Lecture 2

Characterisation of generators of C_0-semigroups

Generators of C_0-semigroups have special spectral properties. They will be studied in the Section 2.2. The main goal of this lecture is the Hille-Yosida theorem which characterises generators of C_0-semigroups. The exponential formula for C_0-semigroups presented in the Section 2.3 will be useful for applications. We start with an interlude on spectral theory of operators as well as on more integration.

2.1 Interlude: the resolvent of operators, and some more integration

2.1.1 Resolvent set, spectrum and resolvent

Let X be a Banach space over $\mathbb{K}$, and let A be an operator in X.

We define the **resolvent set of** A,

$$\rho(A) := \bigl\{\lambda \in \mathbb{K};\ \lambda I - A\colon\ \operatorname{dom}(A) \to X \text{ bijective},\ (\lambda I - A)^{-1} \in \mathcal{L}(X)\bigr\}.$$

The operator $R(\lambda, A) := (\lambda I - A)^{-1}$ is called the **resolvent of** A **at** λ, and the mapping

$$R(\cdot, A)\colon \rho(A) \to \mathcal{L}(X)$$

is called the **resolvent of** A. The set

$$\sigma(A) := \mathbb{K} \setminus \rho(A)$$

is called the **spectrum of** A.

2.1 Remarks. (a) If $\rho(A) \neq \varnothing$ and $\lambda \in \rho(A)$, then $(\lambda I - A)^{-1} \in \mathcal{L}(X)$ is closed – note that every operator belonging to $\mathcal{L}(X)$ is closed. Hence $\lambda I - A$ is closed, and therefore A is closed, by the reasoning presented subsequently in part (b).

(b) If A is a closed operator and $B \in \mathcal{L}(X)$, then the sum $A + B$ is a closed operator. Indeed, if $\bigl((x_n, y_n)\bigr)$ is a sequence in $A + B$, $(x_n, y_n) \to (x, y)$ in $X \times Y$ $(n \to \infty)$, then $Bx_n \to Bx$, and therefore $Ax_n = (A + B)x_n - Bx_n \to y - Bx$, and the hypothesis that A is closed implies that $(x, y - Bx) \in A$, i.e., $x \in \operatorname{dom}(A)$ and $y = Ax + Bx$.

(c) Let A be a closed operator. Assume that $\lambda \in \mathbb{K}$ is such that $\lambda I - A\colon \operatorname{dom}(A) \to X$ is bijective. Then the inverse $(\lambda I - A)^{-1}$ is a closed operator which is defined on all of X. Therefore the closed graph theorem (for which we refer to [Yos68; II.6, Theorem 1], [Bre83; Theorem II.7]) implies that $(\lambda I - A)^{-1} \in \mathcal{L}(X)$. This implies that

$$\rho(A) = \{\lambda \in \mathbb{K};\ \lambda I - A\colon \operatorname{dom}(A) \to X \text{ bijective}\}.$$

(d) Usually, in treatments of operator theory the above notions are only defined for the case of complex Banach spaces. The reason is that many important results of spectral theory depend on complex analysis of one variable. For our purpose it is – for the moment – possible and convenient to include the case of real scalars.

The following theorem contains the basic results concerning the resolvent.

2.2 Theorem. *Let A be a closed operator in X.*

(a) *If $\lambda \in \rho(A)$, $x \in \operatorname{dom}(A)$, then $AR(\lambda, A)x = R(\lambda, A)Ax$.*

(b) *For all $\lambda, \mu \in \rho(A)$ one has the* ***resolvent equation***

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

(c) *For $\lambda \in \rho(A)$ one has $B\big(\lambda, \frac{1}{\|R(\lambda,A)\|}\big) \subseteq \rho(A)$, and for $\mu \in B\big(\lambda, \frac{1}{\|R(\lambda,A)\|}\big)$ one has*

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

As a consequence, $\rho(A)$ is an open subset of $\mathbb{K}$, and $R(\cdot, A)\colon \rho(A) \to \mathcal{L}(X)$ is analytic.

2.3 Remarks. (a) For the proof of Theorem 2.2 we recall the **Neumann series**: If $B \in \mathcal{L}(X)$ satisfies $\|B\| < 1$, then $I - B$ is invertible in $\mathcal{L}(X)$, and the inverse is given by $(I - B)^{-1} = \sum_{n=0}^{\infty} B^n$, with absolute convergence of the series.

(b) If $A \in \mathcal{L}(X)$ and $\lambda \in \mathbb{K}$ with $|\lambda| > \|A\|$, then part (a) implies that $\lambda I - A = \lambda\big(I - \frac{1}{\lambda}A\big)$ is invertible in $\mathcal{L}(X)$, with inverse

$$(\lambda I - A)^{-1} = \sum_{j=0}^{\infty} \frac{1}{\lambda^{j+1}} A^j.$$

As a consequence, $\{\lambda \in \mathbb{K};\ |\lambda| > \|A\|\} \subseteq \rho(A)$.

Proof of Theorem 2.2. (a) $AR(\lambda, A)x - \lambda R(\lambda, A)x = -x = R(\lambda, A)Ax - R(\lambda, A)\lambda x$.

(b) Multiplying the equation

$$(\mu I - A) - (\lambda I - A) = (\mu - \lambda)I|_{\operatorname{dom}(A)}$$

from the right by $R(\lambda, A)$ and from the left by $R(\mu, A)$, one obtains the resolvent equation.

(c) Let $\lambda \in \rho(A)$ and $\mu \in B\big(\lambda, \frac{1}{\|R(\lambda,A)\|}\big)$. Then the operator $I - (\lambda - \mu)R(\lambda, A)$ is invertible in $\mathcal{L}(X)$ since $|\lambda - \mu|\|R(\lambda, A)\| < 1$ (Neumann series). Therefore the equality

$$\mu I - A = (\lambda I - A) - (\lambda - \mu)I = \big(I - (\lambda - \mu)R(\lambda, A)\big)(\lambda I - A)$$

shows that the mapping $\mu I - A\colon \operatorname{dom}(A) \to X$ is bijective, and hence $\mu \in \rho(A)$. Moreover,

$$R(\mu, A) = R(\lambda, A)\big(I - (\lambda - \mu)R(\lambda, A)\big)^{-1},$$

and the formula (2.1) for the resolvent is then a consequence of the Neumann series. Now it follows that $\rho(A)$ is an open subset of $\mathbb{K}$, and the analyticity of $R(\cdot, A)$ (meaning that $R(\cdot, A)$ can be written as a power series about every point of $\rho(A)$) is a consequence of (2.1) as well. □

2.4 Remarks. (a) Theorem 2.2(c) shows that $\|R(\lambda, A)\| \geqslant \operatorname{dist}(\lambda, \sigma(A))^{-1}$ for all $\lambda \in \rho(A)$. This implies that the norm of the resolvent has to blow up if λ approaches $\sigma(A)$.

(b) As in first year analysis, the analyticity of $R(\cdot, A)$ implies that $R(\cdot, A)$ is infinitely differentiable, and from the power series (2.1) one can read off the derivatives,

$$\Big(\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big)^n R(\lambda, A) = (-1)^n\, n!\, R(\lambda, A)^{n+1} \qquad (\lambda \in \rho(A),\ n \in \mathbb{N}_0).$$

2.1.2 Integration of operator valued functions, and improper integrals

2.5 Proposition. *Let X, Y be Banach spaces, $a, b \in \mathbb{R}$, $a < b$. Let $F\colon [a, b] \to \mathcal{L}(X, Y)$ be strongly continuous, and assume that $h\colon [a, b] \to [0, \infty)$ is an integrable function such that $\|F(t)\| \leqslant h(t)$ $(a \leqslant t \leqslant b)$. Then the mapping*

$$X \ni x \mapsto \int_a^b F(t)x \,\mathrm{d}t \in Y$$

belongs to $\mathcal{L}(X, Y)$ and has norm less or equal $\int_a^b h(t)\,\mathrm{d}t$.

Some comments: Writing $\mathcal{L}(X, Y)$ we tacitly assume that the two Banach spaces are over the same scalar field. **Strongly continuous** means that $t \mapsto F(t)x$ is continuous for all $x \in X$. (In other words, it means that F is continuous with respect to the **strong operator topology** on $\mathcal{L}(X, Y)$, which is defined as the initial topology with respect to the family of mappings $\big(\mathcal{L}(X, Y) \ni A \mapsto Ax \in Y\big)_{x \in X}$.)

Proof of Proposition 2.5. The linearity of the mapping is obvious. For $x \in X$ we estimate

$$\Big\| \int_a^b F(t)x \,\mathrm{d}t \Big\| \leqslant \int_a^b \|F(t)x\| \,\mathrm{d}t \leqslant \int_a^b h(t)\,\mathrm{d}t\, \|x\|,$$

and this shows the norm estimate for the mapping. □

Abbreviating, we will write $\int_a^b F(t)\,\mathrm{d}t$ for the mapping defined in Proposition 2.5. This integral is called the **strong integral**; one has to keep in mind that, in general, it is not an integral of the $\mathcal{L}(X, Y)$-valued function as treated in Subsection 1.3.2.

We will also need 'improper integrals' of continuous Banach space valued functions. For simplicity we restrict our attention to integrals over $[0, \infty)$ (because this is what will be needed next).

2.6 Proposition. *Let X be a Banach space, $f\colon [0,\infty) \to X$ continuous, and assume that there exists an integrable function $g\colon [0,\infty) \to [0,\infty)$ such that $\|f(t)\| \leqslant g(t)$ $(0 \leqslant t < \infty)$. Then*

$$\int_0^\infty f(t)\,\mathrm{d}t := \lim_{c\to\infty} \int_0^c f(t)\,\mathrm{d}t$$

exists.

We omit the (easy) proof of this proposition and mention that Proposition 2.5 has its analogue for these improper integrals.

2.2 Characterisation of generators of C_0-semigroups

In this section let X be a Banach space.

2.7 Theorem. *Let T be a C_0-semigroup on X, and let A be its generator. Let $M \geqslant 1$, $\omega \in \mathbb{R}$ be such that*

$$\|T(t)\| \leqslant M\mathrm{e}^{\omega t} \qquad (t \geqslant 0)$$

(see Proposition 1.4).

Then $\{\lambda \in \mathbb{K};\ \operatorname{Re}\lambda > \omega\} \subseteq \rho(A)$, and for all $\lambda \in \mathbb{K}$ with $\operatorname{Re}\lambda > \omega$ one has

$$R(\lambda, A) = \int_0^\infty \mathrm{e}^{-\lambda t} T(t)\,\mathrm{d}t \quad (\textit{strong improper integral; see Subsection 2.1.2}),$$

$$\|R(\lambda, A)^n\| \leqslant \frac{M}{(\operatorname{Re}\lambda - \omega)^n} \qquad (n \in \mathbb{N}).$$

In the proof we will use the concept of rescaling. If T is a C_0-semigroup on X with generator A, and $\lambda \in \mathbb{K}$, then it is easy to see that T_λ, defined by

$$T_\lambda(t) := \mathrm{e}^{-\lambda t} T(t) \qquad (t \geqslant 0),$$

is also a C_0-semigroup, called a **rescaled semigroup**, and that the generator of T_λ is given by $A - \lambda I$; see Exercise 2.2.

Proof of Theorem 2.7. Let $\lambda \in \mathbb{K}$, $\operatorname{Re}\lambda > \omega$. Observe that the rescaled semigroup T_λ obeys the estimate

$$\|T_\lambda(t)\| \leqslant M\mathrm{e}^{(\omega - \operatorname{Re}\lambda)t} \qquad (t \geqslant 0)$$

and that the resolvent of A at λ corresponds to the resolvent of $A - \lambda I$ at 0. This means that it is sufficient to prove the existence and the formula of the resolvent for the case $\lambda = 0$ and $\omega < 0$.

The estimate $\|T(t)\| \leqslant M\mathrm{e}^{\omega t}$ $(t \geqslant 0)$ implies that the strong improper integral

$$R := \int_0^\infty T(t)\,\mathrm{d}t$$

defines an operator $R \in \mathcal{L}(X)$. Let $x \in \operatorname{dom}(A)$. Then

$$RAx = \int_0^\infty T(t)Ax\,\mathrm{d}t = \lim_{c\to\infty} \int_0^c \frac{\mathrm{d}}{\mathrm{d}t} T(t)x\,\mathrm{d}t = \lim_{c\to\infty} (T(c)x - x) = -x.$$

Further, $\|T(t)x\| \leqslant M\mathrm{e}^{\omega t}\|x\|$ and $\|AT(t)x\| \leqslant M\mathrm{e}^{\omega t}\|Ax\|$ $(t \geqslant 0)$, and therefore Theorem 1.8(b) (Hille's theorem), which also holds in the present context, implies that $Rx \in \operatorname{dom}(A)$ and

$$ARx = \int_0^\infty AT(t)x\,\mathrm{d}t = \int_0^\infty T(t)Ax\,\mathrm{d}t = RAx = -x.$$

If $x \in X$, and (x_n) is a sequence in $\operatorname{dom}(A)$ with $x = \lim_{n\to\infty} x_n$, then $Rx_n \to Rx$ and $ARx_n = -x_n \to -x$ $(n \to \infty)$, and because A is closed we conclude that $Rx \in \operatorname{dom}(A)$ and $ARx = -x$. The two equations $RA = -I|_{\operatorname{dom}(A)}$, $AR = -I$ imply that $0 \in \rho(A)$ and $R = (-A)^{-1}$.

For the powers of $R(\lambda, A)$ we now obtain (recall Remark 2.4(b))

$$\begin{aligned} R(\lambda,A)^n &= (-1)^{n-1}\frac{1}{(n-1)!}\Big(\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big)^{n-1}\int_0^\infty \mathrm{e}^{-\lambda t}T(t)\,\mathrm{d}t \\ &= \frac{1}{(n-1)!}\int_0^\infty t^{n-1}\mathrm{e}^{-\lambda t}T(t)\,\mathrm{d}t. \end{aligned} \tag{2.2}$$

(The last equality is obtained by differentiation under the integral; Exercise 2.1. See also the subsequent Remark 2.8.) By Proposition 2.5 this yields the estimate

$$\begin{aligned} \|R(\lambda,A)^n\| &\leqslant \frac{1}{(n-1)!}\,M\int_0^\infty t^{n-1}\mathrm{e}^{(\omega-\operatorname{Re}\lambda)t}\,\mathrm{d}t \\ &= \frac{1}{(n-1)!}\,M\Big(\frac{\mathrm{d}}{\mathrm{d}\omega}\Big)^{n-1}\int_0^\infty \mathrm{e}^{(\omega-\operatorname{Re}\lambda)t}\,\mathrm{d}t \\ &= \frac{1}{(n-1)!}\,M\Big(\frac{\mathrm{d}}{\mathrm{d}\omega}\Big)^{n-1}\frac{1}{\operatorname{Re}\lambda-\omega} = \frac{M}{(\operatorname{Re}\lambda-\omega)^n}\,. \end{aligned}$$

□

2.8 Remark. In the following we will mainly be interested in the case where $M = 1$ in the estimate for the C_0-semigroup, in which case the C_0-semigroup is called **quasi-contractive**. For such semigroups it is sufficient to prove the estimate for the resolvent in Theorem 2.7 for $n = 1$ (because then taking powers one obtains the estimate for all $n \in \mathbb{N}$). For $n = 1$ the second equality in (2.2) is trivial.

In Exercise 2.3 one can find a method how to reduce the proof for the estimates for the resolvents to the case of contractive C_0-semigroups.

Next, we are going to show that the necessary conditions for the generator are also sufficient. We restrict ourselves to the quasi-contractive case and delegate the general case to Exercise 2.4.

2.9 Theorem. *(Theorem of Hille-Yosida, quasi-contractive case) Let A be a closed, densely defined operator in X. Assume that there exists $\omega \in \mathbb{R}$ such that $(\omega,\infty) \subseteq \rho(A)$ and*

$$\|R(\lambda,A)\| \leqslant \frac{1}{\lambda-\omega} \qquad (\lambda \in (\omega,\infty)).$$

Then A is the generator of a C_0-semigroup T satisfying the estimate

$$\|T(t)\| \leqslant \mathrm{e}^{\omega t} \qquad (t \geqslant 0).$$

As a preliminary remark we note that it is sufficient to treat the case $\omega = 0$. Indeed, defining $\tilde{A} := A - \omega I$ we see that $\tilde{A}$ satisfies the conditions of Theorem 2.9 with $\omega = 0$. Having obtained the contractive C_0-semigroup $\tilde{T}$ with generator $\tilde{A}$ one obtains the C_0-semigroup generated by $A = \tilde{A} + \omega I$ as the rescaled semigroup $\tilde{T}_{-\omega}$.

We now define the **Yosida approximations**

$$A_n := A\Big(I - \frac{1}{n}A\Big)^{-1} = nAR(n, A) = n^2R(n, A) - nI \in \mathcal{L}(X) \qquad (n \in \mathbb{N})$$

of A. The proof of Theorem 2.9 will consist of three steps:

In the first step we show that the semigroups generated by A_n are contractive.

In the second step we show that these semigroups converge strongly to a C_0-semigroup.

In the third step we show that A is the generator of the limiting semigroup.

The following lemma justifies that the operators A_n can be considered as approximations of A.

2.10 Lemma. *Let A be a closed operator in X. Assume that there exists $\lambda_0 > 0$ such that $(\lambda_0, \infty) \subseteq \rho(A)$ and that $M := \sup_{\lambda > \lambda_0} \|\lambda R(\lambda, A)\| < \infty$. Then:*

(a) *$\lambda R(\lambda, A)x \to x$ ($\lambda \to \infty$) for all $x \in X$.*

(b) *$A\big(\lambda R(\lambda, A)\big)x \to Ax$ ($\lambda \to \infty$) for all $x \in \mathrm{dom}(A)$.*

Proof. (a) If $x \in \mathrm{dom}(A)$, then

$$\lambda R(\lambda, A)x = (\lambda I - A + A)R(\lambda, A)x = x + R(\lambda, A)Ax \to x \qquad (\lambda \to \infty).$$

As $\|\lambda R(\lambda, A)\| \leqslant M$ ($\lambda > \lambda_0$) and $\mathrm{dom}(A)$ is dense, the convergence carries over to all $x \in X$, by Proposition 1.6.

(b) For $x \in \mathrm{dom}(A)$ the convergence proved in part (a) implies

$$A\big(\lambda(\lambda I - A)^{-1}\big)x = \lambda(\lambda I - A)^{-1}Ax \to Ax \qquad (\lambda \to \infty). \qquad \square$$

Proof of Theorem 2.9. Recall that, without loss of generality, we only treat the case $\omega = 0$.

(i) For $n \in \mathbb{N}$, $t \geqslant 0$ we obtain the estimate

$$\|\mathrm{e}^{tA_n}\| = \|\mathrm{e}^{t(n^2R(n,A)-nI)}\| = \mathrm{e}^{-tn}\Big\|\sum_{k=0}^{\infty}\frac{(tn^2R(n, A))^k}{k!}\Big\| \leqslant \mathrm{e}^{-tn}\sum_{k=0}^{\infty}\frac{(tn)^k}{k!} = 1.$$

(ii) For $x \in X$, $t > 0$ and $m, n \in \mathbb{N}$ we compute

$$\begin{aligned}\mathrm{e}^{tA_m}x - \mathrm{e}^{tA_n}x &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\big(\mathrm{e}^{(t-s)A_n}\mathrm{e}^{sA_m}x\big)\,\mathrm{d}s = \int_0^t \mathrm{e}^{(t-s)A_n}\big(A_m - A_n\big)\mathrm{e}^{sA_m}x\,\mathrm{d}s \\ &= \int_0^t \mathrm{e}^{(t-s)A_n}\mathrm{e}^{sA_m}\big(A_m - A_n\big)x\,\mathrm{d}s\end{aligned}$$

(where in the last equality we have used that A_m, A_n as well as the generated semigroups commute). Recalling part (i) we obtain the estimate

$$\big\|\mathrm{e}^{tA_m}x - \mathrm{e}^{tA_n}x\big\| \leqslant t\big\|\big(A_m - A_n\big)x\big\|. \tag{2.3}$$

Let $c > 0$. For $n \in \mathbb{N}$ we define the operator $\mathcal{T}_n^c\colon X \to C([0,c];X)$ (where $C([0,c];X)$ denotes the Banach space of continuous X-valued functions, equipped with the supremum norm) by

$$\mathcal{T}_n^c x := [t \mapsto \mathrm{e}^{tA_n}x] \qquad (x \in X).$$

Then part (i) of the proof shows that $\mathcal{T}_n^c$ is a contraction, and inequality (2.3) shows that

$$\bigl\|\mathcal{T}_m^c x - \mathcal{T}_n^c x\bigr\| \leqslant c\|A_m x - A_n x\| \qquad (m, n \in \mathbb{N}),$$

for all $x \in \operatorname{dom}(A)$, which implies that $\bigl(\mathcal{T}_n^c x\bigr)_{n\in\mathbb{N}}$ is a Cauchy sequence, because $(A_n x)_{n\in\mathbb{N}}$ is convergent (to Ax). Applying Proposition 1.6 we conclude that there exists $\mathcal{T}^c \in \mathcal{L}\bigl(X, C([0,c];X)\bigr)$ such that $\mathcal{T}_n^c \to \mathcal{T}^c$ $(n \to \infty)$ strongly.

Clearly, if $0 < c < c'$, then $\mathcal{T}^{c'}x|_{[0,c]} = \mathcal{T}^c x$ for all $x \in X$, and therefore we can define $T\colon [0,\infty) \to \mathcal{L}(X)$ by

$$T(t)x := \mathcal{T}^c x(t) \qquad (0 \leqslant t < c,\ x \in X).$$

From $T(\cdot)x|_{[0,c]} = \mathcal{T}^c x$ $(c > 0,\ x \in X)$ we infer that T is strongly continuous. Since $T(t) = \text{s-lim}_{n\to\infty} \mathrm{e}^{tA_n}$ $(t \geqslant 0)$, the semigroup property carries over from the semigroups $\bigl(\mathrm{e}^{tA_n}\bigr)_{t\geqslant 0}$ to T. As a result, T is a C_0-semigroup of contractions.

(iii) Let B be the generator of T. Let $x \in \operatorname{dom}(A)$. Using the notation of part (ii), with $c := 1$, we see that $\mathcal{T}_n^1 x \to \mathcal{T}^1 x$ and $(\mathcal{T}_n^1 x)' = \mathcal{T}_n^1 A_n x \to \mathcal{T}^1 Ax$ $(n \to \infty)$ in $C([0,1];X)$ (recall Lemma 1.11). This implies that $T(\cdot)x|_{[0,1]} = \mathcal{T}^1 x$ is differentiable with continuous derivative $\mathcal{T}^1 Ax$, and therefore $x \in \operatorname{dom}(B)$ and $Bx = Ax$.

So far we have shown that $A \subseteq B$. We also know that $(0,\infty) \subseteq \rho(B)$, by Theorem 2.7, and that $(0,\infty) \subseteq \rho(A)$, by hypothesis. Now from $I - A \subseteq I - B$, the injectivity of $I - B$ and $\operatorname{ran}(I - A) = X$ we obtain $I - A = I - B$, and hence $A = B$. $\square$

2.11 Remarks. (a) The proof of the Hille-Yosida theorem for the general case can be given along the same lines; see Exercise 2.4.

(b) The Yosida approximations for the general case are defined in the same way as for the contractive case as treated in our proof. We point out that the proof given above does not show that the semigroups generated by the Yosida approximations for the non-rescaled semigroup approximate the non-rescaled semigroup (although it is true!). This might seem cumbersome, because sometimes one wants to prove properties of the semigroup using properties of the approximating semigroups. However, in the following section we will show another way of approximation which is similarly (or even more) effective.

(c) As an iteresting feature in the proof of Theorem 2.9 we point out that the approximating semigroups are norm continuous, whereas in general the resulting semigroup is only strongly continuous. The norm continuity is lost, obviously, because the convergence of the approximating semigroups is only strong (even though the strong convergence is uniform on bounded intervals).

2.3 An exponential formula

Given $a \in \mathbb{K}$, besides the exponential series there are two well-known ways to approximate e^{ta}, namely

$$\mathrm{e}^{ta} = \lim_{n\to\infty}\Big(1 + \frac{t}{n}a\Big)^{n} \qquad \text{and} \qquad \mathrm{e}^{ta} = \lim_{n\to\infty}\Big(1 - \frac{t}{n}a\Big)^{-n}.$$

Trying to replace a by an unbounded generator in the first formula leads to hopeless problems with the domains of the powers of the operators involved, whereas the second formula looks more promising because the occuring inverses are just those whose existence is guaranteed by Theorem 2.7. (In fact, the resulting formulas are those known in numerical mathematics as 'backward Euler method'.)

We mention that the proof of the 'exponential formula' presented in this section does not depend on the Hille-Yosida theorem.

2.12 Theorem. *Let X be a Banach space, T a C_0-semigroup on X, and A its generator. Then*

$$T(t)x = \lim_{n\to\infty}\Big(I - \frac{t}{n}A\Big)^{-n}x$$

for all $x \in X$, with uniform convergence on compact intervals.

2.13 Remarks. (a) If M and ω are such that $\|T(t)\| \leqslant M\mathrm{e}^{\omega t}$ $(t \geqslant 0)$, then $(\omega,\infty) \subseteq \rho(A)$, by Theorem 2.7. Assume that $\omega > 0$, and let $c > 0$. If $0 < t \leqslant c$, then $\frac{n}{t} \geqslant \frac{n}{c}$, therefore $\frac{n}{t} \in \rho(A)$ if $n > c\omega$, and then

$$\Big(I - \frac{t}{n}A\Big)^{-1} = \frac{n}{t}\Big(\frac{n}{t} - A\Big)^{-1} \in \mathcal{L}(X).$$

This means that the operator $(I - \frac{t}{n}A)^{-n}$ is defined for all $t \in [0,c]$ only for sufficiently large n, depending on c.

This problem does not occur if $\omega \leqslant 0$.

(b) From elementary analysis we will need the fact that the convergence $(1-t/n)^{-n} \to \mathrm{e}^t$ as $n \to \infty$ is uniform for t in compact subsets of $\mathbb{R}$.

(c) As a third remark we note that the expressions $(I - rA)^{-1} = (1/r)\big((1/r)I - A\big)^{-1}$, for small $r > 0$ correspond to expressions $\lambda(\lambda I - A)^{-1}$ for large $\lambda > 0$. This implies that the behaviour of $(I - rA)^{-1}$ for $r \to 0$ is the same as that of $\lambda(\lambda I - A)^{-1}$ for $\lambda \to \infty$.

Proof of Theorem 2.12. Let $M \geqslant 1$, $\omega \in \mathbb{R}$, $c > 0$ and $n > c\omega$ be as in Remark 2.13(a). For $0 < t \leqslant c$ we compute, using Remark 2.4(b),

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Big(I - \frac{t}{n}A\Big)^{-1} &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{n}{t}\Big(\frac{n}{t}I - A\Big)^{-1}\Big) = -\frac{n}{t^2}\Big(\frac{n}{t}I - A\Big)^{-1} + \frac{n}{t}\cdot\frac{n}{t^2}\Big(\frac{n}{t}I - A\Big)^{-2}\\ &= \Big(-\frac{n^2}{t^3} + \frac{n}{t^2}A + \frac{n^2}{t^3}\Big)\Big(\frac{n}{t}I - A\Big)^{-2} = \frac{1}{n}A\Big(I - \frac{t}{n}A\Big)^{-2},\end{aligned}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(I - \frac{t}{n}A\Big)^{-n} = n\Big(\Big(I - \frac{t}{n}A\Big)^{-1}\Big)^{n-1}\frac{1}{n}A\Big(I - \frac{t}{n}A\Big)^{-2} = A\Big(I - \frac{t}{n}A\Big)^{-n-1}.$$

(In the last computation the product rule for differentiation is used. The result of this computation should not be too surprising; if A is a number, then it is a consequence of the chain rule.)

Let $x \in \operatorname{dom}(A)$. Then the function $[0,t] \ni s \mapsto T(t-s)\bigl(I - \frac{s}{n}A\bigr)^{-n}x$ is continuous as well as continuously differentiable on $(0,t)$, with

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\Bigl(T(t-s)\Bigl(I-\frac{s}{n}A\Bigr)^{-n}x\Bigr) &= T(t-s)\Bigl(-A + A\Bigl(I-\frac{s}{n}A\Bigr)^{-1}\Bigr)\Bigl(I-\frac{s}{n}A\Bigr)^{-n}x \\ &= T(t-s)\Bigl(I-\frac{s}{n}A\Bigr)^{-n}\Bigl(\Bigl(I-\frac{s}{n}A\Bigr)^{-1} - I\Bigr)Ax.\end{aligned}$$

By the fundamental theorem of calculus (see Theorem 1.9) it follows that

$$\begin{aligned}&\bigl\|(I-\tfrac{t}{n}A)^{-n}x - T(t)x\bigr\| \\ &= \Bigl\|\int_0^t T(t-s)(I-\tfrac{s}{n}A)^{-n}\bigl((I-\tfrac{s}{n}A)^{-1}-I\bigr)Ax\,\mathrm{d}s\Bigr\| \\ &\leqslant c \sup_{0<s<t\leqslant c}\bigl\|T(t-s)(I-\tfrac{s}{n}A)^{-n}\bigr\| \sup_{0<s<c}\bigl\|\bigl((I-\tfrac{s}{n}A)^{-1}-I\bigr)Ax\bigr\|.\end{aligned} \tag{2.4}$$

The first of these suprema is estimated by

$$\begin{aligned}\bigl\|T(t-s)(I-\tfrac{s}{n}A)^{-n}\bigr\| &\leqslant M^2\mathrm{e}^{\omega(t-s)}(1-\tfrac{s}{n}\omega)^{-n} \\ &\leqslant M_0 := M^2\max\bigl\{1,\mathrm{e}^{\omega c}\bigr\}\sup_{0<s<c,\,n>c\omega}(1-\tfrac{s}{n}\omega)^{-n} < \infty,\end{aligned}$$

where the finiteness of the last term is a consequence of Remark 2.13(b). Now (2.4) yields

$$\sup_{0\leqslant t\leqslant c}\bigl\|(I-\tfrac{t}{n}A)^{-n}x - T(t)x\bigr\| \leqslant cM_0\sup_{0<s<c}\bigl\|\bigl((I-\tfrac{s}{n}A)^{-1}-I\bigr)Ax\bigr\|,$$

which tends to 0 as $n \to \infty$, by Lemma 2.10(a) (recall also Remark 2.13(c)).

Now the proof is completed pretty much as in part (ii) of the proof of Theorem 2.9. For $n > c\omega$ we define $\mathcal{T}_n^c \colon X \to C([0,c];X)$,

$$\mathcal{T}_n^c x := [t \mapsto (I-\tfrac{t}{n}A)^{-n}x] \qquad (x \in X).$$

Then $\|\mathcal{T}_n^c\| \leqslant M_0$ for all $n > c\omega$, and $\mathcal{T}_n^c x \to T(\cdot)x|_{[0,c]}$ $(n \to \infty)$ for all $x \in \operatorname{dom}(A)$, as shown above. Using that $\operatorname{dom}(A)$ is dense and applying Proposition 1.6 we obtain $\mathcal{T}_n^c x \to T(\cdot)x|_{[0,c]}$ $(n \to \infty)$ for all $x \in X$. □

2.14 Remark. An important application of the exponential formula will be the invariance of a closed subset $M \subseteq X$ under a C_0-semigroup. If M is invariant under $(I-rA)^{-1}$ for small $r > 0$, then Theorem 2.12 implies that M is invariant under T.

Notes

The section on resolvents etc. is pretty standard and can be found in any book on functional analysis treating fundamentals of operator theory. The Hille-Yosida theorem is basic for

C_0-semigroups and can be found (with varying proofs) in any treatment of C_0-semigroups. Our proof is Yosida's original proof in [Yos48]. The exponential formula, Theorem 2.12, can be found in [HP57; Theorem 11.6.6] or in [Paz83; Theorem 1.8.3], with proofs different from ours. A proof of the Hille-Yosida theorem can also be given using the exponential formula; see [Kat80; IX, Sections 1.2 and 1.3]. It was found by Hille, independently of Yosida, and published in [Hil48]. It is worth mentioning that the exponential formula can also be derived from the Chernoff product formula, for which we refer to [EN00; III, Theorem 5.2]. The proof of the Chernoff product formula requires results from the perturbation theory of C_0-semigroups.

Exercises

2.1 Show the second equality in (2.2).

2.2 Let T be a C_0-semigroup on the Banach space X, with generator A. Let $\lambda \in \mathbb{K}$. Show that

$$T_\lambda(t) := \mathrm{e}^{-\lambda t} T(t) \qquad (t \geqslant 0),$$

defines a C_0-semigroup (the rescaled semigroup), and that the generator of T_λ is given by $A - \lambda I$.

2.3 Let T be a bounded C_0-semigroup on the Banach space X, $M := \sup_{t \geqslant 0} \|T(t)\|$, and let A be its generator.

(a) Show that

$$|||x||| := \sup_{t \geqslant 0} \|T(t)x\| \qquad (x \in X)$$

defines a norm $||| \cdot |||$ on X which is equivalent to $\| \cdot \|$, and that T is a C_0-semigroup of contractions on $(X, ||| \cdot |||)$.

(b) For any $\alpha_1, \ldots, \alpha_n > 0$, show that

$$\|(I - \alpha_1 A)^{-1} \cdots (I - \alpha_n A)^{-1}\| \leqslant M.$$

2.4 Show the Hille-Yosida theorem for the general case:

Let A be a closed, densely defined operator in the Banach space X. Assume that there exist $M \geqslant 1$ and $\omega \in \mathbb{R}$ such that $(\omega, \infty) \subseteq \rho(A)$ and

$$\|R(\lambda, A)^n\| \leqslant \frac{M}{(\lambda - \omega)^n} \qquad (\lambda \in (\omega, \infty),\ n \in \mathbb{N}).$$

Then A is the generator of a C_0-semigroup T satisfying the estimate

$$\|T(t)\| \leqslant M \mathrm{e}^{\omega t} \qquad (t \geqslant 0).$$

(Hint: Proceed as in the proof of Theorem 2.9, with adapted estimates.)

2.5 Let T be a C_0-semigroup on the Banach space X. For $h > 0$ we define $A_h := h^{-1}(T(h) - I)$. Show that $\mathrm{e}^{tA_h} x \to T(t)x$ for all $x \in X$ as $h \to 0$, uniformly for t in compact subsets of $[0, \infty)$. (Hint: Use a procedure similar to the proof of the exponential formula, Theorem 2.12.)

References

[Bre83] H. Brezis: *Analyse fonctionelle – Théorie et applications*. Masson, Paris, 1983.

[EN00] K.-J. Engel and R. Nagel: *One-Parameter Semigroups for Linear Evolution Equations*. Springer, New York, 2000.

[Hil48] E. Hille: *Functional Analysis and Semi-Groups*. Amer. Math. Soc., New York, 1948.

[HP57] E. Hille and R. S. Phillips: *Functional Analysis and Semi-Groups*. Revised edition. Amer. Math. Soc., Providence, Rh. I., 1957.

[Kat80] T. Kato: *Perturbation Theory for Linear Operators*. Corrected Printing of the Second edition. Springer, Berlin, 1980.

[Paz83] A. Pazy: *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, New York, 1983.

[Yos48] K. Yosida: On the differentiability and the representation of one-parameter semi-groups of linear operators. *J. Math. Soc. Japan* **1**, 15–21 (1948).

[Yos68] K. Yosida: *Functional Analysis*. 2nd edition. Springer, New York, 1968.