\int_0^t T(s)\Delta\varphi\,ds + T(t)\varphi$, $T(t)$ = heat semigroup, and use the estimates $\|D_iT(t)f\|_\infty \le Ct^{-1/2+\alpha/2}\|f\|_{C_b^\alpha}$, $\|D_{ij}T(t)f\|_\infty \le Ct^{-1+\alpha/2}\|f\|_{C_b^\alpha}$].

Exercise 11.4.5. Prove that $C_b^1(\mathbb{R})$ is of class $J_{1/4}$ between $C_b(\mathbb{R})$ and $C_b^4(\mathbb{R})$.

Exercise 11.4.6. Let $X = C^\alpha([0,1])$ for some $\alpha \in (0,1)$. Prove that the operator $A : D(A) = \{u \in C^{2+\alpha}([0,1]) : u(0) = u(1) = 0\} \to X$, defined by $Au = u''$ for $u \in D(A)$, is not sectorial.

Exercise 11.4.7. Let $X = C_0([0,1];X) = \{f \in C([0,1];\mathbb{C}) : f(0) = f(1) = 0\}$, endowed with the sup-norm. Further, let $A : D(A) = \{u \in C^2([0,1]) : u(0) = u''(0) = u(1) = u''(1) = 0\} \to X$ be the operator defined by $Au = u''$ for every $u \in D(A)$.

(a) Prove that A is sectorial;

(b) Prove that for every $\alpha \in (0,1)\setminus\{1/2\}$, $D_A(\alpha,\infty) = C^{2\alpha}(\mathbb{R}) \cap X$ and that $D_A(\alpha+1,\infty) = \{u \in C^{2+2\alpha}([0,1];\mathbb{C}) : u(0) = u''(0) = u(1) = u''(1) = 0\}$, with equivalence of the corresponding norms.

Bibliography

[1] A. Belleni-Morante, Applied Semigroups and Evolution Equations, Oxford University Press, 1979.

[2] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[3] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[4] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[5] K.-J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Graduate Text in Mathematics **194**, Springer-Verlag, 2000.

[6] K.-J. Engel and R. Nagel, Short Course on Operator Semigroups, Universitext, Springer-Verlag, 2006.

[7] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[8] J.A. Goldstein, Semigroups of Operators and Applications, Oxford University Press, 1985.

[9] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[10] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[11] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981.

[12] L. Lorenzi and A. Rhandi, Semigroups of Bounded Operators and Second-Order Elliptic and Parabolic Partial Differential Equations, CRC Press, Taylor & Francis Group, 2021.

[13] A. Lunardi, Analytic Semigroups and Optimal Regularity in Parabolic Problems, Birkhäuser Verlag, 1995.

[14] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Appl. Math.Sci. **44**, Springer-Verlag,1983.

[15] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[16] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[17] W. Rudin, Real and Complex Analysis, 3rd ed., McGraw-Hill, 1986.

[18] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[19] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958.

[20] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861.

[21] A. Zygmund, Trigonometric Series, Cambridge Univ. Press., 2nd Edition Reprinted (1968).

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

January 27, 2022

Lecture 12

Spectral mapping theorem for semigroups

In this lecture, we investigate the relations which occur between the spectrum of the infinitesimal generator of a strongly continuous semigroup $\{T(t)\}$ and the spectrum of $\{T(t)\}$. We also consider the case when the semigroup $\{T(t)\}$ is analytic without being strongly continuous.

Based on the basic notions of functional calculus presented in Lecture 4 and on the analogy of a semigroup with the exponential function[1] of an operator A, one might think that the spectra of the operators belonging to a strongly continuous semigroup and the spectrum of its infinitesimal generator A are related each other by the formula

$$\sigma(T(t)) = e^{t\sigma(A)} = \{e^{t\lambda} : \lambda \in \sigma(A)\}, \qquad t > 0.$$

As it is easily seen, this formula does not hold in general, since it would imply that 0 lies in the resolvent set of the operator $T(t)$ for every $t > 0$, so that actually $\{T(t)\}$ should embed in a group (see Exercise 12.3.1). But, it is easy to find examples of strongly continuous semigroups which are not groups.

Consider now the semigroup of left translations on $X = L^p((0,1))$ $(1 \le p < \infty)$, defined by

$$T(t)f(x) = \begin{cases} f(x+t), & \text{if } x+t \le 1 \\ 0. & \text{otherwise.} \end{cases}$$

One can see that $\{T(t)\}$ is a strongly continuous semigroup on X and its generator A is the realization of the first-order derivative in $L^p((0,1))$ with domain $D(A) := \{f \in X : f \text{ absolutely continuous}, f' \in X, f(1) = 0\}$. From the definition of $\{T(t)\}$ one deduces that $\|T(t)\| = 0$ for any $t \ge 1$. So, $r(T(t)) = \lim_{n\to\infty} \|T(nt)\|^{1/n} = 0$ for all $t > 0$. This implies that $\sigma(T(t)) = \{0\}$ for all $t > 0$. On the other hand, it is easy to see that $\sigma(A) = \varnothing$.

Based on this example, instead of asking the validity of (12.1) one asks if the formula

$$\sigma(T(t)) \setminus \{0\} = e^{t\sigma(A)}, \qquad t > 0, \tag{12.1}$$

holds.

[1] Indeed, it is rather common to denote by $\{e^{tA}\}$ the strongly continuous semigroup generated by the operator A and, similarly, quite often the analytic semigroup associated with the sectorial operator A is denoted in the same way.

Aim of this lecture is to determine sufficient condition for the spectral mapping theorem (i.e., formula (12.1) to hold).

In this lecture, which is strongly inspired by [6], X will denote a complex Banach space.

12.1 Preliminary results

To begin with, we recall the following definitions.

Definition 12.1.1. Let $A : D(A) \subset X \to X$ be a closed operator. Then,

(i) the point spectrum of A ($\sigma_{\rm p}(A)$) is the set of all $\lambda \in \mathbb{C}$ such that the operator $\lambda I - A$ is not injective, i.e., there exists $0 \neq x \in D(A)$ such that $Ax = \lambda x$;

(ii) the approximate spectrum of A ($\sigma_{\rm a}(A)$) is the set of all $\lambda \in \mathbb{C}$ such that there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset D(A)$ with $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $\lambda x_n - Ax_n$ converges to 0 as n tends to ∞;

(iii) the residual spectrum of A ($\sigma_{\rm res}(A)$) is the set of all $\lambda \in \mathbb{C}$ such that the range of the operator $\lambda I - A$ is not dense in X.

Definition 12.1.2. Let $\{T(t)\}$ be a strongly continuous semigroup on X. Then, the spectral bound $s(A)$ of its infinitesimal generator A and the growth bound ω_0 of the semigroup $\{T(t)\}$ are defined as follows

$$
\begin{aligned}
&s(A) = \sup\{{\rm Re}\lambda : \lambda \in \sigma(A)\},\\
&\omega_0 = \inf\{\omega \in \mathbb{R} : \exists M \geq 1 \text{ s.t. } \|T(t)\| \leq Me^{\omega t} \ \forall t > 0\}.
\end{aligned}
$$

For a general strongly continuous semigroup it holds that $s(A) \leq \omega_0$ (see Exercise 12.3.2).

As the following example shows, in general $s(A)$ is strictly smaller than ω_0.

Example 12.1.3. Let X be the set of all functions $f \in C_0([0,\infty))$ such that $\displaystyle\int_0^\infty |f(x)|e^x dx < \infty$. It is a Banach space when endowed with the norm

$$
\|f\|_X = \|f\|_\infty + \int_0^\infty |f(x)|e^x dx
$$

On X we consider the semigroup $\{T(t)\}$ of left-translations.

We claim that $\{T(t)\}$ is strongly continuous. First of all, we observe that $T(t)$ maps X into itself. Indeed, it is obvious that $T(t)$ maps $C_0([0,\infty))$ into itself. Moreover, if $f \in X$, then

$$
\int_0^\infty |(T(t)f)(x)|e^x dx = \int_0^\infty |f(t+x)|e^x dx = e^{-t}\int_t^\infty |f(y)|e^y dy \leq \int_0^\infty |f(y)|e^y dy.
$$

Clearly, the sup-norm of $T(t)f$ converges to f as t tends to 0 uniformly in $[0,\infty)$ for every $f \in C_0([0,\infty))$, since this space is contained in $BUC([0,\infty))$. On the other hand, if $f \in X$, then we can estimate

$$
\int_0^\infty |(T(t)f)(x) - f(x)|e^x dx
$$

$$
\begin{aligned}
&=\int_0^{\infty} |f(t+x)-f(x)|e^x dx\\
&=\int_0^{M} |f(t+x)-f(x)|e^x dx+\int_M^{\infty} |f(t+x)-f(x)|e^x dx\\
&\le \|f(\cdot+t)-f\|_\infty(e^M-1)+\int_M^{\infty}|f(x+t)|e^x dx+\int_M^\infty |f(x)|e^x dx\\
&\le \|f(\cdot+t)-f\|_\infty(e^M-1)+e^{-t}\int_{M+t}^{\infty}|f(y)|e^y dx+\int_M^\infty |f(x)|e^x dx\\
&\le \|f(\cdot+t)-f\|_\infty(e^M-1)+2\int_M^\infty |f(x)|e^x dx
\end{aligned}
$$

for every $M>0$. Hence,

$$
\begin{aligned}
\limsup_{t\to 0^+}\int_0^\infty |(T(t)f)(x)-f(x)|e^x dx \le&(e^M-1)\lim_{t\to 0^+}\|f(t+\cdot)-f\|_\infty+2\int_M^\infty |f(x)|e^x dx\\
=&2\int_M^\infty |f(x)|e^x dx.
\end{aligned}
$$

Since the previous inequality holds true for every $M>0$, letting M tend to ∞, we conclude that

$$\limsup_{t\to 0^+}\int_0^\infty |(T(t)f)(x)-f(x)|e^x dx=0.$$

The strong continuity of $\{T(t)\}$ follows at once.

We now prove that $\omega_0=0$. The computations in the first part show that $\|T(t)\|\le 1$ for every $t>0$. Hence, $\omega_0\le 0$. To prove that, actually, $\omega_0=0$, it is enough to prove that $\|T(t)\|$ does not vanish as t tends to ∞. To this aim, for every $t>0$, we set $\delta_t=\log(1+e^{-t})$ and consider a continuous function f_t which vanishes in $[0,t]\cup[t+\delta_t,\infty)$, linearly increases to $1/2$ in $[t,t+\delta_t/2]$ and linearly decreases to 0 in $[t+\delta_t/2,t+\delta]$. Clearly, for every $t>0$, f_t belongs to $C_0([0,\infty))$ and $\|f_t\|_\infty=1/2$. Moreover,

$$\int_0^\infty |f_t(x)|e^x dx=\int_t^{t+\delta_t}|f(x)|e^x dx\le\frac{1}{2}\int_t^{t+\delta_t}e^x dx=\frac{1}{2}e^t(e^{\delta_t}-1)=\frac{1}{2},\qquad t>0.$$

Hence, f_t belongs to X for every $t>0$ and $\|f_t\|_X\le 1$. Moreover, $\|T(t)f_t\|_\infty=\|f_t(t+\cdot)\|_\infty=\|f_t\|_\infty=1/2$, since f_t vanishes in $[0,t]$. Hence,

$$\frac{1}{2}\le\|T(t)f_t\|_X\le\|T(t)\|\|f_t\|_X\le\|T(t)\|,\qquad t>0,$$

and this shows that the operator norm of $T(t)$ does not vanish as t tends to ∞.

Next, we observe that the infinitesimal generator of $\{T(t)\}$ is the operator $A: D(A)\subset X\to X$, defined by $Au=u'$ for every $u\in D(A)$ which is the set of all functions $u\in C_0^1([0,\infty))$ such that

$$\int_0^\infty (|u(x)|+|u'(x)|)e^x dx<\infty. \tag{12.2}$$

Indeed, if $u\in D(A)$, then the ratio $t^{-1}(u(\cdot+t)-u)$ should admit Au as limit in $C_0([0,\infty))$ as t tends to 0. This means that u is differentiable in $[0,\infty)$ and $Au=u'$.

Vice versa, let us assume that $u \in C_0^1([0,\infty))$ satisfies (12.2). Then,

$$\frac{(T(t)u)(x)-u(x)}{t}-u'(x)=\frac{1}{t}\int_0^t (u'(x+r)-u'(x))dr, \qquad x\geq 0,\ t>0.$$

From this formula and taking into account that u' is uniformly continuous in $[0,\infty)$, we easily deduce that the ratio $t^{-1}(T(t)u-u)$ converges to u' uniformly in $[0,\infty)$.

On the other hand,

$$\begin{aligned}\int_0^\infty \left|\frac{(T(t)u)(x)-u(x)}{t}-u'(x)\right|e^x dx =&\frac{1}{t}\int_0^\infty \left|\int_0^t (u'(x+r)-u'(x))dr\right|e^x dx\\ \leq&\frac{1}{t}\int_0^\infty e^x dx\int_0^t |u'(x+r)-u'(x)|dr\\ \leq&\frac{1}{t}\int_0^t dr\int_0^\infty |(T(r)u')(x)-u'(x)|e^x dx\\ \leq&\frac{1}{t}\int_0^t \|T(r)u'-u'\|_X dr\end{aligned}$$

and the last side of the previous chain of inequalities vanishes as t tends to 0, due to the strong continuity of the semigroup $\{T(t)\}$. We have so proved that $u \in D(A)$ as claimed.

Let us prove that $s(A)=-1$. For this purpose, we fix $\lambda\in\mathbb{C}$, with $\mathrm{Re}\lambda<-1$. Then, the function φ_λ, defined by $\varphi_\lambda(x)=e^{\lambda x}$ for every $x\in[0,\infty)$, belongs to $D(A)$ and $A\varphi_\lambda=\lambda\varphi_\lambda$, so that $\lambda\in\sigma_{\rm p}(A)$. On the other hand, if $\mathrm{Re}\lambda>-1$ then $\lambda\in\rho(A)$. This is clear if $\mathrm{Re}\lambda>0$ in view of the Hille-Yosida theorem (see Theorem 8.3.1), since $\{T(t)\}$ is a strongly continuous semigroup of contractions. Hence, let us assume that $\mathrm{Re}\lambda\in(-1,0]$ and consider the operator R_λ, defined by $R_\lambda f=\displaystyle\int_0^\infty e^{-\lambda t}T(t)f dt$ for every $f\in X$. It is well defined and bounded. Indeed,

$$|(R_\lambda f)(x)|\leq\int_0^\infty e^{-\mathrm{Re}\lambda t}|f(x+t)|dt=e^{\mathrm{Re}\lambda x}\int_x^\infty e^s|f(s)|ds\leq e^{\mathrm{Re}\lambda x}\|f\|_X$$

so that $\|R_\lambda f\|_\infty\leq\|f\|_X$.

Similarly,

$$\begin{aligned}\int_0^\infty |(R_\lambda f)(x)|e^x dx=&\int_0^\infty\left|\int_0^\infty e^{-\lambda t}f(t+x)dt\right|e^x dx\\ \leq&\int_0^\infty e^x dx\int_0^\infty e^{-\mathrm{Re}\lambda t}|f(t+x)|dt\\ =&\int_0^\infty e^{-\mathrm{Re}\lambda t}dt\int_0^\infty e^x|f(x+t)|dx\\ =&\int_0^\infty e^{-\mathrm{Re}\lambda t}dt\int_t^\infty e^{y-t}|f(y)|dy\\ \leq&\|f\|_X\int_0^\infty e^{-(1+\mathrm{Re}\lambda)x}dx=\frac{1}{1+\mathrm{Re}\lambda}\|f\|_X.\end{aligned}$$

We have so proved that R_λ belongs to $\mathcal{L}(X)$. Moreover, it is easy to check that R_λ is the inverse of the operator $\lambda I-A$ (see Exercise 12.3.1). Hence, $\lambda\in\rho(A)$.

Summing up, we have shown that $\sigma(A)=\{\lambda\in\mathbb{C}:\mathrm{Re}\lambda\leq-1\}$, so that $s(A)=-1$. Since $\omega_0=0$, the strict inequality $s(A)<\omega_0$ holds true.

For further use in this lecture, we state and prove the following equivalent characterization of the growth bound of a strongly continuous or analytic semigroup.

Lemma 12.1.4. *Let $\{T(t)\}$ be a strongly continuous or analytic semigroup with growth bound ω_0. Then,*

$$\omega_0 = \inf_{t>0} \frac{1}{t}\log(\|T(t)\|) = \lim_{t\to+\infty} \frac{1}{t}\log(\|T(t)\|), \tag{12.3}$$

where we set $\log(0) := -\infty$.

Proof. The proof is immediate if there exists $t_0 > 0$ such that $\|T(t_0)\| = 0$. So we assume that $\|T(t)\| > 0$ for every $t > 0$.

We begin by proving the first equality in (12.3). For this purpose, we denote by ω the infimum of the function $t^{-1}\log(\|T(t)\|)$ over $(0,\infty)$ and first prove that $\omega \le \omega_0$. Clearly, we just need to assume that $\omega \in \mathbb{R}$. Then, $\omega t \le \log(\|T(t)\|)$ for every $t > 0$ or, equivalently, $e^{\omega t} \le \|T(t)\|$ for every $t > 0$. If $r > \omega_0$, we know that $\|T(t)\| \le M_r e^{rt}$ for every $t > 0$ and some positive constant M_r. As a consequence, $e^{\omega t} \le M_r e^{rt}$ for every $t > 0$, which shows, first, that $\omega \le r$ and, then, that $\omega \le \omega_0$.

We now prove the inequality $\omega_0 \le \omega$. For this purpose, we fix $\tau > \omega$ and $t_1 > 0$ such that $\log(\|T(t_1)\|) \le \tau t_1$, i.e., $\|T(t_1)\| \le e^{\tau t_1}$. Take $t > t_1$ and let $n \in \mathbb{N}$ be such that $t = nt_1 + \sigma$ for some $\sigma \in [0, t_1)$. Then, using the semigroup law we can estimate

$$\begin{aligned}\|T(t)\| =&\|T(nt_1+\sigma)\| \le \|T(t_1)\|^n \|T(\sigma)\|\\ \le& Me^{\tau n t_1} = Me^{-\tau\sigma}e^{\tau t} \le Me^{|\tau| t_1}e^{\tau t},\end{aligned} \tag{12.4}$$

where $M = \sup_{t\in[0,t_1]} \|T(t)\|$. This shows that $\tau > \omega_0$. The arbitrariness of $\tau > \omega$ yields the inequality $\omega \ge \omega_0$.

To prove the other equality in (12.3), we observe that formula (12.4) also shows that

$$\limsup_{t\to\infty} \frac{1}{t}\log(\|T(t)\|) \le \omega_0. \tag{12.5}$$

On the other hand, since $\omega_0 t \le \log(\|T(t)\|)$ for every $t > 0$, we can also infer that

$$\liminf_{t\to\infty} \frac{1}{t}\log(\|T(t)\|) \ge \omega_0. \tag{12.6}$$

From (12.5) and (12.6), we conclude that the ratio $t^{-1}\log(\|T(t)\|)$ admits limit as t tends to ∞ and the limit is ω_0. □

Throughout this lecture we will use the following integral formula.

Lemma 12.1.5. *Let $\{T(t)\}$ be a strongly continuous (resp. analytic) semigroup on X with infinitesimal generator (resp. associated to the sectorial operator) A. Then, for every $\lambda \in \mathbb{C}$, for every $x \in X$ and $t > 0$ it holds that*

$$e^{-\lambda t}T(t)x - x = (A-\lambda I)\int_0^t e^{-\lambda s}T(s)x\,ds. \tag{12.7}$$

Further, if $x \in D(A)$, then

$$e^{-\lambda t}T(t)x - x = \int_0^t e^{-\lambda s}T(s)(Ax-\lambda x)\,ds. \tag{12.8}$$

Proof. It suffices to apply Propositions 8.2.2(ii) and 9.2.7 to the semigroup $\{e^{-\lambda t}T(t)\}$, which is a strongly continuous (resp. analytic) semigroup on X with infinitesimal generator (resp. associated to the sectorial operator) $A - \lambda$. □

12.1.1 Periodic semigroups

Definition 12.1.6. A strongly continuous semigroup $\{T(t)\}$ on X is a periodic semigroup if there exists $t_0 > 0$ such that $T(t_0) = I$.

The periodicity of the function $t \mapsto T(t)$ follows from the semigroup law since $T(t+t_0) = T(t)T(t_0) = T(t)I = T(t)$ for every $t > 0$.

Actually, periodic semigroups embed in groups since, by the definition, $T(t_0/2)$ is clearly invertible (see Exercise 12.3.1).

The periodicity of the function $t \mapsto \|T(t)\|$ implies that the growth bound ω_0 is equal to zero.

The Hille-Yosida theorem (see Theorem 8.3.1), now shows that the spectrum of the infinitesimal generator A of a periodic semigroup is contained in $i\mathbb{R}$. Actually, more can be said on the spectrum of A.

Lemma 12.1.7. *Let $\{T(t)\}$ be a periodic semigroup, with period τ, and let A be its infinitesimal generator. Then, the spectrum of A is contained in $2\pi\tau^{-1}i\mathbb{Z}$ and*

$$R(\lambda, A) = \frac{1}{1 - e^{-\lambda\tau}} \int_0^\tau e^{-\lambda s} T(s) ds$$

for every $\lambda \notin 2\pi\tau^{-1}i\mathbb{Z}$. In particular, $\sigma(A)$ is a discrete subset of $\mathbb{C}$.

Proof. Using formulas (12.7) and (12.8) with $t = \tau$ and recalling that $T(\tau) = I$, we can write

$$(e^{-\lambda\tau} - 1)x = (A - \lambda I) \int_0^\tau e^{-\lambda s} T(s) x ds,$$

if $x \in X$ and, similarly,

$$(e^{-\lambda\tau} - 1)x = \int_0^\tau e^{-\lambda s} T(s)(Ax - \lambda x) ds,$$

if $x \in D(A)$. Since, for every $\lambda \notin 2\pi\tau^{-1}i\mathbb{Z}$ the operator $R_\lambda : X \to X$, defined by

$$R_\lambda x = \frac{1}{1 - e^{-\lambda\tau}} \int_0^\tau e^{-\lambda s} T(s) x ds, \qquad x \in X,$$

is bounded, from the previous two formulas we can infer that, if $\lambda \notin 2\pi\tau^{-1}i\mathbb{Z}$, then R_λ is the bounded inverse of the operator $\lambda I - A$. Hence, $\lambda \in \rho(A)$ and the proof is complete. □

Remark 12.1.8. Since the spectrum of the infinitesimal generator of a periodic strongly continuous semigroup is discrete, we can enumerate its elements of $\sigma(A)$ and set $\sigma(A) = \{\lambda_n : n \in J\}$, where J is a subset of $\mathbb{N}$, possibly coinciding with the whole $\mathbb{N}$. In view of Lemma 12.1.7, for every $n \in J$ we can set

$$P_n x = \lim_{\lambda \to \lambda_n} (\lambda - \lambda_n) R(\lambda, A) = \lim_{\lambda \to \lambda_n} \frac{\lambda - \lambda_n}{1 - e^{-\lambda\tau}} \int_0^\tau e^{-\lambda s} T(s) x ds = \frac{1}{\tau} \int_0^\tau e^{-\lambda_n s} T(s) x ds.$$

Each operator P_n is a projection. Indeed,

$$\begin{aligned}P_n^2x =&\frac{1}{\tau}\int_0^\tau e^{-\lambda_n s}T(s)P_nxds = \frac{1}{\tau^2}\int_0^\tau e^{-\lambda_n s}T(s)\bigg(\int_0^\tau e^{-\lambda_n t}T(t)xdt\bigg)ds\\ =&\frac{1}{\tau^2}\int_0^\tau ds\int_s^{\tau+s} e^{-\lambda_n\sigma}T(\sigma)xdt = \frac{1}{\tau^2}\int_0^\tau ds\int_0^\tau e^{-\lambda_n\sigma}T(\sigma)xdt\\ =&\frac{1}{\tau}\int_0^\tau e^{-\lambda_n\sigma}T(\sigma)xdt = P_nx,\end{aligned}$$

where we have used the fact that the map $\sigma \mapsto e^{-\lambda_n\sigma}T(\sigma)x$ is τ-periodic.

Clearly, the range of P_n is contained in $D(A)$. Actually, $P_n(X) = \text{Ker}(\lambda_n I - A)$ (see Exercise 12.3.5).

Finally, we observe that, if $n, m \in J$, with $m \neq n$, then $P_nP_m = 0$. For this purpose, we write

$$\begin{aligned}P_nP_mx =&\frac{1}{\tau}\int_0^\tau e^{-\lambda_n s}T(s)P_mxds = \frac{1}{\tau^2}\int_0^\tau e^{(\lambda_m-\lambda_n)s}ds\int_0^\tau e^{-\lambda_m(t+s)}T(s+t)xdt\\ =&\frac{1}{\tau^2}\int_0^\tau e^{(\lambda_m-\lambda_n)s}ds\int_s^{\tau+s} e^{-\lambda_m\sigma}T(\sigma)xdt = \bigg(\frac{1}{\tau}\int_0^\tau e^{(\lambda_m-\lambda_n)s}ds\bigg)P_mx = 0\end{aligned}$$

for every $x \in X$, since $\int_0^\tau e^{(\lambda_m-\lambda_n)s}ds = 0$ for $m \neq n$.

Based on this remark, we can show that, for a periodic semigroup, the spectrum coincides with the point spectrum.

Theorem 12.1.9. *Let $\{T(t)\}$ be a strongly continuous semigroup and let $A : D(A) \subset X \to X$ be its infinitesimal generator. The following properties are equivalent.*

(i) *The semigroup is periodic;*

(ii) *$\sigma(A) = \sigma_{\rm p}(A) \subset 2\pi\alpha i\mathbb{Z}$ for some $\alpha > 0$. Moreover, the closure of the linear span of all the eigenvectors of A is dense in X.*

Proof. (ii)⇒(i) We fix $n \in \mathbb{N}$ such that $2\pi\alpha ni \in \sigma_{\rm p}(A)$ and let $0 \neq x \in D(A)$ be such that $Ax = 2\pi\alpha nix$. As it is easily seen, $T(t)x = e^{2\pi\alpha nit}x$ for every $t > 0$. We claim that $T(\alpha^{-1}) = I$. For this purpose, we observe that, since the span of all the eigenvectors of A is dense in X, for every $x \in X$ we can determine a sequence $\{x_n\}_{n\in\mathbb{N}} \subset D(A)$ converging to x. Each x_n has the form $x_n = \sum_{k=1}^n \beta_k e_k$ for some constant $\beta_k \in \mathbb{C}$, $e_1, \ldots, e_k \in D(A)$ such that $Ae_k = 2\pi\alpha n_k ie_k$ for every k and a suitable $n_k \in \mathbb{N}$. Since $T(t)$ is a bounded operator and x_n converges to x, it follows that $T(t)x_n$ converges to $T(t)x$ as n tends to ∞ for every $t > 0$. In particular, $T(\alpha)x_n$ converges to $T(\alpha)x$. Since

$$T(\alpha^{-1})x_n = \sum_{k=1}^n \beta_k T(\alpha^{-1})e_k = \sum_{k=1}^n \beta_k e^{2i\pi\alpha n_k\alpha^{-1}}e_k = \sum_{k=1}^n \beta_k e_k = x_n, \qquad n \in \mathbb{N},$$

we conclude that $T(\alpha^{-1}) = I$.

(i)⇒(ii) We begin by showing that $\sigma(A) = \sigma_{\rm p}(A)$. Let τ be the period of the semigroup. By Lemma 12.1.7, we know that the spectrum of A is discrete and contained in $2\pi\tau^{-1}i\mathbb{Z}$. Fix $m \in \mathbb{Z}$ such that $\lambda_m := 2\pi\tau^{-1}mi$ belongs to $\sigma(A)$ and consider the projection P_m, which has $\text{Ker}(\lambda_m I - A)$ as image (see Exercise 12.3.5).

Let us prove that the span of all the eigenvectors of A is dense in X. By contradiction, we assume that this is not the case. Then, there exists a nontrivial functional $x' \in X'$ such that $x'(P_n x) = 0$ for every $n \in \mathbb{N}$ (note that if $2\pi\tau^{-1}ni \notin \sigma(A)$, then P_n is the null operator). Therefore,

$$0 = \langle x', P_n x\rangle = \left\langle x', \frac{1}{\tau}\int_0^\tau e^{-2\pi\tau^{-1}nis}T(s)x ds\right\rangle = \frac{1}{\tau}\int_0^\tau e^{-2\pi\tau^{-1}nis}\langle x', T(s)x\rangle ds$$

for every $n \in \mathbb{N}$. This shows that all the Fourier coefficients of the τ-periodic vector-valued function $s \mapsto \langle x', T(s)x\rangle$ vanish, showing that, $\langle x', T(s)x\rangle = 0$ for every $s \in [0,\tau]$. In particular, $\langle x', x\rangle = 0$ for every $x \in X$, which, clearly, is a contradiction since x' is not the trivial functional. □

We can now prove a spectral representation formula for periodic semigroups and their infinitesimal generators.

Theorem 12.1.10. *Let $\{T(t)\}$ be a strongly continuous periodic semigroup with period $\tau > 0$ and let $A : D(A) \subset X \to X$ be its infinitesimal generator. Then,*

$$T(t)x = \sum_{n\in\mathbb{Z}} e^{\lambda_n t}P_n x, \qquad x \in D(A),\ t > 0, \tag{12.9}$$

$$Ax = \sum_{n\in\mathbb{Z}} \lambda_n P_n x, \qquad x \in D(A^2), \tag{12.10}$$

where $\lambda_n = 2\pi\tau^{-1}in$ for every $n \in \mathbb{N}$.

Proof. Let us prove that, for every $x \in D(A)$, the sequence $\{\sum_{|k|\le n} P_k x\}_{n\in\mathbb{N}}$ converges in X (and it converges to x). For this purpose, we fix $x \in D(A)$ and observe that

$$P_n Ax = AP_n x = \lambda_n P_n x = (2\pi n\tau^{-1}i)P_n x.$$

Let us fix $m \in \mathbb{N}$, $x' \in X'$, with $\|x'\| = 1$ and estimate

$$\bigg|\sum_{|n|>m}\langle x', P_n x\rangle\bigg| = \bigg|\sum_{|n|>m}\frac{\tau i}{2\pi n}\langle x', P_n Ax\rangle\bigg| \le \frac{\tau}{2\pi}\bigg(\sum_{|n|>m} n^{-2}\bigg)^{\frac12}\bigg(\sum_{|n|>m}|\langle x', P_n Ax\rangle|^2\bigg)^{\frac12}.$$

We observe that $\{\langle x', P_n x\rangle\}_{n\in\mathbb{N}}$ is the sequence of the Fourier coefficients of the function $s \mapsto \langle x', T(s)x\rangle$. Hence, applying Parseval's inequality, we can infer that

$$\sum_{|n|>m}\frac{1}{\tau}\int_0^\tau e^{-2\pi\tau^{-1}nis}\langle x', T(s)Ax\rangle ds \le \frac{1}{\tau}\int_0^\tau |\langle x', T(s)Ax\rangle|^2 ds,$$

so that

$$\begin{aligned}\bigg|\sum_{|n|>m}\langle x', P_n x\rangle\bigg| \le& \frac{\tau}{2\pi}\bigg(\sum_{|n|>m} n^{-2}\bigg)^{\frac12}\bigg(\frac{1}{\tau}\int_0^\tau |\langle x', T(s)Ax\rangle|^2 ds\bigg)^{\frac12}\\ \le& \frac{\tau}{2\pi}\bigg(\sum_{|n|>m} n^{-2}\bigg)^{\frac12}\bigg(\frac{1}{\tau}\int_0^\tau |T(s)Ax|^2 ds\bigg)^{\frac12}.\end{aligned}$$

Taking the supremum with respect to all $x' \in X'$ with $\|x'\| = 1$, we conclude that

$$\Big\| \sum_{|n|>m} P_n x \Big\| \le \frac{\tau}{2\pi} \bigg(\sum_{|n|>m}^{\infty} n^{-2} \bigg)^{\frac{1}{2}} \bigg(\frac{1}{\tau} \int_0^\tau |T(s)Ax|^2 ds \bigg)^{\frac{1}{2}}$$

and the convergence of $\sum_{k=-n}^{n} P_k x$ follows. We now prove that it actually converges to x. For this purpose, we set $z = \sum_{n\in\mathbb{Z}} P_n x$ and observe that, since $P_n P_m x = 0$ for every $n \ne m$ (see Remark 12.1.8), it follows that $P_n x = P_n z$ for every $n \in \mathbb{N}$. Thus, for every $x' \in X'$, the functions $\langle x', T(\cdot)x\rangle$ and $\langle x', T(\cdot)z\rangle$ have the same Fourier coefficients and, as a byproduct, they coincide. In particular $\langle x', x\rangle = \langle x', T(0)x\rangle = \langle x', T(0)z\rangle = \langle x', z\rangle$. It thus follows that $x = z$.

We can now complete the proof. For this purpose, we observe that

$$T(t) \sum_{k=-n}^{n} P_k x = \sum_{k=-n}^{n} T(t) P_n x = \sum_{k=-n}^{n} e^{\lambda_n t} P_n x,$$

$$A \sum_{k=-n}^{n} P_k x = \sum_{k=-n}^{n} A P_n x = \sum_{k=-n}^{n} \lambda_n P_n x$$

for every $t > 0$ and $n \in \mathbb{N}$. Since $\sum_{k=-n}^{n} P_k x$ converges to x as n tends to ∞ and the operator $T(t)$ is bounded for every $t > 0$, formula (12.9) follows at once.

As far as (12.10) is concerned, we observe that, if $x \in D(A^2)$, then $P_n A^2 x = A^2 P_n x = \lambda_n A P_n x = \lambda_n^2 P_n x$ for every $n \in \mathbb{N}$. Hence, $A P_n x = \lambda_n^{-1} P_n A^2 x$ for every $n \in \mathbb{N}$. Now the arguments we used to prove the convergence of the series $\sum_{n\in\mathbb{Z}} \langle x', P_n x\rangle$ show that $A \sum_{|k|\le n} P_k x$ converges is X as n tends to ∞. Since A is closed, formula (12.10) follows at once. This completes the proof. □

12.1.2 Dual semigroups

Let $\{T(t)\}$ be a strongly continuous semigroup in X with infinitesimal generator A. The family $\{T'(t)\}$, where $T'(t)$ is the dual operator of $T(t)$ (see Section 1.1), is a semigroup of operators on X'. Indeed, for every $s, t > 0$, $x' \in X'$ and $x \in X$, it holds that

$$\langle T'(t+s)x', x\rangle = \langle x', T(t+s)x\rangle = \langle x', T(t)T(s)x\rangle = \langle T'(t)x', T(s)x\rangle = \langle T'(s)T'(t)x', x\rangle.$$

The semigroup $\{T'(t)\}$ is called the dual semigroup of $\{T(t)\}$ and,in general, is not strongly continuous in X'.

Example 12.1.11. Let $\{T(t)\}$ the semigroup of left translations on $X = L^1(\mathbb{R})$ (see Example 8.1.8). We know that this semigroup is strongly continuous Via the identification of X' with $L^\infty(\mathbb{R})$, the dual semigroup $\{T'(t)\}$ is the semigroup of right translations on $L^\infty(\mathbb{R})$. Indeed, for every $f \in L^\infty(\mathbb{R})$, $g \in L^1(\mathbb{R})$ and $t > 0$, we can write

$$\langle T'(t)f, g\rangle = \int_{\mathbb{R}} f(x)(T(t)g)(x)dx = \int_{\mathbb{R}} f(y-t)g(y)dx.$$

The arbitrariness of $g \in L^1(\mathbb{R})$ implies that $T'(t)f = f(\cdot - t)$.

Note that the semigroup of right-translation is not strongly continuous in $L^\infty(\mathbb{R})$. Indeed, fix a bounded and continuous function $f : \mathbb{R} \to \mathbb{C}$. Then $\|T'(t)f - f\|_\infty$ converges to 0 as t tends to 0 if and only if f is uniformly continuous. (Actually, Lotz proved in [13] that a strongly continuous semigroup on a L^∞-space is forced to be uniformly continuous).

The semigroup $\{T'(t)\}$ becomes strongly continuous if we restrict it to a suitable subset of X' as the following lemma shows.

Lemma 12.1.12. *Let $\widetilde{X}' = \{x' \in X' : \lim_{t\to 0} T'(t)x' = x' \text{ in } X'\}$, endowed with the norm of X'. Then, each operator $T'(t)$ maps $\widetilde{X}'$ into itself and the restriction of the semigroup $\{T'(t)\}$ to $\widetilde{X}'$ (still denoted by $\{T'(t)\}$) is strongly continuous. Moreover, $\widetilde{X}'$ is a closed subspace of X'. Finally, $D(A') \subset \widetilde{X}'$ (see 8.3.7 for the definition of A').*

Proof. Fix $x' \in \widetilde{X}'$ and $s > 0$. Then,

$$\lim_{t\to 0}(T'(t)T'(s)x' - T'(s)x') = T'(s)\left(\lim_{t\to 0} T'(t)x' - x'\right) = 0,$$

where the limits are taken in X'. This shows that $T(s)x'$ belongs to $\widetilde{X}'$. It is clear from the definition, that $\widetilde{X}'$ is the largest subspace of X' of strong continuity of the semigroup $\{T'(t)\}$.

Let us now prove that $\widetilde{X}'$ is a closed subspace of X'. We fix $x' \in X'$ and a sequence $\{x'_n\}_{n\in\mathbb{N}} \subset \widetilde{X}'$ such that $\lim_{n\to\infty} x'_n = x'$. By assumptions, for every $n \in \mathbb{N}$, $T'(t)x'_n$ converges to x'_n in X' as t tends to 0. Based on this fact, we split $T'(t)x' - x' = T'(t)(x' - x'_n) + (T'(t)x'_n - x'_n) + (x'_n - x')$ and estimate

$$\begin{aligned}\|T'(t)x' - x'\|_{X'} \le& \|T'(t)(x' - x'_n)\|_{X'} + \|T'(t)x'_n - x'_n\|_{X'} + \|x'_n - x'\|_{X'}\\ \le&(\|T(t)\| + 1)\|x'_n - x'\|_{X'} + \|T'(t)x'_n - x'_n\|_{X'}\\ \le&(Me^{\omega t} + 1)\|x'_n - x'\|_{X'} + \|T'(t)x'_n - x'_n\|_{X'}\end{aligned}$$

so that

$$\begin{aligned}\limsup_{t\to 0} \|T'(t)x' - x'\|_{X'} \le&(M + 1)\|x'_n - x'\|_{X'} + \lim_{t\to 0} \|T'(t)x'_n - x'_n\|_{X'}\\ =&(M + 1)\|x'_n - x'\|_{X'}\end{aligned}$$

for every $n \in \mathbb{N}$. Letting n tend to ∞, we obtain that $\limsup_{t\to 0} \|T'(t)x' - x'\|_{X'} = 0$ so that $T'(t)x'$ converges to x' in X' as t tends to 0. Hence, $x' \in \widetilde{X}'$.

To conclude the proof, let us show that $D(A') \subset \widetilde{X}'$. For this purpose, we fix $x' \in D(A')$, $x \in X$ and observe that

$$\langle T'(t)x' - x', x\rangle = \langle x', T(t)x - x\rangle = \left\langle x', A\int_0^t T(s)xds\right\rangle = \left\langle A'x', \int_0^t T(s)xds\right\rangle.$$

Hence,

$$|\langle T'(t)x' - x', x\rangle| \le \sup_{t\in[0,1]} \|T(t)\|\|A'x'\|_{X'}\|x\|t, \qquad t \in (0, 1],$$

or, equivalently,

$$\|T'(t)x' - x'\|_{X'} \le \sup_{t\in[0,1]} \|T(t)\|\|A'x'\|_{X'}t, \qquad t \in (0, 1].$$

Hence, $T'(t)x'$ converges to x' in X' as t tends to 0. The proof is complete. □

A natural question arises: which operator is the infinitesimal generator of the restriction of the dual semigroup to $\widetilde{X}'$. This question is answered by the following proposition.

Proposition 12.1.13. *The infinitesimal generator of the restriction of the semigroup* $\{T'(t)\}$ *to* $\widetilde{X}'$ *is the part of* A' *in* $\widetilde{X}'$, *i.e., the operator* $\widetilde{A}: D(\widetilde{A}) \subset X' \to X'$, *defined by* $\widetilde{A}x' = A'x'$ *for every* $x' \in D(\widetilde{A}) = \{x' \in D(A') : A'x' \in \widetilde{X}'\}$.

Proof. Fix $x' \in D(\widetilde{A})$. Then,

$$\lim_{t\to 0} \frac{T'(t)x' - x'}{t} = \widetilde{A}x',$$

where the limit is taken in X', so that

$$\lim_{t\to 0} \left\langle \frac{T'(t)x' - x'}{t}, x \right\rangle = \langle \widetilde{A}x', x\rangle$$

for every $x \in X$. In particular, if $x \in D(A)$, then

$$\lim_{t\to 0} \left\langle \frac{T'(t)x' - x'}{t}, x \right\rangle = \lim_{t\to 0} \left\langle x', \frac{T(t)x - x}{t} \right\rangle = \langle x', Ax\rangle.$$

From the previous two formulas, it follows that $\langle \widetilde{A}x', x\rangle = \langle x', Ax\rangle$ for every $x \in D(A)$. Hence, $x' \in D(A')$ and $A'x' = \widetilde{A}x'$. The inclusion $D(\widetilde{A}) \subset \{x' \in D(A') : A'x' \in \widetilde{X}'\}$ follows.

Let us now suppose that $x' \in D(A')$ is such that $A'x \in \widetilde{X}'$. Then, for every $t > 0$ and $x \in X$ we can write

$$\begin{aligned}\langle T'(t)x' - x', x\rangle =& \langle x', T(t)x - x\rangle = \left\langle x', A\int_0^t T(s)xds \right\rangle = \left\langle A'x', \int_0^t T(s)xds \right\rangle \\ &= \int_0^t \langle A'x', T(s)x\rangle ds = \int_0^t \langle T'(s)A'x', x\rangle ds = \left\langle \int_0^t T'(s)A'x'ds, x \right\rangle.\end{aligned}$$

The arbitrariness of $x \in X$, shows that

$$T'(t)x' - x' = \int_0^t T'(s)A'x'ds, \qquad t > 0. \tag{12.11}$$

Since, by assumptions, the function $T(\cdot)A'x'$ is continuous at 0 and, hence, in $[0, \infty)$, dividing both sides of (12.11) by t and, then, letting t tend to 0, we conclude that $x' \in D(\widetilde{A})$ and $\widetilde{A}x' = A'x'$. This completes the proof. □

Remark 12.1.14. Since $D(A') \subset \widetilde{X}'$ and $\overline{D(\widetilde{A})} = \widetilde{X}'$, it follows that

$$\overline{D(A')} = \widetilde{X}'.$$

The following result, which will be used in the next section, relates the residual spectrum of operator A to the point spectrum of operator $\widetilde{A}$.

Theorem 12.1.15. *The residual spectrum of the operator* A *coincides with the point spectrum of the operator* $\widetilde{A}$.

Proof. Let us begin by proving the inclusion $\sigma_{\rm p}(\widetilde{A}) \subseteq \sigma_{\rm res}(A)$. For this purpose, we fix $\lambda \in \sigma_{\rm p}(\widetilde{A})$ and $0 \neq x' \in D(\widetilde{A})$ such that $\widetilde{A}x' = A'x' = \lambda x'$. Then,

$$0 = \langle \lambda x' - A'x', x\rangle = \langle x', \lambda x - Ax\rangle, \qquad x \in D(A),$$

and, we conclude that x' identically vanishes on $\overline{(\lambda I - A)(X)}$. Since $x' \neq 0$, $\overline{(\lambda I - A)(X)}$ should be a proper subset of X and, consequently, $\lambda \in \sigma_{\rm res}(A)$.

Vice versa, fix $\lambda \in \sigma_{\rm res}(A)$. Since $\overline{(\lambda I - A)(X)}$ is a proper subspace of X, Hahn-Banach theorem shows that there exists a functional $0 \neq x' \in X'$ which identically vanishes on $\overline{(\lambda I - A)(X)}$. In particular, $\langle x', \lambda x - Ax\rangle = 0$ for every $x \in D(A)$. Arguing as in the first part of the proof, we conclude that $x' \in D(A')$ and $A'x' = \lambda x'$. Since $D(A') \subset \widetilde{X}'$ (see Lemma 12.1.12), we deduce that $A'x' \in D(\widetilde{A}')$. □

Remark 12.1.16. We stress that $\sigma_{\rm p}(\widetilde{A}) = \sigma_{\rm p}(A')$ (see Exercise 12.3.7).

12.2 Spectral mapping theorem

For a general strongly continuous semigroup the following is true.

Proposition 12.2.1. *Let $\{T(t)\}$ be a strongly continuous semigroup on X and let $A : D(A) \subset X \to X$ be its infinitesimal generator. Then, $\sigma(T(t)) \supset e^{t\sigma(A)}$ for every $t \geq 0$. Moreover,*

$$\sigma_{\rm p}(T(t)) \setminus \{0\} = e^{t\sigma_{\rm p}(A)}, \tag{12.12}$$

$$\sigma_{\rm res}(T(t)) \setminus \{0\} = e^{t\sigma_{\rm res}(A)}, \tag{12.13}$$

$$\sigma_{\rm a}(T(t)) \supseteq e^{t\sigma_{\rm a}(A)} \tag{12.14}$$

for every $t > 0$.

Proof. We split the proof into three steps. In the first one we prove that $e^{t\sigma(A)} \subseteq \sigma(T(t))$ for every $t > 0$ and that (12.14) and the inclusions "$\supseteq$" in (12.12) and (12.13) are satisfied. Then, in Steps 2 and 3, we complete the proof of (12.12) and (12.13).

Step 1. Suppose that $\lambda \in \sigma(A)$. Then, using (12.7), (12.8) and the fact that the operator $R_{\lambda,t}$, defined by $R_{\lambda,t}x = \displaystyle\int_0^t e^{\lambda(t-s)}T(s)x ds$ for every $x \in X$, and $A - \lambda I$ commute on $D(A)$, it follows immediately that the operator $e^{\lambda t} - T(t)$ is not invertible, so that $e^{\lambda t}$ belongs to $\sigma(T(t))$.

Next, we observe that, if $0 \neq x \in D(A)$ is such that $Ax - \lambda x = 0$, then $e^{\lambda t}x - T(t)x = 0$. This shows that $e^{t\sigma_{\rm p}(A)} \subset \sigma_{\rm p}(T(t))$ for every $t > 0$.

To prove (12.14), we fix $\lambda \in \sigma_{\rm a}(A)$. Then, there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset D(A)$ such that $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $Ax_n - \lambda x_n$ converges to 0 as n tends to ∞. As a consequence, $e^{\lambda t}x_n - T(t)x_n = R_{\lambda,t}(Ax_n - \lambda x_n)$ converges to 0 as n tends to ∞. The inclusion $e^{t\sigma_{\rm a}(A)} \subset \sigma_{\rm a}(T(t))$ follows for every $t > 0$.

Finally, suppose that $\lambda \in \sigma_{\rm res}(A)$. Then, the range of the operator $A - \lambda I$ is not dense in X. Since formula (12.7) shows that the range of the operator $e^{\lambda t}I - T(t)$ is contained in the range of the operator $A - \lambda I$, the range of $e^{\lambda t}I - T(t)$ is not dense in X, so that $\lambda \in \sigma_{\rm res}(T(t))$.

Step 2. Here, we complete the proof of (12.12). We fix $t_0 \in (0,\infty)$, $\lambda \in \sigma_{\rm p}(T(t_0)) \setminus \{0\}$ and consider the rescaled semigroup $\{S_{t_0,-\log(\lambda)}(t)\}$, where $\log(\lambda)$ denotes any arbitrarily fixed determination of the complex logarithm of λ and $S_{\alpha,\mu}(t) := e^{\mu t}T(\alpha t)$ for $t \geq 0$, $\mu \in \mathbb{C}$ and $\alpha > 0$. The generator of this rescaled semigroup is the operator $A_{t_0,-\log\lambda} = t_0 A - \log\lambda$ (see Exercise 12.3.3). Moreover, since $T(t_0)x = \lambda x$ for some $x \in X \setminus \{0\}$, it follows that $S_{t_0,-\log\lambda}(1)x = e^{-\log\lambda}T(t_0)x = x$. Hence, 1 is an eigenvalue of $S_{t_0,-\log\lambda}(1)$.

Let us set $Y = \mathrm{Ker}(S_{t_0,-\log\lambda}(1) - I)$. As it is easily seen, the semigroup $\{S_{t_0,-\log\lambda}\}$ leaves Y invariant. Indeed, if $y \in Y$ and $t > 0$, then $S_{t_0,-\log\lambda}(1)S_{t_0,-\log\lambda}(t)y = S_{t_0,-\log\lambda}(t)S_{t_0,-\log\lambda}(1)y = S_{t_0,-\log\lambda}(t)y$. Moreover, since $S_{t_0,-\log\lambda}(1) = I_Y$, $\{S_{t_0,-\log\lambda}\}$ is a periodic strongly continuous semigroup on Y. Let τ denote the period of this semigroup. Then, there exists $n \in \mathbb{N}$ such that $\tau = n^{-1}$. From Theorem 12.1.9, it follows that there exists $m \in \mathbb{N}$ such that $2\pi\tau^{-1}mi = 2\pi mni$ belongs to the point spectrum of the part of the operator $A_{t_0,-\log\lambda}$ in Y, i.e., the set of all $x \in D(A)$ such that $t_0Ax - (\log\lambda)x$ belongs to Y. Since $1 = e^{2\pi mni}$ and the point spectrum of the part of $A_{t_0,-\log\lambda}$ is contained in $\sigma_{\mathrm{p}}(A_{t_0,-\log\lambda})$, we conclude that $1 \in e^{\sigma_{\mathrm{p}}(A_{t_0,-\log\lambda})}$. This is enough to conclude that $\lambda \in e^{t_0\sigma_{\mathrm{p}}(A)}$.

Step 3. Here, we complete the proof, by showing (12.13). For this purpose, we introduce the restriction $\{\widetilde{T}'(t)\}$ of the adjoint semigroup $\{T'(t)\}$ to $\widetilde{X}'$. By Step 2, we know that $\sigma_{\mathrm{p}}(\widetilde{T}'(t)) \setminus \{0\} = e^{t\sigma_{\mathrm{p}}(\widetilde{A})}$, where $\widetilde{A}$ is the infinitesimal generator of $\{\widetilde{T}'(t)\}$. By Theorem 12.1.15, $\sigma_{\mathrm{p}}(\widetilde{A}) = \sigma_{\mathrm{res}}(A)$ and, applying the same arguments in the proof of the quoted theorem, we can easily verify that $\sigma_{\mathrm{p}}(\widetilde{T}'(t)) = \sigma_{\mathrm{res}}(T(t))$ for every $t > 0$. Combining all these fact, the proof of (12.13) follows. □

From Proposition 12.2.1 it follows that the approximate spectrum of the operator A is responsible for the validity of the spectral mapping theorem. Clearly, the spectral mapping theorem holds true if one of the following conditions is satisfied:

(i) the spectrum of $T(t)$ consists only of the point and the residual spectra for every $t > 0$;

(ii) $\sigma_{\mathrm{a}}(T(t)) \setminus \{0\} = e^{t\sigma_{\mathrm{a}}(A)}$ for every $t > 0$.

To investigate the validity of the second property listed here above, we consider the following proposition.

Proposition 12.2.2. *Let $\{T(t)\}$ be a strongly continuous semigroup on X and let $\lambda \in \mathbb{C}\setminus\{0\}$ belong to the approximate spectrum of the operator $T(t_0)$ for some $t_0 > 0$. The following properties are equivalent.*

(i) *There exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$ such that*

  (a) $\|x_n\| = 1$ *for every* $n \in \mathbb{N}$*;*

  (b) $T(t_0)x_n - \lambda x_n$ *converges to zero in* X *as* n *tends to* ∞*;*

  (c) $\sup_{n\in\mathbb{N}} \|T(t)x_n - x_n\|$ *vanishes as* t *tends to* 0*;*

(ii) *there exists $\mu \in \sigma_{\mathrm{a}}(A)$ such that $\lambda = e^{\mu t_0}$.*

Proof. "(ii)⇒(i)". Fix a sequence $\{x_n\}_{n\in\mathbb{N}} \subset D(A)$ such that $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $Ax_n - \mu x_n$ vanishes as n tends to ∞. Applying formula (12.8) with x, t and λ being replaced by x_n, t_0 and μ respectively, we get

$$\lambda x_n - T(t_0)x_n = \int_0^{t_0} e^{\mu(t_0-s)}T(s)(Ax_n - \mu x_n)ds, \qquad n \in \mathbb{N}.$$

From this formula we can infer that $\|\lambda x_n - T(t_0)x_n\| \le C\|Ax_n - \mu x_n\|$ for some positive constant C and every $n \in \mathbb{N}$. We have so proved that λ is an approximate eigenvalue of

$T(t_0)$ and properties (i)-(a) and (i)-(b) follows. To prove property (i)-(c), we still use formula (12.8), with t instead of t_0 to infer that

$$T(t)x_n - x_n = (e^{t\mu} - 1)x_n - \int_0^t e^{\mu(t-s)}T(s)(Ax_n - \mu x_n)ds, \qquad t > 0, \ n \in \mathbb{N},$$

so that

$$\|T(t)x_n - x_n\| \leq |e^{t\mu} - 1|\|x_n\| + \int_0^t e^{\mathrm{Re}\mu(t-s)}\|T(s)(Ax_n - \mu x_n)\|ds \tag{12.15}$$

$$\leq |e^{t\mu} - 1| + \sup_{t\in[0,1]}\|T(t)\|\sup_{n\in\mathbb{N}}\|Ax_n - \mu x_n\|\int_0^t e^{\mathrm{Re}\mu s}ds \tag{12.16}$$

and the right-hand side of (12.16) converges to zero as t tends to 0, yielding property (i)-(c).

"(i)⇒(ii)". Up to properly rescaling the semigroup, if needed, we can assume that $\lambda = 1$ and $t_0 = 1$. Let $\{x_n\}_{n\in\mathbb{N}} \subset X$ be a sequence such that $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $T(1)x_n - x_n$ converges to 0 as n tends to ∞.

For every $n \in \mathbb{N}$, fix a functional $x_n' \in X'$ such that $\|x_n'\| \leq 1$ and $\langle x_n', x_n\rangle \geq 1/2$. Further, consider the functions $f_n : [0,1] \to C$, defined by $f_n(t) = \langle x_n', T(t)x_n\rangle$ for every $t \in [0,1]$ and $n \in \mathbb{N}$. We claim that the set $\mathcal{F} = \{f_n : n \in \mathbb{N}\}$ is an equicontinuous and equibounded subset of $C([0,1]; X)$. To prove the claim, we observe that

$$|f_n(t)| \leq \|x_n'\|\|T(t)x_n\| \leq \sup_{t\in[0,1]}\|T(t)\|\|x_n'\|\|x_n\| \leq \sup_{t\in[0,1]}\|T(t)\|, \qquad t \in [0,1],$$

for every $n \in \mathbb{N}$. Moreover, for every $t_0 \in [0,1)$ and $t > t_0$, we can estimate

$$\begin{aligned}|f_n(t) - f_n(t_0)| \leq &\|x_n'\|\|T(t)x_n - T(t_0)x_n\| \leq \|T(t_0)\|\|T(t-t_0)x_n - x_n\| \\ \leq &\|T(t_0)\|\sup_{n\in\mathbb{N}}\|T(t-t_0)x_n - x_n\|.\end{aligned}$$

Similarly, if $t_0 \in (0,1]$ and $t < t_0$, we can estimate

$$|f_n(t) - f_n(t_0)| \leq \|T(t)\|\|T(t_0 - t)x_n - x_n\| \leq (\sup_{s\in[0,1]}\|T(s)\|)\sup_{n\in\mathbb{N}}\|T(t_0-t)x_n - x_n\|.$$

From the previous two chain of inequalities, the equicontinuity of $\mathcal{F}$ follows.

We can thus invoke Arzelà-Ascoli theorem to infer that there exists $g \in C([0,1]; \mathbb{C})$ such that, up to a subsequence, $\{f_n\}_{n\in\mathbb{N}}$ converges to g uniformly on $[0,1]$. Note that g is not the trivial function since $f_n(0) = \langle x_n', T(0)x_n\rangle = \langle x_n', x_n\rangle \geq 1/2$. Therefore, there exists $m \in \mathbb{Z}$ such that the m-th Fourier coefficient of g differs from zero, i.e.,

$$\int_0^1 e^{-2\pi mit}g(t)dt \neq 0.$$

Let us set $z_n = \displaystyle\int_0^1 e^{-2\pi mit}T(t)x_n dt$ for every $n \in \mathbb{N}$. Clearly, $z_n \in D(A)$ for every $n \in \mathbb{N}$ and

$$(2\pi mi - A)z_n = (I - T(1))x_n,$$

so that $(2\pi mi - A)z_n$ vanishes as n tends to ∞. To conclude that $2\pi mi$ is an approximate eigenvalue of A, we show that there exists a positive constant C' such that $\|z_n\| \geq C'$ for every $n \in \mathbb{N}$. For this purpose, we observe that

$$\liminf_{n\to\infty} \|z_n\| \geq \liminf_{n\to\infty} |\langle x_n', z_n\rangle| = \liminf_{n\to\infty} \left|\int_0^1 e^{-2\pi mit}\langle x_n', T(t)x_n\rangle dt\right| \geq \left|\int_0^1 e^{-2\pi mit} g(t)dt\right| > 0$$

and the estimate $\|z_n\| \geq C'$ follows. Now, if we set $y_n = \|z_n\|^{-1}z_n$, then $\|y_n\| = 1$ for every $n \in \mathbb{N}$ and $Ay_n - 2\pi miy_n$ converges to 0 as n tends to ∞. This shows that $2\pi mi$ belongs to $\sigma_{\rm a}(A)$ and this is enough to conclude that property (ii) holds true since $1 = e^{2\pi mi}$. □

Definition 12.2.3. A semigroup $\{T(t)\}$ on X is called eventually norm continuous if there exists $t_0 > 0$ such that the function $t \mapsto T(t)$ is continuous in $[t_0, \infty)$ with values in $\mathcal{L}(X)$.

Remark 12.2.4. By Theorem 9.2.2, any analytic semigroup is eventually norm continuous. Moreover, if $\{T(t)\}$ is a strongly continous semigroup in X such that for some $t_0 > 0$ the operator $T(t_0)$ is compact, then, the semigroup is eventually norm continuous. Indeed, the semigroup law implies that $T(t)$ is a compact operator for every $t > t_0$. Now, fix $t_1 > t_0$ and $x \in \overline{B(0,1)}$ (the closed unit ball of X) and suppose that $\|T(t) - T(t_1)\|$ does not vanish as t tends to t_1. Then, there exist $\varepsilon > 0$, a sequence $\{\bar{t}_n\}_{n\in\mathbb{N}}$, converging to t_1 as $n \to \infty$, and a sequence $\{x_n\} \subset \overline{B(0,1)}$ such that

$$\varepsilon < \|T(\bar{t}_n)x_n - T(t_1)x_n\| = \|(T(\bar{t}_n - t_0) - T(t_1 - t_0))T(t_0)x_n\|$$

for every $n \in \mathbb{N}$. Since $T(t_0)$ is a compact operator, up to a subsequence $T(t_0)x_n$ converges to some $y \in X$ as n tends to ∞. Writing

$$\begin{aligned}
&\|T(\bar{t}_n - t_0)y - T(t_1 - t_0)y\| \\
\geq&\|(T(\bar{t}_n - t_0) - T(t_1 - t_0))T(t_0)x_n\| - \|(T(\bar{t}_n - t_0) - T(t_1 - t_0))(y - T(t_0)x_n)\| \\
\geq&\|(T(\bar{t}_n - t_0) - T(t_1 - t_0))T(t_0)x_n\| - C\|y - T(t_0)x_n\| \\
\geq&\varepsilon - C\|y - T(t_0)x_n\|,
\end{aligned}$$

for a constant $C > 0$, independent of n, and letting n tend to ∞, we conclude that $T(\bar{t}_n - t_0)y$ does not converge to $T(t_1 - t_0)y$ as n tends to ∞, contradicting the strong continuity of the semigroup.

Theorem 12.2.5 (Spectral mapping theorem for eventually norm continuous semigroups)**.** *Let $\{T(t)\}$ be an eventually norm continuous semigroup. Then, $\sigma(T(t)) \setminus \{0\} = e^{\sigma(A)}$ for every $t > 0$.*

Proof. In view of all the above results, it suffices to show that $\sigma_{\rm a}(T(t)) \setminus \{0\} \subseteq e^{\sigma_{\rm a}(A)}$ for every $t > 0$. We fix $\lambda \in \sigma_{\rm a}(T(t_0)) \setminus \{0\}$ for some $t_0 > 0$. As in the proof of Proposition 12.2.1, we can assume that $\lambda = t_0 = 1$. We have to prove that there exists $m \in \mathbb{Z}$ such that $2\pi mi \in \sigma_{\rm a}(A)$. For this purpose, we fix a sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$ such that $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $T(1)x_n - x_n$ converges to 0 as n tends to ∞. Moreover, we fix an integer k such that the function $t \mapsto T(t)$ is continuous in $[k, \infty)$ with values in $\mathcal{L}(X)$. Then, we split

$$\|T(k)x_n - x_n\| \leq \sum_{j=1}^k \|T(j)x_n - T(j-1)x_n\| \leq \sum_{j=1}^k \|T(j-1)\|\|T(1)x_n - x_n\|,$$

so that $T(k)x_n - x_n$ converges to 0 as n tends to ∞. Further, we observe that

$$\|T(t)T(k)x_n - T(k)x_n\| = \|T(t+k)x_n - T(k)x_n\| \le \|T(t+k) - T(k)\|,$$

so that $\sup_{n\in\mathbb{N}} \|T(t)T(k)x_n - T(k)x_n\|$ converges to 0 as t tends to 0. Similarly, we can show that $\sup_{n\in\mathbb{N}} \|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\|$ vanishes as t tends to 0. Indeed, since the sequence $\{T(k)x_n - x_n\}_{n\in\mathbb{N}}$ is null, for every arbitrarily fixed $\varepsilon > 0$, we can determine $n_0 \in \mathbb{N}$ such that $\|T(k)x_n - x_n\| \le \varepsilon$ for every $n > n_0$. Then, we can write

$$\begin{aligned}&\sup_{n\in\mathbb{N}} \|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\|\\ \le &\max_{n\le n_0} \|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\| + \sup_{n>n_0} \|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\|.\end{aligned}$$

Since

$$\begin{aligned}\|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\| \le &\Big(\sup_{t\in[0,1]} \|T(t)\| + 1\Big)\|T(k)x_n - x_n\|\\ \le &\Big(\sup_{t\in[0,1]} \|T(t)\| + 1\Big)\varepsilon =: C\varepsilon, \qquad t \in [0,1],\ n > n_0,\end{aligned}$$

we obtain that

$$\sup_{n\in\mathbb{N}} \|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\| \le \max_{n\le n_0} \|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\| + C\varepsilon,$$

so that

$$\begin{aligned}&\limsup_{t\to 0} \sup_{n\in\mathbb{N}} \|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\|\\ \le &\lim_{t\to 0} \max_{n\le n_0} \|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\| + C\varepsilon = C\varepsilon.\end{aligned}$$

The arbitrariness of $\varepsilon > 0$ shows that $\sup_{n\in\mathbb{N}} \|T(t)(T(k)x_n - x_n) - (T(k)x_n - x_n)\|$ vanishes as t tends to 0.

Finally, writing $T(t)x_n - x_n = (T(t)T(k)x_n - T(k)x_n) + [(T(t)(x_n - T(k)x_n - (x_n - T(k)x_n)]$, we conclude that property (i)-(c) in Proposition 12.2.1 is satisfied. Hence, using that proposition, we can easily complete the proof. □

The following easy consequence of the spectral mapping theorem is usually referred to as "*spectral bound equal growth bound condition*".

Corollary 12.2.6. *Whenever the spectral mapping theorem holds, the equality $s(A) = \omega_0$ holds true.*

Proof. We already know that the spectral bound does not exceed the spectral growth of a strongly continuous or analytic semigroup. To complete the proof, we show that

$$t_0\omega_0 = \log(r(T(t_0)) \tag{12.17}$$

for every $t_0 > 0$ such that $T(t_0) \neq 0$, where, we recall that $r(T(t_0))$ denote the spectral radius of the operator $T(t_0)$. Once the previous formula is proved, the spectral mapping theorem allows us to easily complete the proof. Indeed, $\sigma(T(t_0)) \setminus \{0\} = e^{t_0\sigma(A)}$ for every fixed $t_0 > 0$ so that

$$r(T(t_0)) = \sup_{\lambda\in\sigma(T(t_0))} |\lambda| = \sup_{\mu\in\sigma(A)} e^{t_0 \mathrm{Re}\mu} = e^{s(A)t_0}.$$

So, let us prove formula (12.17). For this purpose, we fix $t_0 > 0$ and, recall that $r(T(t_0)) = \lim_{n\to\infty} \|T(t_0)^n\|^{\frac{1}{n}}$ (see Theorem 2.1.3). Using, Lemma 12.1.4, we can infer that

$$r(T(t_0)) = \lim_{n\to\infty} e^{\frac{1}{n}\|\log(T(nt_0))\|} = \exp\left(t_0 \lim_{n\to\infty} \frac{1}{nt_0}\|\log(T(nt_0))\|\right) = e^{\omega_0 t_0}. \quad \square$$

12.3 Exercises

Exercise 12.3.1. Prove that a semigroup of operators $\{T(t)\}_{t\geq 0}$ on a Banach space X embeds in a group $\{T(t)\}_{t\in\mathbb{R}}$ if and only if there exists $t_0 > 0$ such that $T(t_0)$ is invertible. Further, find an example of semigroup that does not embed in a group.

Exercise 12.3.2. Let $\{T(t)\}$ be a strongly continuous semigroup with infinitesimal generator A. Prove that $s(A) \leq \omega_0$.

Exercise 12.3.3. Let $\{T(t)\}$ be a strongly continuous semigroup on X with infinitesimal generator $A : D(A) \subset X \to X$. Fix $\mu \in \mathbb{C}$ and $\alpha > 0$ and prove that the family of bounded operators $\{S_{\alpha,\mu}(t)\}$, defined by $S_{\alpha,\mu}(t) = e^{\mu t}T(\alpha t)$ for every $t \geq 0$, is a strongly continuous semigroup on X. Further prove that

(i) the growth bound ω_0 of the semigroup $\{S_{\alpha,\mu}(t)\}$ is given by $\alpha\omega_0 + \mathrm{Re}\mu$, where ω_0 is the spectral bound of the semigroup $\{T(t)\}$;

(ii) the infinitesimal generator of $\{S_{\alpha,\mu}\}$ is the operator $A_{\alpha,\mu} = \alpha A + \mu$ and use this fact to deduce that $\sigma(A_{\alpha,\mu}) = \alpha\sigma(A) + \mu$;

(iii) $R(\lambda, A_{\alpha,\mu}) = \alpha^{-1}R(\alpha^{-1}(\lambda - \mu), A)$ for every $\lambda \in \rho(A_{\alpha,\mu})$.

Exercise 12.3.4. For every $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda \in (-1, 0]$ let R_λ be the operator in Example (12.1.3). Prove that R_λ is the inverse of the operator $\lambda I - A$.

Exercise 12.3.5. Let P_n be the projection introduced in Lemma 12.1.8. Prove that $P_n(X) = \mathrm{Ker}(\lambda_n I - A)$ for every $n \in \mathbb{N}$.

Exercise 12.3.6. Let $\{T(t)\}$ be the semigroup of left translations on $X = L^1(\mathbb{R})$ and let $\{T'(t)\}$ be its dual semigroup on $L^\infty(\mathbb{R})$. Prove that $\widetilde{X}' = BUC(\mathbb{R})$.

Exercise 12.3.7. Let $\{T(t)\}$ be a semigroup on X and let $\{\widetilde{T}'(t)\}$ be the restriction to $\widetilde{X}'$ of its dual semigroup. Prove that $\sigma_{\mathrm{p}}(\widetilde{A}) = \sigma_{\mathrm{p}}(A')$.

Exercise 12.3.8. Let $1 < p < q < \infty$ and $X := L^p[1, \infty) \cap L^q[1, \infty)$ be the Banach space endowed with the norm

$$\|f\| := \|f\|_p + \|f\|_q, \quad f \in X.$$

Consider the family of operators

$$T(t)f(s) = f(se^t), \quad s \geq 1,\ t \geq 0.$$

(i) Show that $\{T(t)\}$ is a C_0-semigroup on X with generator

$$\begin{aligned}(Af)(s) =& sf'(s),\ s \geq 1,\ \text{and} \\ f \in D(A) =& \{f \in X\ :\ f \text{ absolutely continuous and } Af \in X\}.\end{aligned}$$

(ii) Prove that

$$s(A) = -\frac{1}{p} < -\frac{1}{q} = \omega_0.$$

Bibliography

[1] A. Belleni-Morante, Applied Semigroups and Evolution Equations, Oxford University Press, 1979.

[2] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[3] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[4] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[5] K.-J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Graduate Text in Mathematics **194**, Springer-Verlag, 2000.

[6] K.-J. Engel and R. Nagel, Short Course on Operator Semigroups, Universitext, Springer-Verlag, 2006.

[7] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[8] J.A. Goldstein, Semigroups of Operators and Applications, Oxford University Press, 1985.

[9] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[10] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[11] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981.

[12] L. Lorenzi and A. Rhandi, Semigroups of Bounded Operators and Second-Order Elliptic and Parabolic Partial Differential Equations, CRC Press, Taylor & Francis Group, 2021.

[13] H.P. Lotz, Uniform convergence of operators on L^∞ and similar spaces, Math. Z. **190** (1985), 207–220.

[14] A. Lunardi, Analytic Semigroups and Optimal Regularity in Parabolic Problems, Birkhäuser Verlag, 1995.

[15] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Appl. Math.Sci. **44**, Springer-Verlag,1983.

[16] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[17] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[18] W. Rudin, Real and Complex Analysis, 3^{rd} ed., McGraw-Hill, 1986.

[19] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[20] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958.

[21] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861.

[22] A. Zygmund, Trigonometric Series, Cambridge Univ. Press., 2nd Edition Reprinted (1968).

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

February 2, 2022

Lecture 13

Asymptotic behavior for linear problems

In this lecture, we analyze the long time behaviour of the solutions to the differential equation $u' = Au + f$, when A is a sectorial operator and f is a continuous function defined in $[0, \infty)$ or in $(-\infty, 0]$, i.e., we consider both forward and backward (in time) abstract Cauchy problems. The starting point is the analysis of the homogeneous Cauchy problem

$$\begin{cases} u'(t) = Au(t), & t \in I, \\ u(0) = x, \end{cases} \tag{13.1}$$

when $I = (0, \infty)$ or $I = (-\infty, 0)$, looking for conditions on the datum x, which guarantee that the solution u is bounded in I. By the results of Lecture 9, we know that the solution to such a problem is the function $u = T(\cdot)x$.

When $I = (-\infty, 0)$, problem (13.1) is usually ill-posed. Indeed, due to smoothing effects of analytic semigroups x should be "very smooth" for problem (13.1) to admit a backward solution.

If the spectral bound $s(A)$ is negative, then, by Corollary 12.2.6, the operator norm of the semigroup $\{T(t)\}$ decreases to zero exponentially as t tends to ∞ so that all the solutions to the equation $u' = Au$ are bounded in $[0, \infty)$. Of particular interest is the case when the spectrum of A does not intersect the imaginary axis. In such a case (the so-called hyperbolic case) a spectral projection P is introduced and used to characterize the set of $x \in X$ for which problem (13.1) admits a bounded forward (resp. backward) solution. Based on these results, we then move our attention to the nonhomogeneous Cauchy problem

$$\begin{cases} u'(t) = Au(t) + f(t), & t \in I, \\ u(0) = x, \end{cases} \tag{13.2}$$

in the hyperbolic case and provide necessary and sufficient conditions for the mild solution to be bounded in I. The definition of mild solution to the forward problem (13.2) has been introduced in Lectures 8 and 9. A similar definition can be provided in the case of backward solutions. Indeed, if u is a solution to (13.2) in $(-\infty, 0]$, then, for every $a < 0$, the function $v = u(\cdot + a)$ solves the equation $v' = Av + f(\cdot + a)$ in $[0, -a)$. Hence, it can be written via the variation-of-constants formula and this leads to the definition of mild solutions to backward Cauchy problems. We concentrate our attention on mild solutions since in view of the results in Lecture 10, we know sufficient conditions for mild solutions to be classical or even strict solutions.

Finally, we relax the hyperbolicity assumption on operator A and assume that $\omega \in \mathbb{R}$ is such that $\sigma(A)$ does not contain elements with real part equal to ω. We characterize the data x and f such that problem (13.2) admits solutions u such that $\|u(t)\| \le Me^{\omega t}$ for every $t \in I$ and some positive constant M.

13.1 A preliminary result

As we have seen in Lecture 12, the spectral mapping theorem is satisfied by analytic semigroups, so that the spectral bound equals the growth bound (see Corollary 12.2.6). Based on this result, the following proposition can be easily proved.

Proposition 13.1.1. *For every $n \in \mathbb{N} \cup \{0\}$ and $\varepsilon > 0$ there exists a positive constant $M_{n,\varepsilon} > 0$ such that*

$$\|t^n A^n T(t)\| \le M_{n,\varepsilon} e^{(s(A)+\varepsilon)t}, \qquad t > 0. \tag{13.3}$$

Proof. As it has been already remarked, for every $\varepsilon > 0$, there exists a positive constant M_ε such that

$$\|T(t)\| \le M_\varepsilon e^{(s(A)+\varepsilon)t}, \qquad t > 0. \tag{13.4}$$

Moreover, by Remark 9.2.3, we know that $\|AT(t)\| \le N_1 t^{-1}$ for every $t \in (0, 1]$ and some positive constant N_1, so that (13.3) follows with $n = 1$ and $t \in (0, 1]$. If $t > 1$, we can split $AT(t) = AT(1)T(t-1)$ and estimate $\|AT(t)\| \le \|AT(1)\|\|T(t-1)\|$. Applying (13.4), with ε being replaced by $\varepsilon/2$, we obtain

$$\|AT(t)\| \le N_1\|T(t-1)\| \le N_1 M_{\varepsilon/2} e^{(s(A)+\varepsilon/2)(t-1)} = \overline{M}_\varepsilon e^{(s(A)+\varepsilon/2)t}, \qquad t \ge 1. \tag{13.5}$$

Finally, observing that there exists a positive constant $\overline{N}_\varepsilon$ such that $t^{-1}e^{\varepsilon t/2} \ge \overline{N}_\varepsilon$ for every $t \ge 1$, from (13.5) we deduce estimate (13.3) with $n = 1$ for every $t \ge 1$.

For a general value of $n \in \mathbb{N}$, it suffices to split $A^nT(t) = (AT(t/n))^n$ for every $t > 0$ and apply (13.3) with $n = 1$. □

In the case $s(A) = \omega = 0$, where ω is the number in Definition 9.2.1, estimates (9.6) are better than (13.3) for t large, as the following example shows.

Example 13.1.2. Let us consider the operator $A : D(A) \subset C_b(\mathbb{R}) \to C_b(\mathbb{R})$, defined by $Au = u''$ for every $u \in D(A) = C_b^2(\mathbb{R})$. By Exercise 9.4.7, A is a sectorial operator. Moreover, its spectrum is the negative real axis and $\|\lambda R(\lambda, A)\| \le (\cos\theta/2)^{-1}$ for every $\lambda \in \rho(A)$, where $\theta = \arg(\lambda)$ (see Exercise 13.5.2). In this case $\omega = s(A) = 0$, and estimates (13.3) are worse than (9.6) for large t. It is convenient to use estimate (9.6), which yields $\|T(t)\| \le M_0$ and $\|t^kA^kT(t)\| \le M_k$ for every $t > 0$, $k \in \mathbb{N}$ and some positive constants M_k ($k \in \mathbb{N} \cup \{0\}$). Therefore, for every initial datum $u_0 \in C_b(\mathbb{R})$, the function $t \mapsto T(t)u_0$ is bounded. Moreover, the k-th derivative with respect to time and the $2k$-th derivative with respect to x decay at least like t^{-k}, as $t \to \infty$, in the sup norm.

As a straightforward consequence of Proposition 13.1.1, it follows that, if $s(A) < 0$, then the function $t \mapsto T(t)x$ is bounded in $[0, \infty)$ for every $x \in X$. Actually, its operator norm decreases exponentially to 0 as t tends to ∞. This means that, for every $x \in X$, the classical solution to the Cauchy problem

$$\begin{cases} u'(t) = Au(t), & t \in (0, \infty), \\ u(0) = x \end{cases}$$

decreases to zero as t tends to ∞, i.e., the null solution to the equation $u' = Au$ "attracts" all the other solutions.

It is interesting to understand how the picture changes when $s(A) > 0$. Clearly, we do not expect that all the solutions to (13.1) decreases to zero exponentially or even that all the solutions are bounded in $[0, \infty)$.

13.2 Large time behaviour of the semigroup in the hyperbolic case

Throughout this section, we assume that

$$\sigma(A) \cap i\mathbb{R} = \varnothing. \tag{13.6}$$

In this case A is said to be hyperbolic. Due to condition (13.6), we can split the spectrum of A into the two disjoint subsets $\sigma_-(A)$ and $\sigma_+(A)$, where $\sigma_-(A)$ is the part of $\sigma(A)$ in the left halfplane $\{\lambda \in \mathbb{C} : \text{Re}\lambda < 0\}$, whereas $\sigma_+(A)$ is the part of $\sigma(A)$ in the open right halfplane $\{\lambda \in \mathbb{C} : \text{Re}\lambda > 0\}$. Both $\sigma_-(A)$ and $\sigma_+(A)$ are closed subsets of $\mathbb{C}$.

Remark 13.2.1. Since A is a sectorial operator, the part of its spectrum on the right halfplane (i.e., $\sigma_+(A)$), if not empty, is a bounded subset of the complex plane.

On the other hand, no a priori information is available on $\sigma_-(A)$ which could be bounded or unbounded. For instance, if A is a bounded operator on a Banach space X, then the associated semigroup is uniformly continuous and, in particular, is analytic. By Proposition 2.1.1 its spectrum is bounded, so that also $\sigma_-(A)$ is a closed and bounded subset of $\mathbb{C}$. Differently, the operator $A : D(A) \subset C_b(\mathbb{R}) \to C_b(\mathbb{R})$, defined by $Au = u'' - u$ for every $u \in D(A) = C_b^2(\mathbb{R})$ is sectorial and $\sigma_-(A) = (-\infty, -1]$, so that it is an unbounded subset of the complex plane.

Since $\sigma_-(A)$ and $\sigma_+(A)$ are closed subsets of $\mathbb{C}$, it follows easily that

$$-\omega_- := \sup\{\text{Re}\,\lambda : \lambda \in \sigma_-(A)\} < 0, \qquad \omega_+ := \inf\{\text{Re}\,\lambda : \lambda \in \sigma_+(A)\} > 0.$$

If $\sigma_-(A)$ (resp. $\sigma_+(A)$) is void, we set $\omega_- = \infty$ (resp. $\omega_+ = \infty$).

In the analysis of the semigroup $\{T(t)\}$ for large values of t, the operator P, defined by

$$P = \frac{1}{2\pi i} \int_{\gamma_+} R(\lambda, A) d\lambda \tag{13.7}$$

will play a crucial role. Here, γ_+ is an arbitrarily fixed closed regular curve contained in $\rho(A)$, surrounding $\sigma_+(A)$, oriented counterclockwise, which "covers" its support only once.

Proposition 13.2.2. *The operator P is a projection, i.e., $P^2 = P$. Moreover, the following properties hold true.*

(i) *For each $t > 0$, the operators $T(t)$ and P commute, and*

$$T(t)P = PT(t) = \frac{1}{2\pi i}\int_{\gamma_+} e^{\lambda t}R(\lambda, A)d\lambda. \tag{13.8}$$

As a byproduct, $T(t)(P(X)) \subset P(X)$, $T(t)((I-P)(X)) \subset (I-P)(X)$.

(ii) *$P \in \mathcal{L}(X, D(A^n))$ for every $n \in \mathbb{N}$. Therefore, $P(X) \subset D(A)$ and the operator $A_{|P(X)} : P(X) \to P(X)$ is bounded.*

(iii) *The restriction of the semigroup $\{T(t)\}$ to $P(X)$ is actually a uniformly continuous group, generated by the part of A in $P(X)$. Moreover, for every $\omega \in [0, \omega_+)$ there exists a positive constant $C_\omega > 0$ such that*

$$\|T(t)x\| \le C_\omega e^{\omega t}\|x\|, \qquad t \le 0, \;\; x \in P(X). \tag{13.9}$$

(iv) *The restriction of the semigroup $\{T(t)\}$ to $(I-P)(X)$ is an analytic semigroup and is associated to the part of A in $(I-P)(X)$. Moreover, for every $\omega \in [0, \omega_-)$ there exists a positive constant $C'_\omega > 0$ such that*

$$\|T(t)x\| \le C'_\omega e^{-\omega t}\|x\|, \qquad t \ge 0, \;\; x \in (I-P)(X). \tag{13.10}$$

Proof. To prove that P is a projection, we fix two regular curves (say γ_+ and γ'_+) contained in $\rho(A)$, surrounding $\sigma_+(A)$ and oriented counterclockwise, which "cover" their supports only once. Further, we assume that γ_+ is contained in the bounded connected component of $\mathbb{C}\setminus\gamma'_+$. Using the resolvent identity, we can write.

$$\begin{aligned}
P^2 &= \left(\frac{1}{2\pi i}\right)^2 \int_{\gamma'_+} R(\xi, A)d\xi \int_{\gamma_+} R(\lambda, A)d\lambda \\
&= \left(\frac{1}{2\pi i}\right)^2 \int_{\gamma'_+\times\gamma_+} [R(\lambda, A) - R(\xi, A)](\xi-\lambda)^{-1}d\xi d\lambda \\
&= \left(\frac{1}{2\pi i}\right)^2 \int_{\gamma_+} R(\lambda, A)d\lambda \int_{\gamma'_+}(\xi-\lambda)^{-1}d\xi - \left(\frac{1}{2\pi i}\right)^2 \int_{\gamma'_+} R(\xi, A)d\xi \int_{\gamma_+}(\xi-\lambda)^{-1}d\lambda = P.
\end{aligned}$$

(i) The proof of this property can be obtained adapting the above arguments. Hence, it is left to the reader as an exercise. See Exercise 13.5.1.

(ii) Since the curve γ_+ is bounded and the function $\lambda \mapsto R(\lambda, A)$ is continuous with values in $\mathcal{L}(X, D(A))$, using Proposition B.0.3 we can infer that $P \in \mathcal{L}(X, D(A))$ and

$$AP = \frac{1}{2\pi i}\int_{\gamma_+} AR(\lambda, A)d\lambda = -\frac{1}{2\pi i}\int_{\gamma_+} d\lambda + \frac{1}{2\pi i}\int_{\gamma_+}\lambda R(\lambda, A)d\lambda = \frac{1}{2\pi i}\int_{\gamma_+}\lambda R(\lambda, A)d\lambda.$$

Therefore, the operator AP is bounded in X with values in $D(A)$. The same argument shows that P and A commute on $D(A)$. Based on these properties, we can easily show, by recurrence, that $P \in \mathcal{L}(X, D(A^n))$ for every $n \in \mathbb{N}$.

(iii) Since the restriction of each operator $T(t)$ to $P(X)$ is given by formula (13.8), it is easy to extend the operator $T(t)$ to all negative times obtaining a group of bounded operators. Its infinitesimal generator is clearly the part of A in $P(X)$, i.e., the operator AP, which is bounded in $P(X)$. It thus follows that the restriction of the semigroup $\{T(t)\}$ to $P(X)$ is a uniformly continuous group.

To prove estimate (13.9), we fix $\omega \in [0,\omega_+)$ and a regular curve γ_+ such that $\inf_{\lambda\in\gamma_+} \text{Re}\lambda = \omega$. Then, we can estimate

$$\|T(t)x\| \le \frac{1}{2\pi}\left|\int_{\gamma_+} |e^{\lambda t}|\,\|R(\lambda,A)\|\,\|x\|d\lambda\right| \le C_\omega \sup_{\lambda\in\gamma_+}|e^{\lambda t}|\,\|x\| = C_\omega e^{\omega t}\|x\|$$

for every $t \le 0$ and $x \in P(X)$, where with $C_\omega = (2\pi)^{-1}|\gamma_+|\sup\{\|R(\lambda,A)\| : \lambda\in\gamma_+\}$ and $|\gamma_+|$ denotes the length of the curve γ_+, Estimate (13.9) is proved.

(iv) Taking formula (13.8) into account, for every $t > 0$ we can write

$$\begin{aligned}T(t)(I-P) =&T(t) - T(t)P = \frac{1}{2\pi i}\int_\gamma e^{\lambda t}R(\lambda,A)d\lambda - \frac{1}{2\pi i}\int_{\gamma_+} e^{\lambda t}R(\lambda,A)d\lambda\\ =&\frac{1}{2\pi i}\int_{\gamma_-} e^{\lambda t}R(\lambda,A)d\lambda,\end{aligned} \tag{13.11}$$

where γ is the same curve used in Section 9.2 to define the operator $T(t)$ (see (9.4)), γ_- is the curve oriented counterclockwise and having the set $\{\lambda\in\mathbb{C} : \lambda = -\omega + re^{\pm i\theta}, r \ge 0\}$ as support, for some $\theta > \pi/2$. From property (i), it follows that the restriction of each operator $T(t)$ to $(I-P)(X)$ maps this space into itself. Moreover, since $\|T(t)(I-P)\| \le \|T(t)\|(1+\|P\|)$, formula (13.10) follows immediately for $t\in(0,1]$. On the other hand, if $t \ge 1$, estimate (13.10) can be obtained, starting from formula (13.11) and arguing as in the proof of Proposition 13.1.1. The details are left to the reader as an exercise (see Exercise 13.5.1). □

Corollary 13.2.3. *Fix $x \in X$. Then, the following properties are satisfied.*

(i) *The function $T(\cdot)x$ is bounded in $[0,\infty)$ if and only if $Px = 0$. In this case, the norm $\|T(t)x\|$ decays exponentially to 0 as t tends to ∞.*

(ii) *The backward Cauchy problem*

$$\begin{cases} u'(t) = Au(t), & t \le 0,\\ u(0) = x,\end{cases} \tag{13.12}$$

admits a bounded solution in $(-\infty,0]$ if and only if $x\in P(X)$. In this case, the bounded solution is unique and is given by $u = T(\cdot)x$. Hence, it decays exponentially to 0 as t tends to $-\infty$.

Proof. (i) Fix $x \in X$ and split it along $P(X)$ and $(I-P)(X)$, i.e., $x = Px + (I-P)x$. Then, $T(t)x = T(t)Px + T(t)(I-P)x$ for every $t > 0$. The norm of the second term in the previous formula decays exponentially to 0 as t tends to ∞, due to Proposition 13.2.2(iv). On the other hand, since the restriction of $\{T(t)\}$ to $P(X)$ is actually a group, we can write $Px = T(-t)T(t)Px$, so that, using (13.9), we can estimate

$$\|Px\| \le \|T(-t)\|_{\mathcal{L}(P(X))}\|T(t)Px\| \le C_\omega e^{-\omega t}\|T(t)Px\|, \qquad t > 0,$$

with $\omega > 0$, which implies that $\|T(t)Px\| \ge C_\omega^{-1}e^{\omega t}\|Px\|$ for every $t > 0$. It thus follows that the map $t \mapsto T(t)x$ is bounded in $[0,\infty)$ if and only if $Px = 0$.

(ii) If $x \in P(X)$, then the function $t \mapsto T(t)x$ is a strict solution to the backward Cauchy problem (13.12) and decays exponentially to zero as t tends to $-\infty$.

Conversely, suppose that problem (13.12) admits a backward bounded solution u. Then, for $t \in [0,-a)$ the function $t \mapsto u(t+a)$ solves the Cauchy problem

$$\begin{cases} v'(t) = Av(t), & t \in [0,-a),\\ v(0) = u(a). \end{cases}$$

Hence, $u(t+a) = T(t)u(a)$ for every $t \in [0,-a)$ or, equivalently,

$$u(t) = T(t-a)u(a) = T(t-a)(I-P)u(a) + T(t-a)Pu(a) = (I-P)u(t) + Pu(t)$$

for every $t \in (a,0]$.

Fix $\omega \in (0,\omega_-)$ and use (13.10) to deduce that $\|T(t-a)(I-P)\| \le C'_\omega e^{-\omega(t-a)}$. Letting a tend to $-\infty$, we conclude that $(I-P)u(t) = 0$ for each $t \le 0$, so that $x \in P(X)$ and $u = T(\cdot)x$. □

Remark 13.2.4. Note that problem (13.12) is ill-posed in general. Changing t to $-t$ it is equivalent to a forward Cauchy problem with A replaced by $-A$, and $-A$ may have very bad spectral properties. If A is sectorial, $-A$ is sectorial if and only if it is bounded (see Exercise 9.4.3).

In view of Proposition 13.2.2, we can now introduce the following definition.

Definition 13.2.5. The operator P, defined by (13.7), is called spectral projection relative to $\sigma_+(A)$. Moreover, the subspaces $(I-P)(X)$ and $P(X)$ are called stable subspace and unstable subspace, respectively.

In the following example, we illustrate the results so far obtained in a concrete case.

Example 13.2.6. Let us consider the Cauchy problem

$$\begin{cases} D_t u(t,x) = D_{xx}u(t,x) + \alpha u(t,x), & t > 0, \quad x \in [0,1],\\ u(t,0) = u(t,1) = 0, & t \ge 0,\\ u(0,x) = u_0(x), & \qquad\quad x \in [0,1], \end{cases} \tag{13.13}$$

where $\alpha \in \mathbb{R}$ is a parameter.

We recast it as the abstract Cauchy problem

$$\begin{cases} u'(t) = Au(t), & t > 0,\\ u(0) = u_0 \end{cases} \tag{13.14}$$

by choosing $X = C([0,1])$, $D(A) = \{f \in C^2([0,1]) : f(0) = f(1) = 0\} \subset X$, $Au = u'' + \alpha u$ for every $u \in D(A)$.

The spectrum of A consists of the sequence of eigenvalues $\{\lambda_n\}_{n\in\mathbb{N}}$, where $\lambda_n = -\pi^2 n^2 + \alpha$ for every $n \in \mathbb{N}$ (see Exercise 13.5.4). We now discuss some cases depending on the value of α.

- If $\alpha < \pi^2$, then $\sigma(A)$ is contained in the halfplane $\{\lambda \in \mathbb{C} : \mathrm{Re}\lambda < 0\}$ and, by Proposition 13.1.1, the solution $u = T(\cdot)u_0$ of (13.14) (and, hence, of problem (13.13)) and all its derivatives decay exponentially as $t \to \infty$, for every initial datum $u_0 \in C([0,1])$.

- If $\alpha = \pi^2$, then, clearly, $s(A) = 0$. Moreover, there exists a positive constant C such that $\|T(t)\| \le C$ for every $t > 0$. To show this last fact, we observe that the operator $A_2 u = u'' + \pi^2 u$ with domain $D(A_2) = \{u \in H^2(0,1) : u(0) = u(1) = 0\}$ is sectorial in $L^2(0,1)$ and the associated semigroup $\{T_2(t)\}$ coincides with $\{T(t)\}$ on $C([0,1])$ (see Exercise 13.5.4). Since the functions $u_k : [0,1] \to \mathbb{R}$, defined by $u_k(x) = \sin(k\pi x)$ for every $x \in [0,1]$, are eigenfunctions of A_2 with eigenvalues $(-k^2+1)\pi^2$ for every $k \in \mathbb{N}$, then (see Exercise 13.5.3) $T_2(t)u_k = e^{-(k^2-1)\pi^2 t}u_k$ for every $t \ge 0$.

  If $f \in C([0,1]) \subset L^2(0,1)$, then we expand it in a sine-series in $L^2((0,1))$

  $$f = \sum_{k=1}^{\infty} c_k u_k, \qquad c_k = 2\int_0^1 f(x)u_k(x)dx. \tag{13.15}$$

  Hence[1],

  $$T(t)f = T_2(t)f = \sum_{k=1}^{\infty} c_k T(t)u_k = \sum_{k=1}^{\infty} c_k e^{-(k^2-1)\pi^2 t}u_k, \qquad t \ge 0,$$

  so that

  $$\|T(t)f\|_\infty \le 2\|f\|_\infty \sum_{k=1}^{\infty} e^{-(k^2-1)\pi^2 t}, \qquad t > 0,$$

  which is bounded in $[1,\infty)$. Since $T(t)$ is an analytic semigroup, then $\|T(t)\|$ is bounded in $[0,1]$. Hence, the function $t \mapsto T(t)$ is bounded in $[0,\infty)$ with values in $\mathcal{L}(X)$.

- If $\alpha > \pi^2$, then the spectrum of A contains elements with positive real part. In the case where $\alpha \ne n^2\pi^2$ for every $n \in \mathbb{N}$, condition (13.6) is satisfied. Let $m \in \mathbb{N}$ be such that $\pi^2 m^2 < \alpha < \pi^2(m+1)^2$. By Corollary 13.2.3, the initial data u_0 such that the solution is bounded are those which satisfy the condition $Pu_0 = 0$. Since $\sigma_+(A)$ is a finite set, the projection P may be split into the sum

  $$P = \sum_{k=1}^{m} \frac{1}{2\pi i}\int_{|\lambda-\lambda_k|<\varepsilon} R(\lambda, A)d\lambda = \sum_{k=1}^{m} P_k, \tag{13.16}$$

  where $\lambda_k = -\pi^2 k^2 + \alpha$ ($k = 1, \ldots, m$) are the eigenvalues of A with positive real part and ε is fixed sufficiently small such that the set $\overline{B(\lambda_k,\varepsilon)} \setminus \{\lambda_k\}$ is contained in $\rho(A)$. Let us show that

  $$(P_k f)(x) = 2\sin(k\pi x)\int_0^1 \sin(k\pi y)\, f(y)dy, \qquad x \in [0,1]. \tag{13.17}$$

[1] To justify the above expansion, it suffices to observe that (13.15) is the Fourier series of the function $\overline{f} : [-1,1] \to \mathbb{R}$ which is the odd extension of f.

For every $\lambda \neq \lambda_k$ expand $f \in C([0,1])$ as in (13.15). Using Exercise 13.5.3, we get

$$R(\lambda, A)f = R(\lambda, A_2)f = \sum_{n=1}^{\infty} \frac{c_n}{\lambda - \lambda_n} u_n.$$

Hence,

$$P_k f = \frac{1}{2\pi i} \int_{|\lambda - \lambda_k| \leq \varepsilon} R(\lambda, A)f\, d\lambda = c_k u_k.$$

Consequently, from (13.16) and (13.17) it follows that the solution of (13.13) is bounded in $[0,\infty)$ if and only if the compatibility conditions

$$\int_0^1 \sin(k\pi y) u_0(y) dy = 0, \qquad k = 1, \ldots, m,$$

are satisfied.

13.3 Bounded solutions to nonhomogeneous problems in unbounded intervals

In this section, we consider nonhomogeneous Cauchy problems in halflines. To begin with, we consider the forward Cauchy problem

$$\begin{cases} u'(t) = Au(t) + f(t), & t > 0, \\ u(0) = x, \end{cases} \tag{13.18}$$

where $f : [0,\infty) \to X$ is a continuous function and $x \in X$. Throughout the section, we assume that condition (13.6) is satisfied. Further, we consider the sets $\sigma_-(A)$, $\sigma_+(A)$ and the constants ω_-, ω_+ defined in Section 13.2. Similarly, by P we still denote the spectral projection defined by (13.7).

Fix once and for all a positive number ω such that $-\omega_- < -\omega < \omega < \omega_+$, and let C_ω, C'_ω be the constants introduced in Proposition 13.2.2(iv)-(v).

We recall that the mild solution to problem (13.18) is the function $u : [0,\infty) \to X$ defined by

$$u(t) = T(t)x + \int_0^t T(t-s)f(s)ds, \qquad t > 0.$$

Proposition 13.3.1. *Fix $f \in C_b([0,\infty);X)$ and $x \in X$. Then, the mild solution u to problem (13.18) is bounded in $[0,\infty)$ with values in X if and only if*

$$Px = -\int_0^\infty T(-s)Pf(s)ds. \tag{13.19}$$

In this case,

$$u(t) = T(t)(I-P)x + \int_0^t T(t-s)(I-P)f(s)ds - \int_t^\infty T(t-s)Pf(s)ds, \qquad t \geq 0. \tag{13.20}$$

Proof. For every $t \geq 0$ we split $u(t) = (I-P)u(t) + Pu(t)$, where

$$(I-P)u(t) = T(t)(I-P)x + \int_0^t T(t-s)(I-P)f(s)ds$$

and

$$\begin{aligned} Pu(t) =& T(t)Px + \int_0^t T(t-s)Pf(s)ds` \\ =& T(t)Px + \int_0^\infty T(t-s)Pf(s)ds - \int_t^\infty T(t-s)Pf(s)ds \\ =& T(t)\Big(Px + \int_0^\infty T(-s)Pf(s)ds\Big) - \int_t^\infty T(t-s)Pf(s)ds. \end{aligned} \tag{13.21}$$

Using (13.10) we can estimate

$$\begin{aligned} \|(I-P)u(t)\| \leq& C'_\omega e^{-\omega t}\|(I-P)x\| + C'_\omega\Big(\sup_{0\leq s\leq t}\|(I-P)f(s)\|\Big)\int_0^t e^{-\omega(t-s)}ds \\ \leq& C'_\omega\|I-P\|(\|x\| + \omega^{-1}\|f\|_\infty) \end{aligned}$$

for every $t > 0$, so that $(I-P)u$ is bounded in $[0,\infty)$ with values in X. The last integral term in the last side of (13.21) is bounded as well, and

$$\left\|\int_t^\infty T(t-s)Pf(s)ds\right\| \leq C_\omega\Big(\sup_{s\geq 0}\|Pf(s)\|\Big)\int_t^\infty e^{\omega(t-s)}ds = C_\omega\omega^{-1}\|P\|\|f\|_\infty, \qquad t \geq 0.$$

Hence, u is bounded if and only if the function

$$t \mapsto T(t)\left(Px + \int_0^\infty T(-s)Pf(s)ds\right) =: T(t)y,$$

is bounded. By Corollary 13.2.3, $T(t)y$ is bounded if and only if $Py = 0$, i.e., if and only if condition (13.19) is satisfied. In this case, u is given by (13.20). □

Corollary 13.3.2. *Suppose that $\sigma(A) \subset \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda < \delta\}$ for some $\delta < 0$. Then, for every $x \in X$ and $f \in C_b([0,\infty);X)$, the mild solution to Cauchy problem* (13.18) *is bounded in $[0,\infty)$ with values in X.*

Proof. It suffices to observe that $\sigma(A) = \sigma_-(A)$, so that $P = 0$. Hence, for every $x \in X$ and $f \in C_b([0,\infty);X)$, condition (13.19) is trivially satisfied. Applying Proposition 13.3.1, we can complete the proof. □

Next, we consider the backward problem,

$$\begin{cases} u'(t) = Au(t) + f(t), & t \leq 0, \\ u(0) = x, \end{cases} \tag{13.22}$$

where $f : (-\infty, 0] \to X$ is a bounded and continuous function and $x \in X$.

Problem (13.22) is in general ill-posed, and to guarantee the existence of a solution we need to assume rather restrictive conditions on f and x. Such conditions will also ensure good regularity properties of the solutions.

Definition 13.3.3. A function $u \in C((-\infty,0];X)$ is said to be a mild solution to problem (13.22) in $(-\infty,0]$ if $u(0)=x$ and

$$u(t) = T(t-a)u(a) + \int_a^t T(t-s)f(s)ds, \qquad t \in [a,0],$$

for every $a<0$.

In other words, u is a mild solution of (13.22) if and only if for every $a<0$ u is a mild solution of the problem

$$\begin{cases} u'(t) = Au(t) + f(t), & t \in (a,0], \\ u(a) = y, \end{cases} \tag{13.23}$$

where $y=u(a)$, and moreover $u(0)=x$.

Proposition 13.3.4. *Fix $x \in X$ and $f \in C_b((-\infty,0];X)$. Then, problem* (13.22) *admits a mild solution $u \in C_b((-\infty,0];X)$ if and only if*

$$(I-P)x = \int_0^\infty T(s)(I-P)f(-s)ds. \tag{13.24}$$

In such a case, the bounded mild solution u is unique and

$$u(t) = T(t)Px - \int_t^0 T(t-s)Pf(s)ds + \int_{-\infty}^t T(t-s)(I-P)f(s)ds, \qquad t \le 0. \tag{13.25}$$

Proof. Assume that problem (13.22) admits a bounded mild solution u. Then, for every $a<0$ and $t\in[a,0]$ we can write $u(t)=(I-P)u(t)+Pu(t)$, where

$$\begin{aligned}(I-P)u(t) =& T(t-a)(I-P)u(a) + \int_a^t T(t-s)(I-P)f(s)ds \\ =& T(t-a)\Big((I-P)u(a) - \int_{-\infty}^a T(a-s)(I-P)f(s)ds\Big) + u_1(t) \\ =& T(t-a)\big((I-P)u(a) - u_1(a)\big) + u_1(t),\end{aligned}$$

for every $t\in[a,0]$ and

$$u_1(t) = \int_{-\infty}^t T(t-s)(I-P)f(s)ds, \qquad t \le 0.$$

The function u_1 is bounded in $(-\infty,0]$. Indeed,

$$\|u_1(t)\| \le C'_\omega \Big(\sup_{s\le 0}\|(I-P)f(s)\|\Big)\int_{-\infty}^t e^{-\omega(t-s)}ds \le C'_\omega \omega^{-1}\|I-P\|\|f\|_\infty \tag{13.26}$$

for every $t\le 0$. Since u is bounded by assumptions, it follows that $\sup_{a\le 0}\|(I-P)u(a)\|<\infty$. Letting a tend to $-\infty$ and using estimate (13.10) we get $(I-P)u(t)=u_1(t)$ for every $t\le 0$. Taking $t=0$, condition (13.24) follows at once. On the other hand, Pu is a strict solution to the equation $w' = Aw + Pf$ and, since $Pu(0) = Px$, it is given by the variation-of-constants formula, i.e.,

$$Pu(t) = T(t)Px - \int_t^0 T(t-s)Pf(s)ds, \qquad t \le 0.$$

Summing up, u is given by (13.25).

Conversely, let us assume that condition (13.24) holds true and define the function $u = u_1 + u_2$, where u_1 is defined above and

$$u_2(t) = T(t)Px - \int_t^0 T(t-s)Pf(s)ds, \qquad t \le 0.$$

Then, u_1 is bounded by estimate (13.26) and u_2 is bounded due to estimate (13.9), so that function u is bounded as well in $(-\infty, 0]$.

As it is easily seen, u is a mild solution of (13.23) for every $a < 0$, and since condition (13.24) holds true, it follows that

$$u(0) = Px + \int_{-\infty}^0 T(-s)(I-P)f(s)ds = Px + (I-P)x = x.$$

Hence, u is a bounded mild solution to the problem (13.22). □

Corollary 13.3.5. *Suppose that $\sigma(A) \subset \{\lambda \in \mathbb{C} : \mathrm{Re}\,\lambda < \delta\}$ for some $\delta < 0$. Then, for every $f \in C_b((-\infty, 0]; X)$, problem* (13.22) *admits a mild solution if and only if*

$$x = \int_0^\infty T(s)f(-s)ds.$$

In such a case the mild solution u is unique and

$$u(t) = \int_{-\infty}^t T(t-s)f(s)ds, \qquad t \le 0.$$

Proof. It suffices to argue as in the proof of Corollary 13.3.2, taking Proposition 13.3.4 into account. □

13.4 Solutions with exponential growth or exponential decay

In this section, we relax a bit assumption (13.6) and assume that

$$\sigma(A) \cap \{\lambda \in \mathbb{C} : \mathrm{Re}\lambda = \omega\} = \varnothing \tag{13.27}$$

for some $\omega \in \mathbb{R}$. Clearly, condition (13.27) is satisfied if we take $\omega > s(A)$.

For every unbounded interval I and $\omega \in \mathbb{R}$, we denote by $C_\omega(I; X)$ the set of all continuous functions $f : I \to X$ such that $\|f\|_{C_\omega(I;X)} := \sup_{t \in I} \|e^{-\omega t} f(t)\| < \infty$. Clearly, $C_\omega(I; X)$ is a Banach space when endowed with the norm $\|\cdot\|_{C_\omega(I;X)}$.

Fix $f \in C_\omega([0, \infty); X)$ and $x \in X$. As it is easily seen, problem (13.18) admits a mild solution $u \in C_\omega([0, \infty); X)$ if and only if the problem

$$\begin{cases} \widetilde{u}'(t) = (A - \omega I)\widetilde{u}(t) + e^{-\omega t} f(t), & t > 0, \\ \widetilde{u}(0) = x, \end{cases} \tag{13.28}$$

admits a mild solution $\widetilde{u} \in C_b([0,\infty);X)$. In this case, $u(t) = e^{\omega t}\widetilde{u}(t)$ for every $t \geq 0$. Similarly, if $x \in X$ and $f \in C_\omega((-\infty,0];X)$, then problem (13.22) admits a mild solution $u \in C_\omega((-\infty,0];X)$ if and only if the problem

$$\begin{cases} \widetilde{u}'(t) = (A-\omega I)\widetilde{u}(t) + e^{-\omega t}f(t), & t \leq 0, \\ \widetilde{u}(0) = x, \end{cases} \tag{13.29}$$

admits a mild solutions $\widetilde{u} \in C_b((-\infty,0];X)$ and in this case $u(t) = e^{\omega t}\widetilde{u}(t)$ for every $t \leq 0$. Clearly, the operator $\widetilde{A} = A - \omega I : D(A) \subset X \to X$ is sectorial and satisfies condition (13.6). Hence, all the results of Sections 13.2 and 13.3 may be applied to problems (13.28) and (13.29). Note that the spectral projection P associated with the operator $\widetilde{A}$, introduced in Section 13.2, is given by

$$P = \frac{1}{2\pi i}\int_{\gamma_+} R(\lambda, A-\omega I)d\lambda = \frac{1}{2\pi i}\int_{\gamma_+ + \omega} R(z,A)dz, \tag{13.30}$$

where the smooth enough curve $\gamma_+ + \omega$ surrounds $\sigma_+^\omega(A) = \{\lambda \in \sigma(A) : \text{Re}\lambda > \omega\}$ and is contained in the halfplane $\{\lambda \in \mathbb{C} : \text{Re}\,\lambda > \omega\}$. Set moreover $\sigma_-^\omega(A) = \{\lambda \in \sigma(A) : \text{Re}\,\lambda < \omega\}$. Note that if $\omega > s(A)$ then P is the trivial operator.

Applying the results of the previous section we can prove the following result.

Theorem 13.4.1. *Under assumption* (13.27) *let P be the operator defined by* (13.30). *Then, the following properties are satisfied.*

(i) *If $f \in C_\omega([0,\infty);X)$ and $x \in X$, then the mild solution u to problem* (13.18) *belongs to $C_\omega([0,\infty);X)$ if and only if*[2]

$$Px = -\int_0^\infty T(-s)Pf(s)ds.$$

In this case u is given by formula (13.20), *and there exists a positive constant C_1, depending on ω, such that*

$$\|u\|_{C_\omega([0,\infty);X)} \leq C_1(\|x\| + \|f\|_{C_\omega([0,\infty);X)}).$$

(ii) *If $f \in C_\omega((-\infty,0];X)$ and $x \in X$, then problem* (13.22) *admits a mild solution $u \in C_\omega((-\infty,0];X)$ if and only if condition* (13.24) *holds. In this case, the solution is unique in $C_\omega((-\infty,0];X)$ and is given by formula* (13.25). *Finally, there exists a positive constant C_2, depending on ω, such that*

$$\|u\|_{C_\omega((-\infty,0];X)} \leq C_2(\|x\| + \|f\|_{C_\omega((-\infty,0];X)}).$$

Remark 13.4.2. All the results in this and in the previous sections, require that X is a complex Banach space. Anyway, they can be extended to the case when X is a real Banach space. Indeed, in that case, we can introduce the complexification of X as in Remark 9.2.9. If $A : D(A) \subset X \to X$ is a linear operator such that the complexification $\widetilde{A}$ is sectorial in $\widetilde{X}$, then the projection P maps X into itself. To prove this claim, it is convenient to choose as

[2] Note that, since $\sigma_+^\omega(A)$ is bounded, $T(t)P$ is well defined also for $t < 0$, and the results of Proposition 13.2.2 hold, with obvious modifications.

γ_+ the smooth curve $\gamma_+(\eta) = \omega' + re^{i\eta}$ for every $\eta \in [0, 2\pi]$, with $\omega' \in \mathbb{R}$. For every $x \in X$, we can write

$$\begin{aligned} Px =& \frac{1}{2\pi}\int_0^{2\pi} re^{i\eta}R(\omega' + re^{i\eta}, A)x\, d\eta \\ =& \frac{r}{2\pi}\int_0^{\pi} \left(e^{i\eta}R(\omega' + re^{i\eta}, A) + e^{-i\eta}R(\omega' + re^{-i\eta}, A)\right) x\, d\eta, \end{aligned}$$

and the imaginary part of the function under the integral identically vanishes. Therefore, $P(X) \subset X$, and consequently $(I - P)(X) \subset X$.

In the following example, we illustrate in a concrete case the results of this section.

Example 13.4.3. Consider the nonhomogeneous heat equation

$$\begin{cases} D_t u(t,x) = D_{xx}u(t,x) + f(t,x), & t > 0, \quad 0 \le x \le 1, \\ u(t,0) = u(t,1) = 0, & t \ge 0, \\ u(0,x) = u_0(x), & \qquad\quad 0 \le x \le 1, \end{cases} \tag{13.31}$$

where $f : [0,\infty) \times [0,1] \to \mathbb{R}$ is bounded and continuous, $u_0 : [0,1] \to \mathbb{R}$ is continuous and u_0 vanishes at the endpoints 0 and 1. As in Example 13.2.6, we choose $X = C([0,1])$, $A : D(A) = \{u \in C^2([0,1]) : u(0) = u(1) = 0\} \to X$, $Au = u''$. Since $s(A) = -\pi^2$, A is hyperbolic, and in this case $P = 0$. Proposition 13.3.1 implies that the mild solution to problem (13.31) is bounded. Note that $u_0(0) = u_0(1) = 0$ is a compatibility condition for the solution of problem (13.31) to be continuous up to $t = 0$ and to satisfy the condition $u(0,\cdot) = u_0$.

As far as exponentially decaying solutions are concerned, we can take advantage of Theorem 13.4.1(i). Fix $\omega \notin \{\pi^2 n^2 : n \in \mathbb{N}\}$ and a continuous function $f : [0,\infty) \times [0,1] \to \mathbb{R}$ such that

$$\sup_{t\ge 0,\, 0\le x\le 1} |e^{\omega t}f(t,x)| < \infty.$$

Then, the mild solution u to problem (13.31) satisfies the condition

$$\sup_{t\ge 0,\, 0\le x\le 1} |e^{\omega t}u(t,x)| < \infty$$

if and only if (13.19) holds. This is equivalent to requiring that (see Example 13.2.6)

$$\int_0^1 u_0(x)\sin(k\pi x)\, dx = -\int_0^\infty e^{k^2\pi^2 s} ds \int_0^1 f(s,x)\sin(k\pi x) dx,$$

for every natural number k such that[3] $\pi^2 k^2 < \omega$.

Let us now consider the backward problem

$$\begin{cases} D_t u(t,x) = D_{xx}u(t,x) + f(t,x), & t < 0, \quad 0 \le x \le 1, \\ u(t,0) = u(t,1) = 0, & t \le 0, \\ u(0,x) = u_0(x), & \qquad\quad 0 \le x \le 1, \end{cases} \tag{13.32}$$

[3] Recall that, since the function $x \mapsto u_k(x) = \sin(k\pi x)$ is an eigenfunction of the operator A, with eigenvalue $-k^2\pi^2$, it follows that $T(t)u_k = e^{-t\pi^2 k^2}u_k$, for every $t \in \mathbb{R}$.

to which we apply Proposition 13.3.4. Since $P = 0$, it follows that, if $f : (-\infty, 0] \times [0, 1] \to \mathbb{R}$ is bounded and continuous, then there exists only a final datum u_0 for which the mild solution to problem (13.32) is bounded in $(-\infty, 0)$. Such a datum is given by the following formula (see (13.24)):

$$u_0 = \int_{-\infty}^{0} T(-s) f(s, \cdot) ds.$$

13.5 Exercises

Exercise 13.5.1. Prove property (i) and complete the proof of property (iii) in Proposition 13.2.2 .

Exercise 13.5.2. Let $A : D(A) \subset C_b(\mathbb{R}) \to C_b(\mathbb{R})$ be the sectorial operator defined by $Au = u''$ for every $u \in D(A) = C_b^2(\mathbb{R})$. Prove that $\sigma(A)$ is the negative real axis and $\|\lambda R(\lambda, A)\| \le (\cos\theta/2)^{-1}$ for every $\lambda \in \rho(A)$, where $\theta = \arg(\lambda)$.

Exercise 13.5.3. Let $A : D(A) \subset X \to X$ be a sectorial operator and let x be an eigenvector of A with eigenvalue λ. Prove that $T(t)x = e^{\lambda t}x$ for every $t > 0$ and $R(\mu, A)x = (\mu - \lambda)^{-1}x$ for every $\mu \in \rho(A)$.

Exercise 13.5.4. Determine the spectrum of the operator A in Example 13.2.6 and prove that the operator A_2 is sectorial in $L^2((0, 1))$. Finally, show that the semigroups $\{T(t)\}$ and $\{T_2(t)\}$ coincide on $C([0, 1])$.

Exercise 13.5.5. Suppose that A is a hyperbolic sectorial operator. Prove that the spectrum of the restrictions A_+ and A_- of A to $P(X)$ and to $(I - P)(X)$ are, respectively, $\sigma_+(A)$ and $\sigma_-(A)$.
[**Hint:** prove that

$$R(\lambda, A_+) = \frac{1}{2\pi i} \int_{\gamma_+} \frac{R(\xi, A)}{\lambda - \xi} d\xi,$$

if $\lambda \notin \sigma_+$, where γ_+ is the same curve defined at the beginning of Section 13.2, and

$$R(\lambda, A_-) = -\frac{1}{2\pi i} \int_{\gamma} \frac{R(\xi, A)}{\lambda - \xi} d\xi,$$

if $\lambda \notin \sigma_-$, where $\gamma = \gamma_r$ is the usual curve used to define the operator $T(t)$ (see Figure 9.1)].

Exercise 13.5.6. Let α, $\beta \in \mathbb{R}$, and let A be the realization of the second order derivative in $C([0, 1])$, with domain $\{f \in C^2([0, 1]) : \alpha f(i) + \beta f'(i) = 0,\ i = 0, 1\}$. Determine $s(A)$.

Exercise 13.5.7. Let A satisfy (13.6), and let $T > 0$, $f : [-T, 0] \to P(X)$ be a continuous function. Prove that for every $x \in P(X)$ the backward problem

$$\begin{cases} u'(t) = Au(t) + f(t), & t \in [-T, 0], \\ u(0) = x, \end{cases}$$

admits a unique strict solution in the interval $[-T,0]$ with values in $P(X)$, given by the variation of constants formula

$$u(t) = T(t)x - \int_t^0 T(t-s)f(s)ds, \qquad t \in [-T,0].$$

Prove that for each $\omega \in [0,\omega_+)$ it holds that

$$\|u(t)\| \le C_\omega\bigg(\|x\| + \omega^{-1}\sup_{-T<t<0}\|f(t)\|\bigg).$$

Exercise 13.5.8 (A generalization of Proposition 13.2.2))**.** Let A be a sectorial operator such that $\sigma(A) = \sigma_1 \cup \sigma_2$, where σ_1 is compact, σ_2 is closed and $\sigma_1 \cap \sigma_2 = \varnothing$. Define the operator Q by setting

$$Q = \frac{1}{2\pi i}\int_\gamma R(\lambda, A)d\lambda,$$

where γ is any regular closed curve in $\rho(A)$, around σ_1, with index 1 with respect to each point in σ_1 and with index 0 with respect to each point in σ_2.

Prove that

(i) Q is a projection;

(ii) the part A_1 of A in $Q(X)$ is a bounded operator;

(iii) the group $\{T_1(t)\}$ generated by operator A_1 in $Q(X)$ is given by

$$T_1(t) = \frac{1}{2\pi i}\int_\gamma e^{\lambda t}R(\lambda, A)d\lambda.$$

Exercise 13.5.9. Let A be a hyperbolic sectorial operator. Using Propositions 13.3.1 and 13.3.4, prove that for every $f \in C_b(\mathbb{R};X)$ the equation

$$u'(t) = Au(t) + f(t), \qquad t \in \mathbb{R}, \tag{13.33}$$

admits a unique mild solution[4] $u \in C_b(\mathbb{R};X)$, given by

$$u(t) = \int_{-\infty}^t T(t-s)(I-P)f(s)ds - \int_t^\infty T(t-s)Pf(s)ds, \qquad t \in \mathbb{R}.$$

Further, prove that

(i) if f is constant, then u is constant as well;

(ii) if $\lim_{t\to\infty} f(t) = f_\infty$ (resp., $\lim_{t\to-\infty} f(t) = f_{-\infty}$) then

$$\lim_{t\to\infty} u(t) = \int_0^\infty T(s)(I-P)f_\infty ds - \int_{-\infty}^0 T(s)Pf_\infty ds$$

(resp., the same formula but with f_∞ replaced by $f_{-\infty}$);

(iii) if f is T-periodic, then u is T-periodic as well.

Exercise 13.5.10. Prove that the spectrum of the realization of the Laplacian in $C_b(\mathbb{R}^N)$ and in $L^p(\mathbb{R}^N)$ ($p \in [1,\infty)$) is $(-\infty, 0]$.
[**Hint:** to prove that $\lambda \le 0$ belongs to $\sigma(\Delta)$, use or approximate the functions $f(x_1,\ldots,x_n) = e^{i\sqrt{-\lambda}x_1}$].

[4] The definition of a mild solution of (13.33) is similar to the definition of mild solution to (13.22).

Bibliography

[1] A. Belleni-Morante, Applied Semigroups and Evolution Equations, Oxford University Press, 1979.

[2] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[3] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[4] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[5] K.-J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Graduate Text in Mathematics **194**, Springer-Verlag, 2000.

[6] K.-J. Engel and R. Nagel, Short Course on Operator Semigroups, Universitext, Springer-Verlag, 2006.

[7] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[8] J.A. Goldstein, Semigroups of Operators and Applications, Oxford University Press, 1985.

[9] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[10] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[11] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981.

[12] L. Lorenzi and A. Rhandi, Semigroups of Bounded Operators and Second-Order Elliptic and Parabolic Partial Differential Equations, CRC Press, Taylor & Francis Group, 2021.

[13] H.P. Lotz, Uniform convergence of operators on L^∞ and similar spaces, Math. Z. **190** (1985), 207–220.

[14] A. Lunardi, Analytic Semigroups and Optimal Regularity in Parabolic Problems, Birkhäuser Verlag, 1995.

[15] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Appl. Math.Sci. **44**, Springer-Verlag,1983.

[16] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[17] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[18] W. Rudin, Real and Complex Analysis, 3$^{\text{rd}}$ ed., McGraw-Hill, 1986.

[19] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[20] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958.

[21] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861.

[22] A. Zygmund, Trigonometric Series, Cambridge Univ. Press., 2nd Edition Reprinted (1968).

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

February 10, 2022

Lecture 14

Nonlinear abstract Cauchy problems

In this lecture, based on the results in Lecture 10, we study nonlinear evolution problems, associated to sectorial operators, on a Banach space X. The problems that we study here have the form

$$\begin{cases} u'(t) = Au(t) + F(t, u(t)), & t \in (0, T], \\ u(0) = u_0, \end{cases} \tag{14.1}$$

where $A : D(A) \subset X \to X$ is a sectorial operator and $F : [0, T] \times Y \to X$ is a continuous function. Here, Y is either the whole space X or an intermediate space of class J_α, for some $\alpha \in (0, 1)$, between X and $D(A)$, see Definition 10.2.6. Using a fixed-point technique, we prove the existence and uniqueness of the mild solution to problem (14.1), which can be extended to a maximal interval $[0, \tau(u_0))$, using arguments close to those used in the classical case of ordinary differential equations. In general, $\tau(u_0) < T$ and this should be surprising, if one thinks to the ordinary differential equations case. A general sufficient condition is provided which guarantees the existence in the large (i.e., $\tau(u_0) = T$) of the mild solution to problem (14.1). The results in Lecture 10 can be used to regularize the mild solution and show that it is actually a classical or strict solution (see the forthcoming definition 14.1.1).

Finally, the results are explained by some concrete examples, which allow us to introduce a less restrictive condition for the solution to problem (14.1) to be defined in the large.

Throughout the lecture, for every $T > 0$ we set

$$M_{0,T} = \sup_{0 \le t \le T} \|T(t)\|.$$

14.1 Nonlinearities defined in X

Consider the initial value problem (14.1) when $F : [0, T] \times X \mapsto X$ is a continuous function. As in the case of linear problems, we can give the following definition.

Definition 14.1.1. A function u defined in an interval $I = [0, \tau)$ or $I = [0, \tau]$, with $\tau \le T$, is said to be

- a strict solution of problem (14.1) in I if it is continuous with values in $D(A)$ and differentiable with values in X in the interval I, and it satisfies problem (14.1);

- a classical solution if it is continuous with values in $D(A)$ and differentiable with values in X in the interval $I \setminus \{0\}$, it is continuous in I with values in X, and it satisfies problem (14.1);
- a mild solution if it is continuous in $I \setminus \{0\}$ with values in X and it satisfies the equation

$$u(t) = T(t)u_0 + \int_0^t T(t-s)F(s,u(s))ds, \qquad t \in I. \tag{14.2}$$

Thanks to Proposition 10.2.11 every strict or classical solution satisfies (14.2).

14.1.1 Local existence, uniqueness, regularity

It is natural to solve problem (14.2) via a fixed point argument, applied to equation (14.2). This will allow us to guarantee the existence and uniqueness of a mild solution. Under more restrictive assumptions on F and u_0, such a solution turns out to be classical or even strict.

We assume that $F : [0,T] \times X \to X$ is continuous and locally Lipschitz continuous with respect to the second variable, uniformly with respect to time, i.e., for every $R > 0$ there exists a positive constant L_R such that

$$\|F(t,x) - F(t,y)\| \le L_R\|x-y\|, \qquad t \in [0,T],\ x,y \in B(0,R). \tag{14.3}$$

In the proof of Theorems 14.1.3 and 14.2.1, we use the following version of the Gronwall lemma, whose proof may be found for instance in [10, p. 188].

Lemma 14.1.2. *Let $0 \le a < b \le T$, and let $u : [a,b] \to \mathbb{R}$ be a nonnegative function, bounded in any interval $[a, b-\varepsilon]$, integrable and such that*

$$u(t) \le k + h\int_a^t (t-s)^{-\alpha}u(s)ds, \qquad t \in [a,b],$$

for some $\alpha \in [0,1)$, $h > 0$ and $k \ge 0$. Then, there exists a positive constant C_1, independent of a, b, k, such that $u(t) \le C_1 k$ for every $t \in [a,b)$.

Theorem 14.1.3. *Let $F : [0,T] \times X \to X$ be a continuous function satisfying* (14.3)*. Then, the following properties are satisfied.*

(i) *If $u, v \in C_b((0,a];X)$ are mild solutions to problem* (14.1) *for some $a \in (0,T]$, then u and v coincide.*

(ii) *For every $\overline{u} \in X$ there exist r, $\delta > 0$, $K > 0$ such that for $\|u_0 - \overline{u}\| \le r$ problem* (14.1) *admits a mild solution $u = u(\cdot;u_0) \in C_b((0,\delta];X)$. This function belongs to $C([0,\delta];X)$ if and only if $u_0 \in \overline{D(A)}$. Moreover, u depends continuously on the datum u_0, i.e., for every u_0, $u_1 \in B(\overline{u}, r)$ there exists a positive constant K such that*

$$\|u(t;u_0) - u(t;u_1)\| \le K\|u_0 - u_1\|, \qquad t \in [0,\delta]. \tag{14.4}$$

Proof. (i) Let $u, v \in C_b((0,a];X)$ be mild solutions to (14.1) and set $w = v - u$. By (14.2), the function w satisfies

$$w(t) = \int_0^t T(t-s)[F(s,v(s)) - F(s,u(s))]ds, \qquad t \in (0,a].$$

Using estimate (14.3) with $R = \max\{\sup_{0<t<a} \|u(t)\|, \sup_{0<t<a} \|v(t)\|\} + 1$ we can estimate

$$\|w(t)\| \leq L_R M_{0,T} \int_0^t \|w(s)\| ds.$$

Gronwall's lemma (with $\alpha = 0$) implies that $w = 0$ in $[0, a]$.

(ii) Fix $\overline{u} \in X$ and $R > 0$ such that $R \geq 8M_{0,T}\|\overline{u}\|$. If $u_0 \in \overline{B(\overline{u}, r)}$, where $r = (8M_{0,T})^{-1}R$, then we get $\sup_{t\in[0,T]} \|T(t)u_0\| \leq 4^{-1}R$.

We look for a mild solution to problem (14.1) in the closed ball Y of $C_b((0, \delta]; X)$, centered at zero and with radius R, where $\delta \in (0, T]$ has to be chosen properly. As it is easily seen, the function $F(\cdot, v(\cdot))$ belongs to $C_b((0, \delta]; X)$ for every $v \in Y$.

We define the operator Γ in Y, by setting

$$(\Gamma(v))(t) = T(t)u_0 + \int_0^t T(t-s)F(s, v(s))ds, \qquad t \in [0, \delta].$$

Clearly, a function $v \in Y$ is a mild solution of (14.1) in $[0, \delta]$ if and only if it is a fixed point of the operator Γ. Note that the function $\Gamma(v)$ is continuous in $(0, \delta]$ with values in X, if $v \in C_b((0, \delta]; X)$, as it can be easily checked arguing as in the proof of Proposition 10.2.12.

We shall show that Γ is a contraction and maps Y, equipped with the sup-norm, into itself provided δ is sufficiently small. For this purpose, we fix $v_1, v_2 \in Y$ and observe that, taking (14.3) into account, we can estimate

$$\|\Gamma(v_1) - \Gamma(v_2)\|_\infty \leq \delta M_{0,T}\|F(\cdot, v_1(\cdot)) - F(\cdot, v_2(\cdot))\|_\infty \leq \delta M_{0,T} L_R \|v_1 - v_2\|_\infty.$$

Therefore, if $\delta \leq \delta_0 = (2M_{0,T}L_R)^{-1}$, then Γ is a 1/2-contraction in Y. Moreover,

$$\begin{aligned}\|\Gamma(v)\|_\infty \leq & \|\Gamma(v) - \Gamma(0)\|_\infty + \|\Gamma(0)\|_\infty \leq \frac{R}{2} + \|T(\cdot)u_0\|_\infty + M_{0,T}\delta\|F(\cdot, 0)\|_\infty \\ \leq & \frac{3}{4}R + M_{0,T}\delta\|F(\cdot, 0)\|_\infty\end{aligned}$$

for every $v \in Y$ and $\delta \leq \delta_0$. Up to replacing δ_0 with a smaller value, if needed, we can assume that $M_{0,T}\delta\|F(\cdot, 0)\|_\infty \leq R/4$, so that Γ maps Y into itself and we conclude that it admits a unique fixed point.

Concerning the continuity of u up to $t = 0$, the function $u - T(\cdot)u_0$ is continuous up to $t = 0$ (as it can be seen adapting the arguments in the proof of Proposition 10.2.12), whereas the function $T(\cdot)u_0$ is continuous at $t = 0$ if and only if $u_0 \in \overline{D(A)}$ (see Proposition 9.2.5). Therefore, $u \in C([0, \delta]; X)$ if and only if $u_0 \in \overline{D(A)}$.

To conclude the proof, let us check estimate (14.4). For this purpose, we fix u_0 and u_1 in the ball of $C_b((0, \delta]; X)$ centered at $\overline{u}$ and with radius r. Since Γ is a 1/2-contraction in Y and both $u(\cdot; u_0)$, $u(\cdot; u_1)$ belong to Y, it follows that $\|u(\cdot; u_0) - u(\cdot; u_1)\|_\infty \leq 2\|T(\cdot)(u_0 - u_1)\|_\infty \leq 2M_{0,T}\|u_0 - u_1\|$, so that estimate (14.4) follows with $K = 2M_{0,T}$. □

14.1.2 The maximally defined solution

Now we can construct a maximally defined solution as follows. Set

$$\begin{cases} \tau(u_0) = \sup\{a > 0 : \text{problem (14.1) admits a mild solution } u_a \text{ in } [0, a]\} \\ u(t; u_0) = u_a(t), \quad \text{if } t \leq a. \end{cases}$$

Function $u(t; u_0)$ is well defined in the interval $I(u_0) = [0, \tau(u_0))$, thanks to Theorem 14.1.3(i). Moreover, if $\tau(u_0) < T$, then u is not defined at $\tau(u_0)$ because, otherwise, the solution could be extended to a larger interval, contradicting the definition of $\tau(u_0)$ (see Exercise 14.3.2).

We now discuss the regularity of $u(\cdot; u_0)$ and provide sufficiently conditions for the mild solutions to be classic or even strict.

Proposition 14.1.4. *Let F satisfy condition* (14.3) *and assume in addition that there exists a constant $\theta \in (0, 1)$, and, for every $R > 0$, a positive constant C_R such that*

$$\|F(t, x) - F(s, x)\| \leq C_R(t - s)^{\theta}, \qquad 0 \leq s \leq t \leq T, \quad \|x\| \leq R. \tag{14.5}$$

Then, for every $u_0 \in X$, the function u belongs to $C^{\theta}([\varepsilon, \tau(u_0) - \varepsilon]; D(A)) \cap C^{1+\theta}([\varepsilon, \tau(u_0) - \varepsilon]; X)$ and u' is bounded in $[\varepsilon, \tau(u_0) - \varepsilon]$ with values in $D_A(\theta, \infty))$ for every $\varepsilon \in (0, \tau(u_0)/2)$. Moreover the following statements hold.

(i) *If $u_0 \in \overline{D(A)}$, then $u(\cdot; u_0)$ is a classical solution of problem* (14.1).

(ii) *If $u_0 \in D(A)$ and $Au_0 + F(0, u_0) \in \overline{D(A)}$, then $u(\cdot; u_0)$ is a strict solution of problem* (14.1).

Proof. Fix $a \in (0, \tau(u_0))$ and $\varepsilon \in (0, a)$. Since the function $F(\cdot, u(\cdot))$ belongs to $C_b((0, a]; X)$, adapting the argument in the proof of Proposition 10.2.12, it can be easily checked that the integral term (say the function v) in the definition of mild solution (see (14.2)) belongs to $C^{\theta}([0, a]; X)$. Moreover, the function $T(\cdot)u_0$ belongs to $C^{\infty}([\varepsilon, a]; X)$. Summing up, we conclude that u belongs to $C^{\theta}([\varepsilon, a]; X)$. Now, conditions (14.3) and (14.5) imply that the function $F(\cdot, u(\cdot))$ belongs to $C^{\theta}([\varepsilon, a]; X)$. Since

$$u(t) = T(t - \varepsilon)u(\varepsilon) + \int_{\varepsilon}^{t} T(t - s)F(s, u(s))ds, \qquad t \in [\varepsilon, a],$$

we may apply Theorem 10.2.14 in the interval $[\varepsilon, a]$ (see Remark 10.2.17) and deduce that u belongs to $C^{\theta}([2\varepsilon, a]; D(A)) \cap C^{1+\theta}([2\varepsilon, a]; X)$ for each $\varepsilon \in (0, a/2)$, and $u'(t) = Au(t) + F(t, u(t))$ for every $t \in (\varepsilon, a]$. Exercise 10.4.8 implies that u' is bounded with values in $D_A(\theta, \infty)$ in $[\varepsilon, a]$. Since a and ε are arbitrary, we conclude that $u \in C^{\theta}([\varepsilon, \tau(u_0) - \varepsilon]; D(A)) \cap C^{1+\theta}([\varepsilon, \tau(u_0) - \varepsilon]; X)$ for each $\varepsilon \in (0, \tau(u_0)/2)$. If $u_0 \in \overline{D(A)}$, then the function $T(\cdot)u_0$ is continuous up to 0, and statement (i) follows.

(ii). The arguments in the proof of Proposition 10.2.12 show that the function v is θ-Hölder continuous up to $t = 0$ with values in X. Since $u_0 \in D(A) \subset D_A(\theta, \infty)$, then the function $T(\cdot)u_0$ is θ-Hölder continuous up to $t = 0$, as well. Hence, u is θ-Hölder continuous up to $t = 0$ with values in X, so that the function $F(\cdot, u(\cdot))$ is θ-Hölder continuous in $[0, a]$ with values in X. Statement (ii) follows now from Theorem 10.2.14(ii). □

Proposition 14.1.5. *Assume that F satisfies* (14.3). *Let u_0 be such that $I(u_0) \neq [0, T]$. Then, the function $t \mapsto \|u(t)\|$ is unbounded in $I(u_0)$.*

Proof. Assume by contradiction that u is bounded in $[0, \tau(u_0))$. Then, the function $F(\cdot, u(\cdot; u_0))$ is bounded and continuous in the interval $(0, \tau(u_0))$ with values in X. Since u satisfies the variation of constants formula (14.2), it may be continuously extended at $t = \tau(u_0)$, in such a way that the extension is Hölder continuous in every interval $[\varepsilon, \tau(u_0)]$, with $0 < \varepsilon < \tau(u_0)$.

Indeed, the function $T(\cdot)u_0$ is well defined and analytic in the whole $(0,\infty)$, and the function $u - T(\cdot)u_0$ belongs to $C^\alpha([0,\tau(u_0)];X)$ for each $\alpha \in (0,1)$, due to Proposition 10.2.12.

By Theorem 14.1.3, the problem

$$\begin{cases} v'(t) = Av(t) + F(t, v(t)), & t \geq \tau(u_0), \\ v(\tau(u_0)) = u(\tau(u_0)), \end{cases}$$

admits a unique mild solution $v \in C([\tau(u_0), \tau(u_0)+\delta]; X)$ for some $\delta > 0$. Note that the function v is continuous up to $t = \tau(u_0)$ because $u(\tau(u_0)) \in \overline{D(A)}$.

The function $w : [0, \tau(u_0)+\delta] \to X$, defined by

$$w(t) = \begin{cases} u(t), & t \in [0, \tau(u_0)), \\ v(t), & t \in [\tau(u_0), \tau(u_0)+\delta], \end{cases}$$

is a mild solution to problem (14.1) in $[0, \tau(u_0)+\delta]$ (see Exercise 14.3.2). This contradicts the definition of $\tau(u_0)$. Therefore, $u(\cdot; u_0)$ cannot be bounded. □

Proposition 14.1.5 provides us with a very useful tool to prove the existence in the large of the mild solution to problem (14.1). Indeed, if we are able to prove that the norm of $u(t)$ stays bounded when t runs in the interval $I(u_0)$, then we conclude that the mild solution actually exists in the large. Such an a priori estimate on the sup-norm of u can be easily proved if the nonlinear term F grows not more than linearly as $\|x\|$ tends to ∞, as in the case of ordinary differential equations.

Proposition 14.1.6. *Assume that there exists a positive constant C such that*

$$\|F(t,x)\| \leq C(1 + \|x\|), \qquad t \in [0,T],\ x \in X. \tag{14.6}$$

Then, the mild solution $u : I(u_0) \to X$ to problem (14.1) *is bounded in $I(u_0)$ with values in X, so that $I(u_0) = [0,T]$.*

Proof. Fix $t \in I(u_0)$. Using condition (14.6), we obtain

$$\|u(t)\| \leq M_{0,T}\|u_0\| + M_{0,T}C\left(T + \int_0^t \|u(s)\| ds\right).$$

Applying Lemma 14.1.2 (with $\alpha = 0$) to the real valued function $t \mapsto \|u(t)\|$, we conclude that

$$\|u(t)\| \leq C_1(M_{0,T}\|u_0\| + M_{0,T}CT), \qquad t \in I(u_0),$$

and the assertion follows. □

Remark 14.1.7. Condition (14.6) is satisfied if condition (14.3) holds true with a constant L, independent of R, i.e., if F is globally Lipschitz continuous with respect to x, uniformly with respect to t.

14.1.3 Examples: reaction-diffusion equations and systems

Consider the problem

$$(14.7)\qquad \begin{cases} D_t u(t,x) = D\Delta u(t,x) + f(t,x,u(t,x)), & t>0, \quad x\in\Omega;\\ u(t,x)=0, & t>0, \quad x\in\partial\Omega,\\ u(0,x)=u_0(x), & \qquad\quad x\in\Omega, \end{cases}$$

in the unknown $u = (u_1,\dots,u_m)$, where $D = \mathrm{diag}(d_1,\dots,d_m)$ is a diagonal matrix with positive entries on the main diagonal and $\Omega\subset\mathbb{R}^d$ is a bounded and sufficiently smooth domain.

This type of problems are often encountered as mathematical models in chemistry and biology. The part $D\Delta u$ in the system is called diffusion part, the numbers d_i are called diffusion coefficients, f is the reaction part of the equation. We refer the reader e.g., to the monographs [16, 19, 22] for further details.

For simplicity, we confine our analysis to the case $d = 1$ and $\Omega = [0,1]$.

On $X = C([0,1];\mathbb{R}^m)$ consider the linear operator $A : D(A) = \{u\in C^2([0,1];\mathbb{R}^m) : u(0) = u(1) = 0\}\to X$, defined by $Au = Du''$, which is sectorial in X and $\overline{D(A)} = C_0([0,1];\mathbb{R}^m) = \{u\in C([0,1];\mathbb{R}^m) : u(0) = u(1) = 0\}$ (see Exercises 9.4.7(ii) and 14.3.5). Assume that the function $f : [0,T]\times[0,1]\times\mathbb{R}^m\to\mathbb{R}^m$ is continuous and there exists a constant $\theta\in(0,1)$ and, for every $R > 0$ a positive constant K_R, such that

$$(14.8)\qquad \|f(t,x,u) - f(s,y,v)\| \le K_R(|t-s|^\theta + \|u-v\|),$$

for every $s,t\in[0,T]$, $x,y\in[0,1]$, $u,v\in\overline{B(0,R)}\subset\mathbb{R}^m$. Then, we may apply the previous theorems to obtain a smooth solution of problem (14.7).

Proposition 14.1.8. *If f satisfies the condition(14.8) and $u_0\in C([0,1],\mathbb{R}^m)$, then there exist a maximal interval $I(u_0)$ and a unique solution u to (14.7) in $I(u_0)\times[0,1]$, such that $u\in C(I(u_0)\times[0,1];\mathbb{R}^m)$, $D_t u$ and $D_x u$, $D_{xx}u$ are bounded and continuous in $[\varepsilon,\tau(u_0)-\varepsilon]\times[0,1]$ for every $\varepsilon\in(0,\tau(u_0)/2)$, where $\tau(u_0) = \sup I(u_0)$.*

Proof. To begin with, we observe that the function $F : [0,T]\times X\to X$, defined by $F(t,\varphi)(x) = f(t,x,\varphi(x))$ for every $t\in[0,T]$, $x\in[0,1]$ and $\varphi\in X$, is continuous and satisfies the conditions (14.3) and (14.5). Indeed, fix any φ_1, $\varphi_2\in B(0,R)\subset X$. From (14.8), we deduce that $\|F(t,\varphi_1) - F(s,\varphi_2)\|_\infty \le K(|t-s|^\theta + \|\varphi_1-\varphi_2\|_\infty)$ for every $s,t\in[0,T]$. Theorem 14.1.3 guarantees the existence and uniqueness of a mild solution $u\in C_b((0,\delta];X)$ to problem (14.1), that may be extended to a maximal time interval $I(u_0)$.

By Proposition 14.1.4, u, u' and Au are continuous in $(0,\tau(u_0))$ with values in X (in fact, Hölder continuous in each compact subinterval). Then, the function $(t,x)\mapsto u(t,x) := u(t)(x)$ is bounded and continuous in $[0,a]\times[0,1]$ for each $a\in I(u_0)$ and is continuously differentiable with respect to t in $I(u_0)\times[0,1]$. Moreover, u is twice continuously differentiable with respect to the spatial variables in $I(u_0)\times[0,1]$ as well. □

A sufficient condition for u to be bounded, and, hence for u to be defined in the large, is the following:

$$(14.9)\qquad \|f(t,x,u)\| \le C(1+\|u\|), \qquad t\in[0,T],\ x\in[0,1],\ u\in\mathbb{R}^m.$$

Indeed, in this case the nonlinear function $F : [0,T] \times X \to X$, introduced in the proof of Proposition 14.1.8, satisfies condition (14.6).

Up to now, in this subsection we have dealt with real valued functions just because in most mathematical models the unknown u is real valued. But we could replace $C([0,1];\mathbb{R}^m)$ by $C([0,1];\mathbb{C}^m)$ as well without any modification in the proofs, getting the same results in the case of complex valued data. On the contrary, the results in the rest of this subsection hold true only for real valued functions.

Using the well known properties of the first and second order derivatives of real valued functions at relative maximum or minimum points it is possible to find estimates on the solutions to several first- or second-order partial differential equations. Such techniques are called maximum principles.

Proposition 14.1.9. *Let $f : [0,T] \times [0,1] \times \mathbb{R}^m \to \mathbb{R}^m$ be a continuous function satisfying the condition $\|f(t,x,u) - f(s,x,v)\| \leq K_R(|t-s|^\theta + \|u-v\|)$ for every $s,t \in [0,T]$, $x \in [0,1]$, $u,v \in \overline{B(0,R)} \subset \mathbb{R}^m$, $R > 0$ and some positive constant K_R. Further, assume that*

$$\langle y, f(t,x,y)\rangle \leq C(1+y^2), \qquad t \in [0,T],\ x \in [0,1],\ y \in \mathbb{R}^m, \tag{14.10}$$

for some nonnegative constant C. Then, for every $u_0 \in C([0,1])$, the solution to problem (14.7) is bounded in $I(u_0) \times [0,1]$ and, therefore, it exists in the large. If $C = 0$ in (14.10), then

$$\sup_{(t,x)\in I(u_0)\times[0,1]} \|u(t,x)\| = \|u_0\|_\infty.$$

Proof. Fix $\lambda > C$, $a < \tau(u_0)$ and set $v(t,x) = e^{-\lambda t}u(t,x)$ for every $t \in [0,a]$ and $x \in [0,1]$. Such a function solves the problem

$$\begin{cases} D_t v(t,x) = \Delta v(t,x) + f(e^{\lambda t}v(t,x))e^{-\lambda t} - \lambda v(t,x), & t > 0, \quad x \in [0,1], \\ v(t,0) = v(t,1) = 0, & t > 0, \\ v(0,x) = u_0(x), & \qquad\quad x \in [0,1]. \end{cases}$$

Let $w : [0,a] \times [0,1] \to \mathbb{R}$ be the scalar function defined by $w(t,x) = \|v(t,x)\|^2 = \sum_{i=1}^m v_i(t,x)^2$ for every $(t,x) \in [0,a] \times [0,1]$. As it is easily seen, $D_t w = 2\langle D_t v, v\rangle$ and $D_{xx}w = 2\sum_{i=1}^m \|D_x v_i\|^2 + 2\langle v, D_{xx}v\rangle$.

Since w is continuous, there exists a point $(t_0,x_0) \in [0,a] \times [0,1]$ such that $w(t_0,x_0) = \pm\|w\|_{C([0,a]\times[0,1])}$. The point (t_0,x_0) is either a point of positive maximum or of negative minimum for w. Assume for instance that (t_0,x_0) is a maximum point. If $t_0 = 0$, then it follows immediately that $\|v\|_\infty \leq \|u_0\|_\infty$. If $t_0 > 0$, then, clearly, we can assume that $x_0 \in (0,1)$. Then, $D_t w(t_0,x_0) \geq 0$, $D_{xx}w(t_0,x_0) \leq 0$ and, hence, $\langle w(t_0,x_0), D_{xx}w(t_0,x_0)\rangle \leq 0$. Writing the differential system at (t_0,x_0) and taking the inner product with $w(t_0,x_0)$, we get

$$\begin{aligned} 0 \leq &\langle D_t v(t_0,x_0), v(t_0,x_0)\rangle \\ = &\langle D_{xx}v(t_0,x_0), v(t_0,x_0)\rangle + \langle f(e^{\lambda t_0}v(t_0,x_0)), v(t_0,x_0)e^{-\lambda t_0}\rangle - \lambda|v(t_0,x_0)|^2 \\ \leq &C(1 + \|v(t_0,x_0)\|^2) - \lambda\|v(t_0,x_0)\|^2, \end{aligned}$$

so that $\|v\|_\infty^2 \leq C/(\lambda - C)$. Therefore, $\|v\|_\infty \leq \max\{\|u_0\|_\infty, \sqrt{C/(\lambda-C)}\}$, and consequently $\|u\|_{C_b([0,a]\times[0,1])} \leq e^{\lambda T}\max\{\|u_0\|_\infty, \sqrt{C/(\lambda-C)}\}$. If (t_0,x_0) is a minimum point, then the proof is completely analogous and leads us to the same conclusion. The first part of the statement follows. by letting a tend to $\tau(u_0)$.

To complete the proof, we observe that, if $C = 0$, then from the previous estimate we obtain that $\|u\|_\infty \le e^{\lambda T}\|u_0\|_\infty$ for every $\lambda > 0$ and letting λ tend to 0, we conclude that $\|u\|_\infty \le \|u_0\|_\infty$. □

Remark 14.1.10. The results in this subsection can be applied also to the case when homogeneous Neumann boundary conditions are prescribed on the boundary of $[0,1]$, which can be replaced also by a bounded and smooth enough open subset of $\mathbb{R}^d$. They can also be extended to the case when the equation is set in the whole $\mathbb{R}^d$, of course with no boundary conditions.

Remark 14.1.11. Inequality (14.9) is a growth condition at infinity, while (14.10) is only an algebraic condition. For instance, it is satisfied when $f(t,x,u) = \lambda u - u^{2k+1}$ for some $k \in \mathbb{N}$ and $\lambda \in \mathbb{R}$. The sign $-$ is extremely important. Consider the problem

$$\begin{cases} D_t u(t,x) = \Delta u(t,x) + |u(t,x)|^{1+\varepsilon}, & t > 0, \quad x \in \mathbb{R}, \\ u(0,x) = \overline{u}(x), & \qquad\quad\; x \in \mathbb{R}, \end{cases}$$

with $\varepsilon > 0$. If $\overline{u}$ is constant in $\mathbb{R}^n$, then the solution u is independent of x and it coincides with the solution to the ordinary differential equation

$$\begin{cases} \xi'(t) = |\xi(t)|^{1+\varepsilon}, & t > 0, \\ \xi(0) = \overline{u}, \end{cases}$$

which blows up in a finite time if $\overline{u} > 0$.

14.2 Nonlinearities defined in intermediate spaces

In this section, X_α is any space of class J_α between X and $D(A)$, for some $\alpha \in (0,1)$. Consider the Cauchy problem

$$\begin{cases} u'(t) = Au(t) + F(t,u(t)), & t \in (0,T], \\ u(0) = u_0 \in X_\alpha, \end{cases} \tag{14.11}$$

where $F : [0,T] \times X_\alpha \to X$ is a continuous function, for some $T > 0$. The definition of strict, classical, or mild solution to (14.11) is similar to the definition in Section 14.1.

The Lipschitz condition (14.3) is replaced by a similar assumption: for each $R > 0$ there exists a positive constant L_R such that

$$\|F(t,x) - F(t,y)\| \le L_R \|x - y\|_{X_\alpha}, \qquad t \in [0,T],\ x, y \in \overline{B(0,R)} \subset X_\alpha, \tag{14.12}$$

Since $D(A) \hookrightarrow X_\alpha \hookrightarrow X$, the function $T(\cdot)$ is analytic in $(0,+\infty)$ with values in $\mathcal{L}(X_\alpha)$. Nevertheless, the norm $\|T(t)\|_{\mathcal{L}(X_\alpha)}$ could blow up as $t \to 0$ (see Exercise 14.3.6). To avoid this situation, we assume the following condition

$$\limsup_{t \to 0} \|T(t)\|_{\mathcal{L}(X_\alpha)} < \infty, \tag{14.13}$$

which implies that the function $t \mapsto \|T(t)\|_{\mathcal{L}(X_\alpha)}$ is bounded on every compact interval contained in $[0,\infty)$.

In this section we set

$$M_{\alpha,T} = \sup_{0 \le t \le T} \|T(t)\|_{\mathcal{L}(X_\alpha)}, \qquad M_{k,\alpha,T} = \sup_{t \in (0,T]} t^\alpha \|T(t)\|_{\mathcal{L}(X,X_\alpha)}. \tag{14.14}$$

14.2.1 Local existence, uniqueness, regularity

As in the case of nonlinearities defined in the whole X, it is convenient to look for a local mild solution at first, and then to see that under reasonable assumptions the solution is classical or strict.

The proof of the local existence and uniqueness theorem for mild solutions is quite similar to the proof of Theorem 14.1.3, but we need an extension of Proposition 10.2.12.

Theorem 14.2.1. *Let $F : [0,T] \times X_\alpha \to X$ be a continuous function satisfying* (14.12). *Then, the following properties are satisfied.*

(i) *If $u, v \in C_b((0,a];X)$ are mild solutions for some $a \in (0,T]$, then $u \equiv v$.*

(ii) *For each $\overline{u} \in X_\alpha$ there are r, $\delta > 0$, $K > 0$ such that if $\|u_0 - \overline{u}\|_{X_\alpha} \le r$ then problem* (14.11) *has a mild solution $u = u(\cdot;u_0) \in C_b((0,\delta];X_\alpha)$. The function u belongs to $C([0,\delta];X_\alpha)$ if and only if $u_0 \in \overline{D(A)}^{X_\alpha}$ (=closure of $D(A)$ in X_α). Moreover, for u_0, $u_1 \in B(\overline{u},r)$ it holds that*

$$\|u(t;u_0) - u(t;u_1)\|_{X_\alpha} \le K\|u_0 - u_1\|_{X_\alpha}, \qquad t \in [0,\delta].$$

Proof. (i) The proof can be obtained arguing as in the proof of Proposition 14.1.3(i), using Lemma 14.1.2.

(ii) Let $M_{\alpha,T}$ be the constant defined by (14.14). Fix $R \ge 8M_{\alpha,T}\|\overline{u}\|_{X_\alpha}$ and observe that $\sup_{t\in[0,T]} \|T(t)u_0\|_{X_\alpha} \le R/4$ for every u_0 in the closed ball of X_α, centered at $\overline{u}$ and with radius $r = (8M_{\alpha,T})^{-1}R$.

We look for a local mild solution in the space $\mathcal{Y} = \{u \in C_b((0,\delta];X_\alpha) : \|u\|_{\infty,\alpha} := \sup_{t\in(0,\delta]} \|u(t)\|_{X_\alpha} \le R\}$, where $\delta \in (0,T]$ will be chosen later. As it is easily seen, for each function $v \in \mathcal{Y}$ the function $F(\cdot, v(\cdot))$ belongs to $C_b((0,\delta];X)$. Thus, we can define a nonlinear operator Γ in $\mathcal{Y}$ by setting

$$\Gamma(v)(t) = T(t)u_0 + \int_0^t T(t-s)F(s,v(s))ds, \qquad t \in [0,\delta], \;\; v \in \mathcal{Y}.$$

A function $v \in \mathcal{Y}$ is a mild solution to (14.11) in $[0,\delta]$ if and only if it is a fixed point of Γ.

We shall show that Γ is a contraction, and it maps $\mathcal{Y}$ into itself, provided δ is small enough.

Let v_1, $v_2 \in \mathcal{Y}$. By Exercise 10.4.7, $\Gamma(v_1)$ and $\Gamma(v_2)$ belong to $C_b((0,\delta];X_\alpha)$ and

$$\begin{aligned}\|\Gamma(v_1) - \Gamma(v_2)\|_{\infty,\alpha} \le& \frac{M_{0,\alpha,T}}{1-\alpha}\delta^{1-\alpha}\|F(\cdot,v_1(\cdot)) - F(\cdot,v_2(\cdot))\|_\infty \\ \le& \frac{M_{0,\alpha,T}}{1-\alpha}\delta^{1-\alpha}L\|v_1 - v_2\|_{\infty,\alpha}.\end{aligned}$$

Therefore, if

$$\delta \le \delta_0 = \left(\frac{2M_{0,\alpha,T}L_R}{1-\alpha}\right)^{-1/(1-\alpha)},$$

then Γ is a 1/2-contraction in $\mathcal{Y}$. Moreover, for each $v \in \mathcal{Y}$ and $t \in [0,\delta]$, with $\delta \le \delta_0$, we can estimate

$$\|\Gamma(v)\|_{\infty,\alpha} \le \|\Gamma(v) - \Gamma(0)\|_{\infty,\alpha} + \|\Gamma(0)\|_{\infty,\alpha} \le \frac{R}{2} + \|T(\cdot)u_0\|_{\infty,\alpha} + C\delta^{1-\alpha}\|F(\cdot,0)\|_\infty$$

$$\leq \frac{3}{4}R + C\delta^{1-\alpha}\|F(\cdot,0)\|_\infty.$$

Therefore, if $\delta \leq \delta_0$ is such that $C\delta^{1-\alpha}\|F(\cdot,0)\|_{C([0,\delta];X)} \leq R/4$, then Γ maps $\mathcal{Y}$ into itself, and it admits a unique fixed point in $\mathcal{Y}$.

Concerning the continuity of u up to $t = 0$, we observe that the function $v = u - T(\cdot)u_0$ belongs to $C([0,\delta];X_\alpha)$, while the function $T(\cdot)u_0$ belongs to $C([0,\delta];X_\alpha)$ if and only if $u_0 \in \overline{D(A)}^{X_\alpha}$ (see Exercise 14.3.8). Therefore, $u \in C([0,\delta];X_\alpha)$ if and only if $u_0 \in \overline{D(A)}^{X_\alpha}$.

The statements about continuous dependence on the initial data may be proved precisely as in Theorem 14.1.3. □

The local mild solution to problem (14.11) is extended to a maximal time interval $[0,\tau(u_0))$ as in Section 14.1.

Without important modifications in the proofs it is also possible to deal with regularity and behavior of the solution near $\tau(u_0)$, obtaining results similar to the ones of Propositions 14.1.4 and 14.1.5.

Proposition 14.2.2. *Let F satisfy condition* (14.12)*. Further, assume that there exists $\theta \in (0,1)$ and, for every $R > 0$, a positive constant C_R such that*

$$\|F(t,x) - F(s,x)\| \leq C_R|t-s|^\theta, \qquad s,t \in [0,T],\ \|x\|_{X_\alpha} \leq R.$$

Then, u belongs to $C^\theta([\varepsilon,\tau(u_0)-\varepsilon];D(A)) \cap C^{1+\theta}([\varepsilon,\tau(u_0)-\varepsilon];X)$, and u' belongs to $B([\varepsilon,\tau-\varepsilon];D_A(\theta,\infty))$ for each $\varepsilon \in (0,\tau(u_0)/2)$. Moreover, if $u_0 \in X_\alpha \subset \overline{D(A)}$ then $u(\cdot;u_0)$ is a classical solution to problem (14.11)*. Finally, if $u_0 \in D(A)$ and $Au_0 + F(0,u_0) \in \overline{D(A)}$ then u is a strict solution to* (14.11)*.*

Proposition 14.2.3. *Assume that the function F satisfies* (14.12) *and let u_0 be such that $I(u_0) \neq [0,T]$. Then, the function $t \mapsto \|u(t)\|_{X_\alpha}$ is unbounded in $I(u_0)$.*

The simplest situation in which it is possible to show that the X_α-norm of u is bounded in $I(u_0)$ for each initial datum u_0 is again the case when F grows not more than linearly with respect to x as $\|x\|_{X_\alpha}$ tends to ∞.

Proposition 14.2.4. *Assume that condition* (14.12) *holds and there exists a positive constant C such that*

$$\|F(t,x)\| \leq C(1+\|x\|_{X_\alpha}), \qquad t \in [0,T],\ x \in X_\alpha. \tag{14.15}$$

Let $u : I(u_0) \to X_\alpha$ be the mild solution to problem (14.11)*. Then, u is bounded in $I(u_0)$ with values in X_α and, hence, $I(u_0) = [0,T]$.*

Proof. For each $t \in I(u_0)$ we can estimate

$$\begin{aligned}\|u(t)\|_{X_\alpha} \leq& M_{\alpha,T}\|u_0\|_{X_\alpha} + M_{0,\alpha,T}\int_0^t (t-s)^{-\alpha}C(1+\|u(s)\|_{X_\alpha})ds \\ \leq& M_{\alpha,T}\|u_0\|_{X_\alpha} + CM_{0,\alpha,T}\left(\frac{T^{1-\alpha}}{1-\alpha} + \int_0^t \frac{\|u(s)\|_{X_\alpha}}{(t-s)^\alpha}ds\right),\end{aligned}$$

where the constants $M_{\alpha,T}$ and $M_{0,\alpha,T}$ are defined in (14.14) Lemma 14.1.2 implies that

$$\|u(t)\|_{X_\alpha} \leq C_1\left(M_{\alpha,T}\|u_0\|_{X_\alpha} + \frac{CM_{0,\alpha,T}T^{1-\alpha}}{1-\alpha}\right), \qquad t \in I(u_0),$$

and the statement follows. □

The growth condition (14.15) sounds rather restrictive. If we have some a priori estimate for the solution to (14.11) in the X-norm (this happens in several applications to PDE's), it is possible to find *a priori* estimates in the $D_A(\theta,\infty)$-norm if F satisfies suitable growth conditions, less restrictive than (14.15). Since $D_A(\theta,\infty)$ is continuously embedded in X_α for $\theta>\alpha$ by Proposition 10.2.7, we get an a priori estimate for the solution in the X_α-norm, that yields existence in the large.

Proposition 14.2.5. *Assume that condition* (14.12) *holds and there exists an increasing function* $\mu:[0,+\infty)\to[0,+\infty)$ *such that*

$$\|F(t,x)\|\le \mu(\|x\|)(1+\|x\|_{X_\alpha}^\gamma),\qquad t\in[0,T],\ x\in X_\alpha,\tag{14.16}$$

for some $\gamma\in(1,\alpha^{-1})$. *Let* $u:I(u_0)\to X_\alpha$ *be the mild solution to problem* (14.11). *If u is bounded in $I(u_0)$ with values in X, then it is bounded in $I(u_0)$ with values in X_α. Hence,* $I(u_0)=[0,T]$.

Proof. Let us fix $0<a<I(u_0)$. Set $I_a=\{t\in I(u_0),\ t\ge a\}$ and denote by K the supremum of $\|u(t)\|$ over $I(u_0)$. Since $u\in C_b((0,a];X_\alpha)$, it suffices to show that u is bounded in I_a with values in X_α. We show that it is bounded in I_a with values in $D_A(\theta,\infty)$, when $\theta=\alpha\gamma$. This will conclude the proof due to Proposition 10.2.7.

Observe that $u(a)\in D_A(\theta,\infty)$ and that it satisfies the variation of constants formula

$$u(t)=T(t-a)u(a)+\int_a^t T(t-s)F(s,u(s))ds,\qquad t\in I(a).$$

Using the interpolatory estimate $\|x\|_{X_\alpha}\le c\|x\|^{1-\alpha/\theta}\|x\|_{D_A(\theta,\infty)}^{\alpha/\theta}$, with $c=c(\alpha,\theta)$, that holds for every $x\in D_A(\theta,\infty)$, see Exercise 10.4.6(b). We get

$$\|u(s)\|_{X_\alpha}^\gamma\le c\|u(s)\|^{\gamma(1-\alpha/\theta)}\|u(s)\|_{D_A(\theta,\infty)}^{\alpha\gamma/\theta}\le cK^{\gamma(1-\alpha/\theta)}\|u(s)\|_{D_A(\theta,\infty)},$$

so that $\|F(s,u(s))\|\le\mu(K)(1+cK^{\gamma(1-\alpha/\theta)}\|u(s)\|_{D_A(\theta,\infty)})$. Let $M_{\theta,T}$ be a positive constant such that $\|t^\theta T(t)x\|_{D_A(\theta,\infty)}\le M_{\theta,T}\|x\|$ and $\|T(t)y\|_{D_A(\theta,\infty)}\le M_{\theta,T}\|y\|_{D_A(\theta,\infty)}$ for every $x\in X$, $y\in D_A(\theta,\infty)$ and $t\in(0,T]$. Then, for $t\in I_a$ we can estimate

$$\begin{aligned}\|u(t)\|_{D_A(\theta,\infty)}\le& M_\theta\|u(a)\|_{D_A(\theta,\infty)}\\ &+M_\theta\mu(K)\int_a^t(t-s)^{-\theta}(1+cK^{\gamma(1-\alpha/\theta)}\|u(s)\|_{D_A(\theta,\infty)})ds,\end{aligned}\tag{14.17}$$

and Lemma 14.1.2 implies that u is bounded in I_a with values in $D_A(\theta,\infty)$. □

The exponent $\gamma=\alpha^{-1}$ is said to be *critical growth exponent*. If $\gamma=\alpha^{-1}$, then the above method does not work: one should replace $D_A(\alpha\gamma,\infty)$ by $D(A)$, and the integral in (14.17) would equal ∞. We already know that in general we cannot estimate the $D(A)$-norm of the function $v=T(\cdot)*\varphi$ in terms of the sup-norm of $\varphi(t)$.

14.2.2 An example: the Cahn-Hilliard equation

Let us consider a one dimensional Cahn-Hilliard equation

$$\begin{cases}D_tu(t,x)=D_{xx}(-u_{xx}(t,x)+f(u(t,x))), & t>0,\quad x\in[0,1],\\ D_xu(t,0)=D_xu(t,1)=D_{xxx}u(t,0)=D_{xxx}u(t,1)=0, & t>0,\\ u(0,x)=u_0(x), & x\in[0,1],\end{cases}\tag{14.18}$$

assuming that $f \in C^3(\mathbb{R})$ has a nonnegative primitive Φ and $u_0 \in C^2([0,1])$ is such that $u_0'(0) = u_0'(1) = 0$. The smoothness of f and the assumptions on u_0 are sufficient to obtain a local solution. The positivity of a primitive of f will be used to get a priori estimates on the solution that guarantee existence in the large.

Set $X = C([0,1])$ and consider the operators $A : D(A) = \{\varphi \in C^4([0,1]) : \varphi'(0) = \varphi'(1) = \varphi'''(0) = \varphi'''(1) = 0\} \to X$, defined by $A\varphi = -\varphi''''$ for every $\varphi \in D(A)$, and $B : D(B) = \{\varphi \in C^2([0,1]) : \varphi'(0) = \varphi'(1) = 0\} \to X$, defined by $B\varphi = \varphi''$ for every $\varphi \in D(B)$. Operator A has a very special form; specifically $A = -B^2$, where B is sectorial by Exercise 9.4.7(iii) and $\Sigma_{0,\theta} \subset \rho(A)$ for every $\theta \in (\pi/2, \pi)$ (see Definition 9.2.1. Then, A is sectorial in X by Exercise 14.3.11, and $D(B)$ is of class $J_{1/2}$ between X and $D(A)$ by Exercise 10.4.5. So, that we may choose $\alpha = 1/2$, $X_{1/2} = D(B)$. Note that both $D(B)$ and $D(A)$ are dense in X. Since B commutes with $R(\lambda, A)$ on $D(B)$ for each $\lambda \in \rho(A)$, then it commutes with $T(t)$ on $D(B)$, and

$$\|T(t)\varphi\|_{D(B)} = \|T(t)\varphi\|_\infty + \left\|\frac{d^2}{dx^2}T(t)\varphi\right\|_\infty = \|T(t)\varphi\|_\infty + \|T(t)\varphi''\|_\infty \le M_{0,T}\|\varphi\|_{D(B)}.$$

Hence, condition (14.13) is satisfied.

The function $F : X_{1/2} \to X$, defined by $F(\varphi) = \dfrac{d^2}{dx^2}f(\varphi) = f'(\varphi)\varphi'' + f''(\varphi){\varphi'}^2$ is Lipschitz continuous on each bounded subset of $X_{1/2}$, provided f is a C^2 function and f'' is locally Lipschitz continuous.

Theorem 14.2.1 implies that, for each $u_0 \in D(B)$, there exists $\tau(u_0) > 0$ such that problem (14.18) has a unique solution $u : [0, \tau(u_0)) \times [0,1] \to \mathbb{R}$, with u, $D_x u$, $D_{xx}u$ continuous in $[0, \tau(u_0)) \times [0,1]$ and $D_t u$, $D_{xxx}u$, $D_{xxxx}u$ continuous in $(0, \tau(u_0)) \times [0,1]$.

Since we have a fourth-order differential equation, the maximum principles cannot be applied to show that u is bounded. We shall prove that the norm $\|u(t,\cdot)\|_{H^1}$ is bounded in $I(u_0)$; through the Sobolev embedding this will imply that u is bounded in $I(u_0)$.

Since $D_t u = D_{xx}(-D_{xx}u + f(u))$ for each $t > 0$, it follows that

$$\int_0^1 (u(t,x) - u(\varepsilon,x))dx = \int_\varepsilon^t dt \int_0^1 D_t u(s,x)dx = 0, \qquad t \in [\varepsilon, \tau(u_0)) \tag{14.19}$$

for every $\varepsilon \in (0, \tau(u_0))$. Letting ε tend to 0, we get

$$\int_0^1 u(t,x)dx = \int_0^1 u_0(x)dx, \qquad t \in (0, \tau(u_0)),$$

so that the mean value of $u(t,\cdot)$ is a constant, independent[1] of t.

Fix again $\varepsilon \in (0, \tau(u_0))$, multiply both sides of the differential equation by $-D_{xx}u + f(u)$, and integrate over $[\varepsilon, t] \times [0,1]$ for $t \in (\varepsilon, \tau(u_0))$. We get

$$\begin{aligned} &-\int_\varepsilon^t \int_0^1 D_t u D_{xx}u\,dsdx + \int_\varepsilon^t \int_0^1 D_t u f(u)dsdx \\ &= \int_\varepsilon^t \int_0^1 (-D_{xx}u + f(u))D_{xx}(-D_{xx}u + f(u))dsdx. \end{aligned}$$

[1]We take $\varepsilon > 0$ in (14.19) because our solution is just classical and it is not strict in general, so that it is not obvious that $D_t u$ is in $L^1((0,t) \times (0,1))$.

Note that we may integrate by parts with respect to x in the first integral, because the derivative $D_{tx}u$ exists and it is continuous in $[\varepsilon, t] \times [0,1]$ (see Exercise 14.3.10). So, we integrate by parts in the first integral, we rewrite the second integral, recalling that $f = \Phi'$, and integrate by parts in the third integral too. We get

$$\begin{aligned}
&\int_\varepsilon^t \int_0^1 u_x(s,x)u_{tx}(s,x)dsdx + \int_\varepsilon^t \frac{d}{ds}\int_0^1 \Phi(u(s,x))dxds \\
&= -\int_\varepsilon^t \int_0^1 D_x[-u_{xx}(s,x) + \varphi(u(s,x))]^2 dxds
\end{aligned}$$

so that

$$\frac{1}{2}\int_0^1 u_x(t,x)^2dx - \frac{1}{2}\int_0^1 u_x(\varepsilon,x)^2dx + \int_0^1 [\Phi(u(t,x) - \Phi(u(\varepsilon,x))]dx \le 0,$$

and, letting ε tend to 0, we get

$$\|u_x(t,\cdot)\|_{L^2}^2 + 2\int_0^1 \Phi(u(t,x))dx \le \|u_0'\|_{L^2}^2 + 2\int_0^1 \Phi(u_0(x))dx, \qquad t \in (0, \tau(u_0)).$$

Since Φ is nonnegative, then $D_xu(t,\cdot)$ is bounded in $L^2((0,1))$ for $t \in I(u_0)$. The Poincaré inequality and the fact that $u(t,\cdot)$ has constant mean value, yield that $u(t,\cdot)$ is bounded in $H^1(0,1)$ for $t \in I(u_0)$. Since $H^1(0,1)$ is continuously embedded in $C([0,1])$, u is bounded in the sup norm.

Now we may use Proposition 14.2.5, because F satisfies condition (14.16) with $\gamma = 1$. Indeed,

$$\begin{aligned}
\|F(\varphi)\| \le & \Big(\sup_{|\xi|\le\|\varphi\|_\infty} |f'(\xi)|\Big)\|\varphi''\|_\infty + \Big(\sup_{|\xi|\le\|\varphi\|_\infty} |f''(\xi)|\Big)\|\varphi'\|_\infty^2 \\
\le & \sup_{|\xi|\le\|\varphi\|_\infty} |f'(\xi)| \cdot \|\varphi''\|_\infty + \sup_{|\xi|\le\|\varphi\|_\infty} |f''(\xi)| \cdot C\|\varphi\|_\infty\|\varphi''\|_\infty \\
\le & \mu(\|\varphi\|)\|\varphi\|_{D(B)},
\end{aligned}$$

where $\mu(s) = \max\{\sup_{|\xi|\le s}|f'(\xi)|,\ C s \sup_{|\xi|\le s}|f''(\xi)|\}$ and F has subcritical growth (the critical growth exponent is 2). By Proposition 14.2.5, the solution exists in the large.

14.3 Exercises

Exercise 14.3.1. Prove that

(i) if F satisfies (14.3) and $u \in C_b((0,\delta]; X)$ with $0 < \delta \le T$, then the composition $\varphi(t) := F(t,u(t))$ belongs to $C_b((0,\delta]; X)$;

(ii) if F satisfies (14.5) and $u \in C^\theta([a,b]; X)$ with $0 \le a < b \le T$, then the composition $\varphi(t) := F(t,u(t))$ belongs to $C^\theta([a,b]; X)$.

These properties have been used in the proofs of Theorem 14.1.3 and of Proposition 14.1.4.

Exercise 14.3.2. Prove that if u is a mild solution to (14.1) in an interval $[0, t_0]$ and v is a mild solution to

$$\begin{cases} v'(t) = Av(t) + F(t, v(t)), & t \in (t_0, t_1), \\ v(t_0) = u(t_0), \end{cases}$$

then the function $z : [0, t_1] \to X$, defined by $z(t) = u(t)$ for $0 \le t \le t_0$, $z(t) = v(t)$ for $t_0 \le t \le t_1$, is a mild solution to (14.1) in the interval $[0, t_1]$.

Exercise 14.3.3. Under the assumptions of Theorem 14.1.3, for $t_0 \in (0, T)$ let $u(\cdot; t_0, x) : [t_0, \tau(t_0, x)) \to X$ be the maximally defined solution to problem $u' = Au + F(t, u)$, $t > t_0$, $u(t_0) = x$.

(i) Prove that for each $a \in (0, \tau(0, x))$ it holds that $\tau(a, u(a; 0, x)) = \tau(0, x)$ and for $t \in [a, \tau(0; x))$ it holds that $u(t; a, u(a; 0, x)) = u(t; 0, x)$.

(ii) Prove that if F does not depend on t, then $\tau(0, u(a; 0, x)) = \tau(0, x) - a$, and for $t \in [0, \tau(0, x) - a)$ we have $u(t; 0, u(a; 0, x)) = u(a + t; 0, x)$.

and $\|u(t; u_0) - u(t; u_1)\| \le K\|u_0 - u_1\|$ for each $t \in [0, b]$.

Exercise 14.3.4. Prove that if F satisfy (14.3), then for every $u_0 \in X$, the mild solution u of problem (14.1) is bounded with values in $D_A(\beta, \infty)$ in the interval $[\varepsilon, \tau(u_0) - \varepsilon]$, for each $\beta \in (0, 1)$ and $\varepsilon \in (0, \tau(u_0)/2)$.

Exercise 14.3.5. Prove that the closure, with respect to the sup-norm, of the set $\{u \in C^2([0, 1]) : u(0) = u(1) = 0\}$ is the set of all functions $u \in C([0, 1])$ which vanish at 0 and at 1.

Exercise 14.3.6. Let A be the operator in Exercise 14.3.5 and let $\{T(t)\}$ be the associated analytic semigroup. Prove that the function $\|T(\cdot)\|_{\mathcal{L}(C^\alpha([0,1])}$ is unbounded in (0.1].
[**Hint:** Compute $u = R(\lambda, A_\infty)1$ for $\lambda > 0$ and show that $\limsup_{\lambda\to\infty} \lambda^{1-\alpha/2} u(\lambda^{-1/2}) = \infty$, so that $\lambda[R(\lambda, A_\infty)f]_{C^\alpha}$ is unbounded as $\lambda \to \infty$. Taking into account the behaviour of $R(\lambda, A)1$, deduce that the function $\|T(\cdot)\|_{\mathcal{L}(C^\alpha((0,1)))}$ is unbounded in $(0, 1]$].

Exercise 14.3.7. Let u be the solution to

$$\begin{cases} D_t u(t, x) = D_{xx} u(t, x) + (u(t, x))^2, & t \ge 0, \quad x \in [0, 1], \\ u(t, 0) = u(t, 1) = 0, & t \ge 0, \\ u(0, x) = u_0(x), & \qquad\quad x \in [0, 1] \end{cases}$$

with $u_0(0) = u_0(1) = 0$. Prove the following properties.

(i) If $0 \le u_0(x) \le \pi^2 \sin(\pi x)$, then u exists in the large.
[**Hint:** compare with $v(t, x) = \pi^2 \sin(\pi x)$].

(ii) Set $h(t) = \displaystyle\int_0^1 u(t, x) \sin(\pi x) dx$ and prove that $h'(t) \ge (\pi/2)h^2 - \pi^2 h(t)$. Deduce that if $h(0) > 2\pi$ then u blows up in finite time.

Exercise 14.3.8. Show that the function $t \mapsto T(t)u_0$ belongs to $C([0, \delta]; X_\alpha)$ if and only if $u_0 \in \overline{D(A)}^{X_\alpha}$. This fact has been used in Proposition 14.2.1.

Exercise 14.3.9. Prove Propositions 14.2.2 and 14.2.3.
[**Hint:** concerning Proposition 14.2.2 show by induction that the statement holds with θ replaced by $\theta_k = \min\{\theta, (1-\alpha)(1+\alpha+\ldots+\alpha^{k-1})\}$ for every $k \in \mathbb{N}$].

Exercise 14.3.10. Referring to Subsection 14.2.2, prove that the function $D_{tx}u$ exists and it is continuous in $[\varepsilon, \tau-\varepsilon] \times [0,1]$ for each $\varepsilon \in (0, \tau/2)$.
[**Hint:** use Proposition 14.2.2 to deduce that the function $D_t u$ is bounded in $[\varepsilon, \tau-\varepsilon]$ with values in $D_A(\theta, \infty)$ for every $\theta \in (0,1)$ ...].

Exercise 14.3.11. Let $A : D(A) \subset X \to X$ be a sectorial operator in the sector $\Sigma_{0,\theta}$, with $\theta > 3\pi/4$. Show that $-A^2$ is sectorial.

Bibliography

[1] A. Belleni-Morante, Applied Semigroups and Evolution Equations, Oxford University Press, 1979.

[2] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[3] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[4] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[5] K.-J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Graduate Text in Mathematics **194**, Springer-Verlag, 2000.

[6] K.-J. Engel and R. Nagel, Short Course on Operator Semigroups, Universitext, Springer-Verlag, 2006.

[7] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[8] J.A. Goldstein, Semigroups of Operators and Applications, Oxford University Press, 1985.

[9] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[10] D. Henry, Geometric theory of semilinear parabolic equations, Lect. Notes in Math. **840**, Springer-Verlag, New York, 1981.

[11] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[12] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981.

[13] L. Lorenzi and A. Rhandi, Semigroups of Bounded Operators and Second-Order Elliptic and Parabolic Partial Differential Equations, CRC Press, Taylor & Francis Group, 2021.

[14] H.P. Lotz, Uniform convergence of operators on L^∞ and similar spaces, Math. Z. **190** (1985), 207–220.

[15] A. Lunardi, Analytic Semigroups and Optimal Regularity in Parabolic Problems, Birkhäuser Verlag, 1995.

[16] C.-V. Pao, Nonlinear parabolic and elliptic equations, Plenum Press (1992).

[17] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Appl. Math.Sci. **44**, Springer-Verlag,1983.

[18] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[19] F. Rothe, Global Solutions of Reaction-Diffusion Systems, Lect. Notes in Math. **1072**, Springer Verlag, Berlin (1984).

[20] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[21] W. Rudin, Real and Complex Analysis, 3^{rd} ed., McGraw-Hill, 1986.

[22] J. Smoller, Shock Waves and Reaction-Diffusion Equations, Springer Verlag, Berlin (1983).

[23] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[24] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958.

[25] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861.

[26] A. Zygmund, Trigonometric Series, Cambridge Univ. Press., 2nd Edition Reprinted (1968).

25th Internet Seminar on Spectral Theory for Operators and Semigroups

Angela A. Albanese, Luca Lorenzi, Elisabetta M. Mangino, Abdelaziz Rhandi

February 17, 2022

Lecture 15

Behavior near stationary solutions

In this last lecture, we complete our study and combine the spectral analysis on the sectorial operators, with the results in Lecture 14 to investigate on the stability of steady state solutions to nonlinear problems. More precisely, we consider the nonlinear equation

$$u'(t) = Au(t) + F(u(t)), \qquad t > 0, \tag{15.1}$$

where $A : D(A) \subset X \to X$ is a sectorial operator, $F : X \to X$, or $F : X_\alpha \to X$ satisfies the assumptions of the local existence Theorems 14.1.3 or 14.2.1. Then for every $x \in X_\alpha$, the initial value problem

$$\begin{cases} u'(t) = Au(t) + F(u(t)), & t > 0, \\ u(0) = x, \end{cases} \tag{15.2}$$

admits a unique solution, that we denote by $u(\cdot; x)$, defined in a maximal time interval $[0, \tau(x))$.

Assuming that $F(0) = 0$, the equation (15.1) admits the stationary (i.e., constant in time) solution $u \equiv 0$. We are interested in studying the stability of the null solution to equation (15.1). Throughout this lecture, we use the same notation as in Lecture 14, so X_α is a space of class J_α between X and $D(A)$, for some $\alpha \in (0, 1)$, that satisfies the condition

$$\limsup_{t \to 0} \|T(t)\|_{\mathcal{L}(X_\alpha)} < \infty.$$

This condition implies that the function $t \mapsto \|T(t)\|_{\mathcal{L}(X_\alpha)}$ is bounded on every compact interval contained in $[0, \infty)$. Moreover, we assume that the function F satisfies the condition

$$\lim_{\rho \to 0} K(\rho) = 0, \tag{15.3}$$

where

$$K(\rho) = \sup\left\{ \frac{\|F(x) - F(y)\|}{\|x - y\|_{X_\alpha}} : x, y \in B(0, \rho) \subset X_\alpha \right\}.$$

Definition 15.0.1. The null solution of (15.1) is said to be

- stable (in X_α) if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$x \in X_\alpha, \ \ \|x\|_{X_\alpha} \le \delta \Longrightarrow \tau(x) = \infty, \ \ \|u(t; x)\|_{X_\alpha} \le \varepsilon \ \ \forall t \ge 0;$$

- asymptotically stable if it is stable and $\lim_{t\to\infty} \|u(t;x)\|_{X_\alpha} = 0$ uniformly for x in a neighborhood of 0.
- unstable if it is not stable.

Here, $u(\cdot, x) : [0, \tau(x)) \to X_\alpha$ is the maximally defined solution of (15.1) with initial datum x.

From the point of view of the stability, the case when F is defined in X_α does not differ substantially from the case when it is defined in the whole space X, and they will treated together, setting $X_0 := X$ and considering $\alpha \in [0, 1)$.

Remark 15.0.2. We confine our analysis to the case when 0 is a stationary solution to equation (15.1). The study of the stability of other possible stationary solutions $\overline{u} \in D(A)$ can be reduced to the case of the null stationary solution $\overline{u}$, if F is Frechét differentiable at $\overline{u}$ and condition (15.3) is satisfied, with 0 being replaced by $\overline{u}$, and the operator $\widetilde{A} = A + F'(\overline{u})$ is sectorial. It suffices to study the problem $v' = \widetilde{A}v + \widetilde{F}(v)$ solved by the new unknown $v = u - \overline{u}$. Here, $\widetilde{F}(v) = F(v + \overline{u}) - F'(\overline{u})v$.

As in Lecture 13 we consider the Banach space $C_\omega(I; X_\alpha)$ ($\alpha \in [0, 1)$) consisting of all functions $f : I \to X_\alpha$ such that $\|f\|_{C_\omega(I;X_\alpha)} = \sup_{t\in I}(e^{-\omega t}\|f(t)\|_{X_\alpha}) < \infty$. Here, I is either the interval $(0, \infty)$ or the interval $(-\infty, 0)$.

15.1 The principle of linearized stability

The principle of linearized stability states that the null solution to the nonlinear equation (15.1) has the same stability properties of the null solution to the linear problem $u' = Au$. Note that, by assumption (15.3), the linear part of $Ax + F(x)$ near $x = 0$ is Ax, so that the nonlinear part $F(u)$ in problem (15.1) looks like a small perturbation of the linear part $u' = Au$, at least for solutions close to 0. In the next two subsections we make this argument rigorous.

15.1.1 Linearized stability

In this subsection, we assume that the spectral bound $s(A)$ of the operator A is negative. Under this condition, the results in Lecture 13 show that, for every $x \in X$, the solution to the homogeneous linear problem

$$\begin{cases} u'(t) = Au(t), & t > 0, \\ u(0) = x, \end{cases}$$

converges to the null solution (which is a stationary solution to such problem) with exponential rate. As we will see, also problem (15.1) exhibits the same behaviour.

In the proof of the linearized stability theorem we shall use the next lemma, which is a consequence of Proposition 13.1.1.

Lemma 15.1.1. *Fix $\omega \in [0, -s(A))$. If $f \in C_{-\omega}((0, \infty); X)$ and $x \in X_\alpha$ then the function*

$$v(t) = T(t)x + \int_0^t T(t-s)f(s)ds, \qquad t > 0,$$

belongs to $C_{-\omega}((0,\infty);X_\alpha)$*, and there exists a constant* $C = C(\omega)$ *such that*

$$\|v\|_{C_{-\omega}((0,\infty);X_\alpha)} \le C(\|x\|_{X_\alpha} + \|f\|_{C_{-\omega}((0,\infty);X)}).$$

Proof. By Proposition 13.1.1, for each $\omega \in [0, -s(A))$ there exists a positive constant $M(\omega)$ such that $\|T(t)\|_{\mathcal{L}(X)} \le M(\omega)e^{-\omega t}$ for every $t > 0$. Therefore, splitting $T(t) = T(1)T(t-1)$, we can estimate

$$\|T(t)\|_{\mathcal{L}(X,X_\alpha)} \le \|T(1)\|_{\mathcal{L}(X,X_\alpha)}\|T(t-1)\|_{\mathcal{L}(X)} \le Ce^{-\omega t}, \qquad t > 1,$$

with $C = M(\omega)e^{\omega}\|T(1)\|_{\mathcal{L}(X,X_\alpha)}$, while for $0 < t < 1$ we have $\|T(t)\|_{\mathcal{L}(X,X_\alpha)} \le Ct^{-\alpha}$ for some constant $C > 0$, by Proposition 10.2.7. By changing $\omega \in [0, -s(A))$, if necessary, we get that

$$C_\omega := \sup_{t>0}(e^{\omega t}t^\alpha\|T(t)\|_{\mathcal{L}(X,X_\alpha)}) < \infty. \tag{15.4}$$

On the other hand, since X_α is continuously embedded in X, it follows that $\|T(t)\|_{\mathcal{L}(X_\alpha)} \le \widehat{C}e^{-\omega t}$ for $t \ge 1$ and some positive constant $\widehat{C}$. By assumptions, the function $t \mapsto \|T(t)\|_{\mathcal{L}(X_\alpha)}$ is bounded in $(0,1)$. Summing up, we conclude that

$$\widetilde{C}_\omega := \sup_{t>0}(e^{\omega t}\|T(t)\|_{\mathcal{L}(X_\alpha)}) < \infty.$$

Therefore, for every fixed $\omega' \in (\omega, -s(A))$ and $t > 0$, using (15.4), we can estimate

$$\begin{aligned}\left\|e^{\omega t}\int_0^t T(t-s)f(s)ds\right\|_{X_\alpha} =& \left\|e^{\omega t}\int_0^t T(s)f(t-s)ds\right\|_{X_\alpha}\\ \le& \widetilde{C}_{\omega'}e^{\omega t}\int_0^t s^{-\alpha}e^{-\omega' s}\|f(t-s)\|ds \le \frac{\widetilde{C}_{\omega'}\Gamma(1-\alpha)}{(\omega'-\omega)^{1-\alpha}}\|f\|_{C_{-\omega}((0,\infty);X)},\end{aligned}$$

where, here, Γ is the Gamma function. and the statement follows. □

Theorem 15.1.2. *For every* $\omega \in [0, -s(A))$ *there exist two positive constants* $M = M(\omega)$ *and* $r = r(\omega)$ *such that, if* $x \in \overline{B(0,r)} \subset X_\alpha$*, then* $\tau(x) = \infty$ *and*

$$\|u(t;x)\|_{X_\alpha} \le Me^{-\omega t}\|x\|_{X_\alpha}, \qquad t \ge 0. \tag{15.5}$$

As a byproduct, the null solution to equation (15.1) *is asymptotically stable.*

Proof. Let Y be the closed ball centered at 0 in the space $C_{-\omega}((0,\infty); X_\alpha)$, with radius ρ to be fixed later on. We look for the solution to problem (15.2) as a fixed point of the operator Γ defined on Y by

$$(\Gamma(u))(t) = T(t)x + \int_0^t T(t-s)F(u(s))ds, \qquad t \ge 0. \tag{15.6}$$

If $u \in Y$, then

$$\|F(u(t))\| = \|F(u(t)) - F(0)\| \le K(\rho)\|u(t)\|_{X_\alpha} \le K(\rho)\rho e^{-\omega t}, \qquad t \ge 0, \tag{15.7}$$

so that $F(u(\cdot)) \in C_{-\omega}((0,\infty); X)$. Using Lemma 15.1.1 we get

$$\|\Gamma(u)\|_{C_{-\omega}((0,\infty);X_\alpha)} \le C\left(\|x\|_{X_\alpha} + \|F \circ u\|_{C_{-\omega}((0,\infty);X)}\right) \le C\left(\|x\|_{X_\alpha} + \rho K(\rho)\right). \tag{15.8}$$

If ρ is sufficiently small so that $K(\rho) \le (2C)^{-1}$ and $\|x\|_{X_\alpha} \le r := (2C)^{-1}\rho$, then $\Gamma(u)$ belongs to Y. Moreover, again by Lemma 15.1.1 it follows that

$$\|\Gamma(u_1) - \Gamma(u_2)\|_{C_{-\omega}((0,\infty);X_\alpha)} \le C\|F \circ u_1 - F \circ u_2\|_{C_{-\omega}((0,\infty);X)}$$

for every $u_1, u_2 \in Y$, where

$$\|F(u_1(t)) - F(u_2(t))\| \le K(\rho)\|u_1(t) - u_2(t)\|_{X_\alpha}, \qquad t > 0.$$

Taking ρ sufficiently small, it follows that

$$\|\Gamma(u_1) - \Gamma(u_2)\|_{C_{-\omega}((0,\infty);X_\alpha)} \le \frac{1}{2}\|u_1 - u_2\|_{C_{-\omega}((0,\infty);X_\alpha)},$$

so that Γ is a 1/2-contraction. Consequently, there exists a unique fixed point of Γ in Y, which is the solution of the equation (15.1) and satisfies the condition $u(0) = x$. Moreover, from (15.7) and (15.8) we get

$$\begin{aligned}\|u\|_{C_{-\omega}((0,\infty);X_\alpha)} =&\|\Gamma(u)\|_{C_{-\omega}((0,\infty);X_\alpha)} \le C(\|u_0\|_{X_\alpha} + K(\rho)\|u\|_{C_{-\omega}((0,\infty);X_\alpha)})\\ \le& C\|x\|_{X_\alpha} + \frac{1}{2}\|u\|_{C_{-\omega}((0,\infty);X_\alpha)}\end{aligned}$$

which yields estimate (15.5) with $M = 2C$. □

Remark 15.1.3. Note that the mild solution to problem (15.1) is smooth for $t > 0$. Indeed, since u satisfies (15.6) and it is continuous with values in X_α, by Theorem 9.2.2 and Proposition 10.2.12, it follows that $u \in C^\theta([a,b];X)$ for every $\theta \in (0,1)$ and $0 < a < b < \infty$. For $t > a$, we can write

$$u(t) = T(t-a)u(a) + \int_0^{t-a} T(t-a-s)F(u(s+a))ds, \qquad t > a.$$

Since F is Lipschitz continuous, then the function $s \mapsto F(u(s+a))$ is θ-Hölder continuous in $[0,b]$ for any $b > 0$. Theorem 10.2.14 implies that $u \in C^{1+\theta}([a+\varepsilon,b];X) \cap C^\theta([a+\varepsilon,b];D(A))$, for every ε, b such that $a + \varepsilon < b$.

15.1.2 Linearized instability

Throughout this subsection, we assume that the following conditions hold:

$$\text{(15.9)} \qquad \begin{cases} \sigma_+(A) = \sigma(A) \cap \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 0\} \ne \varnothing, \\ \inf\{\operatorname{Re}\lambda : \lambda \in \sigma_+(A)\} = \omega_+ > 0. \end{cases}$$

Then it is possible to prove an instability result for the null solution. We shall use the projection P defined by (13.7), i.e.

$$P = \frac{1}{2\pi i}\int_{\gamma_+} R(\lambda, A)d\lambda,$$

where γ_+ is any closed regular curve with range in $\{\operatorname{Re}\lambda > 0\}$, oriented counterclockwise, which "covers" its support only once.

For the proof of the instability theorem, we need the next lemma, which is a corollary of Theorem 13.4.1(ii). It is a counterpart of Lemma 15.1.1 for the unstable case.

Lemma 15.1.4. *Fix $\omega \in [0,\omega_+)$. If $g \in C_\omega((-\infty,0];X)$ and $x \in P(X)$, then the function $v:(-\infty,0]\to X$, defined by*

$$v(t) = T(t)x - \int_t^0 T(t-s)Pg(s)ds + \int_{-\infty}^t T(t-s)(I-P)g(s)ds, \qquad t \le 0, \tag{15.10}$$

is a mild solution to the equation $v' = Av+g$ in $(-\infty,0]$. Moreover, it belongs to $C_\omega((-\infty,0];X_\alpha)$ and there exists a positive constant $C = C(\omega)$ such that

$$\|v\|_{C_\omega((-\infty,0];X_\alpha)} \le C(\|x\| + \|g\|_{C_\omega((-\infty,0];X)}). \tag{15.11}$$

Finally, the function $t \mapsto e^{-\omega t}u(t)$ is bounded with values in $D_A(\beta,\infty)$, for each $\beta \in (0,1)$ and

$$\sup_{t\le 0} e^{-\omega t}\|v(t)\|_{D_A(\beta,\infty)} \le C(\|x\| + \|g\|_{C_\omega((-\infty,0];X)}). \tag{15.12}$$

Conversely, if v is a mild solution to the equation $v' = Av+g$ belonging to $C_\omega((-\infty,0];X_\alpha)$ then there exists $x \in P(X)$ such that v is given by (15.10).

Proof. Set $y_0 = x + \int_{-\infty}^0 T(-s)(I-P)g(s)ds$ and observe that condition (13.24) is satisfied since $x \in P(X)$ and

$$(I-P)y_0 = \int_{-\infty}^0 T(-s)(I-P)g(s)ds = \int_0^\infty T(s)(I-P)g(-s)ds.$$

Hence, Theorem 13.4.1(ii) implies that v is a mild solution to the equation $v' = Av+g$, which satisfies the initial condition $v(0) = y_0$, since the vertical line $\mathrm{Re}\,\lambda = \omega$ does not intersect the spectrum of A.

Conversely, if v is a mild solution to the equation $v' = Av+g$ in $C_\omega((-\infty,0];X_\alpha)$, then it belongs in $C_\omega((-\infty,0];X)$ and Theorem 13.4.1(ii) shows that it is given by formula (15.10).

Let us prove estimate (15.11). For this purpose, we set $w(t) = e^{-\omega t}v(t)$ for every $t \le 0$. Then,

$$\begin{aligned} w(t) =& e^{-\omega t}T(t)x - \int_t^0 e^{-\omega(t-s)}T(t-s)Pg(s)e^{-\omega s}ds \\ &+ \int_{-\infty}^t e^{-\omega(t-s)}T(t-s)(I-P)g(s)e^{-\omega s}ds \end{aligned}$$

for every $t \le 0$, and the operator $A-\omega$, which is associated to the semigroup $\{e^{-\omega t}T(t)\}$, is hyperbolic with $\sigma_+(A-\omega) = \sigma_+(A)-\omega$ and $\sigma_-(A-\omega) = (\sigma(A)\setminus\sigma_+(A))-\omega$. By Proposition 13.2.2, we can fix $\sigma > 0$ small enough such that

$$\|e^{-\omega t}T(t)(I-P)\|_{\mathcal{L}(X)} \le C_1 e^{-\sigma t}, \quad t \ge 0, \qquad \|e^{-\omega t}T(t)P\|_{\mathcal{L}(X)} \le C_1 e^{\sigma t}, \qquad t \le 0,$$

for some positive constant C_1 that, as the other constants C_j appearing in the proof, is independent of t. Since A is bounded in $P(X)$, it follows that

$$\|e^{-\omega t}T(t)P\|_{\mathcal{L}(X,D(A))} \le C_2 e^{\sigma t}, \qquad t \le 0.$$

Hence,

$$\|e^{-\omega t}T(t)P\|_{\mathcal{L}(X,X_\alpha)} \le C_3 e^{\sigma t}, \qquad t \le 0.$$

Moreover, if $t \ge 1$, we can estimate

(15.13)
$$\|e^{-\omega t}T(t)(I-P)\|_{\mathcal{L}(X,X_\alpha)} \le \|e^{-\omega}T(1)\|_{\mathcal{L}(X,X_\alpha)}\|e^{-\omega(t-1)}T(t-1)(I-P)\|_{\mathcal{L}(X)} \le C_4 e^{-\sigma t}$$

whereas, for $t \in (0,1]$ we can estimate

$$\|e^{-\omega t}T(t)(I-P)\|_{\mathcal{L}(X,X_\alpha)} \le C_5 t^{-\alpha}. \tag{15.14}$$

From (15.13) and (15.14) it follows that

$$\|e^{-\omega t}T(t)(I-P)\|_{\mathcal{L}(X,X_\alpha)} \le C_6 t^{-\alpha}e^{-\sigma t}, \qquad t \ge 0.$$

Therefore, for $t \le 0$ we get

$$\begin{aligned}\|w(t)\|_{X_\alpha} \le & C_1 e^{\sigma t}\|x\| + C_1\|P\|\|g\|_{C_\omega((-\infty,0];X)}\int_t^0 e^{\sigma s}ds \\ & + C_6\|I-P\|\sup_{t\le 0}\int_{-\infty}^t e^{-\sigma(t-s)}(t-s)^{-\alpha}ds\end{aligned}$$

and estimate (15.11) follows easily.

Replacing X_α with $D_A(\beta,\infty)$ ($\beta \in (0,1)$) and repeating the same arguments as above, one can see that u is bounded in $[0,\infty)$ with values in $D_A(\beta,\infty)$ and it satisfies (15.12). □

Theorem 15.1.5. *There exists $r_+ > 0$ such that, for every $x \in P(X)$ with $\|x\| \le r_+$, the Cauchy problem*

$$\begin{cases} v'(t) = Av(t) + F(v(t)), & t \le 0, \\ Pv(0) = x, \end{cases} \tag{15.15}$$

admits a backward solution v such that $\lim_{t\to-\infty} v(t) = 0$.

Proof. Let Y_+ be the closed ball in $C_\omega((-\infty,0];X_\alpha)$ centered at 0 with radius ρ_+ to be fixed later on. In view of Lemma 15.1.4, we look for a solution to (15.15) as a fixed point of the operator Γ_+ defined on Y_+ by

$$(\Gamma_+(v))(t) = T(t)x - \int_t^0 T(t-s)PF(v(s))ds + \int_{-\infty}^t T(t-s)(I-P)F(v(s))ds, \qquad t \le 0.$$

If $v \in Y_+$, then the function $F \circ v$ belongs to $C_\omega((-\infty,0];X)$, see (15.7). Lemma 15.1.4 implies $\Gamma_+(v) \in C_\omega((-\infty,0];X_\alpha)$ and

$$\|\Gamma_+(v)\|_{C_\omega((-\infty,0];X_\alpha)} \le C\left(\|x\| + \|F\circ v\|_{C_\omega((-\infty,0];X)}\right).$$

The rest of the proof is quite similar to that of Theorem 15.1.2 and it is left as an exercise (see Exercise 15.4.1). □

Remark 15.1.6. The existence of a backward mild solution u to problem (15.15) implies that the null solution to equation (15.1) is unstable. Indeed, for every $n \in \mathbb{N}$ set $x_n = v(-n)$. Of course x_n converges to 0 as n tends to ∞. Consider, for any $n \in \mathbb{N}$ the Cauchy problem $u(0) = x_n$ for the equation (15.1), and denote by $v(\cdot; x_n)$ its mild solution. If there exists a subsequence $\{x_{n_k}\}$ such that $v(\cdot; x_{n_k})$ does not exist in $(-\infty, 0]$, of course, the null solution to equation (15.1) is unstable. On the other hand, if $v(\cdot; x_n)$ exists in $(-\infty, 0]$ for every $n \in \mathbb{N}$, then, $u(t; x_n) = v(t - n, x_n)$ for every $t \in [0, n]$. Hence

$$\sup_{t \geq 0} \|u(t; x_n)\|_{X_\alpha} \geq \sup_{0 \leq t \leq n} \|u(t; x_n)\|_{X_\alpha} = \sup_{-n \leq t \leq 0} \|v(t)\|_{X_\alpha} \geq \|v(0)\|_{X_\alpha},$$

which implies that the null solution is unstable since $\sup_{t \geq 0} \|u(t; x_n)\|_{X_\alpha}$ does not converge to 0 as n tends to ∞.

15.2 The saddle point property

If A is hyperbolic, i.e., $\sigma(A) \cap i\mathbb{R} = \varnothing$, we may prove a saddle point property, constructing the so called stable and unstable manifolds. As in Lecture 13, we set

$$-\omega_- := \sup\{\mathrm{Re}\,\lambda : \lambda \in \sigma_-(A)\} < 0, \qquad \omega_+ := \inf\{\mathrm{Re}\,\lambda : \lambda \in \sigma_+(A)\} > 0. \tag{15.16}$$

We shall consider the forward equation (15.1) and the backward problem

$$\begin{cases} v'(t) = Av(t) + F(v(t)), & t \leq 0, \\ v(0) = v_0. \end{cases} \tag{15.17}$$

The stable and the unstable manifolds are the nonlinear analogous of $(I - P)(X)$ and $P(X)$. For the linear problem

$$\begin{cases} u'(t) = Au(t), & t > 0, \\ u(0) = x. \end{cases}$$

Corollary 13.2.3 implies that the solution $u = T(\cdot)x$ is bounded (in fact, it decays exponentially to 0 as $t \to \infty$) if and only if $x \in (I - P)(X)$, and that there exists a backward bounded solution (which in addition decays exponentially to 0 as t tends to $-\infty$) if and only if $x \in P(X)$. We are going to prove that a similar behavior occurs in the nonlinear problem

$$\begin{cases} u'(t) = Au(t) + F(u(t)), & t > 0, \\ u(0) = x, \end{cases}$$

at least for small initial data x. In the case when F is defined in X, the role of subspaces $(I - P)(X)$ and $P(X)$ are played by the stable and the unstable manifold, that are graphs of regular functions from a neighborhood of the origin in $(I - P)(X)$ to $P(X)$ and from a neighborhood of the origin in $P(X)$ to $(I-P)(X)$, respectively. In the case where F is defined in X_α with $\alpha \in (0, 1)$ the behavior is still similar, with $(I - P)(X)$ replaced by $(I - P)(X_\alpha)$ (note that $P(X)$ coincides with $P(X_\alpha)$).

For the construction of the stable manifold we need the next lemma, which is a corollary of Theorem 13.4.1(i).

Lemma 15.2.1. *Let A be a hyperbolic operator and fix $\omega \in [0, \omega_-)$. If $f \in C_{-\omega}((0,\infty); X)$ and $x \in (I-P)(X_\alpha)$ then the function $u : [0,\infty) \to X$, defined by*

$$u(t) = T(t)x + \int_0^t T(t-s)(I-P)f(s)ds - \int_t^\infty T(t-s)Pf(s)ds, \qquad t \geq 0, \tag{15.18}$$

is a mild solution to the equation $u' = Au + f$ in $[0,\infty)$, it belongs to $C_{-\omega}((0,\infty); X_\alpha)$, and there exists a positive constant $C = C(\omega)$ such that

$$\|u\|_{C_{-\omega}(0,\infty);X_\alpha)} \leq C(\|x\|_{X_\alpha} + \|f\|_{C_{-\omega}(0,\infty);X)}). \tag{15.19}$$

Moreover, the function $t \mapsto e^{\omega t}u(t)$ is bounded in $[\varepsilon, \infty)$ with values in $D_A(\beta,\infty)$ for every $\varepsilon > 0$ and $\beta \in (0,1)$. If, in addition $x \in D_A(\beta, \infty)$ for some $\beta \in (0,1)$, then

$$\sup_{t\geq 0}(e^{\omega t}\|u(t)\|_{D_A(\beta,\infty)}) \leq C(\|x\|_{D_A(\beta,\infty)} + \|f\|_{C_{-\omega}((0,\infty);X)}).$$

Conversely, if $u \in C_{-\omega}((0,\infty); X_\alpha)$ is a mild solution to the equation $u' = Au + F(u)$, then u is given by (15.18) *for some $x \in (I-P)(X_\alpha)$.*

Proof. Since the vertical line $\operatorname{Re}\lambda = -\omega$ does not intersect the spectrum of A we may apply Theorem 13.4.1(i), which implies that u is a mild solution. The estimates on $\|u(t)\|_{X_\alpha}$ and $\|u(t)\|_{D_A(\beta,\infty)}$ may be shown as in Lemmas 15.1.1 and 15.1.4, and they are left as an exercise (see Exercise 15.4.2). If u is a mild solution in $C_{-\omega}((0,\infty); X_\alpha)$, then it is in $C_{-\omega}((0,\infty); X)$, and by Theorem 13.4.1(i) it is given by (15.18). □

Theorem 15.2.2. *Assume that A is a hyperbolic operator and fix $\omega \in [0, \min\{\omega_+, \omega_-\})$, where ω_-, ω_+ are defined in* (15.16). *Then, there exist two positive constants r, ρ and two continuous functions*

$$k : \{x_- \in (I-P)(X_\alpha) : \|x_-\| \leq r\} \to P(X),$$

$$h : \{x_+ \in P(X) : \|x_+\| \leq r\} \to (I-P)(X_\alpha),$$

such that, setting

$$\mathcal{M}_S = \mathcal{M}_S(\omega) = \{(x_-, k(x_-)) : x_- \in (I-P)(X_\alpha),\ \|x_-\| \leq r\},$$
$$\mathcal{M}_U = \mathcal{M}_U(\omega) = \{(x_+, h(x_+)) : x_+ \in P(X),\ \|x_+\| \leq r\},$$

the following statements hold true.

- (i) *For every $u_0 \in \mathcal{M}_S$ the solution u of* (15.1) *exists in the large, it belongs to the space $C_{-\omega}((0,\infty); X_\alpha)$, and $\|u\|_{C_{-\omega}(0,\infty);X_\alpha)} \leq \rho$. Conversely, if $u_0 \in X_\alpha$ is such that $\|(I-P)u_0\|_{X_\alpha} \leq r$ and the solution of* (15.1) *belongs to $C_{-\omega}((0,\infty), X_\alpha)$ with norm not exceeding ρ, then $u_0 \in \mathcal{M}_S$.*
- (ii) *For every $v_0 \in \mathcal{M}_U$, problem* (15.17) *has a solution $v \in C_\omega((-\infty,0]; X_\alpha)$, such that $\|v\|_{C_\omega((-\infty,0);X_\alpha)} \leq \rho$. Conversely, if v_0 is such that $\|Pv_0\| \leq r$ and problem* (15.17) *admits a solution belonging to $C_\omega((-\infty,0]; X_\alpha)$, with norm not exceeding ρ, then $v_0 \in \mathcal{M}_U$.*

Proof. (i) Fix $x_- \in (I-P)(X_\alpha)$ and set $Y_- = B(0,\rho_-) \subset C_{-\omega}((0,\infty);X_\alpha)$, with small $\rho_- > 0$ to be chosen later. Note that for each $u \in Y_-$, the function $F \circ u$ belongs to $C_{-\omega}((0,\infty);X)$ if ρ_- is small enough. Therefore, taking Lemma 15.2.1 into account, we look for an exponentially decaying forward solution to $u' = Au + F(u)$ as a fixed point of the operator Γ_- defined on Y_- by

$$(\Gamma_-(u))(t) = T(t)x_- + \int_0^t T(t-s)(I-P)F(u(s))ds - \int_t^\infty T(t-s)PF(u(s))ds$$

for $t \geq 0$. Arguing as in the proof of Theorem 15.1.2 and using Lemma 15.2.1 we see that Γ_- is a 1/2-contraction if ρ_- is small enough and that, if $\|x\|_{X_\alpha} \leq r_- := (2C(\omega))^{-1}\rho$, where $C(\omega)$ is the constant in (15.19), then Γ_- maps Y into itself, so that it admits a unique fixed point $u_- \in Y$. The fixed point satisfies the estimate

$$\|u_-\|_{C_{-\omega}((0,\infty);X_\alpha)} \leq 2C(-\omega)\|x_-\|_{X_\alpha}. \tag{15.20}$$

Moreover, the function $K : \{x \in (I-P)(X_\alpha) : \|x\|_{X_\alpha} \leq r_-\} \times Y_- \to C_{-\omega}((0,\infty);X_\alpha)$, defined by $K(x_-,u) = \Gamma_- u$ for every $(x_-,u) \in \{x \in (I-P)(X_\alpha) : \|x\|_{X_\alpha} \leq r_-\} \times Y_- \to C_{-\omega}((0,\infty);X_\alpha)$, is continuous, so that the fixed point of Γ depends continuously on x_-, thanks to the contraction theorem depending on a parameter. It follows that the function

$$\begin{cases} k : \{x \in (I-P)(X_\alpha) : \|x\|_{X_\alpha} \leq r_-\} \to P(X), \\ k(x) = Pu_-(0), \end{cases}$$

is continuous. The solution of (14.1) with initial datum $u_0 = u_-(0)$ coincides with u_-, so that it belongs to $C_{-\omega}((0,\infty);X_\alpha)$ and its norm is not greater than ρ_-.

Conversely, fix $u_0 \in X_\alpha$ such that $\|(I-P)u_0\|_{X_\alpha} \leq r_-$, and such that the solution of the Cauchy problem (15.2) belongs to $C_{-\omega}((0,\infty);X_\alpha)$ with norm that does not exceed ρ_-. Then, since $F \circ u$ belongs to $C_{-\omega}((0,\infty);X_\alpha)$, by Lemma 15.2.1 it follows that

$$u(t) = T(t)(I-P)u_0 + \int_0^t T(t-s)(I-P)F(u(s))ds - \int_t^\infty T(t-s)F(u(s))ds$$

for $t \geq 0$. Hence, u is a fixed point of the operator Γ_- if we choose $x_- = (I-P)u_0$. Since there exists a unique fixed point of Γ_- with norm less or equal to ρ_-, then $Pu_0 = k((I-P)u_0)$, namely $u_0 \in \mathcal{M}_S$. Statement (i) is proved.

(ii) The proof is quite similar to that of property (i): arguing as in the proof of Theorem 15.1.5 one sets

$$\begin{cases} h : P(X) \cap B(0,r_+) \to (I-P)(X_\alpha), \\ h(x) = (I-P)v(0), \end{cases}$$

where v is the fixed point of the operator Γ_+ in Y_+ given by Theorem 15.1.5, and r_+ is given by Theorem 15.1.5 too.

We take eventually $r = \min\{r_-, r_+\}$, $\rho = \min\{\rho_-, \rho_+\}$, and the statement follows. □

Remark 15.2.3. The stable manifold $\mathcal{M}_S$ (respectively, the unstable manifold $\mathcal{M}_U$) is tangent at the origin to $(I-P)(X_\alpha)$ (resp., to $P(X)$), in the sense that k (resp., h) is

Fréchet differentiable at 0 with derivative $k'(0) = 0$ (respectively, $h'(0) = 0$). Indeed, since $\|u_-\|_{C_{-\omega}((0,\infty);X_\alpha)} \le 2C\|x_-\|$ by (15.20), then (see estimate (15.7))

$$\|F \circ u_-\|_{C_{-\omega}((0,\infty);X)} \le K(\rho_-)\|u_-\|_{C_{-\omega}((0,\infty);X_\alpha)} \le 2CK(\rho_-)\|x_-\|.$$

Consequently,

$$\begin{aligned}\|k(x_-)\| =&\|Pu_-(0)\| = \left\| \int_0^\infty T(-s)PF(u_-(s))ds \right\|_{X_\alpha} \\ \le& c\|F \circ u_-\|_{C_{-\omega}((0,\infty);X_\alpha)} \le C'K(\rho_-)\|x_-\|.\end{aligned}$$

Fix $\varepsilon > 0$ and let $\rho_1 > 0$ be such that $C'K(\rho_1) < \varepsilon$. Then, for every $x_- \in (I-P)(X_\alpha)$, with $\|x_-\|_{X_\alpha}$ small enough, it follows that

$$\frac{\|k(x_-) - k(0)\|}{\|x_-\|_{X_\alpha}} = \frac{\|k(x_-)\|}{\|x_-\|_{X_\alpha}} \le \varepsilon,$$

so that k is differentiable at 0 with null derivative. The proof of the statement concerning the function h is similar.

Remark 15.2.4. The proof of Theorem 15.2.2 works also for $\omega = 0$ and this implies that, if $u : [0,\infty) \to X_\alpha$ is a solution to (15.1) with $\sup_{t\ge 0} \|u(t)\|_{X_\alpha}$ sufficiently small, then, in fact, u decays exponentially to 0, and $u_0 \in \mathcal{M}_S(\omega)$ with $\omega > 0$. Indeed, if $\sup_{t\ge 0} \|u(t)\|_{X_\alpha}$ is small, then also $\|(I-P)u_0\|_{X_\alpha}$ is small, and, hence, u is the fixed point of the operator Γ, with $\omega = 0$ and $x_- = (I-P)u_0$. On the other hand, for the same choice of x_-, Γ admits also a fixed point in $C_{-\omega}((0,\infty);X_\alpha)$, and the two fixed points do coincide.

Similarly, if $v : (-\infty, 0] \to X_\alpha$ is a backward mild solution of (14.1) and $\sup_{t\le 0} \|v(t)\|$ is sufficiently small, then v decays exponentially to 0, as t tends to $-\infty$, and $v(0) \in \mathcal{M}_U(\omega)$.

15.3 A Cauchy-Dirichlet problem

In order to give some examples of PDE's to which the results of this lecture can be applied, we need some comments on the spectrum of the Laplacian with Dirichlet boundary conditions.

Let Ω be a bounded open set in $\mathbb{R}^N$ with C^2 boundary $\partial\Omega$. We choose $X = C(\overline{\Omega})$ and define

$$D(A) = \Big\{\varphi \in \bigcap_{1\le p<\infty} W^{2,p}(\Omega) : \Delta\varphi \in C(\overline{\Omega}), \varphi_{|\partial\Omega} = 0\Big\}$$

and $A\varphi = \Delta\varphi$ for $\varphi \in D(A)$. This operator is sectorial, its spectrum consists of isolated eigenvalues and $s(A)$ is negative (this is a classical result, see e.g., [13, Section 14.3]). In order to give an explicit estimate of $s(A)$ we recall the so called Poincaré inequality

$$\int_\Omega |\varphi|^2 dx \le C_\Omega \int_\Omega |\nabla\varphi|^2 dx \tag{15.21}$$

which holds true for every $\varphi \in W_0^{1,2}(\Omega)$. A proof of (15.21) as well as the inequality $C_\Omega \le 4d^2$, where d is the diameter of Ω, is outlined in the Exercise **??**.

If $\varphi \in D(A)$ and $-\lambda\varphi - \Delta\varphi = 0$, then $\varphi \in W_0^{1,2}(\Omega)$. Multiplying the equation by $\overline{\varphi}$ and integrating over Ω we find

$$\int_\Omega |\nabla\varphi|^2 dx = \lambda \int_\Omega |\varphi|^2 dx,$$

and, therefore, $\lambda \geq C_\Omega^{-1}$, that is $s(A) \leq -C_\Omega^{-1}$.

It can be proved that the spectrum of A consists of a sequence of negative eigenvalues $(-\lambda_n)$ tending to $-\infty$ (we assume $\lambda_n \leq \lambda_{n+1}$) so that $s(A) = -\lambda_1$. Observe that the above argument shows that the eigenvalues of A are negative but does not prove their existence, hence $s(A)$ could be $-\infty$. To see that this is not the case one can work in $L^2(\Omega)$ instead of $C(\overline{\Omega})$ and show that the Laplacian with domain $H^2(\Omega) \cap H_0^1(\Omega)$ is a negative self-adjoint operator with compact resolvent. The general theory then shows that the L^2-spectrum of the Laplacian consists of a sequence $\{-\lambda_n\}_{n\in\mathbb{N}}$ of negative eigenvalues diverging to $-\infty$ as n tends to ∞. Using the elliptic regularity theory it follows that the L^2-eigenfunctions belong to the domain in $C(\overline{\Omega})$ (the converse is obvious), so that the spectrum in $C(\overline{\Omega})$ coincides with the spectrum in $L^2(\Omega)$.

We now study the stability of the null solution of

$$\begin{cases} D_t u(t,x) = \Delta u(t,x) + f(u(t,x), Du(t,x)), & t > 0, \quad x \in \overline{\Omega}, \\ u(t,x) = 0, & t > 0, \quad x \in \partial\Omega, \end{cases} \tag{15.22}$$

where $f = f(x,p) : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ is continuously differentiable and $f(0,0) = 0$. The local existence and uniqueness Theorem 14.2.1 may be applied to the initial value problem for equation (15.22),

$$u(0,x) = u_0(x), \quad x \in \overline{\Omega}, \tag{15.23}$$

choosing $X = C(\overline{\Omega})$, $X_\alpha = C_0^{2\alpha}(\overline{\Omega})$ with $1/2 < \alpha < 1$. The function $F : X_\alpha \to X$, defined by $(F(\varphi))(x) = f(\varphi(x), \nabla\varphi(x))$ for every $x \in \Omega$, is continuously differentiable, vanishes at 0 and $(F'(0)\varphi)(x) = a\varphi(x) + b \cdot \nabla\varphi(x)$ for every $x \in \Omega$ and $\varphi \in X_\alpha$. Here, $a = D_x f(0,0)$, $b = D_p f(0,0)$. Then, set $D(B) = D(A)$ and $B\varphi = \Delta\varphi + b \cdot \nabla\varphi + a\varphi$. The operator B is sectorial (see e.g., [13, Section 14.3]) and

$$s(B) \leq -C_\Omega^{-1} + a. \tag{15.24}$$

Indeed, we observe that the resolvent of B is compact and therefore its spectrum consists of isolated eigenvalues. Moreover, if $\lambda \in \sigma(B)$ and $\varphi \in D(B)$ is such that $\lambda\varphi - \Delta\varphi - b\cdot\nabla\varphi - a\varphi = 0$, then, multiplying by $\overline{\varphi}$ and integrating over Ω, we get

$$\int_\Omega \left((\lambda - a)|\varphi|^2 + |\nabla\varphi|^2 - \frac{1}{2} b \cdot \nabla\varphi\, \overline{\varphi} \right) dx = 0.$$

Taking the real part

$$\int_\Omega \left((\mathrm{Re}\,\lambda - a)|\varphi|^2 + |\nabla\varphi|^2 - \frac{1}{2} b \cdot D|\varphi|^2 \right) dx = \int_\Omega \left[(\mathrm{Re}\,\lambda - a)|\varphi|^2 + |\nabla\varphi|^2 \right] dx = 0$$

and hence $\mathrm{Re}\,\lambda - a \leq -C_\Omega^{-1}$. Therefore, (15.24) holds.

Since $u_0 \in X_\alpha \subset \overline{D(A)}$, Theorem 14.2.1 guarantees the existence of a unique classical solution $u : [0, \tau(u_0)) \to X_\alpha$ of the abstract problem (14.1) having the regularity properties of Proposition 14.2.2. Setting as usual $u(t,x) := u(t)(x)$ for every $t \in [0, \tau(u_0))$ and $x \in \overline{\Omega}$, the function u is continuous in $[0, \tau(u_0)) \times \overline{\Omega}$, continuously differentiable with respect to time for $t > 0$, and it satisfies (15.22), (15.23).

Concerning the stability of the null solution, Theorem 15.1.2 implies that if $s(B) < 0$, in particular if $a < C_\Omega^{-1}$, then the null solution of (15.22) is exponentially stable, i.e., for every $\omega \in (0, -s(B))$ there exist two positive constant r, C such that if $\|u_0\|_{X_\alpha} \leq r$, then $\tau(u_0) = \infty$ and $\|u(t)\|_{X_\alpha} \leq Ce^{-\omega t}\|u_0\|_{X_\alpha}$ for every $t > 0$.

On the contrary, if $s(B) \geq 0$, then there exist elements in the spectrum of A with positive real part. Since they are isolated they satisfy condition (15.9). Theorem 15.1.5 implies that the null solution of (15.22) is unstable: there exist $\varepsilon > 0$ and initial data u_0 with $\|u_0\|_{X_\alpha}$ arbitrarily small, but $\sup_{t\geq 0} \|u(t)\|_{X_\alpha} \geq \varepsilon$.

Finally, we remark that, if f is independent of p, i.e. the nonlinearity in (15.22) does not depend on ∇u, then we can work directly in X.

15.4 Exercises

Exercise 15.4.1. Complete the proof of Theorem 15.1.5.

Exercise 15.4.2. Complete the proof of Lemma 15.2.1.

Exercise 15.4.3. Prove that the stationary solution $(u \equiv 1, v \equiv 0)$ to system (**??**) is asymptotically stable in $C(\overline{\Omega}) \times C(\overline{\Omega})$.

Exercise 15.4.4. Let Ω be a bounded set in $\mathbb{R}^N$ and let d be the diameter of Ω. Prove Poincaré inequality with $C_\Omega \leq 4d^2$.
[Hint: assume that $\Omega \subset B(0,d)$ and for $\varphi \in C_c^\infty(\Omega)$ write

$$\varphi(x_1, \dots, x_N) = \int_{-d}^{x_1} \frac{\partial \varphi}{\partial x_1}(s, x_2, \dots, x_N)ds\Big].$$

Exercise 15.4.5. Let X be a Banach space and Ω be an open set in $\mathbb{R}$ (or in $\mathbb{C}$). Moreover, let $\Gamma : X \times \Omega \to X$ be such that

$$\|\Gamma(y, \lambda) - \Gamma(x, \lambda)\| \leq C(\lambda)\|y - x\|$$

for any $\lambda \in \Omega$, any $x, y \in X$ and some continuous function $C : \Omega \to [0, 1)$. Further, suppose that the function $\lambda \mapsto \Gamma(\lambda, x)$ is continuous in Ω for any $x \in X$. Prove that for any $\lambda \in \Omega$ the equation $x = \Gamma(x, \lambda)$ admits a unique solution $x = x(\lambda)$ and that the function $\lambda \mapsto x(\lambda)$ is continuous in Ω.

Exercise 15.4.6. Let u be the solution to

$$\begin{cases} D_t u = D_{xx}u + u^2, & t \geq 0, \quad x \in [0,1], \\ u(t,0) = u(t,1) = 0, & t \geq 0, \\ u(0,x) = u_0(x), & \qquad\quad x \in [0,1] \end{cases}$$

with $u_0(0) = u_0(1) = 0$. Show that if $\|u_0\|_\infty$ is sufficiently small, then u exists in the large.
[**Hint:** use the exponential decay of the heat semigroup in the variation of constants formula].

Bibliography

[1] A. Belleni-Morante, Applied Semigroups and Evolution Equations, Oxford University Press, 1979.

[2] H. Brezis, Functional analysis, Sobolev Spaces and Partial Differential Equations, Universitext, Springer, New York, 2011

[3] J. Diestel and J.J. Uhl Jr., Vector measures, Mathematical Surveys, No. 15, American Mathematical Society, Providence, R.I., 1977.

[4] N. Dunford and J. Schwartz, Linear Operators, Vol. I, Interscience Publisher, 1958,

[5] K.-J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Graduate Text in Mathematics **194**, Springer-Verlag, 2000.

[6] K.-J. Engel and R. Nagel, Short Course on Operator Semigroups, Universitext, Springer-Verlag, 2006.

[7] M. Fabian, P. Habala, V. Montesinos Santalucia, J. Pelant and V. Zizler, Functional Analysis and Infinite-Dimensional Geometry, Springer-Verlag 2001.

[8] J.A. Goldstein, Semigroups of Operators and Applications, Oxford University Press, 1985.

[9] P.R. Halmos, *What does the Spectral Theorem say?*, The American Mathematical Monthly 70, no. 3 (1963), 241–247.

[10] D. Henry, Geometric theory of semilinear parabolic equations, Lect. Notes in Math. **840**, Springer-Verlag, New York, 1981.

[11] E. Hille and R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society Colloquium Publications, Vol. XXXI, American Mathematical Society, Providence, R.I., 1974, Third printing of the revised edition of 1957.

[12] H. Jarchow, Locally Convex Spaces, B.G. Teubner, Stuttgard, 1981.

[13] L. Lorenzi and A. Rhandi, Semigroups of Bounded Operators and Second-Order Elliptic and Parabolic Partial Differential Equations, CRC Press, Taylor & Francis Group, 2021.

[14] H.P. Lotz, Uniform convergence of operators on L^∞ and similar spaces, Math. Z. **190** (1985), 207–220.

[15] A. Lunardi, Analytic Semigroups and Optimal Regularity in Parabolic Problems, Birkhäuser Verlag, 1995.

[16] C.-V. Pao, Nonlinear parabolic and elliptic equations, Plenum Press (1992).

[17] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Appl. Math.Sci. **44**, Springer-Verlag,1983.

[18] M. Reed and B. Simon, Methods of modern mathematical physics I, Functional Analysis, Academic Press, 1972.

[19] F. Rothe, Global Solutions of Reaction-Diffusion Systems, Lect. Notes in Math. **1072**, Springer Verlag, Berlin (1984).

[20] W. Rudin, Functional Analysis, McGraw-Hill, 1973.

[21] W. Rudin, Real and Complex Analysis, 3^{rd} ed., McGraw-Hill, 1986.

[22] J. Smoller, Shock Waves and Reaction-Diffusion Equations, Springer Verlag, Berlin (1983).

[23] K. Stromberg, An Introduction to Classical Real Analysis, AMS Chelsea Publisshing, 2015.

[24] A.E. Taylor, Introduction to Functional Analysis, John Wiley and Sons, Inc., New York, 1958.

[25] R. Whitley, *The Spectral Theorem for a normal operator*, Amer. Math. Monthly 75 (1968), 856–861.

[26] A. Zygmund, Trigonometric Series, Cambridge Univ. Press., 2nd Edition Reprinted (1968).