3
Strongly continuous semigroups

The most central part of a well-posed linear system is its semigroup. This chapter is devoted to a study of the properties of C_0 semigroups, both in the time domain and in the frequency domain. Typical time domain issues are the generator of a semigroup, the dual semigroup, and the nonhomogeneous initial value problem. The resolvent of the generator lives in the frequency domain.

3.1 Norm continuous semigroups

We begin by introducing the notion of the *generator* of a C_0 (semi)group (cf. Definition 2.2.2).

Definition 3.1.1

(i) The generator A of a C_0 semigroup $\mathfrak{A}$ is the operator

$$Ax := \lim_{h \downarrow 0} \frac{1}{h}(\mathfrak{A}^h - 1)x,$$

defined for all those $x \in X$ for which this limit exists.

(ii) The generator A of a C_0 group $\mathfrak{A}$ is the operator

$$Ax := \lim_{h \to 0} \frac{1}{h}(\mathfrak{A}^h - 1)x,$$

defined for all those $x \in X$ for which this limit exists.

Before we continue our study of C_0 semigroups and their generators, let us first study the smaller class of uniformly continuous semigroups, i.e., semigroups $\mathfrak{A}$ which satisfy (cf. Definition 2.2.2)

$$\lim_{t \downarrow 0} \|\mathfrak{A}^t - 1\| = 0. \tag{3.1.1}$$

Clearly, every uniformly continuous semigroup is a C_0 semigroup.

We begin by presenting an example of a uniformly continuous (semi)group. (As we shall see in Theorem 3.1.3, every uniformly continuous (semi)group is of this type.)

Example 3.1.2 *Let $A \in \mathcal{B}(X)$, and define*

$$e^{At} := \sum_{n=0}^{\infty} \frac{(At)^n}{n!}, \qquad t \in \mathbb{R}. \tag{3.1.2}$$

Then e^{At} is a uniformly continuous group on X, and its generator is A. This group satisfies $\left\|e^{At}\right\| \le e^{\|A\||t|}$ for all $t \in \mathbb{R}$. In particular, the growth bounds of the semigroups $t \mapsto e^{At}$ and $t \mapsto e^{-At}$ (where $t \ge 0$) are bounded by $\|A\|$ (cf. Definition 2.5.6).

Proof The series in (3.1.2) converges absolutely since

$$\sum_{n=0}^{\infty} \left\| \frac{(At)^n}{n!} \right\| \le \sum_{n=0}^{\infty} \frac{(\|A\||t|)^n}{n!} = e^{\|A\||t|}.$$

This proves that $\|e^{At}\|$ satisfies the given bounds. Clearly $e^{0A} = 1$. Being a power series, the function $t \mapsto e^{At}$ is analytic, hence uniformly continuous for all t. By differentiating the power series (this is permitted since e^{At} is analytic) we find that the generator of e^{At} is A (and the limit in Definition 3.1.1 is uniform). Thus, it only remains to verify the group property $e^{A(s+t)} = e^{As}e^{At}$, which is done as follows:

$$\begin{aligned} e^{A(s+t)} &= \sum_{n=0}^{\infty} \frac{A^n(s+t)^n}{n!} = \sum_{n=0}^{\infty} \frac{A^n}{n!} \sum_{k=0}^{n} \binom{n}{k} s^k t^{n-k} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{A^k s^k}{k!} \frac{A^{n-k} t^{n-k}}{(n-k)!} = \sum_{k=0}^{\infty} \frac{A^k s^k}{k!} \sum_{n=k}^{\infty} \frac{A^{n-k} t^{n-k}}{(n-k)!} \\ &= e^{As} e^{At}. \end{aligned}$$

□

Theorem 3.1.3 *Let $\mathfrak{A}$ be a uniformly continuous semigroup. Then the following claims are true:*

(i) *$\mathfrak{A}$ has a bounded generator A and $\mathfrak{A}^t = e^{At}$ for all $t \ge 0$;*
(ii) *$t \mapsto \mathfrak{A}^t$ is analytic and $\frac{d}{dt}\mathfrak{A}^t = A\mathfrak{A}^t = \mathfrak{A}^t A$ for all $t \ge 0$;*
(iii) *$\mathfrak{A}$ can be extended to an analytic group on $\mathbb{R}$ satisfying $\frac{d}{dt}\mathfrak{A}^t = A\mathfrak{A}^t = \mathfrak{A}^t A$ for all $t \in \mathbb{R}$.*

Remark 3.1.4 Actually a slightly stronger result is true: Every C_0 semigroup $\mathfrak{A}$ satisfying

$$\limsup_{t\downarrow 0} \|\mathfrak{A}^t - 1\| < 1 \tag{3.1.3}$$

has a bounded generator A and $\mathfrak{A}^t = e^{At}$. Alternatively, every C_0 semigroup $\mathfrak{A}$ for which the operator $\int_0^h \mathfrak{A}^s\,ds$ is invertible for some $h > 0$ has a bounded generator A and $\mathfrak{A}^t = e^{At}$. The proof is essentially the same as the one given below (it uses strong integrals instead of uniform integrals).

Proof of Theorem 3.1.3 (i) For sufficiently small positive h, $\left\|1 - 1/h\int_0^h \mathfrak{A}^s\,ds\right\| < 1$, hence $1/h\int_0^h \mathfrak{A}^s\,ds$ is invertible, and so is $\int_0^h \mathfrak{A}^s\,ds$. By the semigroup property, for $0 < t < h$,

$$\begin{aligned}
\frac{1}{t}(\mathfrak{A}^t - 1)\int_0^h \mathfrak{A}^s\,ds &= \frac{1}{t}\Big(\int_0^h \mathfrak{A}^{s+t}\,ds - \int_0^h \mathfrak{A}^s\,ds\Big)\\
&= \frac{1}{t}\Big(\int_t^{t+h} \mathfrak{A}^{s+t}\,ds - \int_0^h \mathfrak{A}^s\,ds\Big)\\
&= \frac{1}{t}\Big(\int_h^{t+h} \mathfrak{A}^{s+t}\,ds - \int_0^t \mathfrak{A}^s\,ds\Big).
\end{aligned}$$

Multiply by $\left(\int_0^h \mathfrak{A}^s\,ds\right)^{-1}$ to the right and let $t \downarrow 0$ to get

$$\lim_{t\downarrow 0}\frac{1}{t}(\mathfrak{A}^t - 1) = (\mathfrak{A}^h - 1)\Big(\int_0^h \mathfrak{A}^s\,ds\Big)^{-1}$$

in the uniform operator norm. This shows that $\mathfrak{A}$ has the bounded generator $A = (\mathfrak{A}^h - 1)\left(\int_0^h \mathfrak{A}^s\,ds\right)^{-1}$. By Example 3.1.2, the group e^{At} has the same generator A as $\mathfrak{A}$. By Theorem 3.2.1(vii) below, $\mathfrak{A}^t = e^{At}$ for $t \ge 0$.

(ii)–(iii) See Example 3.1.2 and its proof. □

3.2 The generator of a C_0 semigroup

We now return to the more general class of C_0 semigroups. We already introduced the notion of the *generator* of a C_0 semigroup in Definition 3.1.1. Some basic properties of this generator are listed in the following theorem.

Theorem 3.2.1 *Let $\mathfrak{A}^t$ be a C_0 semigroup on a Banach space X with generator A.*

(i) *For all $x \in X$,*

$$\lim_{h\downarrow 0}\frac{1}{h}\int_t^{t+h} \mathfrak{A}^s x\,ds = \mathfrak{A}^t x.$$

(ii) *For all* $x \in X$ *and* $0 \le s < t < \infty$, $\int_s^t \mathfrak{A}^v x\,dv \in \mathcal{D}(A)$ *and*

$$\mathfrak{A}^t x - \mathfrak{A}^s x = A \int_s^t \mathfrak{A}^v x\,dv.$$

(iii) *For all* $x \in \mathcal{D}(A)$ *and* $t \ge 0$, $\mathfrak{A}^t x \in \mathcal{D}(A)$, $t \mapsto \mathfrak{A}^t x$ *is continuously differentiable in X, and*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{A}^t x = A\mathfrak{A}^t x = \mathfrak{A}^t A x, \qquad t \ge 0.$$

(iv) *For all* $x \in \mathcal{D}(A)$ *and* $0 \le s \le t < \infty$,

$$\mathfrak{A}^t x - \mathfrak{A}^s x = A \int_s^t \mathfrak{A}^v x\,dv = \int_s^t \mathfrak{A}^v A x\,dv.$$

(v) *For all* $n = 1, 2, 3, \ldots$, *if* $x \in \mathcal{D}(A^n)$, *then* $\mathfrak{A}^t x \in \mathcal{D}(A^n)$ *for all* $t \ge 0$, *the function* $t \mapsto \mathfrak{A}^t x$ *is n times continuously differentiable in X, and for all* $k = 0, 1, 2, \ldots, n$,

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^n \mathfrak{A}^t x = A^k \mathfrak{A}^t A^{n-k} x, \qquad t \ge 0.$$

(vi) *A is a closed linear operator and* $\bigcap_{n=1}^{\infty} \mathcal{D}(A^n)$ *is dense in X. For each* $x \in \bigcap_{n=1}^{\infty} \mathcal{D}(A^n)$ *the function* $t \mapsto \mathfrak{A}^t x$ *belongs to* $C^\infty(\overline{\mathbb{R}}^+; U)$.

(vii) $\mathfrak{A}$ *is uniquely determined by its generator A.*

Proof (i) This follows from the continuity of $s \mapsto \mathfrak{A}^s x$ (see Lemma 2.2.13(ii)).

(ii) Let $x \in X$ and $h > 0$. Then

$$\begin{aligned}\frac{1}{h}(\mathfrak{A}^h - 1)\int_s^t \mathfrak{A}^v x\,dv &= \frac{1}{h}\int_s^t \mathfrak{A}^{v+h} x - \mathfrak{A}^v x)\,dv \\ &= \frac{1}{h}\int_t^{t+h} \mathfrak{A}^v x\,dv - \int_s^{s+h} \mathfrak{A}^v x\,dv.\end{aligned}$$

As $h \downarrow 0$ this tends to $\mathfrak{A}^t x - \mathfrak{A}^s x$.

(iii) Let $x \in \mathcal{D}(A)$ and $h > 0$. Then

$$\frac{1}{h}(\mathfrak{A}^h - 1)\mathfrak{A}^t x = \mathfrak{A}^t \frac{1}{h}(\mathfrak{A}^h - 1)x \to \mathfrak{A}^t A x \text{ as } h \downarrow 0.$$

Thus, $\mathfrak{A}^t x \in \mathcal{D}(A)$, and $A\mathfrak{A}^t x = \mathfrak{A}^t A x$ is equal to the right-derivative of $\mathfrak{A}^t x$ at t. To see that it is also a left-derivative we compute

$$\frac{1}{h}(\mathfrak{A}^t x - \mathfrak{A}^{t-h} x) - \mathfrak{A}^t A x = \mathfrak{A}^{t-h}\Big(\frac{1}{h}(\mathfrak{A}^h x - x) - Ax\Big) + (\mathfrak{A}^{t-h} - \mathfrak{A}^t)Ax.$$

This tends to zero because of the uniform boundedness of $\mathfrak{A}^{t-h}$ and the strong continuity of $\mathfrak{A}^t$ (see Lemma 2.2.13).

(iv) We get (iv) by integrating (iii).

(v) This follows from (iii) by induction.

(vi) The linearity of A is trivial. To prove that A is closed we let $x_n \in \mathcal{D}(A)$, $x_n \to x$, and $Ax_n \to y$ in X, and claim that $Ax = y$. By part (iv) with $s = 0$,

$$\mathfrak{A}^t x_n - x_n = \int_0^t \mathfrak{A}^s A x_n \, ds.$$

Both sides converge as $n \to \infty$ (the integrand converges uniformly on $[0, t]$), hence

$$\mathfrak{A}^t x - x = \int_0^t \mathfrak{A}^s y \, ds.$$

Divide by t, let $t \downarrow 0$, and use part (i) to get $Ax = y$.

We still need to show that $\cap_{n=1}^{\infty} \mathcal{D}(A^n)$ is dense in X. Pick some real-valued C^∞ function η with compact support in $(0, 1)$ and $\int_0^\infty \eta(s) \, ds = 1$. For each $x \in X$ and $k = 1, 2, 3, \ldots$, we define

$$x_k = k \int_0^1 \eta(ks) \mathfrak{A}^s x \, ds.$$

Then, for each $h > 0$,

$$\begin{aligned}
\frac{1}{h}(\mathfrak{A}^h - 1)x_k &= \frac{1}{h} \int_0^1 \eta(ks)[\mathfrak{A}^{s+h} x - \mathfrak{A}^s x] \, ds \\
&= k \int_0^{1+h} \frac{1}{h}[\eta(k(s-h)) - \eta(ks)] \mathfrak{A}^s x \, ds \\
&\to -k^2 \int_0^1 \dot{\eta}(ks) \mathfrak{A}^s x \, ds \text{ as } h \downarrow 0.
\end{aligned}$$

Thus, $x_k \in \mathcal{D}(A)$ and $Ax_k = -k^2 \int_0^1 \dot{\eta}(ks) \mathfrak{A}^s x$. We can repeat the same argument with η replaced by $\dot{\eta}$, etc., to get $x_k \in \mathcal{D}(A^n)$ for every $n = 1, 2, 3 \ldots$. This means that $x_k \in \cap_{n=1}^{\infty} \mathcal{D}(A^n)$.

We claim that $x_k \to x$ as $k \to \infty$, proving the density of $\cap_{n=1}^{\infty} \mathcal{D}(A^n)$ in X. To see this we make a change of integration variable to get

$$x_k = \int_0^1 \eta(s) \mathfrak{A}^{s/k} x \, ds.$$

The function $\mathfrak{A}^{s/k} x$ tends uniformly to x on $[0, 1]$, hence the integral tends to $\int_0^\infty \eta(s) x \, ds = x$ as $k \to \infty$.

That $\mathfrak{A}x \in C^\infty(\overline{\mathbb{R}}^+; U)$ whenever $x \in \cap_{n=1}^{\infty} \mathcal{D}(A^n)$ follows from (iv).

(vii) Suppose that there is another C_0 semigroup $\mathfrak{A}_1$ with the same generator A. Take $x \in \mathcal{D}(A)$, $t > 0$, and consider the function $s \mapsto \mathfrak{A}^{t-s} \mathfrak{A}_1^s x$, $s \in [0, t]$.

We can use part (iii) and the chain rule to compute its derivative in the form

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathfrak{A}^s\mathfrak{A}_1^{t-s}x = A\mathfrak{A}^s\mathfrak{A}_1^{t-s}x - \mathfrak{A}^s A\mathfrak{A}_1^{t-s}x = \mathfrak{A}^s\mathfrak{A}_1^{t-s}Ax - \mathfrak{A}^s\mathfrak{A}_1^{t-s}Ax = 0.$$

Thus, this function is a constant. Taking $s = 0$ and $s = t$ we get $\mathfrak{A}^t x = \mathfrak{A}_1^t x$ for all $x \in \mathcal{D}(A)$. By the density of $\mathcal{D}(A)$ in X, the same must be true for all $x \in X$. □

To illustrate Definition 3.1.1, let us determine the generators of the shift (semi)groups τ^t, τ_+^t, τ_-^t, $\tau_{[0,T)}^t$, and $\tau_{\mathbb{T}_T}^t$ in Examples 2.3.2 and 2.5.3. The domains of these generators are spaces of the following type:

Definition 3.2.2 Let J be a subinterval of $\mathbb{R}$, $\omega \in \mathbb{R}$, and let U be a Banach space.

(i) A function u belongs to $W_{\text{loc}}^{n,p}(J;U)$ if it is an nth order integral of a function $u^{(n)} \in L_{\text{loc}}^p(J;U)$ (i.e., $u^{(n-1)}(t_2) - u^{(n-1)}(t_1) = \int_{t_1}^{t_2} u^{(n)}(s)\,ds$, etc.).[1] It belongs to $W_\omega^{n,p}(J;U)$ if, in addition, $u^{(k)} \in L_\omega^p(J;U)$ for all $k = 0, 1, 2, \ldots, n$.

(ii) The space $W_{c,\text{loc}}^{n,p}(\mathbb{R};U)$ consists of the functions in $W_{\text{loc}}^{n,p}(\mathbb{R};U)$ whose support is bounded to the left, and the space $W_{\omega,\text{loc}}^{n,p}(\mathbb{R};U)$ consists of the functions u in $W_{\text{loc}}^{n,p}(\mathbb{R};U)$ which satisfy $\pi_- u \in W_\omega^{n,p}(\mathbb{R}^-;U)$.

(iii) The spaces $W_{0,\omega}^{n,p}(J;U)$, $W_{0,\omega,\text{loc}}^{n,p}(\mathbb{R};U)$, $BC_\omega^n(J;U)$, $BC_{\omega,\text{loc}}^n(\mathbb{R};U)$, $BC_{0,\omega}^n(J;U)$, $BC_{0,\omega,\text{loc}}^n(\mathbb{R};U)$, $BUC_\omega^n(J;U)$, $BUC_{\omega,\text{loc}}^n(\mathbb{R};U)$, $Reg_\omega^n(J;U)$, $Reg_{\omega,\text{loc}}^n(\mathbb{R};U)$, $Reg_{0,\omega}^n(J;U)$, and $Reg_{0,\omega,\text{loc}}^n(\mathbb{R};U)$ are defined in an analogous way, with L^p replaced by BC, BC_0, BUC, Reg, or Reg_0.

Example 3.2.3 *The generators of the (semi)groups τ^t, τ_+^t, τ_-^t, $\tau_{[0,T)}^t$, and $\tau_{\mathbb{T}_T}^t$ in Examples 2.3.2 and 2.5.3 are the following:*

(i) *The generator of the bilateral left shift group τ^t on $L_\omega^p(\mathbb{R};U)$ is the differentiation operator $\frac{\mathrm{d}}{\mathrm{d}s}$ with domain $W_\omega^{1,p}(\mathbb{R};U)$, and the generator of the left shift group τ^t on $BUC_\omega(\mathbb{R};U)$ is the differentiation operator $\frac{\mathrm{d}}{\mathrm{d}s}$ with domain $BUC_\omega^1(\mathbb{R};U)$. We denote these generators simply by $\frac{\mathrm{d}}{\mathrm{d}s}$.*

(ii) *The generator of the incoming left shift semigroup τ_+^t on $L_\omega^p(\mathbb{R}^+;U)$ is the differentiation operator $\frac{\mathrm{d}}{\mathrm{d}s}$ with domain $W_\omega^{1,p}(\overline{\mathbb{R}}^+;U)$, and the generator of the left shift semigroup τ_+^t on $BUC_\omega(\overline{\mathbb{R}}^+;U)$ is the differentiation operator $\frac{\mathrm{d}}{\mathrm{d}s}$ with domain $BUC_\omega^1(\overline{\mathbb{R}}^+;U)$. We denote these generators by $\frac{\mathrm{d}}{\mathrm{d}s}_+$.*

[1] Our definition of $W_{\text{loc}}^{n,p}$ implies that the functions in this space are locally absolutely continuous together with their derivatives up to order $n-1$. This is true independently of whether U has the Radon–Nikodym property or not.

(iii) *The generator of the outgoing left shift semigroup* τ_-^t *on* $L^p_\omega(\mathbb{R}^-;U)$ *is the differentiation operator* $\frac{d}{ds}$ *with domain* $\{u \in W^{1,p}_\omega(\overline{\mathbb{R}}^-;U) \mid u(0)=0\}$*, and the generator of the left shift semigroup* τ_-^t *on* $\{u \in BUC_\omega(\overline{\mathbb{R}}^-;U) \mid u(0)=0\}$ *is the differentiation operator* $\frac{d}{ds}$ *with domain* $\{u \in BUC^1_\omega(\overline{\mathbb{R}}^-;U) \mid u(0)=\dot{u}(0)=0\}$*. We denote these generators by* $\frac{d}{ds}_-$.

(iv) *The generator of the finite left shift semigroup* $\tau_{[0,T)}^t$ *on* $L^p([0,T);U)$ *is the differentiation operator* $\frac{d}{ds}$ *with domain* $\{u \in W^{1,p}([0,T];U) \mid u(T)=0\}$*, and the generator of the left shift semigroup* $\tau_{[0,T)}^t$ *on* $\{u \in C([0,T];U) \mid u(T)=0\}$ *is the differentiation operator* $\frac{d}{ds}$ *with domain* $\{u \in C^1([0,T];U) \mid u(T)=\dot{u}(T)=0\}$*. We denote these generators by* $\frac{d}{ds}_{[0,T)}$.

(v) *The generator of the circular left shift group* $\tau_{\mathbb{T}_T}^t$ *on* $L^p(\mathbb{T}_T;U)$ *is the differentiation operator* $\frac{d}{ds}$ *with domain* $W^{1,p}(\mathbb{T}_T;U)$ *(which can be identified with* $\{u \in W^{1,p}([0,T];U) \mid u(T)=u(0)\}$*), and the generator of the circular left shift group* $\tau_{\mathbb{T}_T}^t$ *on* $C(\mathbb{T}_T;U)$ *is the differentiation operator* $\frac{d}{ds}$ *with domain* $C^1(\mathbb{T}_T;U)$ *(which can be identified with the set* $\{u \in C^1([0,T];U) \mid u(T)=u(0)$ *and* $\dot{u}(T)=\dot{u}(0)\}$*). We denote these generators by* $\frac{d}{ds}_{\mathbb{T}_T}$.

Proof The proofs are very similar to each other, so let us only prove, for example, (iii). Since the proof for the L^p-case works in the BUC-case, too, we restrict the discussion to the L^p-case. For simplicity we take $\omega=0$, but the same argument applies when ω is nonzero.

Suppose that $u \in L^p(\mathbb{R}^-;U)$, and that $\frac{1}{h}(\tau_+^h u - u) \to g$ in $L^p(\mathbb{R}^-;U)$ as $h \downarrow 0$. If we extend u and g to $L^p(\mathbb{R};U)$ by defining them to be zero on $\overline{\mathbb{R}}^+$, then this can be written as $\frac{1}{h}(\tau^h u - u) \to g$ in $L^p(\mathbb{R};U)$ as $h \downarrow 0$.

Fix some $a \in \mathbb{R}$, and for each $t \in \mathbb{R}$, define

$$f(t) = \int_t^{t+a} u(s)\,ds = \int_0^a u(s+t)\,ds.$$

Then

$$\begin{aligned}\frac{1}{h}(\tau^h f - f) &= \int_0^a \frac{1}{h}(\tau^h u(s+t) - u(s+t))\,ds \\ &= \int_t^{t+a} \frac{1}{h}(\tau^h u(s) - u(s))\,ds \\ &\to \int_t^{t+a} g(s)\,ds \text{ as } h \downarrow 0.\end{aligned}$$

On the other hand

$$\frac{1}{h}(\tau^h f - f) = \frac{1}{h}\int_{t+h}^{t+a+h} u(s)\,ds - \frac{1}{h}\int_t^{t+a} u(s)\,ds$$
$$= \frac{1}{h}\int_{t+a}^{t+a+h} u(s)\,ds - \frac{1}{h}\int_t^{t+h} u(s)\,ds,$$

and as $h \downarrow 0$ this tends to $u(t+a) - u(t)$ whenever both t and $t+a$ are Lebesgue points of u. We conclude that for almost all a and t,

$$u(t+a) = u(t) + \int_t^{t+a} g(s)\,ds.$$

By definition, this means that $u \in W^{1,p}(\mathbb{R}; U)$ and that $\dot{u} = g$. Since we extended u to all of $\mathbb{R}$ by defining u to be zero on $\overline{\mathbb{R}}^+$, we have, in addition $u(0) = 0$ (if we redefine u on a set of measure zero to make it continuous everywhere).

To prove the converse claim it suffices to observe that, if $u \in W^{1,p}(\overline{\mathbb{R}}^-; U)$ and $u(0) = 0$, then we can extend u to a function in $W^{1,p}(\mathbb{R}; U)$ by defining u to be zero on $\mathbb{R}^+$, and that

$$\frac{1}{h}(\tau^h u - u)(t) = \frac{1}{h}(u(t+h) - u(t)) = \frac{1}{h}\int_t^{t+h} \dot{u}(s)\,ds,$$

which tends to $\dot{u}$ in $L^p(\mathbb{R}; U)$ as $h \downarrow 0$ (see, e.g., Gripenberg et al. [1990, Lemma 7.4, p. 67]). □

Let us record the following fact for later use:

Lemma 3.2.4 *For $1 \le p < \infty$, $W^{1,p}_\omega(\mathbb{R}; U) \subset BC_{0,\omega}(\mathbb{R}; U)$, i.e., every $u \in W^{1,p}_\omega(\mathbb{R}; U)$ is continuous and $\mathrm{e}^{-\omega t}u(t) \to 0$ as $t \to \pm\infty$.*

Proof The continuity is obvious. The function $u_{-\omega}(t) = \mathrm{e}^{-\omega t}u(t)$ belongs to L^p, and so does its derivative $-\omega u_{-\omega} + \mathrm{e}_{-\omega}\dot{u}$. This implies that $u_{-\omega}(t) \to 0$ as $t \to \infty$. □

By combining Theorem 3.2.1(vi) with Example 3.2.3 we get the major part of the following lemma:

Lemma 3.2.5 *Let $1 \le p < \infty$, $\omega \in \mathbb{R}$, and $n = 0, 1, 2, \ldots$. Then $C_c^\infty(\mathbb{R}; U)$ is dense in $L^p_\omega(\mathbb{R}; U)$, $L^p_{\mathrm{loc}}(\mathbb{R}; U)$, $W^{n,p}(\mathbb{R}; U)$, $W^{n,p}_{\mathrm{loc}}(\mathbb{R}; U)$, $BC_0^n(\mathbb{R}; U)$, and $C^n(\mathbb{R}; U)$.*

Proof It follows from Theorem 3.2.1(vi) and Example 3.2.3 that $\cap_{k=1}^\infty W^{k,p}(\mathbb{R}; U)$ is dense in $L^p(\mathbb{R}; U)$ and in $W^{n,p}(\mathbb{R}; U)$. Let u belong to this space. Then $u \in C^\infty$. Choose any $\eta \in C_c^\infty(\mathbb{R}; \mathbb{R})$ satisfying $\eta(t) = 1$ for $|t| \le 1$, and define $u_m(t) = \eta(t/m)u(t)$. Then $u_m \in C_c^\infty(\mathbb{R}; U)$, and $u_m \to u$ in

$L^p(\mathbb{R};U)$ and in $W^{n,p}(\mathbb{R};U)$, proving the density of C_c^∞ in L^p and in $W^{n,p}$. The other claims are proved in a similar manner (see also Lemma 2.3.3). □

Example 3.2.6 *Let $\mathfrak{A}^t$ be a C_0 semigroup on a Banach space X with generator A.*

(i) *For each $\alpha \in \mathbb{C}$, the generator of the exponentially shifted semigroup $e^{\alpha t}\mathfrak{A}^t$, $t \ge 0$ (see Example 2.3.5) is $A + \alpha$.*
(ii) *For each $\lambda > 0$, the generator of the time compressed semigroup $\mathfrak{A}^{\lambda t}$, $t \ge 0$ (see Example 2.3.6) is λA.*
(iii) *For each (boundedly) invertible $E \in \mathcal{B}(X_1;X)$, the generator A_E of the similarity transformed semigroup $\mathfrak{A}^t_E = E^{-1}\mathfrak{A}^t E$, $t \ge 0$ (see Example 2.3.7) is $A_E = E^{-1}AE$, with domain $\mathcal{D}(A_E) = E^{-1}\mathcal{D}(A)$.*

We leave the easy proof to the reader.

Theorem 3.2.1 does not say anything about the spectrum and resolvent set of the generator A. These notions and some related ones are defined as follows:

Definition 3.2.7 Let $A\colon X \supset \mathcal{D}(A) \to X$ be closed, and let $\alpha \in \mathbb{C}$.

(i) α belongs to the *resolvent set* $\rho(A)$ of A if $\alpha - A$ is injective, onto, and has an inverse $(\alpha - A)^{-1} \in \mathcal{B}(X)$. Otherwise α belongs to the *spectrum* $\sigma(A)$ of A.
(ii) α belongs to the *point spectrum* $\sigma_p(A)$, or equivalently, α is an *eigenvalue* of A, if $(\alpha - A)$ is not injective. A vector $x \in X$ satisfying $(\alpha - A)x = 0$ is called an *eigenvector* corresponding to the eigenvalue α.
(iii) α belongs to the *residual spectrum* $\sigma_r(A)$ if $(\alpha - A)$ is injective but its range is not dense in X.
(iv) α belongs to the *continuous spectrum* $\sigma_c(A)$ if $(\alpha - A)$ is injective and has dense range, but the range is not closed.
(v) The *resolvent* of A is the operator-valued function $\alpha \mapsto (\alpha - A)^{-1}$, defined on $\rho(A)$.

By the closed graph theorem, $\sigma(A)$ is the disjoint union of $\sigma_p(A)$, $\sigma_r(A)$, and $\sigma_c(A)$. The different parts of the spectrum need not be closed (see Examples 3.3.1 and 3.3.5), but, as the following lemma shows, the resolvent set is always open, hence the whole spectrum is always closed.

Lemma 3.2.8 *Let A be a (closed) operator $X \supset \mathcal{D}(A) \to X$, with a nonempty resolvent set.*

(i) *For each α and β in the resolvent set of A,*

$$(\alpha - A)^{-1} - (\beta - A)^{-1} = (\beta - \alpha)(\alpha - A)^{-1}(\beta - A)^{-1}. \tag{3.2.1}$$

In particular, $(\alpha - A)^{-1}(\beta - A)^{-1} = (\beta - A)^{-1}(\alpha - A)^{-1}$.

(ii) *Let* $\alpha \in \rho(A)$ *and denote* $\|(\alpha - A)^{-1}\|$ *by* κ. *Then every* β *in the circle* $|\beta - \alpha| < 1/\kappa$ *belongs to the resolvent set of A, and*

$$\|(\beta - A)^{-1}\| \le \frac{\kappa}{1 - \kappa|\beta - \alpha|}. \tag{3.2.2}$$

(iii) *Let* $\alpha \in \rho(A)$. *Then* $\delta\|(\alpha - A)^{-1}\| \ge 1$, *where* δ *is the distance from* α *to* $\sigma(A)$.

The identity (3.2.1) in (i) is usually called *the resolvent identity*. Note that the closedness of A is a consequence of the fact that A has a nonempty resolvent set.

Proof of Lemma 3.2.8. (i) Multiply the left hand side by $(\alpha - A)$ to the left and by $(\beta - A)$ to the right to get

$$(\alpha - A)\big[(\alpha - A)^{-1} - (\beta - A)^{-1}\big](\beta - A) = \beta - \alpha.$$

(ii) By part (i), for all $\beta \in \mathbb{C}$,

$$(\beta - A) = \big(1 + (\beta - \alpha)(\alpha - A)^{-1}\big)(\alpha - A). \tag{3.2.3}$$

It follows from the contraction mapping principle that if we take $|\beta - \alpha| < 1/\kappa$, then $\big(1 + (\beta - \alpha)(\alpha - A)^{-1}\big)$ is invertible and

$$\Big\|\big(1 + (\beta - \alpha)(\alpha - A)^{-1}\big)^{-1}\Big\| \le \frac{1}{1 - \kappa|\beta - \alpha|}.$$

This combined with (3.2.3) implies that $\beta \in \rho(A)$ and that (3.2.2) holds.

(iii) This follows from (ii). □

Our next theorem lists some properties of the resolvent $(\lambda - A)^{-1}$ of the generator of a C_0 semigroup. Among others, it shows that the resolvent set of the generator of a semigroup contains a right half-plane.

Theorem 3.2.9 *Let* $\mathfrak{A}^t$ *be a* C_0 *semigroup on a Banach space X with generator A and growth bound* $\omega_{\mathfrak{A}}$ *(see Definition 2.5.6).*

(i) *Every* $\lambda \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$ *belongs to the resolvent set of A, and*

$$(\lambda - A)^{-(n+1)}x = \frac{1}{n!}\int_0^\infty s^n e^{-\lambda s}\mathfrak{A}^s x\, ds$$

for all $x \in X$, $\lambda \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$, *and* $n = 0, 1, 2, \ldots$. *In particular,*

$$(\lambda - A)^{-1}x = \int_0^\infty e^{-\lambda s}\mathfrak{A}^s x\, ds$$

for all $x \in X$ *and* $\lambda \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$.

(ii) *For each $\omega > \omega_{\mathfrak{A}}$ there is a finite constant M such that*

$$\left\|(\lambda - A)^{-n}\right\| \le M(\Re\lambda - \omega)^{-n}$$

for all $n = 1, 2, 3, \ldots$ and $\lambda \in \mathbb{C}^+_\omega$. In particular,

$$\left\|(\lambda - A)^{-1}\right\| \le M(\Re\lambda - \omega)^{-1}$$

for all $\lambda \in \mathbb{C}^+_\omega$.

(iii) *For all $x \in X$, the following limits exist in the norm of X*

$$\lim_{\lambda\to+\infty} \lambda(\lambda - A)^{-1}x = x \text{ and } \lim_{\lambda\to+\infty} A(\lambda - A)^{-1}x = 0.$$

(iv) *For all $t \ge 0$ and all $\lambda \in \rho(A)$,*

$$(\lambda - A)^{-1}\mathfrak{A}^t = \mathfrak{A}^t(\lambda - A)^{-1}.$$

Proof (i) Define $\mathfrak{A}^t_\lambda = \mathrm{e}^{-\lambda t}\mathfrak{A}^t$ and $A_\lambda = A - \lambda$. Then by Example 3.2.6(i), A_λ is the generator of $\mathfrak{A}_\lambda$. We observe that $\mathfrak{A}_\lambda$ has negative growth bound, i.e., for all $x \in X$, $\mathfrak{A}^t_\lambda x$ tends exponentially to zero as $t \to \infty$. More precisely, for each $\omega_{\mathfrak{A}} < \omega < \Re\lambda$ there is a constant M such that for all $s \ge 0$ (cf. Example 2.3.5),

$$\|\mathrm{e}^{\lambda s}\mathfrak{A}^s\| \le M\mathrm{e}^{-(\Re\lambda-\omega)s}. \tag{3.2.4}$$

Apply Theorem 3.2.1(ii) with $s = 0$ and $\mathfrak{A}$ and A replaced by $\mathfrak{A}_\lambda$ and A_λ to get

$$\mathfrak{A}^t_\lambda x - x = A_\lambda \int_0^t \mathfrak{A}^s_\lambda x\, ds.$$

Since A_λ is closed, we can let $t \to \infty$ to get

$$x = -A_\lambda \int_0^\infty \mathfrak{A}^s_\lambda x\, ds.$$

On the other hand, if $x \in \mathcal{D}(A)$, then we can do the same thing starting from the identity in Theorem 3.2.1(iv) to get

$$x = -\int_0^\infty \mathfrak{A}^s_\lambda A_\lambda x\, ds.$$

This proves that λ belongs to the resolvent set of A and that

$$(\lambda - A)^{-1}x = \int_0^\infty \mathrm{e}^{-\lambda s}\mathfrak{A}^s x\, ds, \quad x \in X. \tag{3.2.5}$$

To get a similar formula for iterates of $(\lambda - A)^{-1}$ we differentiate this formula with respect to λ. By the resolvent identity in Lemma 3.2.8(i) with $h = \beta - \lambda$,

$$\lim_{h\to 0} \frac{1}{h}\left[(\lambda + h - A)^{-1}x - (\lambda - A)^{-1}x\right] = -(\lambda - A)^{-2}x.$$

The corresponding limit of the right hand side of (3.2.5) is

$$\lim_{h\to 0}\int_0^\infty \frac{1}{h}\big(\mathrm{e}^{(\lambda+h)s}-\mathrm{e}^{\lambda s}\big)\mathfrak{A}^s x\,ds = \int_0^\infty \frac{1}{h}\big(\mathrm{e}^{hs}-1\big)\mathrm{e}^{\lambda s}\mathfrak{A}^s x\,ds.$$

As $h\to 0$, $\frac{1}{h}\big(\mathrm{e}^{hs}-1\big)\to s$ uniformly on compact subsets of $\overline{\mathbb{R}}^+$, and

$$\big|\tfrac{1}{h}(\mathrm{e}^{hs}-1)\big| = s\frac{1}{|hs|}\left|\int_0^{hs}\mathrm{e}^y\,dy\right| \le s\frac{1}{|hs|}\int_0^{|hs|}\mathrm{e}^{|y|}\,dy \le s\mathrm{e}^{|hs|}.$$

This combined with (3.2.4) shows that we can use the Lebesgue dominated convergence theorem to move the limit inside the integral to get

$$(\lambda-A)^{-2}x = \int_0^\infty s\mathrm{e}^{-\lambda s}\mathfrak{A}^s x\,ds, \qquad x\in X.$$

The same argument can be repeated. Every time we differentiate the right hand side of (3.2.5) the integrand is multiplied by a factor $-s$ (but we can still use the Lebesgue dominated convergence theorem). Thus, to finish the proof of (i) we need to show that

$$\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}(\lambda-A)^{-1}x = (-1)^n n!(\lambda-A)^{-(n+1)}x. \tag{3.2.6}$$

To do this we use induction over n, the chain rule, and the fact that the formula is true for $n=1$, as we have just seen. We leave this computation to the reader.

(ii) Use part (i), (3.2.4), and the fact that (cf. Lemma 4.2.10)

$$\frac{1}{n!}\int_0^\infty s^n\mathrm{e}^{-(\Re\lambda-\omega)s}\,ds = (\Re\lambda-\omega)^{-(n+1)}, \quad \Re\lambda>\omega, \quad n=0,1,2,\ldots.$$

(iii) We observe that the two claims are equivalent to each other since $(\lambda-A)(\lambda-A)^{-1}x=x$. If $x\in\mathcal{D}(A)$, then we can use part (ii) to get

$$\begin{aligned}|\lambda(\lambda-A)^{-1}x-x| &= |A(\lambda-A)^{-1}x|\\ &= |(\lambda-A)^{-1}Ax|\to 0 \text{ as } \lambda\to\infty.\end{aligned}$$

As $\mathcal{D}(A)$ is dense in X and $\limsup_{\lambda\to+\infty}\|\lambda(\lambda-A)^{-1}\|<\infty$ (this, too, follows from part (ii)), it must then be true that $\lambda(\lambda-A)^{-1}x\to x$ for all $x\in X$.

(iv) By Theorem 3.2.1(iii), for all $y\in\mathcal{D}(A)$,

$$\mathfrak{A}^t(\lambda-A)y = (\lambda-A)\mathfrak{A}^t y.$$

Substituting $y=(\lambda-A)^{-1}x$ and applying $(\lambda-A)^{-1}$ to both sides of this identity we find that

$$(\lambda-A)^{-1}\mathfrak{A}^t x = \mathfrak{A}^t(\lambda-A)^{-1}x$$

for all $x\in X$. □

In this proof we used an estimate on $\left|\frac{1}{h}(e^{hs}-1)\right|$ that will be useful later, too, so let us separate this part of the proof into the following slightly more general lemma:

Lemma 3.2.10 *Let $1 \le p \le \infty$, $\omega \in \mathbb{R}$, and $n = 0, 1, 2, \ldots$.*

(i) *For all $\alpha, \beta \in \mathbb{C}$,*

$$|e^{\alpha}-e^{\beta}| \le |\alpha-\beta| \max\{e^{\Re\alpha}, e^{\Re\beta}\},$$
$$|e^{\alpha}-e^{\beta}-(\alpha-\beta)e^{\beta}| \le \frac{1}{2}|\alpha-\beta|^2 \max\{e^{\Re\alpha}, e^{\Re\beta}\}.$$

(ii) *The function $\alpha \mapsto \big(t \mapsto e^{\alpha t},\ t \in \mathbb{R}^-\big)$ is analytic on the half-plane $\mathbb{C}^+_{\omega}$ in the spaces $L^p_{\omega}(\mathbb{R}^-;\mathbb{C})$, $W^{n,p}_{\omega}(\mathbb{R}^-;\mathbb{C})$, and $BC^n_{0,\omega}(\mathbb{R}^-;\mathbb{C})$ (i.e, it has a complex derivative with respect to α in these spaces when $\Re\alpha > \omega$). Its derivative is the function $t \mapsto te^{\alpha t}$, $t \in \mathbb{R}^-$.*

(iii) *The function $\alpha \mapsto \big(t \mapsto e^{\alpha t},\ t > 0\big)$ is analytic on the half-plane $\mathbb{C}^-_{\omega}$ in the spaces $L^p_{\omega}(\mathbb{R}^+;\mathbb{C})$, $W^{n,p}_{\omega}(\overline{\mathbb{R}}^+;\mathbb{C})$, and $BC^n_{0,\omega}(\overline{\mathbb{R}}^+;\mathbb{C})$. Its derivative is the function $t \mapsto te^{\alpha t}$, $t > 0$.*

Proof (i) Define $f(t) = e^{(\alpha-\beta)t}e^{\beta}$. Then $\dot f(t) = (\alpha-\beta)e^{(\alpha-\beta)t}e^{\beta}$ and

$$\begin{aligned}
|e^{\alpha}-e^{\beta}| = |f(1)-f(0)| &= \left|\int_0^1 \dot f(s)\,ds\right| \\
&\le |\alpha-\beta|e^{\Re\beta}\int_0^1 \left|e^{(\alpha-\beta)s}\right| ds \\
&\le |\alpha-\beta|e^{\Re\beta} \sup_{0\le s\le 1} e^{\Re(\alpha-\beta)s} \\
&= |\alpha-\beta|e^{\Re\beta}\max\{e^{\Re(\alpha-\beta)}, 1\} \\
&= |\alpha-\beta|\max\{e^{\Re\alpha}, e^{\Re\beta)}\}.
\end{aligned}$$

The similar proof of the second inequality is left to the reader. It can be based on the fact that $\ddot f(t) = (\alpha-\beta)^2 e^{(\alpha-\beta)t}e^{\beta}$, and that

$$e^{\alpha}-e^{\beta}-(\alpha-\beta)e^{\beta} = f(1)-f(0)-\dot f(0) = \int_0^1\int_0^s \ddot f(v)\,dv\,ds.$$

(ii) It follows from (i) with α replaced by $(\alpha+h)t$ and β replaced by αt that the function $t \mapsto \frac{1}{h}\big(e^{(\alpha+h)t}-e^{\alpha t}\big) - te^{\alpha t}$ tends to zero as $t \to 0$ (as a complex limit) uniformly for t in each bounded interval. Moreover, combining the growth estimate that (i) gives for this function with the Lebesgue dominated convergence theorem we find that it tends to zero in $L^p_{\omega}(\mathbb{R}^-;\mathbb{C})$. A similar argument shows that all t-derivatives of this function also tend to zero in $L^p_{\omega}(\mathbb{R}^-;\mathbb{C})$, i.e., the function itself tends to zero in $W^{n,p}_\omega(\mathbb{R}^-;\mathbb{C})$. The proof of the analyticity in $BC^n_0(\mathbb{R}^-;\mathbb{C})$ is similar (but slightly simpler).

(iii) This proof is completely analogous to the proof of (ii). □

3.3 The spectra of some generators

To get an example of what the spectrum of a generator can look like, let us determine the spectra of the generators of the shift (semi)groups in Examples 2.3.2 and 3.2.3:

Example 3.3.1 *The generators of the (semi)groups τ^t, τ^t_+, τ^t_-, $\tau^t_{[0,T)}$ and $\tau^t_{\mathbb{T}_T}$ in Examples 2.3.2 and 2.5.3 (see Example 3.2.3) have the following spectra:*

(i) *The spectrum of the generator $\frac{\mathrm{d}}{\mathrm{d}s}$ of the left shift bilateral group τ^t on $L^p_\omega(\mathbb{R};U)$ with $1 \le p < \infty$ or on $BUC_\omega(\mathbb{R};U)$ is equal to the vertical line $\{\Re\lambda = \omega\}$. The whole spectrum is a residual spectrum in the L^1-case, a continuous spectrum in the L^p-case with $1 < p < \infty$, and a point spectrum in the BUC-case.*

(ii) *The spectrum of the generator $\frac{\mathrm{d}}{\mathrm{d}s}_+$ of the incoming left shift semigroup τ^t_+ on $L^p_\omega(\mathbb{R}^+;U)$ with $1 \le p < \infty$ or on $BUC_\omega(\overline{\mathbb{R}}^+;U)$ is equal to the closed half-plane $\overline{\mathbb{C}}^-_\omega$. The open left half-plane $\mathbb{C}^-_\omega$ belongs to the point spectrum, and the boundary $\{\Re\lambda = \omega\}$ belongs to the continuous spectrum in the L^p-case with $1 \le p < \infty$ and to the point spectrum in the BUC-case.*

(iii) *The spectrum of the generator $\frac{\mathrm{d}}{\mathrm{d}s}_-$ of the outgoing left shift semigroup τ^t_- on $L^p_\omega(\mathbb{R}^-;U)$ with $1 \le p < \infty$ or on $\{u \in BUC_\omega(\overline{\mathbb{R}}^-;U) \mid u(0) = 0\}$ is equal to the closed half-plane $\overline{\mathbb{C}}^-_\omega$. The open half-plane $\mathbb{C}^-_\omega$ belongs to the residual spectrum, and the boundary $\{\Re\lambda = \omega\}$ belongs to the residual spectrum in the L^1-case and to the continuous spectrum in the other cases.*

(iv) *The spectrum of the generator $\frac{\mathrm{d}}{\mathrm{d}s}_{[0,T)}$ of the finite left shift semigroup $\tau^t_{[0,T)}$ on $L^p([0,T);U)$ with $1 \le p < \infty$ or on $\{u \in C([0,T];U) \mid u(0) = 0\}$ is empty.*

(v) *The spectrum of the generator $\frac{\mathrm{d}}{\mathrm{d}s}_{\mathbb{T}_T}$ of the circular left shift group $\tau^t_{\mathbb{T}_T}$ on $L^p(\mathbb{T}_T;U)$ with $1 \le p < \infty$ or on $C(\mathbb{T}_T;U)$ is a pure point spectrum located at $\{2\pi jm/T \mid m = 0, \pm 1, \pm 2, \ldots\}$.*

Proof For simplicity we take $\omega = 0$. The general case can either be reduced to the case $\omega = 0$ with the help of Lemma 2.5.2(ii), or it can be proved directly by a slight modification of the argument below.

(i) As τ^t is a group, both τ^t and τ^{-t} are semigroups, and $\|\tau^t\| = 1$ for all $t \in \mathbb{R}$. It follows from Theorem 3.2.9(i) that every $\lambda \notin j\mathbb{R}$ belongs to the resolvent set of $\frac{\mathrm{d}}{\mathrm{d}s}$. It remains to show that $j\mathbb{R}$ belongs to the residual spectrum in the L^1-case, to the continuous spectrum in the L^p-case with $1 < p < \infty$, and to the point spectrum in the BUC-case.

Set $\lambda = j\beta$ where $\beta \in \mathbb{R}$, and let $u \in W^{1,p}(\mathbb{R}; U) = \mathcal{D}\left(\frac{\mathrm{d}}{\mathrm{d}s}\right)$. If $j\beta u - \dot{u} = f$ for some $u \in W^{1,p}(\mathbb{R}; U)$ and $f \in L^p(\mathbb{R}; U)$ then, by the variation of constants formula, for all $T \in \mathbb{R}$,

$$u(t) = \mathrm{e}^{j\beta(t-T)}u(T) - \int_T^t \mathrm{e}^{j\beta(t-s)} f(s)\, ds.$$

By letting $T \to -\infty$ we get (see Lemma 3.2.4)

$$u(t) = -\lim_{T \to -\infty} \int_T^t \mathrm{e}^{j\beta(t-s)} f(s)\, ds.$$

In particular, if $f = 0$ then $u = 0$, i.e., $j\beta - \frac{\mathrm{d}}{\mathrm{d}s}$ is injective. By letting $t \to +\infty$ we find that

$$\lim_{t \to \infty} \lim_{T \to -\infty} \int_T^t \mathrm{e}^{-j\beta s} f(s)\, ds = 0.$$

If $p = 1$, then this implies that the range of $j\beta - \frac{\mathrm{d}}{\mathrm{d}s}$ is not dense, hence $j\beta \in \sigma_r(\frac{\mathrm{d}}{\mathrm{d}s})$. If $p > 1$ then it is not true for every $f \in L^p(\mathbb{R}; U)$ that the limits above exist, so the range of $j\beta - \frac{\mathrm{d}}{\mathrm{d}s}$ is not equal to $L^p(\mathbb{R}; U)$, i.e., $j\omega \in \sigma(\frac{\mathrm{d}}{\mathrm{d}s})$. On the other hand, if $f \in C_c^\infty(\mathbb{R}; U)$ with $\int_{-\infty}^\infty \mathrm{e}^{-j\beta s} f(s)\, ds = 0$, and if we define u to be the integral above, then $u \in C_c^\infty(\mathbb{R}; U) \subset W^{1,p}(\mathbb{R}; U)$ and $j\beta y - \dot{u} = f$. The set of functions f of this type is dense in $L^p(\mathbb{R}; U)$ when $1 < p < \infty$. Thus $j\beta - \frac{\mathrm{d}}{\mathrm{d}s}$ has dense range if $p > 1$, and in this case $j\beta \in \sigma_c(\frac{\mathrm{d}}{\mathrm{d}s})$.

In the BUC-case the function $\mathrm{e}_{j\beta}(t) = \mathrm{e}^{j\beta t}$ is an eigenfunction, i.e., $(j\beta - \frac{\mathrm{d}}{\mathrm{d}s})\mathrm{e}_{j\beta} = 0$; hence $j\beta \in \sigma_p(\frac{\mathrm{d}}{\mathrm{d}s})$.[2]

(ii) That $\mathbb{C}^+ \subset \rho(\frac{\mathrm{d}}{\mathrm{d}s}_+)$ follows from Theorem 3.2.9(i). If $\Re\lambda < 0$ then $\lambda \in \sigma_p(\frac{\mathrm{d}}{\mathrm{d}s}_+)$, because then the function $u = \mathrm{e}^{\lambda t}$ belongs to $W^{1,p}(\overline{\mathbb{R}}^+; U)$ and $\lambda u - \dot{u} = 0$. The proof that the imaginary axis belongs either to the singular spectrum in the L^p-case or to the point spectrum in the BUC-case is quite similar to the one above, and it is left to the reader (in the L^p-case, let $T \to +\infty$ to get $u(t) = \int_t^\infty \mathrm{e}^{j\beta(t-s)} f(s)\, ds$, and see also the footnote about the case $p = 1$).

[2] It is easy to show that the range of $j\beta - \frac{\mathrm{d}}{\mathrm{d}s}$ is not closed in the L^1-case and BUC-case either. For example, in the L^1-case the range is dense in $\{f \in L^1(\mathbb{R}; U) \mid \int_{\mathbb{R}} f(s)\, ds = 0\}$, but it is not true for every $f \in L^1(\mathbb{R}; U)$ with $\int_{\mathbb{R}} f(s)\, ds = 0$ that the function $u(t) = -\int_{-\infty}^t \mathrm{e}^{j\beta(t-s)} f(s)\, ds$ belongs to $L^1(\mathbb{R}; U)$.

(iii) That $\mathbb{C}^+ \subset \lambda \in \rho(\frac{d}{ds}_-)$ follows from Theorem 3.2.9(i). If $\Re\lambda < 0$ or if $\Re\lambda \le 0$ and $p = 1$, then every $f \in \mathcal{R}\left(\lambda - \frac{d}{ds}_-\right)$ satisfies

$$\int_{-\infty}^{0} e^{-\lambda s} f(s)\, ds = 0,$$

hence the range is not dense in this case. We leave the proof of the claim that $\sigma_c = \{\lambda \in \mathbb{C} \mid \Re\lambda = 0\}$ in the other cases to the reader (see the proof of (i)).

(iv) This follows from Theorem 3.2.9(i), since the growth bound of $\tau_{[0,T)}$ is $-\infty$.

(v) For each $m \in \mathbb{Z}$, the derivative of the T-periodic function $e^{2\pi jmt/T}$ with respect to t is $(2\pi jm/T)e^{2\pi jmt/T}$, hence $2\pi jm/T$ is an eigenvalue of $\frac{d}{ds}_{\mathbb{T}_T}$ with eigenfunction $e^{2\pi jmt/T}$.

To complete the proof of (v) we have to show that the remaining points λ in the complex plane belong to the resolvent set of $\frac{d}{ds}_{\mathbb{T}_T}$. To do this we have to solve the equation $\lambda u - \dot{u} = f$, where, for example, $f \in L^p(\mathbb{T}_T; U)$. By the variation of constants formula, a solution of this equation must satisfy

$$u(s) = e^{\lambda(s-t)} u(t) - \int_t^s e^{\lambda(s-v)} f(v)\, dv, \qquad s,\ t \in \mathbb{R}.$$

Taking $s = t + T$, and requiring that $u(t+T) = u(t)$ (in order to ensure T-periodicity of u) we get

$$(1 - e^{\lambda T}) u(t) = -\int_t^{t+T} e^{\lambda(t+T-v)} f(v)\, dv.$$

The factor on the left hand side is invertible iff λ does not coincide with any of the points $2\pi jm/T$, in which case we get the following formula for the unique T-periodic solution u of $\lambda u - \dot{u} = f$:

$$\begin{aligned} u(t) &= (1 - e^{-\lambda T})^{-1} \int_t^{t+T} e^{\lambda(t-v)} f(v)\, dv \\ &= (1 - e^{-\lambda T})^{-1} \int_0^T e^{-\lambda s} f(t+s)\, ds. \end{aligned}$$

The right-hand side of this formula maps $L^p(\mathbb{T}_T; U)$ into $W^{1,p}(\mathbb{T}_T; U)$ and $C(\mathbb{T}_T; U)$ into $C^1(\mathbb{T}_T; U)$, and by differentiating this formula we find that, indeed, $\lambda u - \dot{u} = f$. □

Example 3.3.2 *The resolvents of the generators* $\frac{d}{ds}$, $\frac{d}{ds}_+$, $\frac{d}{ds}_-$, $\frac{d}{ds}_{[0,T)}$, *and* $\frac{d}{ds}_{\mathbb{T}_T}$ *in Example 3.2.3 can be described as follows:*

(i) *The resolvent* $(\lambda - \frac{d}{ds})^{-1}$ *of the generator of the bilateral left shift group* τ^t *on* $L^p_\omega(\mathbb{R}; U)$ *and on* $BUC_\omega(\mathbb{R}; U)$ *maps* f *into* $t \mapsto \int_t^\infty e^{\lambda(t-s)} f(s)\, ds$,

$t \in \mathbb{R}$, if $\Re\lambda > \omega$, and it maps f into $t \mapsto -\int_{-\infty}^{t} \mathrm{e}^{\lambda(t-s)} f(s)\,ds$, $t \in \mathbb{R}$, if $\Re\lambda < \omega$.

(ii) *For each $\lambda \in \mathbb{C}_\omega^+$ the resolvent $(\lambda - \frac{\mathrm{d}}{\mathrm{d}s}_+)^{-1}$ of the generator of the incoming left shift semigroup τ_+^t on $L_\omega^p(\mathbb{R}^+;U)$ and on $BUC_\omega(\overline{\mathbb{R}}^+;U)$ maps f into $t \mapsto \int_t^\infty \mathrm{e}^{\lambda(t-s)} f(s)\,ds$, $t \geq 0$.*

(iii) *For each $\lambda \in \mathbb{C}_\omega^+$ the resolvent $(\lambda - \frac{\mathrm{d}}{\mathrm{d}s}_-)^{-1}$ of the generator of the outgoing left shift semigroup τ_-^t on $L_\omega^p(\mathbb{R}^-;U)$ and on $\{u \in BUC_\omega(\overline{\mathbb{R}}^-;U) \mid u(0) = 0\}$ maps f into $t \mapsto \int_t^0 \mathrm{e}^{\lambda(t-s)} f(s)\,ds$, $t \in \overline{\mathbb{R}}^-$.*

(iv) *For each $\lambda \in \mathbb{C}$ the resolvent $(\lambda - \frac{\mathrm{d}}{\mathrm{d}s}_{[0,T)})^{-1}$ of the generator of the finite left shift semigroup $\tau_{[0,T)}^t$ on $L^p([0,T);U)$ and on $\{u \in C([0,T];U) \mid u(T) = 0\}$ maps f into $t \mapsto \int_t^T \mathrm{e}^{\lambda(t-s)} f(s)\,ds$, $t \in [0,T)$.*

(v) *For each $\lambda \in \mathbb{C}$ which is not one of the points $\{2\pi jm/T \mid m = 0, \pm 1, \pm 2, \ldots\}$ the resolvent $(\lambda - \frac{\mathrm{d}}{\mathrm{d}s}_{\mathbb{T}_T})^{-1}$ of the generator of the circular left shift group $\tau_{\mathbb{T}_T}^t$ on $L^p(\mathbb{T}_T;U)$ and on $C(\mathbb{T}_T;U)$ maps f into $t \mapsto (1 - \mathrm{e}^{-\lambda T})^{-1} \int_t^{t+T} \mathrm{e}^{\lambda(t-s)} f(s)\,ds$.*

The proof of this is essentially contained in the proof of Example 3.3.1.

The shift (semi)group examples that we have seen so far have rather exceptional spectra. They play an important role in our theory, but in typical applications one more frequently encounters semigroups of the following type:

Example 3.3.3 *Let $\{\phi_n\}_{n=1}^\infty$ be an orthonormal basis in a separable Hilbert space X, and let $\{\lambda_n\}_{n=1}^\infty$ be a sequence of complex numbers. Then the sum*

$$\mathfrak{A}^t x = \sum_{n=1}^{\infty} \mathrm{e}^{\lambda_n t} \langle x, \phi_n \rangle \phi_n, \quad x \in X, \quad t \geq 0,$$

converges for each $x \in X$ and $t \geq 0$ and defines a C_0 semigroup if and only if

$$\omega_{\mathfrak{A}} = \sup_{n \geq 0} \Re\lambda_n < \infty.$$

The growth bound of this semigroup is $\omega_{\mathfrak{A}}$, and

$$\|\mathfrak{A}^t\| = \mathrm{e}^{\omega_{\mathfrak{A}} t}, \quad t \geq 0.$$

It is a group if and only if

$$\alpha_{\mathfrak{A}} = \inf_{n \geq 0} \Re\lambda_n > -\infty.$$

in which case

$$\|\mathfrak{A}^t\| = \mathrm{e}^{\alpha_{\mathfrak{A}} t}, \quad t \leq 0.$$

In particular, if $\Re\lambda_n = \omega$ *for all* n*, then* $\mathfrak{A}^t$ *is a group, and*

$$\|\mathfrak{A}^t\| = e^{\omega t}, \quad t \in \mathbb{R}.$$

Proof Clearly, the sum converges always if we choose $x = \phi_n$, in which case $\mathfrak{A}^t\phi_n = e^{\lambda_n t}\phi_n$, and $|\mathfrak{A}^t\phi_n| = e^{\Re\lambda_n t}$. If $\mathfrak{A}^t$ is to be a semigroup, then $\|\mathfrak{A}^t\| \geq e^{\Re\lambda_n t}$ for all n, and by Theorem 2.5.4(i), the number $\omega_{\mathfrak{A}}$ defined above must be finite and less than or equal to the growth bound of $\mathfrak{A}$. If $\mathfrak{A}^t$ is to be a group, then $\mathfrak{A}^{-t}$ is also a semigroup, and the same argument with t replaced by $-t$ shows that necessarily $\alpha_{\mathfrak{A}} > -\infty$ in this case.

Let us suppose that $\omega_{\mathfrak{A}} < \infty$. For each $N = 1, 2, 3, \ldots$, define

$$\mathfrak{A}_N^t x = \sum_{n=1}^{N} e^{\lambda_n t}\langle x, \phi_n\rangle\phi_n.$$

Then it is easy to show that each $\mathfrak{A}_N$ is a C_0 group (since $\phi_n \perp \phi_k$ when $n \neq k$). For each $t > 0$, the sum converges as $N \to \infty$ because the norm of the tail of the series tends to zero (the sequence ϕ_n is orthonormal):

$$\begin{aligned}\left|\sum_{n=N+1}^{\infty} e^{\lambda_n t}\langle x, \phi_n\rangle\phi_n\right|^2 &= \sum_{n=N+1}^{\infty} |e^{\lambda_n t}\langle x, \phi_n\rangle\phi_n|^2 \\ &= \sum_{n=N+1}^{\infty} e^{2\Re\lambda_n t}|\langle x, \phi_n\rangle|^2 \\ &\leq e^{2\omega_{\mathfrak{A}} t}\sum_{n=N+1}^{\infty} |\langle x, \phi_n\rangle|^2 \\ &\to 0 \text{ as } N \to \infty.\end{aligned}$$

Thus $\mathfrak{A}^t \in \mathcal{B}(X)$ (as a strong limit of operators in $\mathcal{B}(X)$). The norm estimate $\|\mathfrak{A}^t\| \leq e^{\omega_{\mathfrak{A}} t}$ follows from the fact that (see the computation above)

$$|\mathfrak{A}^t x|^2 = \left|\sum_{n=1}^{\infty} e^{\lambda_n t}\langle x, \phi_n\rangle\phi_n\right|^2 \leq e^{2\omega_{\mathfrak{A}} t}|x|^2.$$

Moreover, for each x the convergence is uniform in t over bounded intervals since

$$|\mathfrak{A}^t x - \mathfrak{A}_N^t x|^2 \leq e^{2\omega_{\mathfrak{A}} t}\sum_{n=N+1}^{\infty} |\langle x, \phi_n\rangle|^2.$$

This implies that $t \mapsto \mathfrak{A}^t x$ is continuous on $\overline{\mathbb{R}}^+$ for each $x \in X$. Since each $\mathfrak{A}_N$ satisfies $\mathfrak{A}_N^0 = 1$ and $\mathfrak{A}_N^{s+t} = \mathfrak{A}_N^s\mathfrak{A}_N^t$, $s, t \geq 0$, the same identities carry over to the limit. We conclude that $\mathfrak{A}$ is a C_0 semigroup.

If $\alpha_{\mathfrak{A}} > -\infty$, then we can repeat the same argument to get convergence also for $t < 0$. □

Definition 3.3.4 A (semi)group of the type described in Example 3.3.3 is called *diagonal*, with *eigenvectors* $\{\phi_n\}_{n=1}^{\infty}$ and *eigenvalues* $\{\lambda_n\}_{n=1}^{\infty}$.

The reason for this terminology is the following:

Example 3.3.5 *The generator A of the (semi)group in Example 3.3.3 is the operator*

$$Ax = \sum_{n=1}^{\infty} \lambda_n \langle x, \phi_n \rangle \phi_n,$$

with domain

$$\mathcal{D}(A) = \left\{ x \in X \;\middle|\; \sum_{n=1}^{\infty} (1 + |\lambda_n|^2) |\langle x, \phi_n \rangle|^2 < \infty \right\}.$$

The spectrum of A is the closure of the set $\{\lambda_n \mid n = 1, 2, 3, \ldots\}$*: every* λ_n *belongs to the point spectrum and cluster points different from all the* λ_n *belong to the continuous spectrum. The resolvent operator is given by*

$$(\alpha - A)^{-1} x = \sum_{n=1}^{\infty} (\alpha - \lambda_n)^{-1} \langle x, \phi_n \rangle \phi_n.$$

Proof Suppose that $\lim_{h \downarrow 0} \frac{1}{h}(\mathfrak{A}^h x - x)$ exists. Taking the inner product with ϕ_n, $n = 1, 2, 3, \ldots$, we get

$$\begin{aligned} \lim_{h \downarrow 0} \frac{1}{h} \langle (\mathfrak{A}^h x - x), \phi_n \rangle &= \lim_{h \downarrow 0} \frac{1}{h} \sum_{k=1}^{\infty} (e^{\lambda_k h} - 1) \langle x, \phi_k \rangle \langle \phi_k, \phi_n \rangle \\ &= \lim_{h \downarrow 0} \frac{1}{h} (e^{\lambda_n h} - 1) \langle x, \phi_n \rangle \\ &= \lambda_n \langle x, \phi_n \rangle. \end{aligned}$$

Thus, for all $x \in \mathcal{D}(A)$, we have $Ax = \sum_{n=1}^{\infty} \lambda_n \langle x, \phi_n \rangle \phi_n$. The norm of this vector is finite as is the norm of $x = \sum_{n=1}^{\infty} \langle x, \phi_n \rangle \phi_n$, so we conclude that

$$\mathcal{D}(A) \subset \left\{ x \in X \;\middle|\; \sum_{n=1}^{\infty} (1 + |\lambda_n|^2) |\langle x, \phi_n \rangle|^2 < \infty \right\}.$$

To prove the opposite inclusion, let us suppose that

$$\sum_{n=1}^{\infty} (1 + |\lambda_n|^2) |\langle x, \phi_n \rangle|^2 < \infty.$$

Then the sum $y = \sum_{n=1}^{\infty} \lambda_n \langle x, \phi_n \rangle \phi_n$ converges in X, and for each $h > 0$ we have

$$\Big|\frac{1}{h}(\mathfrak{A}^h x - x) - y\Big|^2 = \Big|\sum_{n=1}^{\infty}\Big(\frac{1}{h}(e^{\lambda_n h} - 1) - \lambda_n\Big)\langle x, \phi_n\rangle\phi_n\Big|^2$$
$$= \sum_{n=1}^{\infty}\Big|\frac{1}{h}(e^{\lambda_n h} - 1) - \lambda_n\Big|^2 |\langle x, \phi_n\rangle|^2.$$

Take $h \le 1$. Then, by Lemma 3.2.10, $\frac{1}{h}|e^{\lambda_n h} - 1| \le |\lambda_n| M_n(h)$, where

$$M_n(h) = \max\{1, e^{\Re\lambda_n h}\} \le M = \max\{1, e^{\omega_{\mathfrak{A}}}\},$$

hence

$$\Big|\frac{1}{h}(e^{\lambda_n h} - 1) - \lambda_n\Big|^2 |\langle x, \phi_n\rangle|^2 \le (1 + M)^2 |\lambda_n|^2 |\langle x, \phi_n\rangle|^2.$$

Moreover, for each n, $\big(\frac{1}{h}(e^{\lambda_n h} - 1) - \lambda_n\big) \to 0$ as $h \downarrow 0$. This means that we can use the discrete Lebesgue dominated convergence theorem to conclude $\frac{1}{h}(\mathfrak{A}^h x - x) - y \to 0$ as $h \downarrow 0$, i.e., $x \in \mathcal{D}(A)$.

Obviously every λ_n is an eigenvalue since $\phi_n \in \mathcal{D}(A)$ and $A\phi_n = \lambda_n \phi_n$. The spectrum of A contains therefore at least the closure of $\{\lambda_n \mid n = 1, 2, 3, \dots\}$ (the spectrum is always closed).

If $\inf_{n\ge 0}|\alpha - \lambda_n| > 0$, then there exist two positive constants a and b such that $a(1 + |\lambda_n|) \le |\alpha - \lambda_n| \le b(1 + |\lambda_n|)$, and the sum

$$Bx = \sum_{n=1}^{\infty} (\alpha - \lambda_n)^{-1} \langle x, \phi_n\rangle \phi_n$$

converges for every $x \in X$ and defines an operator $B \in \mathcal{B}(X)$ which maps X onto $\mathcal{D}(A)$. It is easy to show that B is the inverse to $(\alpha - A)$, hence $\alpha \in \rho(A)$. Conversely, suppose that $(\alpha - A)$ has an inverse $(\alpha - A)^{-1}$. Then

$$1 = |\phi_n| = |(\alpha - A)^{-1}(\alpha - A)\phi_n|$$
$$= |(\alpha - A)^{-1}(\alpha - \lambda_n)\phi_n| \le \|(\alpha - A)^{-1}\||\alpha - \lambda_n|.$$

This shows that every $\alpha \in \rho(A)$ satisfies $\inf_{n\ge 0}|\alpha - \lambda_n| > 0$.

If $\inf_{n\ge 0}|\alpha - \lambda_n| = 0$ but $\alpha \ne \lambda_n$ for all n, then α is not an eigenvalue because $\alpha x - Ax = \sum_{n=1}^{\infty}(\alpha - \lambda_n)\langle x, \phi_n\rangle\phi_n = 0$ only when $\langle x, \phi_n\rangle = 0$ for all n, i.e., $x = 0$. On the other hand, the range of $\alpha - A$ is dense because it contains all finite linear combinations of the base vectors ϕ_n. Thus $\alpha \in \sigma_c(A)$ in this case. □

Example 3.3.3 can be generalized to the case where A is an arbitrary *normal* operator on a Hilbert space X, whose spectrum is contained in some left half-plane. The proofs remain essentially the same, except for the fact that the sums have to be replaced by integrals over a spectral resolution. We refer the reader to Rudin [1973, pp. 301–303] for a precise description of the spectral resolution used in the following theorem (dual operators and semigroups are discussed in Section 3.5).

Example 3.3.6 *Let A be a closed and densely defined normal operator on a Hilbert space X (i.e., $A^*A = AA^*$), and let E be the corresponding spectral resolution of A, so that*

$$\langle Ax, y\rangle_X = \int_{\sigma(A)} \lambda \langle E(d\lambda)x, y\rangle, \quad x \in \mathcal{D}(A), \quad y \in X.$$

Then the following claims are valid.

(i) *For each $n = 1, 2, 3, \ldots$ the domain of A^n is given be*

$$\mathcal{D}\left(A^n\right) = \left\{x \in X \;\middle|\; \int_{\sigma(A)} (1 + |\lambda|^2)^n \langle E(d\lambda)x, x\rangle\right\} < \infty,$$

and

$$\|A^n x\|_X^2 = \left\{x \in X \;\middle|\; \int_{\sigma(A)} |\lambda|^{2n} \langle E(d\lambda)x, x\rangle\right\}.$$

(ii) *For each $\alpha \in \rho(A)$, $0 \le k \le n \in \{1, 2, 3, \ldots\}$, and $x, y \in X$,*

$$\langle A^k(\alpha - A)^{-n}x, y\rangle_X = \int_{\sigma(A)} \lambda^k(\alpha - \lambda)^{-n}\langle E(d\lambda)x, y\rangle.$$

(iii) *A generates a C_0 semigroup $\mathfrak{A}$ on X if and only if the spectrum of A is contained in some left half-plane, i.e.,*

$$\omega_{\mathfrak{A}} = \sup_{\lambda \in \sigma(A)} \Re\lambda < \infty.$$

In this case,

$$\|\mathfrak{A}^t\| = e^{\omega_{\mathfrak{A}} t}, \quad t \ge 0,$$

and

$$\langle \mathfrak{A}^t x, y\rangle = \int_{\sigma(A)} e^{\lambda t}\langle E(d\lambda)x, y\rangle, \quad t \ge 0, \quad x \in X, \quad y \in X. \quad (3.3.1)$$

(iv) *A generates a C_0 group $\mathfrak{A}$ on X if and only if $\sigma(A)$ is contained in some vertical strip $\alpha \le \Re\lambda \le \omega$. In this case, if we define*

$$\alpha_{\mathfrak{A}} = \inf\{\Re\lambda \mid \lambda \in \sigma(A)\},$$

then

$$\|\mathfrak{A}^t\| = e^{\alpha_{\mathfrak{A}} t}, \quad t \le 0,$$

and (3.3.1) *holds for all $t \in \mathbb{R}$.*

(v) *A C_0 semigroup $\mathfrak{A}$ on X is normal (i.e., $\mathfrak{A}^{*t} = \mathfrak{A}^t$ for all $t \geq 0$) if and only if its generator is normal.*

Proof (i)–(ii) See Rudin (1973, Theorems 12.21, 13.24 and 13.33).

(iii) See Rudin (1973, Theorem 13.37) and the remark following that theorem.

(iv) The proof of this in analogous to the proof of (iii).

(v) See Rudin (1973, Theorem 13.37).

□

Most of the examples of semigroups that we will encounter in this book are either of the type described in Example 2.3.2, 3.3.3, or 3.3.6, or a transformation of these examples of the types listed in Examples 2.3.10–2.3.13.

3.4 Which operators are generators?

There is a celebrated converse to Theorem 3.2.9(i) that gives a complete characterization of the class of operators A that generate C_0 semigroups:

Theorem 3.4.1 (Hille–Yosida) *A linear operator A is the generator of a C_0 semigroup $\mathfrak{A}$ satisfying $\|\mathfrak{A}^t\| \leq M\mathrm{e}^{\omega t}$ if and only if the following conditions hold:*

(i) *$\mathcal{D}(A)$ is dense in X;*
(ii) *every real $\lambda > \omega$ belongs to the resolvent set of A, and*

$$\left\|(\lambda - A)^{-n}\right\| \leq \frac{M}{(\lambda - \omega)^n} \text{ for } \lambda > \omega \text{ and } n = 1, 2, 3, \ldots.$$

Alternatively, condition (ii) can be replaced by

(ii′) *every real $\lambda > \omega$ belongs to the resolvent set of A, and*

$$\left\|\frac{\partial^n}{\partial \lambda^n}(\lambda - A)^{-1}\right\| \leq \frac{Mn!}{(\lambda - \omega)^{n+1}} \text{ for } \lambda > \omega \text{ and } n = 0, 1, 2, \ldots.$$

Note that the assumption implies that A must be closed, since it has a nonempty resolvent set.

Proof The necessity of (i) and (ii) follows from Theorems 3.2.1(vi) and 3.2.9 (i)–(ii) (the exact estimate in Theorem 3.2.9(ii) was derived from (3.2.4), which is equivalent to $\|\mathfrak{A}^t\| \leq M\mathrm{e}^{\omega t}$). The equivalence of (ii) and (ii′) is a consequence of (3.2.6).

Let us start the proof of the converse claim by observing that the conclusion of Theorem 3.2.9(iii) remains valid, since the proof used only (ii) with $n = 1$

and the density of $\mathcal{D}(A)$ in X. This means that if we define

$$A_\alpha = \alpha A(\alpha - A)^{-1} = \alpha^2(\alpha - A)^{-1} - \alpha,$$

then each $A_\alpha \in \mathcal{B}(X)$, and, for each $x \in \mathcal{D}(A)$, $A_\alpha x \to Ax$ in X as $\alpha \to \infty$. Since A_α is bounded, we can define $\mathfrak{A}^t_\alpha = e^{A_\alpha t}$ as in Example 3.1.2. We claim that for each $x \in X$, the limit $\mathfrak{A}^t x = \lim_{\alpha\to\infty} \mathfrak{A}^t_\alpha x$ exists, uniformly in t on any bounded interval, and that $\mathfrak{A}^t$ is a semigroup with generator A.

Define

$$B_\alpha = A_\alpha + \alpha = \alpha^2(\alpha - A)^{-1}.$$

Then, by (ii), for all $n = 1, 2, 3, \dots,$

$$\|B^n_\alpha\| \le \frac{M\alpha^{2n}}{(\alpha - \omega)^n}, \tag{3.4.1}$$

and by Theorem 3.2.9(iii) and Example 3.2.6(i),

$$\mathfrak{A}^t_\alpha = e^{-\alpha t} e^{B_\alpha t} = e^{-\alpha t} \sum_{n=0}^{\infty} \frac{B^n_\alpha t^n}{n!}. \tag{3.4.2}$$

Therefore

$$\begin{aligned} \|\mathfrak{A}^t_\alpha\| &\le e^{-\alpha t} \sum_{n=0}^{\infty} \frac{t^n}{n!} \frac{M\alpha^{2n}}{(\alpha - \omega)^n} \\ &= M e^{-\alpha t} e^{(\alpha^2 t)/(\alpha-\omega)} = M e^{(\alpha\omega t)/(\alpha-\omega)}, \qquad t \ge 0. \end{aligned} \tag{3.4.3}$$

This tends to $Me^{\omega t}$ as $\alpha \to \infty$, and the convergence is uniform in t on any bounded interval. Since $(\alpha - A)^{-1}$ and $(\beta - A)^{-1}$ commute (see Lemma 3.2.8(i)), also A_α and A_β commute, i.e., $A_\alpha A_\beta = A_\beta A_\alpha$, and this implies that $\mathfrak{A}^t_\alpha A_\beta = A_\beta \mathfrak{A}^t_\alpha$ for all $\alpha, \beta > \omega$ and $t \in \mathbb{R}$. Thus, for all $x \in X$ and $t \in \mathbb{R}$,

$$\begin{aligned} \mathfrak{A}^t_\alpha x - \mathfrak{A}^t_\beta x &= \int_0^t \frac{d}{ds}\left[\mathfrak{A}^s_\alpha \mathfrak{A}^{t-s}_\beta\right] ds x \\ &= \int_0^t \mathfrak{A}^s_\alpha (A_\alpha - A_\beta)\mathfrak{A}^{t-s}_\beta x\, ds = \int_0^t \mathfrak{A}^s_\alpha \mathfrak{A}^{t-s}_\beta (A_\alpha - A_\beta)x\, ds, \end{aligned}$$

and

$$\begin{aligned} &|\mathfrak{A}^t_\alpha x - \mathfrak{A}^t_\beta x| \\ &\quad \le M^2 \int_0^t e^{(\alpha\omega s)/(\alpha-\omega)} e^{(\beta\omega(t-s))/(\beta-\omega)} |A_\alpha x - A_\beta x|\, ds. \end{aligned} \tag{3.4.4}$$

Let $\alpha, \beta \to \infty$. Then the products of the exponentials tend to $e^{\omega s}e^{\omega(t-s)} = e^{\omega t}$, uniformly in s and t on any bounded interval, and if $x \in \mathcal{D}(A)$, then

$|A_\alpha x - A_\beta x| \to 0$ since both $A_\alpha x \to x$ and $A_\beta x \to x$. Therefore,

$$\lim_{\alpha,\,\beta\to\infty} |\mathfrak{A}^t_\alpha x - \mathfrak{A}^t_\beta x| = 0, \qquad x \in \mathcal{D}(A),$$

uniformly in t on any bounded interval. In other words, $\alpha \mapsto \mathfrak{A}^t_\alpha x$ is a Cauchy family in $C(\overline{\mathbb{R}}^+; X)$, and it has a limit in $C(\overline{\mathbb{R}}^+; X)$. Since we have a uniform bound on the norm of $\mathfrak{A}^t_\alpha x$ for t in each bounded interval (see (3.4.3)) and $\mathcal{D}(A)$ is dense in X, the limit $\lim_{\alpha\to\infty} \mathfrak{A}^t_\alpha x$ must exist in $C(\overline{\mathbb{R}}^+; X)$ for all $x \in X$, uniformly in t on any bounded interval. Let us denote the limit by $\mathfrak{A}^t x$. For each $t \geq 0$ we have $\mathfrak{A}^t \in \mathcal{B}(X)$ (the strong limit of a family operators in $\mathcal{B}(X)$ belongs to $\mathcal{B}(X)$). By construction $t \mapsto \mathfrak{A}^t x$ is continuous, i.e., $t \mapsto \mathfrak{A}^t$ is strongly continuous. Moreover, $\mathfrak{A}^t$ inherits the semigroup properties $\mathfrak{A}^0 = 1$ and $\mathfrak{A}^{s+t} = \mathfrak{A}^s\mathfrak{A}^t$ from $\mathfrak{A}^t_\alpha$, and it also inherits the bound $\|\mathfrak{A}^t\| \leq M\mathrm{e}^{\omega t}$. We conclude that $\mathfrak{A}^t$ is a C_0 semigroup.

The only thing left to be shown is that the generator of $\mathfrak{A}$ is A. Let $x \in \mathcal{D}(A)$. Then by Theorem 3.2.1(iv)

$$\mathfrak{A}^t x - x = \lim_{\alpha\to\infty} (\mathfrak{A}^t_\alpha x - x) = \lim_{\alpha\to\infty} \int_0^t \mathfrak{A}^s_\alpha A_\alpha x\, ds = \int_0^t \mathfrak{A}^s A x\, ds$$

(the integrand converges uniformly on $[0, t]$ to $\mathfrak{A}^s Ax$). Divide this by t and let $t \downarrow 0$. This shows that, if we (temporarily) denote the generator of $\mathfrak{A}$ by B, then $\mathcal{D}(A) \subset \mathcal{D}(B)$, and $Bx = Ax$ for all $x \in \mathcal{D}(A)$. In other words, B is an extension of A. But this extension cannot be nontrivial, because if we take some common point α in the resolvent sets of A and B (any $\alpha > \omega$ will do), then

$$X = (\alpha - A)\mathcal{D}(A) = (\alpha - B)\mathcal{D}(A)$$

which implies that

$$\mathcal{D}(B) = (\alpha - B)^{-1}X = (\alpha - A)^{-1}X = \mathcal{D}(A).$$

□

Corollary 3.4.2 *A linear operator A is the generator of a C_0 semigroup $\mathfrak{A}$ satisfying $\|\mathfrak{A}^t\| \leq \mathrm{e}^{\omega t}$ if and only if the following conditions hold:*

(i) *$\mathcal{D}(A)$ is dense in X;*
(ii) *Every real $\lambda > \omega$ belongs to the resolvent set of A, and*

$$\left\|(\lambda - A)^{-1}\right\| \leq \frac{1}{(\lambda - \omega)} \text{ for } \lambda > \omega.$$

Proof This follows from Theorem 3.4.1 since $\|(\alpha - A)^{-n}\| \leq \|(\alpha - A)^{-1}\|^n$. □

The case $\omega = 0$ is of special interest:

Definition 3.4.3 By a *bounded semigroup or group* we mean a semigroup or group $\mathfrak{A}$ satisfying $\sup_{t\geq 0}\|\mathfrak{A}^t\| < \infty$ or $\sup_{t\in\mathbb{R}}\|\mathfrak{A}^t\| < \infty$, respectively. By a *contraction semigroup of group* we mean a semigroup or group $\mathfrak{A}$ satisfying $\|\mathfrak{A}^t\| \leq 1$ for all $t \geq 0$ or $t \in \mathbb{R}$, respectively.

Corollary 3.4.4 *Let A be a linear operator $X \supset \mathcal{D}(A) \to X$ with dense domain and let $M < \infty$. Then the following conditions are equivalent:*

(i) *A is the generator of a (bounded) C_0 semigroup $\mathfrak{A}$ satisfying $\|\mathfrak{A}^t\| \leq M$ for all $t \geq 0$;*
(ii) *every positive real λ belongs to the resolvent set of A and*

$$\left\|(\lambda - A)^{-n}\right\| \leq M\lambda^{-n} \text{ for } \lambda > 0 \text{ and } n = 1, 2, 3, \ldots;$$

(iii) *the right half-plane $\mathbb{C}^+$ belongs to the resolvent set of A and*

$$\left\|(\lambda - A)^{-n}\right\| \leq (\Re\lambda)^{-n} \text{ for } \Re\lambda > 0 \text{ and } n = 1, 2, 3, \ldots$$

Proof By Theorem 3.4.1, (i) ⇔ (ii). Obviously (iii) ⇒ (ii). To show that (i) ⇒ (iii) we split $\lambda \in \mathbb{C}$ into $\lambda = \alpha + j\beta$ and apply Theorem 3.4.1 with λ replaced by α, $\mathfrak{A}^t$ replaced by $e^{-j\beta t}\mathfrak{A}^t$ and A replaced by $A - j\beta$. □

Corollary 3.4.5 *Let A be a linear operator $X \supset \mathcal{D}(A) \to X$ with dense domain. Then the following conditions are equivalent*

(i) *A is the generator of a C_0 contraction semigroup;*
(ii) *every positive real λ belongs to the resolvent set of A and*

$$\left\|(\lambda - A)^{-1}\right\| \leq \lambda^{-1} \text{ for } \lambda > 0;$$

(iii) *the right-half plane $\mathbb{C}^+$ belongs to the resolvent set of A, and*

$$\left\|(\lambda - A)^{-1}\right\| \leq (\Re\lambda)^{-1} \text{ for } \Re\lambda > 0.$$

Proof This proof is similar to the proof of Corollary 3.4.4, but we replace Theorem 3.4.1 by Corollary 3.4.2. □

There is also another characterization of the generators of contraction semigroups which is based on *dissipativity*.

Definition 3.4.6 A linear operator $A\colon X \supset \mathcal{D}(A) \to X$ is *dissipative* if for every $x \in \mathcal{D}(A)$ there is a vector $x^* \in X^*$ with $|x^*|^2 = |x|^2 = \langle x^*, x\rangle$ such that $\Re\langle x^*, Ax\rangle \leq 0$ (if X is a Hilbert space, then we take $x^* = x$).[3]

Lemma 3.4.7 *Let $A\colon X \supset \mathcal{D}(A) \to X$ be a linear operator. Then the following conditions are equivalent:*

[3] The dual space X^* is discussed at the beginning of Section 3.5.

(i) *A is dissipative;*
(ii) *$A - j\beta I$ is dissipative for all $\beta \in \mathbb{R}$;*
(iii) *$|(\lambda - A)x| \geq \lambda |x|$ for all $x \in \mathcal{D}(A)$ and all $\lambda > 0$;*
(iv) *$|(\lambda - A)x| \geq \Re\lambda |x|$ for all $x \in \mathcal{D}(A)$ and all $\lambda \in \mathbb{C}^+$.*

Proof (i) $\Rightarrow$ (ii): This follows from Definition 3.4.6 since, with the notation of that definition, $\Re\langle x^*, j\beta x\rangle = \Re(-j\beta\langle x^*, x\rangle) = \Re(-j\beta|x|^2) = 0$.

(ii) $\Rightarrow$ (iv): Suppose that (ii) holds. Let $x \in \mathcal{D}(A)$ and $\lambda = \alpha + j\beta$ with $\alpha > 0$ and $\beta \in \mathbb{R}$. Choose some $x^* \in X^*$ with $|x^*|^2 = |x|^2 = \langle x^*, x\rangle$ such that $\Re\langle x^*, Ax\rangle \leq 0$ (by the Hahn–Banach theorem, this is possible). Then

$$\begin{aligned}|\lambda x - Ax||x| &\geq |\langle x^*, \lambda x - Ax\rangle| \geq \Re\langle x^*, \lambda x - Ax\rangle \\ &= \Re\langle x^*, \alpha x\rangle - \Re\langle x^*, (A - j\beta)x\rangle \geq \alpha|x|^2,\end{aligned}$$

and (iv) follows.

(iv) $\Rightarrow$ (iii): This is obvious.

(iii) $\Rightarrow$ (i): Let $x \in \mathcal{D}(A)$, and suppose that $\lambda|x| \leq |(\lambda - A)x|$ for all $\lambda > 0$. Choose some $z_\lambda^* \in X^*$ with $|z_\lambda^*| = 1$ such that $\langle z_\lambda^*, (\lambda - A)x\rangle = |(\lambda - A)x|$. Then, for all $\lambda > 0$,

$$\begin{aligned}\lambda|x| &\leq |\lambda x - Ax| = \langle z_\lambda^*, \lambda x - Ax\rangle \\ &= \lambda\Re\langle z_\lambda^*, x\rangle - \Re\langle z_\lambda^*, Ax\rangle \leq \lambda|x| - \Re\langle z_\lambda^*, Ax\rangle.\end{aligned}$$

This implies that $\Re\langle z_\lambda^*, Ax\rangle \leq 0$ and that

$$\lambda\Re\langle z_\lambda^*, x\rangle \geq |\lambda x - Ax| \geq \lambda|x| - |Ax|.$$

For all $\lambda > 0$, let Z_λ^* be the weak* closure of the set $\{z_\alpha^* \mid \alpha \geq \lambda\}$. Then each Z_λ^* is a weak* compact subset of the unit ball in X^*, and for all $z^* \in Z_\lambda^*$ we have

$$\Re\langle z^*, Ax\rangle \leq 0, \quad \Re\langle z^*, x\rangle \geq |x| - \lambda^{-1}|Ax|, \quad |z^*| \leq 1$$

(the functionals $z^* \mapsto \Re\langle z^*, x\rangle$ and $z^* \mapsto \Re\langle z^*, Ax\rangle$ are continuous in the weak* topology). The sets Z_λ^* obviously have the finite intersection property and they are weak* compact, so their intersection $\cap_{\lambda>0} Z_\lambda^*$ is nonempty (see, e.g., Rudin [1987, Theorem 2.6, p. 37]). Choose any z^* in this intersection. Then

$$\Re\langle z^*, Ax\rangle \leq 0, \quad \Re\langle z^*, x\rangle \geq |x|, \quad |z^*| \leq 1.$$

The last two inequalities imply that $|z^*| = \langle z^*, x\rangle = |x|$. By taking $y^* = |x|z^*$ in Definition 3.4.6 we find that A is dissipative. □

By using the notion of dissipativity we can add one more condition to the list of equivalent conditions in Corollary 3.4.5.

Theorem 3.4.8 (Lumer–Phillips) *Let A be a linear operator $X \supset \mathcal{D}(A) \to X$ with dense domain. Then the following conditions are equivalent (and they are equivalent to the conditions (ii) and (iii) in Corollary 3.4.5):*

(i) *A is the generator of a C_0 contraction semigroup;*
(iv) *A is dissipative and $\rho(A) \cap \mathbb{C}^+ \neq \emptyset$.*

These conditions are, in particular, true if

(v) *A is closed and densely defined, and both A and A^* are dissipative.*

If X is reflexive, then (v) is equivalent to the other conditions.

Proof (i) $\Rightarrow$ (iv): This follows from Corollary 3.4.5 and Lemma 3.4.7.

(iv) $\Rightarrow$ (i): Suppose that (iv) holds. Then A is closed (since its resolvent set is nonempty). Take some $\lambda = \alpha + j\beta \in \rho(A)$ with $\alpha > 0$ and $\beta \in \mathbb{R}$. If A is dissipative, then we get from Lemma 3.4.7(iv) for all $x \in \mathcal{D}(A)$, $|(\lambda - A)x| \geq \alpha|x|$. This implies that $\|(\lambda - A)^{-1}\| \leq 1/\alpha$. By Lemma 3.2.8, the resolvent set of A contains an open circle with center λ and radius $\alpha = \Re\lambda$. We can repeat this argument with α replaced by first $3/2\alpha$, then $(3/2)^2\alpha$, then $(3/2)^3\alpha$, etc., to show that the whole right-half plane belongs to the resolvent set, and that $\|(\lambda - A)^{-1}\| \leq (\Re\lambda)^{-1}$ for all $\lambda \in \mathbb{C}^+$. By Corollary 3.4.5, A is therefore the generator of a C_0 contraction semigroup.

(v) $\Rightarrow$ (iv): By Lemma 3.4.7, $|(1 - A)x| \geq |x|$ for all $x \in \mathcal{D}(A)$. This implies that $1 - A$ is injective and has closed range (see Lemma 9.10.2(iii)). If $\mathcal{R}(1 - A) \neq X$ then, by the Hahn–Banach theorem, there is some nonzero $x^* \in X^*$ such that $\langle x^*, x - Ax = 0\rangle$, or equivalently, $\langle x^*, Ax\rangle = \langle x^*, x\rangle$ for all $x \in \mathcal{D}(A)$. This implies that $x^* \in \mathcal{D}(A^*)$ and that $A^*x^* = x^*$, i.e., $(1 - A^*)x^* = 0$. By Lemma 3.4.7 and the dissipativity of A^*, $|x^*| \leq |(1 - A^*)x^*| = 0$, contradicting our original choice of x^*. Thus $\mathcal{R}(1 - A) = X$. By the closed graph theorem, $(1 - A)^{-1} \in \mathcal{B}(X)$, so $1 \in \rho(A)$, and we have proved that (iv) holds.

If X is reflexive, then A is a generator of a C_0 contraction semigroup if and only if A^* is the generator of a C_0 contraction semigroup (see Theorem 3.5.6(v)), so (v) follows from (iv) in this case. □

In the Hilbert space case there is still another way of characterizing a generator of a contraction semigroup.

Theorem 3.4.9 *Let X be a Hilbert space, and let A be a linear operator $X \supset \mathcal{D}(A) \to X$ with dense domain. Then the following conditions are equivalent:*

(i) *A is the generator of a C_0 contraction semigroup,*
(ii) *there is some $\lambda \in \mathbb{C}^+ \cap \rho(A)$ for which the operator $\mathbf{A}_\lambda = (\overline{\lambda} + A)(\lambda - A)^{-1}$ is a contraction,*
(iii) *all $\lambda \in \mathbb{C}^+$ belong to $\rho(A)$, and $\mathbf{A}_\lambda = (\overline{\lambda} + A)(\lambda - A)^{-1}$ is a contraction,*

If $\mathbf{A}_\lambda$ is defined as in (vi) and (vii), then -1 is not an eigenvalue of $\mathbf{A}_\lambda$, and $\mathcal{R}(1+\mathbf{A}_\lambda) = \mathcal{D}(A)$. Conversely, if X is a Hilbert space and if $\mathbf{A}$ is a contraction on X such that -1 is not an eigenvalue of $\mathbf{A}$, then $\mathcal{R}(1+\mathbf{A})$ is dense in X, and, for all $\lambda \in \mathbb{C}^+$, the operator A_λ with $\mathcal{D}(A_\lambda) = \mathcal{R}(1+\mathbf{A})$ defined by

$$A_\lambda x = \lambda x - 2\Re\lambda\,(1+\mathbf{A})^{-1}x, \qquad x \in \mathcal{R}(1+\mathbf{A}),$$

is the generator of a C_0 contraction semigroup on X (and the operator $\mathbf{A}_\lambda$ in (vi) and (vii) corresponding to A_λ is $\mathbf{A}$).

Proof (i) $\Rightarrow$ (vii) $\Rightarrow$ (vi) $\Rightarrow$ (i): Let us denote $\lambda = \alpha + j\beta$ where $\alpha > 0$ and $\beta \in \mathbb{R}$. For all $x \in \mathcal{D}(A)$, if we denote $B = A - j\beta$, then

$$|(\lambda - A)x|^2 = |(\alpha - B)x|^2 = \alpha^2|x|^2 - 2\alpha\Re\langle x, Bx\rangle + |Bx|^2$$
$$|(\overline{\lambda} + A)x|^2 = |(\alpha + B)x|^2 = \alpha^2|x|^2 + 2\alpha\Re\langle x, Bx\rangle + |Bx|^2.$$

If (i) holds, then by Lemma 3.4.7 and Theorem 3.4.8, $B = A - j\beta$ is dissipative, and we get $|(\overline{\lambda} + A)x| \le |(\lambda - A)x|$ for all $\lambda \in \mathbb{C}^+$ and all $x \in \mathcal{D}(A)$. By Corollary 3.4.5, $\lambda \in \rho(A)$, and by replacing x by $(\lambda - A)^{-1}x$ we find that $|\mathbf{A}_\lambda x| \le |x|$ for all $x \in X$, i.e, $\mathbf{A}_\lambda$ is a contraction. This proves that (i) $\Rightarrow$ (vii). Obviously (vii) $\Rightarrow$ (vi). If (vi) holds, then for that particular value of λ, we have $|\mathbf{A}_\lambda x| \le |x|$ for all $x \in X$, or equivalently, $|(\overline{\lambda} + A)x| \le |(\lambda - A)x|$ for all $x \in \mathcal{D}(A)$. The preceding argument then shows that $B = A - j\beta$ is dissipative, hence so is A, and (i) follows from Theorem 3.4.8. This proves that (i), (vi), and (vii) are equivalent.

Let us next show that -1 cannot be an eigenvalue of $\mathbf{A}_\lambda$ (although $-1 \in \sigma(\mathbf{A}_\lambda)$ whenever A is unbounded) and that $\mathcal{R}(1+\mathbf{A}_\lambda) = \mathcal{D}(A)$. This follows from the (easily verified) identity that

$$1 + \mathbf{A}_\lambda = 2\Re\lambda\,(\lambda - A)^{-1}.$$

Here the right-hand side is injective, hence so is the left-hand side, and the range of the right-hand side is $\mathcal{D}(A)$, hence so is the range of the left-hand side.

It remains to prove the converse part. Let $\mathbf{A}$ be a contraction on X such that -1 is not an eigenvalue of $\mathbf{A}$. Then the operator A_λ is well-defined on $\mathcal{D}(A_\lambda) = \mathcal{R}(1+\mathbf{A})$, and

$$(\lambda 1 - A_\lambda)x = 2\Re\lambda\,(1+\mathbf{A})^{-1}x, \qquad x \in \mathcal{R}(1+\mathbf{A}),$$

This implies that $\lambda - A_\lambda$ is injective, $\mathcal{R}(\lambda - A_\lambda) = X$, and $(\lambda - A_\lambda)^{-1} = 2\Re\lambda\,(1+\mathbf{A})^{-1}$. In particular, $\lambda \in \rho(A_\lambda)$. Arguing as in the proof of the implication (vi) $\Rightarrow$ (i) we find that A_λ is dissipative since $\mathbf{A}$ is a contraction (note that we have the same relationship between A_λ and $\mathbf{A}$ as we had between A and $\mathbf{A}_\lambda$, namely $\mathbf{A} = (\overline{\lambda} + A_\lambda)(\lambda - A_\lambda)^{-1}$). If we knew that $\mathcal{D}(A)$ is dense in X, then we could conclude from Theorem 3.4.8 that A_λ is the generator of a C_0

contraction semigroup. Thus, to complete the proof, the only thing remaining to be verified is that $\mathcal{R}(1+\mathbf{A})$ is dense in X. This is true if and only if -1 is not an eigenvalue of $\mathbf{A}^*$, so let us prove this statement instead. If $\mathbf{A}^*x = x$ for some $x \in X$, then

$$\langle \mathbf{A}^*x, x\rangle = \langle x, \mathbf{A}x\rangle = \langle x, x\rangle = |x|^2,$$

hence

$$|x - \mathbf{A}x|^2 = |x|^2 - 2\Re\langle x, \mathbf{A}x\rangle + |\mathbf{A}x|^2 = |\mathbf{A}x|^2 - |x|^2 \le 0,$$

and we see that $\mathbf{A}x = x$. This implies that $x = 0$, because -1 was supposed not to be an eigenvalue of $\mathbf{A}$. □

The operator $\mathbf{A}_\lambda$ in Theorem 3.4.9 is called the *Cayley transform* of A with parameter $\alpha \in \mathbb{C}^+$. We shall say much more about this transform in Chapter 11.

3.5 The dual semigroup

Many results in quadratic optimal control rely on the possibility of passing from a system to its dual system. In this section we shall look at the dual of the semigroup $\mathfrak{A}$. The dual of the full system will be discussed in Section 6.2.

In most applications of the duality theory the state space X is a Hilbert space. In this case it is natural to identify the dual X with X itself. This has the effect that the mapping from an operator A on X to its dual A^* becomes conjugate-linear instead of linear, as is the case in the standard Banach space theory. To simplify the passage from the Banach space dual of an operator to the Hilbert space dual we shall throughout use the conjugate-linear dual instead of the ordinary dual of a Banach space.

As usual, we define the dual X^* of the Banach space X to be the space of all bounded linear functionals $x^* \colon X \to \mathbb{C}$. We denote the value of the functional $x^* \in X^*$ acting on the vector $x \in X$ alternatively by

$$x^*x = \langle x, x^*\rangle = \langle x, x^*\rangle_{(X,X^*)}.$$

The norm in X^* is the usual supremum-norm

$$|x^*|_{X^*} := \sup_{|x|_X=1} |\langle x, x^*\rangle|, \tag{3.5.1}$$

and by the Hahn–Banach theorem, the symmetric relation

$$|x|_X = \sup_{|x^*|_{X^*}=1} |\langle x, x^*\rangle| \tag{3.5.2}$$

also holds. On this space we use a nonstandard linear structure, defining the sum of two elements x^* and y^* in X^* and the product of a scalar $\lambda \in \mathbb{C}$ and a vector $x^* \in X^*$ by

$$\begin{aligned} \langle x, x^* + y^* \rangle &:= \langle x, x^* \rangle + \langle x, y^* \rangle, \quad x \in X, \\ \langle x, \lambda x^* \rangle &:= \overline{\lambda} \langle x, x^* \rangle, \qquad\qquad x \in X, \quad \lambda \in C. \end{aligned} \tag{3.5.3}$$

In other words, the mapping $(x, x^*) \mapsto \langle x, x^* \rangle$ is anti-linear (linear in x and conjugate-linear in x^*). All the standard results on the dual of a Banach space and the dual operator remain valid in this conjugate-linear setting, except for the fact that the mapping from an operator A to its dual operator A^* becomes conjugate-linear instead of linear, like in the standard Hilbert space case.

Let A be a closed (unbounded) operator $X \supset \mathcal{D}(A) \to Y$ with dense domain. The domain of the dual A^* of A consists of those $y^* \in Y^*$ for which the linear functional

$$x \mapsto \langle Ax, y^* \rangle_{(Y,Y^*)}, \quad x \in \mathcal{D}(A),$$

can be extended to a bounded linear functional on X. This extension is unique since $\mathcal{D}(A)$ is dense, and it can be written in the form

$$x \mapsto \langle Ax, y^* \rangle_{(Y,Y^*)} = \langle x, x^* \rangle_{(X,X^*)}, \quad x \in \mathcal{D}(A),$$

for some $x^* \in X$. For $y^* \in \mathcal{D}(A^*)$ we define A^* by $A^* y^* = x^*$, where $x^* \in X^*$ is the element above. Thus,

$$\langle Ax, y^* \rangle_{(Y,Y^*)} = \langle x, A^* y^* \rangle_{(X,X^*)}, \quad x \in \mathcal{D}(A), \quad y^* \in \mathcal{D}\left(A^*\right), \tag{3.5.4}$$

and this relationship serves as a definition of A^*.

Lemma 3.5.1 *Let $A\colon X \supset \mathcal{D}(A) \to Y$ be a closed linear operator with dense domain. Then*

(i) *$A^*\colon Y^* \supset \mathcal{D}(A^*) \to U^*$ is a closed linear operator,*
(ii) *If $A \in \mathcal{B}(X;Y)$, then $A^* \in \mathcal{B}(Y^*;X^*)$, and $\|A\| = \|A^*\|$,*
(iii) *$\mathcal{D}(A^*)$ weak*-dense in Y^*,*
(iv) *If Y is reflexive, then $\mathcal{D}(A^*)$ is dense in Y^*.*

Proof (i) It is a routine calculation to show that A^* is linear. Let us show that it is closed. Take some sequence $y_n^* \in \mathcal{D}(A^*)$ such that $y_n^* \to y^* \in Y^*$ and $A^* y_n^* \to x^*$ in X^* as $n \to \infty$. Then, for each $x \in \mathcal{D}(A)$,

$$\langle x, x^* \rangle = \lim_{n\to\infty} \langle x, A y_n^* \rangle = \lim_{n\to\infty} \langle Ax, y_n^* \rangle = \langle Ax, y^* \rangle.$$

This means that the functional $\langle Ax, y^* \rangle$ can be extended to a bounded linear functional on X, hence $y^* \in \mathcal{D}(A^*)$ and $x^* = A^* y^*$. Thus, A^* is closed.

(ii) If $A \in \mathcal{B}(X;Y)$, then it is clear that $\mathcal{D}(A^*) = Y^*$. Moreover,

$$\begin{aligned}\|A^*\|_{\mathcal{B}(Y^*;X^*)} &= \sup_{|y^*|=1} |A^*y^*|_{X^*} = \sup_{\substack{|x|=1\\|y^*|=1}} |\langle x, A^*y^*\rangle_{(X,X^*)}| \\ &= \sup_{\substack{|x|=1\\|y^*|=1}} |\langle Ax, y^*\rangle_{(X,X^*)}| = \sup_{|x|=1} |Ax|_Y \\ &= \|A\|_{\mathcal{B}(X;Y)}.\end{aligned}$$

(iii) Let $y \in Y$, $y \neq 0$. As A is closed, the set $\left\{\left[\begin{smallmatrix} Ax \\ x \end{smallmatrix}\right] \,\middle|\, x \in \mathcal{D}(A)\right\}$ is a closed subspace of $\left[\begin{smallmatrix} Y \\ X \end{smallmatrix}\right]$, and $\left[\begin{smallmatrix} y \\ 0 \end{smallmatrix}\right]$ certainly does not belong to this subspace. By the Hahn–Banach theorem in $\left[\begin{smallmatrix} Y \\ X \end{smallmatrix}\right]^* = \left[\begin{smallmatrix} Y^* \\ X^* \end{smallmatrix}\right]$, there is some $x_1^* \in X^*$ and $y_1^* \in Y^*$ such that $\langle x, x_1^*\rangle + \langle Ax, y_1^*\rangle = 0$ for all $x \in \mathcal{D}(A)$, but $\langle 0, x_1\rangle - \langle y, y_1^*\rangle \neq 0$. The first equation says that $y_1^* \in \mathcal{D}(A^*)$ (and that $A^*y_1^* = -x_1^*$). Thus, for each nonzero $y \in Y$, it is possible to find some $y^* \in \mathcal{D}(A^*)$ such that $\langle y, y^*\rangle \neq 0$, or equivalently, $Y \ni y = 0$ iff $\langle y, y^*\rangle = 0$ for all $y^* \in \mathcal{D}(A^*)$. This shows that $\mathcal{D}(A^*)$ is weak*-dense in Y^* (apply the Hahn–Banach theorem [Rudin, 1973, Theorem 3.5, p. 59] to the weak*-topology).

(iv) If Y is reflexive, then (iii) implies that $\mathcal{D}(A^*)$ is weakly dense in Y^*, hence dense in Y^* Rudin [1973, Corollary 3.12(b), p. 65]).

□

Lemma 3.5.2 *Let $A\colon X \supset \mathcal{D}(A) \to Y$ be closed, densely defined, and injective, and suppose that $\mathcal{R}(A) = Y$. Then $A^{-1} \in \mathcal{B}(Y;X)$, and $(A^{-1})^* = (A^*)^{-1}$. We denote this operator by A^{-*}.*

Proof The operator A^{-1} is closed since A is closed, and by the closed graph theorem, it is bounded, i.e., $A^{-1} \in \mathcal{B}(Y;X)$. By Lemma 3.5.1(ii), $(A^{-1})^* \in \mathcal{B}(X^*;Y^*)$. It remains to show that $(A^{-1})^* = (A^*)^{-1}$.

Take some arbitrary $x \in \mathcal{D}(A)$ and $x^* \in X^*$. Then

$$\langle x, x^*\rangle = \langle A^{-1}Ax, x^*\rangle = \langle Ax, (A^{-1})^*x^*\rangle.$$

This implies that $(A^{-1})^*x^* \in \mathcal{D}(A^*)$ and that $A^*(A^{-1})^*x^* = x^*$. Thus, $(A^{-1})^*$ is a left inverse of A^*. If we instead take some arbitrary $x \in X$ and $x^* \in \mathcal{D}(A^*)$, then

$$\langle x, x^*\rangle = \langle AA^{-1}x, x^*\rangle = \langle A^{-1}x, A^*x^*\rangle = \langle x, (A^{-1})^*A^*x^*\rangle.$$

Thus, $(A^{-1})^*$ is also a right inverse of A^*. This means that A^* is invertible, with $(A^*)^{-1} = (A^{-1})^*$. □

Lemma 3.5.3 *Let $A\colon X \supset \mathcal{D}(A) \to X$ be densely defined, and let $\alpha \in \rho(A)$ (in particular, this means that A is closed). Then $\overline{\alpha} \in \rho(A^*)$, and $((\alpha - A)^*)^{-1} = ((\alpha - A)^{-1})^* = (\alpha - A)^{-*}$.*

Proof By the definition of the dual operator, $(\alpha - A)^* = \overline{\alpha} - A^*$. Therefore Lemma 3.5.3 follows from Lemma 3.5.2, applied to the operator $\alpha - A$. □

Lemma 3.5.4 *Let $A\colon X \supset \mathcal{D}(A) \to Y$ be densely defined, and let $B \in \mathcal{B}(Y;Z)$. Then $(BA)^* = A^*B^*$ (with $\mathcal{D}((BA)^*) = \mathcal{D}(A^*B^*) = \{z^* \in Z^* \mid B^*z \in \mathcal{D}(A^*)\}$).*

Proof Let $x \in \mathcal{D}(A) = \mathcal{D}(BA)$ and $z^* \in \mathcal{D}(A^*B^*) = \{z^* \in Z^* \mid B^*z \in \mathcal{D}(A^*)\}$. Then

$$\langle BAx, z^*\rangle_{(Z,Z^*)} = \langle Ax, B^*z^*\rangle_{(Y,Y^*)} = \langle x, A^*B^*z^*\rangle_{(X,X^*)}.$$

This implies that $z^* \in \mathcal{D}((BA)^*)$, and that $(BA)^*z^* = A^*B^*z^*$. To complete the proof it therefore suffices to show that $\mathcal{D}((BA)^*) \subset \mathcal{D}(A^*B^*)$. Let $z^* \in \mathcal{D}((BA)^*)$. Then, for every $x \in \mathcal{D}(A) = \mathcal{D}(BA)$,

$$\langle Ax, B^*z^*\rangle_{(Y,Y^*)} = \langle BAx, z^*\rangle_{(Z,Z^*)} = \langle x, (BA)^*z^*\rangle_{(X,X^*)}.$$

This implies that $B^*z^* \in \mathcal{D}(A^*)$, and hence $z^* \in \mathcal{D}(A^*B^*)$. □

Lemma 3.5.5 *Let $B \in \mathcal{B}(X;Y)$ be invertible (with an inverse in $\mathcal{B}(Y;X)$), and let $A\colon Y \supset \mathcal{D}(A) \to Z$ be densely defined. Then AB is densely defined (with $\mathcal{D}(AB) = \{x \in X \mid Bx \in \mathcal{D}(A)\}$), and $(AB)^* = B^*A^*$ (with $\mathcal{D}(B^*A^*) = \mathcal{D}(A^*)$).*

Proof The domain of AB is the image under B^{-1} of $\mathcal{D}(A)$ which is dense in Y, and therefore $\mathcal{D}(AB)$ is dense in X (if $x \in X$, and if $y_n \in \mathcal{D}(A)$ and $y_n \to y := Bx$ in y, then $x_n := B^{-1}y_n$ in $\mathcal{D}(AB)$, and $x_n \to B^{-1}y = x$ in X). Thus AB has an adjoint $(AB)^*$.

Let $x \in \mathcal{D}(AB) = \{x \in X \mid Bx \in \mathcal{D}(A)\}$ and $z^* \in \mathcal{D}(A^*)$. Then

$$\langle ABx, z^*\rangle_{(Z,Z^*)} = \langle Bx, A^*z^*\rangle_{(Y,Y^*)} = \langle x, B^*A^*z^*\rangle_{(X,X^*)}.$$

This implies that $z^* \in \mathcal{D}((AB)^*)$, and that $(AB)^*z^* = B^*A^*z^*$. To complete the proof it therefore suffices to show that $\mathcal{D}((AB)^*) \subset \mathcal{D}(A^*)$. Let $z^* \in \mathcal{D}((AB)^*)$. Then, for every $y \in \mathcal{D}(A)$, we have $B^{-1}y \in \mathcal{D}(AB)$, and

$$\begin{aligned}\langle Ay, z^*\rangle_{(Z,Z^*)} &= \langle ABB^{-1}y, z^*\rangle_{(Z,Z^*)} = \langle B^{-1}y, (AB)^*z^*\rangle_{(X,X^*)}\\ &= \langle y, B^{-*}(AB)^*z^*\rangle_{(Y,Y^*)}.\end{aligned}$$

This implies that $z^* \in \mathcal{D}(A^*)$. □

Theorem 3.5.6 *Let $\mathfrak{A}$ be a C_0 semigroup on a Banach space X with generator A.*

(i) $\mathfrak{A}^{*t} = (\mathfrak{A}^t)^*$, $t \geq 0$, *is a locally bounded semigroup on X^* (but it need not be strongly continuous). This semigroup has the same growth bound as $\mathfrak{A}$.*

(ii) *Let* $X^{\odot} = \{x^* \in X^* \mid \lim_{t\downarrow 0} \mathfrak{A}^{*t} x^* = x^*\}$. *Then* $X^{\odot}$ *is a closed subspace of* X^* *which is invariant under* $\mathfrak{A}^*$, *and the restriction* $\mathfrak{A}^{\odot}$ *of* $\mathfrak{A}^*$ *to* $X^{\odot}$ *is a* C_0 *semigroup on* $X^{\odot}$.

(iii) *The generator* $A^{\odot}$ *of the semigroup* $\mathfrak{A}^{\odot}$ *in (ii) is the restriction of* A^* *to* $\mathcal{D}\left(A^{\odot}\right) = \{x^* \in \mathcal{D}(A^*) \mid A^*x^* \in X^{\odot}\}$.

(iv) $X^{\odot}$ *is the closure of* $\mathcal{D}(A^*)$ *in* X^*. *Thus,* $\mathcal{D}(A^*) \subset X^{\odot}$ *and* $\mathcal{D}(A^*)$ *is dense in* $X^{\odot}$.

(v) *If* X *is reflexive, then* $X^{\odot} = X^*$, $A^{\odot} = A^*$, *and* $\mathfrak{A}^*$ *is a* C_0 *semigroup on* X^* *with generator* A^*.

(vi) *If* $A \in \mathcal{B}(X)$, *then* $X^{\odot} = X^*$, $A^{\odot} = A^*$, *and* $\mathfrak{A}^*$ *is a* C_0 *semigroup on* X^* *with generator* A^*.

For an example where $X^{\odot} \neq X^*$, see Example 3.5.11 with $p = 1$.

Proof of Theorem 3.5.6 (i) This follows from Lemmas 3.5.1(ii) and 3.5.4.

(ii) The proof of the claim that $\mathfrak{A}^{*t}$ maps $X^{\odot}$ into $X^{\odot}$ is the same as the proof of Lemma 2.2.13(ii).

To show that $X^{\odot}$ is closed we let $x_n^* \in X^{\odot}$, $x_n^* \to x^* \in X^*$. Write

$$\|\mathfrak{A}^{*s}x^* - x^*\| \le \|\mathfrak{A}^{*s}x^* - \mathfrak{A}^{*s}x_n^*\| + \|\mathfrak{A}^{*s}x_n^* - x_n^*\| + \|x_n^* - x^*\|.$$

Given $\epsilon > 0$, we can make $\|\mathfrak{A}^{*s}x^* - \mathfrak{A}^{*s}x_n^*\| + \|x_n^* - x^*\| < \epsilon/2$ for all $0 \le s \le 1$ by choosing n large enough (since $\|\mathfrak{A}^{*s}\| \le M\mathrm{e}^{\omega s}$ for some $M > 0$ and $\omega \in \mathbb{R}$). Next we choose $t \le 1$ so small that $\|\mathfrak{A}^{*s}x_n^* - x_n^*\| \le \epsilon/2$ for all $0 \le s \le t$. Then $\|\mathfrak{A}^{*s}x^* - x^*\| \le \epsilon$ for $0 \le s \le t$. This proves that $\lim_{t\downarrow 0} \mathfrak{A}^{*t}x^* = x^*$, hence $x^* \in X^{\odot}$. Thus $X^{\odot}$ is closed in X^*.

Since $X^{\odot}$ is closed in X^*, it is a Banach space with the same norm, and by definition, $\mathfrak{A}^{\odot}$ is a C_0 semigroup on $X^{\odot}$.

(iii) Let $A^{\odot}$ be the generator of $\mathfrak{A}^{\odot}$. Choose some $x \in \mathcal{D}(A)$ and $x^* \in \mathcal{D}\left(A^{\odot}\right) \subset X^{\odot} \subset X^*$. Then

$$\begin{aligned}\langle Ax, x^*\rangle_{(X,X^*)} &= \lim_{t\downarrow 0} \left\langle \frac{1}{t}(\mathfrak{A}^t - 1)x, x^*\right\rangle_{(X,X^*)} \\ &= \lim_{t\downarrow 0} \left\langle x, \frac{1}{t}(\mathfrak{A}^{*t} - 1)x^*\right\rangle_{(X,X^*)} \\ &= \langle x, A^{\odot}x^*\rangle_{(X,X^*)}.\end{aligned}$$

This implies that $x^* \in \mathcal{D}(A^*)$ and $A^{\odot}x^* = A^*x^*$. In other words, if we let B be the restriction of A^* to $\mathcal{D}(B) = \{x^* \in \mathcal{D}(A^*) \mid A^*x^* \in X^{\odot}\}$, then $A^{\odot} \subset B$, i.e., $\mathcal{D}\left(A^{\odot}\right) \subset \mathcal{D}(B)$ and $A^{\odot}X^* = Bx^*$ for all $x^* \in \mathcal{D}\left(A^{\odot}\right)$.

It remains to show that $\mathcal{D}(B) = \mathcal{D}\left(A^{\odot}\right)$. Choose some $\alpha \in \rho(A^{\odot}) \cap \rho(A^*)$ (by Theorem 3.2.9(i) and Lemma 3.5.3, any α with $\Re\alpha$ large enough will do). Then $\alpha - A^{\odot}$ maps $\mathcal{D}\left(A^{\odot}\right)$ one-to-one onto $X^{\odot}$, hence $\alpha - B$ maps $\mathcal{D}(B)$

onto $X^{\odot}$, i.e.,

$$X^{\odot} = (\alpha - A^{\odot})\mathcal{D}\left(A^{\odot}\right) = (\alpha - B)\mathcal{D}(B).$$

But $\alpha - B$ is a restriction of $\alpha - A^*$ which is one-to-one on X^*; hence $\alpha - B$ is injective on $\mathcal{D}(B)$, and

$$\mathcal{D}(B) = (\alpha - B)^{-1}X^{\odot} = (\alpha - A^{\odot})^{-1}X^{\odot} = \mathcal{D}(A).$$

(iv) Let $x^* \in \mathcal{D}(A^*)$. Choose α and M such that $\|\mathfrak{A}^s\| \le M\mathrm{e}^{\alpha s}$ for all $s \ge 0$. Then, for all $x \in X$, all $t \ge 0$, and all real $\alpha > \omega_{\mathfrak{A}}$, by Theorem 3.2.1(ii) and Example 3.2.6(i),

$$\begin{aligned}
\left|\left\langle x, (\mathrm{e}^{-\alpha t}\mathfrak{A}^{*t} - 1)x^*\right\rangle\right| &= \left|\left\langle (\mathrm{e}^{-\alpha t}\mathfrak{A}^{t} - 1)x, x^*\right\rangle\right| \\
&= \left|\left\langle (\alpha - A)(\alpha - A)^{-1}(\mathrm{e}^{-\alpha t}\mathfrak{A}^{t} - 1)x, x^*\right\rangle\right| \\
&= \left|\left\langle (\alpha - A)^{-1}(\mathrm{e}^{-\alpha t}\mathfrak{A}^{t} - 1)x, (\alpha - A)x^*\right\rangle\right| \\
&= \left|\left\langle \int_0^t \mathrm{e}^{-\alpha s}\mathfrak{A}^{s}x\,ds, (\alpha - A)x^*\right\rangle\right| \\
&\le Mt\|x\|\|(\alpha - A)x^*\|.
\end{aligned}$$

Taking the supremum over all $x \in X$ with $\|x\| = 1$ an using (3.5.1) we get

$$\|(\mathrm{e}^{-\alpha t}\mathfrak{A}^{*t} - 1)x^*\| \le Mt\|(\alpha - A)x^*\| \to 0 \text{ as } t \downarrow 0,$$

which implies that $\lim_{t\downarrow 0}\mathfrak{A}^{*t}x^* = x^*$. This shows that $\mathcal{D}(A^*) \subset X^{\odot}$. That $\mathcal{D}(A^*)$ is dense in $X^{\odot}$ follows from the fact that $\mathcal{D}\left(A^{\odot}\right) \subset \mathcal{D}(A^*)$ and $\mathcal{D}\left(A^{\odot}\right)$ is dense in $X^{\odot}$.

(v)–(vi) This follows from (iv) and Lemma 3.5.1(ii)–(iv). □

Definition 3.5.7 The C_0 semigroup $\mathfrak{A}^{\odot}$ in Theorem 3.5.6 is the dual of the C_0 semigroup $\mathfrak{A}$, $X^{\odot}$ is the $\odot$-dual of X (with respect to $\mathfrak{A}$), and $A^{\odot}$ is the $\odot$-dual of A.

Example 3.5.8 *The dual $\mathfrak{A}^*$ of the diagonal (semi)group $\mathfrak{A}$ in Example 3.3.3 is another diagonal (semi)group where the eigenvectors $\{\phi_n\}_{n=1}^{\infty}$ stay the same but the sequence of eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ has been replaced by its complex conjugate $\{\overline{\lambda}_n\}_{n=1}^{\infty}$. Thus*

$$\mathfrak{A}^{*t}x = \sum_{n=1}^{\infty} \mathrm{e}^{\overline{\lambda}_n t}\langle x, \phi_n\rangle\phi_n, \qquad x \in H, \qquad t \ge 0.$$

The dual generator A^* has the same domain as A, and is it given by

$$Ax = \sum_{n=1}^{\infty} \overline{\lambda}_n\langle x, \phi_n\rangle\phi_n, \qquad x \in \mathcal{D}(A).$$

In particular, $\mathfrak{A}^t = \mathfrak{A}^{*t}$ *for all* $t \geq 0$ *and* $A = A^*$ *if and only if all the eigenvalues are real.*

We leave the proof to the reader as an exercise.

Let us next look at the duals of the shift (semi)groups in Examples 2.3.2 and 2.5.3. To do this we need to determine the dual of an L^p-space.

Lemma 3.5.9 *Let* U *be a reflexive Banach space*[4]*, let* $1 \leq p < \infty$*,* $1/p + 1/q = 1$ *(with* $1/\infty = 0$*),* $\omega \in \mathbb{R}$*, and* $J \subset \mathbb{R}$ *(with positive measure).*

(i) *The dual of* $L^p_\omega(J;U)$ *can be identified with* $L^q_{-\omega}(J;U^*)$ *in the sense that every bounded linear functional* f *on* $L^p_\omega(J;U)$ *is of the form*

$$\langle u, f\rangle = \int_J \langle u(t), u^*(t)\rangle_{(U,U^*)}\, dt, \quad u \in L^p_\omega(J;U),$$

for some $u^* \in L^q_{-\omega}(J;U^*)$*. The norm of the functional* f *is equal to the* $L^q_{-\omega}(J)$*-norm on* u^**.*

(ii) $L^p_\omega(J;U)$ *is reflexive iff* $1 < p < \infty$*.*

Proof For $\omega = 0$ this lemma is contained in Diestel and Uhl [1977, Theorem 1, p. 98 and Corollary 2, p. 100]. If f is a bounded linear functional on $L^p_\omega(J;U)$ for some $\omega \neq 0$, then $f_\omega \colon v \mapsto \langle v, f_\omega\rangle = \langle \mathrm{e}_\omega v, f\rangle$ (where $\mathrm{e}_\omega(t) = \mathrm{e}^{\omega t}$) is a bounded linear functional on $L^p(J;U)$, hence this functional has a representation of the form

$$\langle v, f_\omega\rangle = \int_J \langle v(t), u^*_\omega(t)\rangle\, dt$$

for some $u^*_\omega \in L^q(J;U^*)$. Replacing $v \in L^p(J;U)$ by $u = \mathrm{e}_\omega v \in L^q_\omega(J;U)$ and u^*_ω by $u^* = \mathrm{e}_{-\omega} u^*_\omega \in L^q_{-\omega}(J;U^*)$ we get the desired representation

$$\begin{aligned}\langle u, f\rangle = \langle \mathrm{e}_{-\omega} u, f_\omega\rangle &= \int_J \langle \mathrm{e}^{-\omega t} u(t), u^*_\omega(t)\rangle\, dt \\ &= \int_J \langle u(t), \mathrm{e}^{-\omega t} u^*_\omega(t)\rangle\, dt = \int_J \langle u(t), u^*(t)\rangle\, dt.\end{aligned}$$

□

The representation in Lemma 3.5.9 is canonical in the sense that it 'independent of p and ω' in the following sense:

Lemma 3.5.10 *Let* U *be a reflexive Banach space. If* f *is a bounded linear functional on* $L^{p_1}_{\omega_1}(J;U) \cap L^{p_2}_{\omega_2}(J;U)$*, where* $1 \leq p_1 < \infty$*,* $1 \leq p_2 < \infty$*,* $\omega_1 \in \mathbb{R}$*, and* $\omega_2 \in \mathbb{R}$*, then we get the same representing function* u^* *for* f *if we use*

[4] In part (i) the reflexivity assumption on U can be weakened to the assumption that U has the Radon–Nikodym property. See Diestel and Uhl (1977, Theorem 1, p. 98).

any combination of p_i *and* ω_j*,* $i,\ j = 1, 2$ *in Lemma 3.5.9. In particular,* $u^* \in L^{q_1}_{-\omega_1}(J;U^*) \cap L^{q_2}_{-\omega_2}(J;U^*)$*, where* $1/p_1 + 1/q_1 = 1$ *and* $1/p_2 + 1/q_2 = 1$*.*

Proof This follows from the fact that the integral $\int_J \langle u(t), u^*(t)\rangle\, dt$ does not depend on either p or ω (as long as it converges absolutely). □

Example 3.5.11 *Let* U *be a reflexive Banach space, let* $1 \le p < \infty$*,* $1/p + 1/q = 1$ *(with* $1/\infty = 0$*), and* $\omega \in \mathbb{R}$*.*

(i) *The dual of the bilateral left shift group* τ^t*,* $t \in \mathbb{R}$*, on* $L^p_\omega(\mathbb{R};U)$ *is the right shift group* τ^{-t}*,* $t \in \mathbb{R}$*, which acts on* $L^q_{-\omega}(\mathbb{R};U^*)$ *if* $1 < p < \infty$ *and on* $BUC_{-\omega}(\mathbb{R};U^*)$ *if* $p = 1$*.*

(ii) *The dual of the incoming left shift semigroup* τ^t_+*,* $t \ge 0$*, on* $L^p_\omega(\mathbb{R}^+;U)$ *is the right shift semigroup*

$$(\tau_+^{-t}u)(t) = (\tau^{-t}\pi_+u)(s) = \begin{cases} u(s-t), & s > t,\\ 0, & \textit{otherwise},\end{cases}$$

which acts on $L^q_{-\omega}(\mathbb{R}^+;U^*)$ *if* $1 < p < \infty$ *and on* $\{u^* \in BUC_{-\omega}(\overline{\mathbb{R}}^+;U^*) \mid u^*(0) = 0\}$ *if* $p = 1$*.*

(iii) *The dual of the outgoing left shift semigroup* τ^t_-*,* $t \ge 0$*, on* $L^p_\omega(\mathbb{R}^-;U)$ *is the right shift semigroup*

$$(\tau_-^{-t}u)(s) = (\pi_-\tau^{-t}u)(s) = \begin{cases} u(s-t), & s \le 0,\\ 0, & \textit{otherwise},\end{cases}$$

which acts on $L^q_{-\omega}(\mathbb{R}^-;U^*)$ *if* $1 < p < \infty$ *and on* $BUC_{-\omega}(\overline{\mathbb{R}}^-;U^*)$ *if* $p = 1$*.*

(iv) *The dual of the finite left shift semigroup* $\tau^t_{[0,T)}$*,* $t \ge 0$*, on* $L^p_\omega((0,T);U)$ *is the right shift semigroup*

$$(\tau_{[0,T)}^{-t}u)(s) = (\pi_{[0,T)}\tau^{-t}\pi_{[0,T)}u)(s) = \begin{cases} u(s-t), & t \le s < T,\\ 0, & \textit{otherwise},\end{cases}$$

which acts on $L^q((0,T);U^*)$ *if* $1 < p < \infty$ *and on* $\{u^* \in C([0,T];U^*) \mid u^*(0) = 0\}$ *if* $p = 1$*.*

(v) *The dual of the circular left shift group* $\tau^t_{\mathbb{T}_T}$*,* $t \ge 0$*, on* $L^p_\omega(\mathbb{T}_T;U)$ *is the circular right shift group*

$$(\tau_{\mathbb{T}_T}^{-t}u)(s) = (\tau^{-t}u)(s) = u(s-t),$$

which acts on $L^q(\mathbb{T}_T;U^*)$ *if* $1 < p < \infty$ *and on* $C(\mathbb{T}_T;U^*)$ *if* $p = 1$*.*

Proof (i) By Lemma 3.5.9, the dual of $L^p_\omega(\mathbb{R};U)$ is $L^q_{-\omega}(\mathbb{R};U^*)$. Let $u \in L^p_\omega(\mathbb{R};U)$ and $u^* \in L^q_{-\omega}(\mathbb{R};U^*)$ and $t \in \mathbb{R}$. Then

$$\begin{aligned}\langle \tau^t u, u^*\rangle &= \int_{-\infty}^{\infty} \langle \tau^t u(s), u^*(s)\rangle\, ds = \int_{-\infty}^{\infty} \langle u(s+t), u^*(s)\rangle\, ds \\ &= \int_{-\infty}^{\infty} \langle u(s), u^*(s-t)\rangle\, ds = \int_{-\infty}^{\infty} \langle u(s), \tau^{-t}u^*(s)\rangle\, ds \\ &= \langle u, \tau^{-t}u^*\rangle.\end{aligned}$$

This shows that $\tau^{*t} = \tau^{-t}$. The rest of (i) follows from Theorem 3.5.6, Definition 3.5.7, and Examples 2.3.2 and 2.5.3.

(ii)–(iv) This follows from (i) and Examples 2.3.2 and 2.5.3. □

The new right shift semigroups that we obtained in Example 3.5.11 are similar to the left shift semigroups that we have encountered earlier. The similarity transform is the *reflection operator* $\mathbf{Я}$ (in one case combined with a shift), which we define as follows.

Definition 3.5.12 Let $1 \le p \le \infty$, and let U be a Banach space.

(i) For each function $u \in L^p_{\text{loc}}(\mathbb{R};U)$ we define the *reflection* $\mathbf{Я}u$ of u by

$$(\mathbf{Я}u)(s) = u(-s), \quad s \in \mathbb{R}. \tag{3.5.5}$$

(ii) For each function $u \in Reg_{\text{loc}}(\mathbb{R};U)$ we define the *reflection* $\mathbf{Я}u$ of u by

$$(\mathbf{Я}u)(s) = \lim_{t\downarrow -s} u(t), \quad s \in \mathbb{R}. \tag{3.5.6}$$

Observe that these two cases are consistent in the sense that in part (ii) we have $(\mathbf{Я}u)(s) = u(-s)$ for all but countably s.

Lemma 3.5.13 *Let $J \subset \mathbb{R}$, $t \in \mathbb{R}$, $\omega \in \mathbb{R}$, and $1 \le p \le \infty$.*

(i) $\mathbf{Я}$ *maps $L^p|Reg_\omega(\mathbb{R};U)$ onto $L^p|Reg_{-\omega}(\mathbb{R};U)$, and*
  (a) $\mathbf{Я}^{-1} = \mathbf{Я}$,
  (b) $\mathbf{Я}\tau^t = \tau^{-t}\mathbf{Я}$,
  (c) $\mathbf{Я}\pi_J = \pi_{\mathbf{Я}J}\mathbf{Я}$,[5] *and*
  (d) $\mathbf{Я}^* = \mathbf{Я}$ *(in $L^p_\omega(J;U)$ with reflexive U and $1 \le p < \infty$).*

(ii) $\pi_J^* = \pi_J$ *(in $L^p_\omega(J;U)$ with reflexive U and $1 \le p < \infty$).*

(iii) *The dual of the time compression operator γ_λ (see Example 2.3.6) is the inverse time compression operator $\gamma_{1/\lambda}$ (in $L^p_\omega(J;U)$ with reflexive U and $1 \le p < \infty$).*

[5] In the *Reg*-well-posed case we require χ_J to be right-continuous and define $\mathbf{Я}J$ to be the set whose characteristic function is $\chi_{\mathbf{Я}J}$

(iv) *The right shift (semi)groups in Example 3.5.11 are similar to the corresponding left shifts (semi)groups in Examples 2.3.2 and 2.5.3 as follows:*
   (a) $\tau^{\odot} = \mathbf{Я}\tau\mathbf{Я}$;
   (b) $\tau_+^{\odot} = \mathbf{Я}\tau_-\mathbf{Я}$;
   (c) $\tau_-^{\odot} = \mathbf{Я}\tau_+\mathbf{Я}$;
   (d) $\tau_{[0,T)}^{\odot} = \tau^{-T}\mathbf{Я}\tau_{[0,T)}\mathbf{Я}\tau^{T}$;
   (e) $\tau_{\mathbb{T}_T}^{\odot} = \mathbf{Я}\tau_{\mathbb{T}_T}\mathbf{Я}$.

We leave the easy proof to the reader.

3.6 The rigged spaces induced by the generator

In our subsequent theory of $L^p|Reg$-well-posed linear systems we shall need a scale of spaces X_n, $n = 0, \pm 1, \pm 2, \ldots$, which are constructed from X by means of the semigroup generator A. In particular, the spaces X_1 and X_{-1} will be of fundamental importance. To construct these spaces we need not even assume that A generates a C_0 semigroup on X; it is enough if A has a nonempty resolvent set and dense domain.

We begin with the case $n \geq 0$, and define

$$X_0 = X, \qquad X_n = \mathcal{D}\left(A^n\right) \text{ for } n = 1, 2, 3, \ldots.$$

Choose an arbitrary number α from the resolvent set of A. Then $(\alpha - A)^{-n}$ maps X one-to-one onto $\mathcal{D}(A^n)$ (this can be proved by induction over n), and we can define a norm in X_n by

$$|x|_n = |x|_{X_n} = \left|(\alpha - A)^n x\right|_X.$$

With this norm each X_n becomes a Banach space, $X_{n+1} \subset X_n$ with a dense injection, and $(A - \alpha)^n$ is an isometric (i.e., norm-preserving) isometry (i.e., bounded linear operator with a bounded inverse) from X_n onto X. If X is a Hilbert space, then so are all the spaces X_n.

If we replace α by some other $\beta \in \rho(A)$, then $(\beta - A)^{-n}$ has the same range as $(\alpha - A)^{-n}$, so if we use β instead of α in the definition of X_n then we still get the same space, but with a different norm. However, the two norms are equivalent since $(\alpha - A)^n(\beta - A)^{-n}$ is an isomorphism (not isometric) on X: for $n = 1$ this follows from the resolvent formula in Lemma 3.2.8(i) which gives

$$(\alpha - A)(\beta - A)^{-1} = 1 + (\alpha - \beta)(\beta - A)^{-1},$$

and by iterating this formula we get the general case. Most of the time the value of $\alpha \in \rho(A)$ which determines the exact norm in X_n is not important.

If A generates a C_0 semigroup $\mathfrak{A}$, then the restriction $\mathfrak{A}_{|X_n}$ of $\mathfrak{A}$ to X_n is a C_0 semigroup on X_n. It follows from Theorem 3.2.1(iii) and Example 3.2.6(i) that $\mathfrak{A}_{|X_n} = (\alpha - A)^{-n}\mathfrak{A}(\alpha - A)^n$, i.e., $\mathfrak{A}_{|X_n}$ and $\mathfrak{A}$ are (isometrically) isometric. Thus, all the important properties of these semigroups are identical. In particular, they all have the same growth bound $\omega_{\mathfrak{A}}$, and the generator of $\mathfrak{A}_n$ is the restriction $A_{|X_{n+1}}$ of A to X_{n+1}. In the sequel we occasionally write (for simplicity) $\mathfrak{A}$ instead of $\mathfrak{A}_{|X_n}$ and A instead of $A_{|X_{n+1}}$ (but we still use the more complicated notions in those cases where the distinction is important).

It is also possible to go in the opposite direction to get spaces X_n with negative index n. This time we first define a sequence of weaker norms in X, namely

$$|x|_{-n} = \left|(\alpha - A)^{-n}x\right|_X \text{ for } n = 1, 2, 3, \ldots,$$

and let X_{-n} be the completion of X with respect to the norm $|\cdot|_{-n}$. Then $(\alpha - A)^n$ has a unique extension to an isometric operator which maps X onto X_{-n}. We denote this operator by $(\alpha - A)^n_{|X}$ and it inverse by $(\alpha - A)^{-n}_{|X_{-n}}$, or sometimes simply by $(\alpha - A)^n$ respectively $(\alpha - A)^{-n}$ if no confusion is likely to arise. In the case $n = 1$ we often write $(\alpha - A_{|X})^{-1}$ instead of $(\alpha - A)^{-1}_{|X_{-1}}$. Thus, for all $n, l = 0, \pm 1, \pm 2, \ldots,$

$$(\alpha - A)^l_{|X_{n+l}} \text{ is an isometry of } X_{n+l} \text{ onto } X_n.$$

If A generates a C_0 semigroup $\mathfrak{A}$ on X, then we can use the formula

$$\mathfrak{A}_{|X_{-n}} = (\alpha - A)^n_{|X}\mathfrak{A}(\alpha - A)^{-n}_{|X_{-n}}$$

to extend (rather than restrict) $\mathfrak{A}$ to a semigroup on each of the spaces X_{-n}. In this way we get a full scale of spaces $X_{n+1} \subset X_n$ for $n = 0, \pm 1, \pm 2, \ldots,$ and a corresponding scale of isometric semigroups $\mathfrak{A}_{|X_{-n}}$. In places where no confusion is likely to arise we abbreviate $\mathfrak{A}_{|X_{-n}}$ to $\mathfrak{A}$. The generator of $\mathfrak{A}_{|X_{-n}}$ is $A_{|X_{-n+1}}$. As in the case of the semigroup itself we sometimes abbreviate $A_{|X_{-n+1}}$ to A.

Above we have defined the norm in X_1 by using the fact that $(\alpha - A)^{-1}$ maps X one-to-one onto X_1 whenever $\alpha \in \rho(A)$. Another commonly used norm in X_1 is the graph norm

$$\|x\|_{X_1} = \left(|x|^2_X + |Ax|^2_X\right)^{1/2}. \tag{3.6.1}$$

This is the restriction of the norm $\|\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\| = \left(|x|^2_X + |y|^2_X\right)^{1/2}$ in $\left[\begin{smallmatrix} X \\ X \end{smallmatrix}\right]$ to the graph $\mathcal{G}(A) = \left\{\left[\begin{smallmatrix} Ax \\ x \end{smallmatrix}\right] \;\middle|\; x \in X\right\}$. This graph is closed since A is closed, so it is a Banach space in itself (or a Hilbert space if X is a Hilbert space). The map which takes $\left[\begin{smallmatrix} Ax \\ x \end{smallmatrix}\right] \in \mathcal{G}(A)$ into x is injective, so we may let $x \in \mathcal{D}(A)$ inherit the norm of $\left[\begin{smallmatrix} Ax \\ x \end{smallmatrix}\right] \in \mathcal{G}(A)$, and this is the norm $\|\cdot\|_{X_1}$ in (3.6.1). This norm is majorized by the earlier introduced norm $|\cdot|_{X_1}$ since

$$|Ax|_X = |(A - \alpha + \alpha)x|_X \le \|x\|_{X_1} + |\alpha||x|_X,$$

and

$$|x|_X = |(\alpha - A)^{-1}(\alpha - A)x|_X \le \|(\alpha - A)^{-1}\||x|_{X_1},$$

so by the open mapping theorem, the two norms $|\cdot|_{X_1}$ and $\|\cdot\|_{X_1}$ are equivalent.

A similar norm can be used in $X_n = \mathcal{D}(A^n)$ for $n = 2, 3, \ldots$, namely

$$\|x\|_{X_n} = \left(|x|_X^2 + |A^n x|_X^2\right)^{1/2}. \tag{3.6.2}$$

To prove that this is a norm in X_n we can argue as above: the operator A^n is closed since it is the restriction of $(A^n)_{|X} := A_{|X_{-n+1}} A_{|X_{-n+2}} \cdots A_{|X} \in \mathcal{B}(X; X_{-n})$ to its natural domain $\mathcal{D}(A^n) = \{x \in X \mid (A^n)_{|X} x \in X\}$, and the above norm is the graph norm of A^n on $\mathcal{D}(A^n)$. To show that it is equivalent to the norm $|\cdot|_{X_n}$ we may argue as follows. Take some $\alpha \in \rho(A)$. Then, for each $x \in \mathcal{D}(A^n)$ we have (from the binomial formula)

$$(\alpha - A)^n x = \sum_{k=0}^{n} \binom{n}{k} \alpha^k A^{n-k} x,$$

or equivalently,

$$\begin{aligned} A^n x &= (\alpha - A)^n x - \sum_{k=1}^{n} \binom{n}{k} \alpha^k A^{n-k} x \\ &= \left(1 - \sum_{k=1}^{n} \binom{n}{k} \alpha^k A^{n-k} (\alpha - A)^{-n}\right)(\alpha - A)^n x. \end{aligned}$$

Thus, $|A^n x|_X \le M|x|_{X_n}$, where M is the norm of the operator $1 - \sum_{k=1}^{n} \binom{n}{k} \alpha^k A^{n-k}(\alpha - A)^{-n} \in \mathcal{B}(X)$, and, of course,

$$|x|_X = |(\alpha - A)^{-n}(\alpha - A)^n x|_X \le \|(\alpha - A)^{-n}\||x|_{X_n}.$$

This shows that the norm $\|\cdot\|_{X_n}$ is majorized by the norm $|\cdot|_{X_n}$, so by the open mapping theorem, the two norms $|\cdot|_{X_n}$ and $\|\cdot\|_{X_n}$ are equivalent.

Let us illustrate these constructions by looking at Example 3.3.5. In this example we have

$$|x|_{X_n}^2 = \sum_{k=1}^{\infty} |\alpha - \lambda_k|^{2n} |\langle x, \phi_k \rangle|_H^2,$$

where $\alpha \in \rho(A)$, and each X_n is a Hilbert space with the orthogonal basis $\{\phi_n\}_{n=1}^{\infty}$ (it becomes orthonormal if we divide ϕ_k by $|\alpha - \lambda_k|$). For $n \ge 1$ we can alternatively use the equivalent norm

$$|x|_{X_n}^2 = \sum_{k=1}^{\infty}(1 + |\lambda_k|)^{2n}|\langle x, \phi_k\rangle|_H^2;$$

cf. Example 3.3.5.

Remark 3.6.1 This remark explains how the spaces X_n interact with duality. Since $X_{n+1} \subset X_n$ for all $n = 0, \pm 1, \pm 2, \ldots$, with dense embeddings, the duals of these embedding maps are injective (see Lemma 9.10.2(ii)), so they define embeddings $(X_n)^* \subset (X_{n+1})^*$ (which need not be dense). Since $(\alpha - A)^n$ is an isometry of X_{n+l} onto X_l, it follows that $(\alpha - A^*)^n$ is an isometry of $(X_l)^*$ onto $(X_{n+l})^*$ for all $n, l = 0, \pm 1, \pm 2, \ldots$. If X is reflexive, then the embeddings $(X_n)^* \subset (X_{n+1})^*$ are dense, and these spaces are the same that we would get by repeating the argument leading to the definition of the spaces X_n, with X replaced by X^*, A replaced by A^*, and using a different subindex (i.e., $-n$ instead of n). When we discuss the causal and anti-causal dual systems Σ^d and $\Sigma^\dagger$ it is convenient to denote the domain of A^* by X_1^*, and accordingly, in the sequel we use the notation

$$X_{-n}^* := (X^*)_{-n} := (X_n)^*, \quad n = 0, \pm 1, \pm 2, \ldots.$$

In particular,

$$\langle A^n x, x^*\rangle_{(X_l, X_{-l}^*)} = \langle x, A^{*n} x^*\rangle_{(X_{n+l}, X_{-(n+l)}^*)}, \quad x \in X_{n+l} \quad x^* \in X_{-l}^*,$$

where by A^{*n} we mean $A^{*n} := (A^*)^n = (A^n)^*$.

In the Hilbert space case one often uses a slightly different construction, which resembles the one described in Remark 3.6.1. Assume that $W \subset X$ are two Hilbert spaces, with a continuous and dense embedding. Then $(x, y) \mapsto \langle x, y\rangle_X$ is a bounded sesquilinear form on W, and therefore (see, e.g., Kato (1980, pp. 256–257)) there is a unique operator $E \in \mathcal{B}(W)$ which is positive and self-adjoint (with respect to the inner product of W) such that

$$\langle x, y\rangle_X = \langle Ex, y\rangle_W = \langle x, Ey\rangle_W = \langle \sqrt{E}x, \sqrt{E}y\rangle_W, \qquad x, y \in W,$$

where $\sqrt{E}$ is the positive self-adjoint square root of E (cf. Lemma A.2.2). For all $x \in W$,

$$|Ex|_X^2 = \langle E\sqrt{E}x, E\sqrt{E}x\rangle_W \le \|E\|_{\mathcal{B}(W)}^2 |\sqrt{E}x|_W^2 = \|E\|_{\mathcal{B}(W)}^2 |x|_X^2,$$

and this implies that E can be extended to a unique operator in $\mathcal{B}(X)$, which we still denote by the same symbol E. This operator is still self-adjoint in X since $\langle x, Ey\rangle_X = \langle Ex, Ey\rangle_W = \langle Ex, y\rangle_X$ for all $x, y \in W$, and W is dense in X. The space X may be regarded as the completion of W with respect to the norm $|x|_X = |\sqrt{E}x|_W$, and this means that the extended version of $\sqrt{E}$ is an isometric isomorphism of W onto X.

Let V be the completion of X with respect to the norm $|x|_V = |\sqrt{E}x|_X$. By repeating the same argument that we gave above with W replaced by X and X replaced by V we find that E can be extended to a self-adjoint operator in V (which we still denote by the same letter), that $\sqrt{E}$ is an isometric isomorphism of V onto X, and that E is an isometric isomorphism of V onto W. Moreover,

$$\langle x, y\rangle_V = \langle Ex, y\rangle_X = \langle x, Ey\rangle_X = \langle \sqrt{E}x, \sqrt{E}y\rangle_X, \qquad x, y \in X,$$
$$\langle x, y\rangle_V = \langle Ex, y\rangle_X = \langle Ex, Ey\rangle_W, \qquad x, y \in W.$$

The space V can be interpreted as the dual of W with X as pivot space as follows. Every $x \in V$ induces a bounded linear functional on W through the formula

$$\langle x, y\rangle_{(V,W)} = \langle Ex, y\rangle_W,$$

and every bounded linear functional on W is of this type since E maps W one-to-one onto V. This is a norm-preserving mapping of the dual of W onto V, since the norm of the above functional is $|Ex|_W = |x|_V$. That X is a pivot space means that for all $x \in X$ and $y \in W$,

$$\langle x, y\rangle_{(V,W)} = \langle x, y\rangle_X,$$

which is true since both sides are equal to $\langle Ex, y\rangle_W$.

If we apply this procedure (in the Hilbert space case) to the space $X_1 \subset X$ described at the beginning of this section, then we get $V = X^*_{-1}$ and the extended version of E is given by $E = (\alpha - A)^{-1}(\overline{\alpha} - A^*_{|X})^{-1}$ if we use the norm $|x|_1 = |(\alpha - A)x|_X$ in X_1. If we instead use the graph norm $|x|_1^2 = |x|_X^2 + |Ax|_X^2$ in X_1, then the extended version of E is given by $E = (1 + A^*_{|X}A)^{-1}$.

In this book we shall usually identify X with its dual, and identify the dual of W with V as described above. However, occasionally it is important to compute the dual of an operator with respect to the inner product in W or in V instead of computing it with respect to the inner produce in X. Here the following result is helpful.

Proposition 3.6.2 *Let U, Y, and $W \subset X \subset V$ be Hilbert spaces, where the embeddings are continuous and dense, let $E \in \mathcal{B}(V)$ be injective, selfadjoint (with respect to the inner product in V), and suppose that $\sqrt{E}$ maps V isometrically onto X and that $\sqrt{E}_{|X}$ maps X isometrically onto W (the operator E and the space V can be constructed starting from W and X as explained above). We identify U and Y with their duals.*

(i) *Let $B \in \mathcal{B}(U; W)$, let $B' \in \mathcal{B}(W; U)$ be the adjoint of B with respect to the inner product in W, and let $B^* \in \mathcal{B}(X; U)$ be the adjoint of B with* *respect to the inner product in* X *(note that* $B \in \mathcal{B}(U; X)$*). Then* $B^* = B'E_{|X}$*. In particular, this formula can be used to extend* B^* *to* $B'E \in \mathcal{B}(V; U)$*, which is the adjoint of* B *when we identify the dual of* W *with* V*.*

(ii) *Let* $B \in \mathcal{B}(U; X)$*, let* $B^* \in \mathcal{B}(X; U)$ *be the adjoint of* B *with respect to the inner product in* X*, and let* $B'' \in \mathcal{B}(V; U)$ *be the adjoint of* B *with respect to the inner product in* V *(note that* $B \in \mathcal{B}(U; V)$*). Then* $B'' = B^*E$*.*

(iii) *Let* $B \in \mathcal{B}(U; V)$*, let* $B^* \in \mathcal{B}(W; U)$ *be the adjoint of* B *when we identify the dual of* V *with* W *(with* X *as pivot space), and let* $B'' \in \mathcal{B}(V; U)$ *be the adjoint of* B *with respect to the inner product in* V*. Then* $B'' = B^*E$*.*

(iv) *Let* $C \in \mathcal{B}(V; Y)$*, let* $C'' \in \mathcal{B}(Y; V)$ *be the adjoint of* C *with respect to the inner product in* V*, and let* $C^* \in \mathcal{B}(Y; X)$ *be the adjoint of* C *with respect to the inner product in* X *(note that* $C \in \mathcal{B}(X; Y)$*. Then* $C^* = EC''$*. In particular,* $C^* \in \mathcal{B}(Y; W)$*.*

(v) *Let* $C \in \mathcal{B}(X; Y)$*, let* $C^* \in \mathcal{B}(Y; X)$ *be the adjoint of* C *with respect to the inner product in* X*, and let* $C' \in \mathcal{B}(Y; W)$ *be the adjoint of* C *with respect to the inner product in* W *(note that* $C \in \mathcal{B}(W; Y)$*. Then* $C' = EC^*$*.*

(vi) *Let* $C \in \mathcal{B}(W; Y)$*, let* $C' \in \mathcal{B}(Y; W)$ *be the adjoint of* C *with respect to the inner product in* W*, and let* $C^* \in \mathcal{B}(Y; V)$ *be the adjoint of* C *when we identify the dual of* W *with* V *(with* X *as pivot space). Then* $C' = EC^*$*.*

(vii) *Let* $A \in \mathcal{B}(V)$*, and suppose that* X *is invariant under* A*. Let* $A'' \in \mathcal{B}(W)$ *be the adjoint of* A *with respect to the inner product of* V*, and let* $A^*_{|X}$ *be the adjoint of* $A_{|X}$ *with respect to the inner produce of* X*. Then* $EA'' = A^*_{|X}E$*. In particular,* W *is invariant under* $A^*_{|X}$*.*

(viii) *Let* $A \in \mathcal{B}(X)$*, and suppose that* W *is invariant under* A*. Let* A^* *be the adjoint of* A *with respect to the inner product of* X*, and let* $A'_{|W}$ *be the adjoint of* $A_{|W}$ *with respect to the inner produce of* W*. Then* $A'_{|W}E_{|X} = EA^*$*. In particular,* EX *is invariant under* $A'_{|W}$*.*

(ix) *Let* $A \in \mathcal{B}(W)$*, let* $A' \in \mathcal{B}(W)$ *be the adjoint of* A *with respect to the inner product in* W*, and let* $A^* \in \mathcal{B}(V)$ *be the adjoint of* A *when we identify the dual of* W *with* V*. Then* $A'E = EA^*$*.*

Proof (i) For all $x \in X$ and $u \in U$,

$$\langle u, B^*x\rangle_U = \langle Bu, x\rangle_X = \langle Bu, Ex\rangle_W = \langle u, B'Ex\rangle_U.$$

Thus, $B^* = B'E_{|X}$. If we instead let B^* stand for the adjoint of B when we identify the dual of W by V, then for all $u \in U$ and $x \in V$,

$$\langle B^*x, u\rangle_U = \langle x, Bu\rangle_{(V,W)} = \langle Ex, Bu\rangle_W = \langle B'Ex, u\rangle_U.$$

Thus, $B^* = B'E$.

(ii) Apply (i) with W replaced by X and X replaced by V.

(iii) For all $x \in V$ and $u \in U$,

$$\langle u, B''x\rangle_U = \langle Bu, x\rangle_V = \langle EBu, Ex\rangle_W = \langle Bu, Ex\rangle_{(V,W)} = \langle u, B^*Ex\rangle_U.$$

Thus $B'' = B^*E$.

(iv) For all $x \in X$ and $y \in Y$,

$$\langle x, C^*y\rangle_X = \langle Cx, y\rangle_Y = \langle x, C''y\rangle_V = \langle x, EC''y\rangle_X.$$

Thus $C^* = EC''$.

(v) Apply (iv) with V replaced by X and X replaced by W.

(vi) For all $x \in W$ and $y \in Y$,

$$\langle C'y, x\rangle_W = \langle y, Cx\rangle_Y = \langle C^*y, x\rangle_{(V,W)} = \langle EC^*y, x\rangle_W.$$

Thus, $C' = EC^*$.

(vii) For all $x \in X$ and $y \in V$,

$$\begin{aligned}\langle x, EA''y\rangle_V &= \langle AEx, y\rangle_V = \langle A_{|X}Ex, Ey\rangle_X = \langle x, EA^*_{|X}Ey\rangle_X \\ &= \langle x, A^*_{|X}Ey\rangle_V.\end{aligned}$$

Thus, $EA'' = A^*_{|X}E$ on V. This implies that W is invariant under $A^*_{|X}$.

(viii) Apply (vii) with V replaced by X and X replaced by W.

(ix) For all $x \in V$ and $y \in W$,

$$\begin{aligned}\langle A'Ex, y\rangle_W &= \langle Ex, Ay\rangle_W = \langle x, Ay\rangle_{(V,W)} = \langle A^*x, y\rangle_{(V,W)} \\ &= \langle EA^*x, y\rangle_W.\end{aligned}$$

Thus $A'E = EA^*$. □

3.7 Approximations of the semigroup

The approximation A_α to A that we used in the proof of the Hille–Yosida Theorem 3.4.1 will be quite useful in the sequel, too. For later use, let us record some of the properties of this and some related approximations:

Theorem 3.7.1 *Let A be the generator of a C_0 semigroup on X. Define the space $X_1 = \mathcal{D}(A)$ as in Section 3.6. For all $\alpha \in \rho(A)$ (in particular, for all $\alpha \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$), define*

$$J_\alpha = \alpha(\alpha - A)^{-1}, \qquad A_\alpha = \alpha A(\alpha - A)^{-1},$$

and for all $h > 0$ and $x \in X$, define

$$J^h x = \frac{1}{h}\int_0^t \mathfrak{A}^s x\,ds, \qquad A^h = \frac{1}{h}(\mathfrak{A}^h - 1)x.$$

Then the following claims are true:

(i) *For all $\alpha \in \rho(A)$ and $h > 0$, $J_\alpha \in \mathcal{B}(X; X_1)$, $J^h \in \mathcal{B}(X; X_1)$, $A_\alpha \in \mathcal{B}(X)$, $A^h \in \mathcal{B}(X)$, and*

$$\begin{aligned} J_\alpha &= \alpha(\alpha - A)^{-1} = 1 + A(\alpha - A)^{-1}, \\ A_\alpha &= AJ_\alpha = \alpha(J_\alpha - 1) = \alpha^2(\alpha - A)^{-1} - \alpha, \\ A^h &= AJ^h = \frac{1}{h}(\mathfrak{A}^h - 1)x = \frac{1}{h}A\int_0^t \mathfrak{A}^s x\,ds, \end{aligned}$$

Moreover, for $\alpha \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$,

$$J_\alpha x = \alpha\int_0^\infty \mathrm{e}^{-\alpha s}\mathfrak{A}^s x\,ds, \qquad x \in X.$$

(ii) *For all α, $\beta \in \rho(A)$ and h, k, $t > 0$, the operators J_α, J_β, J^h, J^k, A_α, A_β, A^h, A^k, and $\mathfrak{A}^t$ commute with each other.*

(iii) *J_α and J^h approximate the identity and A_α and A^h approximate A in the sense that the following limits exist:*

$$\begin{aligned} \lim_{\alpha\to+\infty} J_\alpha x &= \lim_{h\downarrow 0} J^h x = x && \text{in } X \text{ for all } x \in X, \\ \lim_{\alpha\to+\infty} A_\alpha x &= \lim_{h\downarrow 0} A^h x = Ax && \text{in } X \text{ for all } x \in X_1, \\ \lim_{\alpha\to+\infty} \alpha^{-1} J_\alpha x &= \lim_{\alpha\to+\infty} (\alpha - A)^{-1}x = 0 && \text{in } X_1 \text{ for all } x \in X, \\ \lim_{h\downarrow 0} hJ^h x &= \lim_{h\downarrow 0}\int_0^t \mathfrak{A}^s x\,dx = 0 && \text{in } X_1 \text{ for all } x \in X, \\ \lim_{\alpha\to+\infty} \alpha^{-1}A_\alpha x &= \lim_{h\downarrow 0} hA^h x = 0 && \text{in } X \text{ for all } x \in X. \end{aligned}$$

(iv) *$\mathfrak{A}$ is uniformly continuous (hence analytic) iff J_α has a bounded inverse for some $\alpha \in \rho(A)$, or equivalently, iff J^h has a bounded inverse for some $h > 0$.*

Proof (i) Obviously, $J_\alpha \in \mathcal{B}(X; X_1)$, $A_\alpha \in \mathcal{B}(X)$, and $A^h \in \mathcal{B}(X)$. By Theorem 3.2.1(ii), $J^h \in \mathcal{B}(X; X_1)$. The algebraic properties in (i) are easy to verify (see also Theorem 3.2.1(ii)). The integral formula for J_α is found in Theorem 3.2.9(i).

(ii) This is true since $\mathfrak{A}^s$ commutes with $\mathfrak{A}^t$ and with $(\alpha - A)^{-1}$. See also Theorems 3.2.1 and 3.2.9.

(iii) That $\lim_{\alpha\to+\infty} J_\alpha x = \lim_{h\downarrow 0} J^h x = x$ in X for all $x \in X$ follows from Theorems 3.2.1(i) and 3.2.9(iii) and that $\lim_{\alpha\to+\infty} A_\alpha x = \lim_{h\downarrow 0} A^h x = Ax$ in X for all $x \in X_1$ follows from (i), Definition 3.1.1, and Theorem 3.2.9(iii). By Theorem 3.2.9(iii), for all $\beta \in \rho(A)$, $\lim_{\alpha\to+\infty}(\beta - A)(\alpha - A)^{-1}x = 0$ in X for all $x \in X$, and this implies that $\lim_{\alpha\to+\infty} \alpha^{-1} J_\alpha x = 0$ in X_1 for all $x \in X$. To prove that $\lim_{h\downarrow 0} hJ^h x = 0$ in X_1 for all $x \in X$ it suffices to observe that, for all $\beta \in \rho(A)$,

$$(\beta - A)hJ^h x = (\beta - A)\int_0^h \mathfrak{A}^s x\,ds = \beta\int_0^h \mathfrak{A}^s x + x - \mathfrak{A}^h x,$$

and here the right-hand side tends to zero in X for every $x \in X$. That $\lim_{\alpha\to+\infty} \alpha^{-1} A_\alpha x = \lim_{h\downarrow 0} hA^h x = 0$ in X for all $x \in X$ follows from (i) and the fact that $\lim_{\alpha\to+\infty} J_\alpha x = \lim_{h\downarrow 0} J^h x = x$ in X for all $x \in X$.

(iv) Obviously $A \in \mathcal{B}(X)$ iff J_α has a bounded inverse. That J^h has a bounded inverse for some $h > 0$ iff $A \in \mathcal{B}(X)$ follows from Example 3.1.2 and Remark 3.1.4. By Theorem 3.1.3, the boundedness of A is equivalent to the uniform continuity of $\mathfrak{A}$. □

Definition 3.7.2 The operators J_α and A_α in Theorem 3.7.1 are called the Yosida (or Abel) approximations of the identity 1 and of A, respectively (with parameter α). The operators J^h and A^h in Theorem 3.7.1 are called the Cesàro approximations (or order one) of the identity 1 and of A, respectively (with parameter h).

Theorem 3.7.3 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ on X, and let $A_\alpha = \alpha A(\alpha - A)^{-1}$ be the Yosida approximation of A. Then for each $x \in X$ and $t \ge 0$, $\lim_{\alpha\to+\infty} \mathrm{e}^{A_\alpha t}x = \mathfrak{A}^t x$, and the convergence is uniform in t on any bounded interval.*

The proof of this theorem is contained in the proof of Theorem 3.4.1.

The same result is true if we replace the Yosida approximation by the Cesàro approximation:

Theorem 3.7.4 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ on X, and let $A_h = \frac{1}{h}(\mathfrak{A}^h - 1)$ be the Cesàro approximation of A. Then for each $x \in X$ and $t \ge 0$, $\lim_{h\downarrow 0} \mathrm{e}^{A^h t}x = \mathfrak{A}^t x$, and the convergence is uniform in t on any bounded interval.*

Proof The proof follows the same lines as the proof of Theorem 3.4.1 with A_α replaced by A^h, $\mathfrak{A}^t_\alpha$ replaced by $\mathfrak{A}^t_h = \mathrm{e}^{A^h t}$, and B_α replaced by $B^h = A^h + \frac{1}{h} = \frac{1}{h}\mathfrak{A}^h$. We can choose M and ω so that $\|\mathfrak{A}^t\| \le M\mathrm{e}^{\omega t}$ for all $t \ge 0$. Then (3.4.1) is replaced by

$$\|(B^h)^n\| = \Big\|\Big(\frac{1}{h}\mathfrak{A}^h\Big)^n\Big\| \le \frac{M\mathrm{e}^{\omega hn}}{h^n},$$

and (3.4.3) is replaced by

$$\begin{aligned}\|\mathfrak{A}_h^t\| &\le \mathrm{e}^{-t/h}\sum_{n=0}^{\infty}\frac{t^n}{n!}\frac{M\mathrm{e}^{\omega hn}}{h^n}\\ &= M\mathrm{e}^{-t/h}\mathrm{e}^{(t/h)\mathrm{e}^{\omega h}} = M\mathrm{e}^{t/h(\mathrm{e}^{\omega h}-1)}, \qquad t \ge 0.\end{aligned}$$

This tends to $M\mathrm{e}^{\omega t}$ as $h \downarrow 0$, uniformly in t on any bounded interval. The new version of estimate (3.4.4) is (for all $h, k > 0$),

$$|\mathfrak{A}_h^t x - \mathfrak{A}_k^t x| \le M^2\int_0^t \mathrm{e}^{s/h(\mathrm{e}^{\omega h}-1)}\mathrm{e}^{(t-s)/k(\mathrm{e}^{\omega k}-1)}|A^h x - A^k x|\,ds,$$

and the remainder of the proof of Theorem 3.4.1 stays the same. □

Theorem 3.7.5 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ on X. Then, for all $t \ge 0$,*

$$\mathfrak{A}^t x = \lim_{n\to\infty}\Big(1 - \frac{t}{n}A\Big)^{-n} x, \qquad x \in X,$$

and the convergence is uniform in t on each bounded interval.

Proof By Theorem 3.2.9(i), for all $x \in X$ and $(n-1)/t > \omega_{\mathfrak{A}}$,

$$\begin{aligned}\Big(1 - \frac{t}{n}A\Big)^{-(n+1)} x &= \Big(\frac{n}{t}\Big)^{n+1}\Big(\frac{n}{t} - A\Big)^{-(n+1)} x\\ &= \Big(\frac{n}{t}\Big)^{n+1}\frac{1}{n!}\int_0^\infty s^n\mathrm{e}^{-ns/t}\mathfrak{A}^s x\,ds\\ &= \frac{n^{n+1}}{n!}\int_0^\infty \big(v\mathrm{e}^{-v}\big)^n\mathfrak{A}^{tv} x\,dv.\end{aligned}$$

As $\frac{n^{n+1}}{n!}\int_0^\infty (v\mathrm{e}^{-v})^n\,dv = 1$, this implies that

$$\begin{aligned}&\Big|\Big(1 - \frac{t}{n}A\Big)^{-(n+1)} x - \mathfrak{A}^t x\Big|\\ &\quad= \Big|\frac{n^{n+1}}{n!}\int_0^\infty \big(v\mathrm{e}^{-v}\big)^n(\mathfrak{A}^{tv}x - \mathfrak{A}^t x)\,dv\Big|\\ &\quad\le \frac{n^{n+1}}{n!}\int_0^\infty v^n\mathrm{e}^{-nv}\big|\mathfrak{A}^{tv}x - \mathfrak{A}^t x\big|\,dv.\end{aligned}$$

For each $T > 0$, the function $v \mapsto \mathfrak{A}^v$ is uniformly continuous on $[0, T]$. Thus, for every $\epsilon > 0$ it is possible to find a $\delta > 0$ such that $|\mathfrak{A}^{tv}x - \mathfrak{A}^t x| \le \epsilon$ for all $t \in [0, T]$ and $1 - \delta \le v \le 1 + \delta$. We split the integral above into three parts I_1, I_2, and I_3, over the intervals $[0, 1-\delta)$, $[1-\delta, 1+\delta)$, and $[1+\delta, \infty)$, respectively. Then

$$\left|\left(1-\frac{t}{n}A\right)^{-(n+1)}x - \mathfrak{A}^t x\right| = I_1 + I_2 + I_3.$$

The function $v \mapsto v\mathrm{e}^{-v}$ is increasing on $[0, 1]$, so we can estimate for all $t \in [0, T]$ (choose $M > 0$ and $\omega > 0$ so that $\|\mathfrak{A}^t\| \le M\mathrm{e}^{\omega t} \le M\mathrm{e}^{\omega T}$)

$$\begin{aligned}
I_1 &\le \frac{n^{n+1}((1-\delta)\mathrm{e}^{-(1-\delta)})^n}{n!}\int_0^{1-\delta}\left|\mathfrak{A}^{tv}x - \mathfrak{A}^t x\right| dv \\
&\le 2M\mathrm{e}^{\omega T}\ \frac{n^{n+1}((1-\delta)\mathrm{e}^{-(1-\delta)})^n}{n!}, \\
I_2 &\le \epsilon\ \frac{n^{n+1}}{n!}\int_{1-\delta}^{1+\delta}\left(v\mathrm{e}^{-v}\right)^n dv < \epsilon, \\
I_3 &= \frac{n^{n+1}}{n!}\int_{1+\delta}^{\infty}\left(v\mathrm{e}^{-v}\right)^n\left|\mathfrak{A}^{tv}x - \mathfrak{A}^t x\right| dv \\
&\le 2M\ \frac{n^{n+1}}{n!}\int_{1+\delta}^{\infty}\left(v\mathrm{e}^{-v}\right)^n \mathrm{e}^{\omega T v}\, dv \\
&= 2M\ \frac{n^{n+1}}{n!}\int_{1+\delta}^{\infty}\left(v\mathrm{e}^{-(1-(1+\omega T)/n)v}\right)^n \mathrm{e}^{-v}\, dv.
\end{aligned}$$

We recall Stirling's formula

$$\lim_{n\to\infty}\frac{n^{n+\frac{1}{2}}}{n!\ \mathrm{e}^n} = \sqrt{2\pi}. \tag{3.7.1}$$

which together with the fact that $(1-\delta)\mathrm{e}^{1-\delta} < 1/\mathrm{e}$ implies that $I_1 \to 0$ as $n \to \infty$. The function $v \mapsto v\mathrm{e}^{-(1-(1+\omega T)/n)v}$ is decreasing for $v \ge (1-(1+\omega T)/n)^{-1}$, so for n large enough, we can estimate I_3 by

$$\begin{aligned}
I_3 &\le 2M\ \frac{n^{n+1}}{n!}\left((1+\delta)\mathrm{e}^{-(1-(1+\omega T)/n)(1+\delta)}\right)^n\int_{1+\delta}^{\infty}\mathrm{e}^{-v}\, dv \\
&\le 2M\ \frac{n^{n+1}}{n!}\left((1+\delta)\mathrm{e}^{-(1-(1+\omega T)/n)(1+\delta)}\right)^n.
\end{aligned}$$

Since

$$\lim_{n\to\infty}(1+\delta)\mathrm{e}^{-(1-(1+\omega T)/n)(1+\delta)} = (1+\delta)\mathrm{e}^{-(1+\delta)} < 1/\mathrm{e},$$

we can use Stirling's formula (3.7.1) once more to conclude that I_3 tends to zero as $n \to \infty$. Thus,

$$\lim_{n\to\infty}\left(1-\frac{t}{n}A\right)^{-(n+1)}x = \mathfrak{A}^t x,$$

uniformly in $t \in [0, T]$. As furthermore

$$\lim_{n\to\infty}\left(1 - \frac{t}{n}A\right)^{-1}x = x$$

uniformly in $t \in [0, T]$ (see Theorem 3.7.1(iii)), this implies that

$$\lim_{n\to\infty}\left(1 - \frac{t}{n}A\right)^{-n}x = \mathfrak{A}^t x,$$

uniformly in t on any bounded interval. □

3.8 The nonhomogeneous Cauchy problem

It is time to study of the relationship between the differential equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + f(t), \qquad t \geq s, \\ x(s) &= x_s, \end{aligned} \tag{3.8.1}$$

and the *variation of constants formula*

$$x(t) = \mathfrak{A}^{t-s}x_s + \int_s^t \mathfrak{A}^{t-v}f(v)\,dv. \tag{3.8.2}$$

It is possible to do this in several different settings, but we choose a setting that is relevant for the full system $\left[\begin{smallmatrix}\mathfrak{A} & \mathfrak{B}\\ \mathfrak{C} & \mathfrak{D}\end{smallmatrix}\right]$. Here the spaces X_n and the extended semigroups $\mathfrak{A}_{|X_n}$ and generators $A_{|X_{n+1}}$ (with $n \leq 0$) introduced in Section 3.6 become important.

Definition 3.8.1 Let $s \in \mathbb{R}$, $x_s \in X$, $n = 0, \pm 1, \pm 2, \ldots,$ and $f \in L^1_{\mathrm{loc}}([s, \infty); X_{n-1})$. A function x is a strong solution of (3.8.1) in X_n (on the interval $[s, \infty)$) if $x \in C([s, \infty); X_n) \cap W^{1,1}_{\mathrm{loc}}([s, \infty); X_{n-1})$, $x(s) = x_s$, and $\dot{x}(t) = A_{|X_n}x(t) + f(t)$ in X_{n-1} for almost all $t \geq s$. By a strong solution of (3.8.1) (without any reference to a space X_n) we mean a strong solution of (3.8.1) in X $(= X_0)$.

Below we shall primarily look for sufficient conditions which imply that we have a strong solution (in X). This means that we must take $x_s \in X$ and $f \in L^1_{\mathrm{loc}}([s, \infty); X_{-1})$, and that (3.8.1) should be interpreted as an equation in X_{-1} (valid for almost all $t \geq s$). Thus, it should really be written in the form (recall that $A_{|X}$ maps $X = X_0$ into X_{-1})

$$\begin{aligned} \dot{x}(t) &= A_{|X}x(t) + f(t), \qquad t \geq s, \\ x(s) &= x_s. \end{aligned} \tag{3.8.3}$$

The integration in (3.8.2) should be carried out in X_{-1}, so that this identity should really be written in the form

$$x(t) = \mathfrak{A}^{t-s} x_s + \int_s^t \mathfrak{A}^{t-v}_{|X_{-1}} f(v)\, dv. \tag{3.8.4}$$

In order for (3.8.1) and (3.8.2) (or more precisely, (3.8.3) and (3.8.4)) to be equivalent we need some sort of smoothness assumptions on f: it should be either smooth in time or smooth in the state space (see parts (iv) and (v) below).

Theorem 3.8.2 *Let $s \in \mathbb{R}$, $x_s \in X$, and $f \in L^1_{\mathrm{loc}}([s,\infty); X_{-1})$.*

(i) *The function x given by* (3.8.4) *is a strong solution of* (3.8.1) *in X_{-1} (hence in X_n for every $n \le -1$).*
(ii) *Equation* (3.8.1) *has at most one strong solution x in X, namely the function x given by* (3.8.4).
(iii) *The function x given by* (3.8.4) *is a strong solution of* (3.8.1) *in X_n for some $n \ge 0$ if and only if $x \in C([s,\infty); X_n)$ and $f \in L^1_{\mathrm{loc}}([s,\infty); X_{n-1})$. (In particular, this implies that $x_s \in X_n$.)*
(iv) *If $f \in L^1_{\mathrm{loc}}([s,\infty); X)$ then the function x given by* (3.8.2) *is a strong solution of* (3.8.1) *in X.*
(v) *If $f \in W^{1,1}_{\mathrm{loc}}([s,\infty); X_{-1})$ then the function x given by* (3.8.4) *is a strong solution of* (3.8.1) *in X, $x \in C^1([s,\infty); X_{-1})$, and $z = \dot{x}$ is a strong solution of the equation*

$$\begin{aligned} \dot{z}(t) &= Az(t) + \dot{f}(t), \qquad t \ge s, \\ z(s) &= Ax_s + f(s) \end{aligned} \tag{3.8.5}$$

in X_{-1}. In particular, $\dot{x}(t) = A_{|X}x(t) + f(t)$ in X_{-1} for all $t \ge s$ (and not just almost all $t \ge s$).
(vi) *If $f = \pi_{[\alpha,\beta)} f_1$, where $s \le \alpha < \beta \le \infty$ and $f_1 \in W^{1,1}_{\mathrm{loc}}([s,\infty); X_{-1})$ then the function x given by* (3.8.4) *is a strong solution of* (3.8.1) *in X.*
(vii) *If f is any finite linear combination of functions of the type presented in (iv)–(vi), then the function x given by* (3.8.4) *is a strong solution of* (3.8.1) *in X.*

Proof (i) Define x by (3.8.4). The term $t \mapsto \mathfrak{A}^{t-s} x_s$ belongs to $C([s,\infty); X) \cap C^1([s,\infty); X_{-1}) \cap C^2([s,\infty); X_{-2})$ and it is a strong solution of (3.8.1) with $f = 0$ in X. Subtracting this term from x we reduce the problem to the case where $x_s = 0$. (The same reduction is valid in the proofs of (ii)–(vii), too.)

That $x \in C([s,\infty); X)$ follows from Proposition 2.3.1 with X replaced by X_{-1}, $\mathfrak{C} = 0$, and $\mathfrak{D} = 0$.

Suppose for the moment that $f \in C([s,\infty); X_{-1})$. Since $A_{|X_{-1}} \in \mathcal{B}(X_{-1}; X_{-2})$, we can then easily justify the following computation for $t \ge 0$

(the double integrals are computed in X_{-1}, and the other integrals in X_{-2} or X_{-1}; see Theorem 3.2.1(ii) for the last step)

$$\begin{aligned}
\int_s^t A_{|X_{-1}} x(v)\,dv &= A_{|X_{-1}} \int_s^t \int_s^v \mathfrak{A}^{v-w}_{|X_{-1}} f(w)\,dw\,dv \\
&= A_{|X_{-1}} \int_s^t \int_w^t \mathfrak{A}^{v-w}_{|X_{-1}} f(w)\,dv\,dw \\
&= A_{|X_{-1}} \int_s^t \int_0^{t-w} \mathfrak{A}^{v}_{|X_{-1}} f(w)\,dv\,dw \\
&= \int_s^t A_{|X_{-1}} \int_0^{t-w} \mathfrak{A}^{v}_{|X_{-1}} f(w)\,dv\,dw \\
&= \int_s^t (\mathfrak{A}^{t-w}_{|X_{-1}} - 1) f(w)\,dw.
\end{aligned}$$

As the set of continuous functions is dense in L^1, the same identity must then be true for all $f \in L^1_{\mathrm{loc}}([s,\infty); X_{-1})$. Rewriting this in terms of the function x in (3.8.4) (with $x_s = 0$) we get

$$x(t) = \int_s^t (A_{|X_{-1}} x(v) + f(v))\,dv.$$

Thus, $x \in W^{1,1}_{\mathrm{loc}}([s,\infty); X_{-2})$ and $\dot{x}(t) = A_{|X_{-1}} x(t) + f(t)$ in X_{-2} for almost all $t \geq s$. Clearly $x(s) = 0$. This implies that x is a strong solution of (3.8.1) in X_{-1} with $x_s = 0$.

(ii) If z is an arbitrary function in $C^1([s,\infty); X)$, then it is easy to show (using Theorem 3.2.1(ii)) that, for each $t > s$, the function $v \mapsto \mathfrak{A}^{t-v} z(v)$ is continuously differentiable in X_{-1}, with derivative $\mathfrak{A}^{t-v}(\dot{z}(v) - A_{|X} z(v))$. Integrating this identity (in X_{n-1}) we get

$$z(t) = \mathfrak{A}^{t-v} z(s) + \int_s^t \mathfrak{A}^{t-v}(\dot{z}(v) - A_{|X} z(v))\,dv.$$

Since $C^1([s,\infty); X)$ is dense in $W^{1,1}_{\mathrm{loc}}([s,\infty); X_{-1}) \cap C([s,\infty); X)$, and since both sides of the above identity depend continuously in X_{-1} on z in the norm of $W^{1,1}_{\mathrm{loc}}([s,\infty); X_{-1}) \cap C([s,\infty); X)$, the same identity must hold for every $z \in W^{1,1}_{\mathrm{loc}}([s,\infty); X_{-1}) \cap C([s,\infty); X)$. In particular, it is true whenever z is a strong solution of (3.8.1) in X_n, in which case we furthermore have $\dot{z}(v) - A_{|X} z(v) = f(v)$ for almost all $v \geq s$. This means that z is given by (3.8.4).

(iii) The necessity of the condition $x \in C([s,\infty); X_n)$ is part of the definition of a strong solution in X_n. The necessity of the condition $f \in L^1_{\mathrm{loc}}([s,\infty); X_{n-1})$ follows from the fact that $f = \dot{x} - A_{|X_n} x$, where $\dot{x} \in L^1_{\mathrm{loc}}([s,\infty); X_{n-1})$ and $A_{|X_n} x \in C([s,\infty); X_{n-1})$.

Conversely, suppose that $x \in C([s,\infty); X_n)$ and that $f \in L^1_{\mathrm{loc}}([s,\infty); X_{n-1})$. By (i), we have $\dot{x} = A_{|X} x + f = A_{|X_n} x + f$ in X_{-2}; in particular, the derivative $\dot{x}$ is computed in X_{-2}. However, the right-hand side of this identity belongs to $L^1_{\text{loc}}([s,\infty);X_{n-1})$, so its integral (which is x) belongs to $W^{1,1}_{\text{loc}}([s,\infty);X_{n-1})$, and the same identity is true a.e. in X_{n-1}. Thus, x is a strong solution in X_n.

(iv)–(vii) In the remainder of the proof we take $x_s = 0$, without loss of generality (see the proof of (i)).

(iv) The proof of (iv) identical to the proof of (i), with X_{-1} replaced by X.

(v) Since $\mathfrak{A}_{|X_{-1}} \in C^1(\overline{\mathbb{R}}^+;\mathcal{B}(X_{-1};X_{-2}))$ and $f \in C([s,\infty);X_{-1})$, we can differentiate under the integral sign to get (as an identity in X_{-2})

$$\dot{x}(t) = f(t) + A_{|X_{-1}} \int_s^t \mathfrak{A}^{t-v}_{|X_{-1}} f(v)\,dv, \qquad t \geq s.$$

Integrate by parts (or alternatively, write $f(v) = f(s) + \int_s^v \dot{f}(w)\,dw$ and use Fubini's theorem) to show that we can write this (still as an identity in X_{-2})

$$\dot{x}(t) = \mathfrak{A}^{t-s}_{|X_{-1}} f(s) + \int_s^t \mathfrak{A}^{t-v}_{|X_{-1}} \dot{f}(v)\,dv, \qquad t \geq s.$$

By (i), the right-hand side of this expression is the strong solution of (3.8.5) in X_{-1}, so from the definition of a strong solution we conclude that $\dot{x} \in C([s,\infty);X_{-1}) \cap W^{1,1}_{\text{loc}}([s,\infty);X_{-2})$. The continuity of $\dot{x}$ in X_{-1} implies that, although we originally computed the derivative $\dot{x}$ of x as a limit in the norm of X_{-2}, this limit actually exists in the norm of X_{-1} (i.e., x is differentiable in the stronger norm of X_{-1}), and that $x \in C^1([s,\infty);X_{-1}) \cap W^{2,1}_{\text{loc}}([s,\infty);X_{-2})$.

We proceed to show that $x \in C([s,\infty);X)$ and that $\dot{x}(t) = A_{|X}x(t) + f(t)$ in X_{-1} for all $t \geq s$. We know from (i) that $\dot{x}(t) = A_{|X_{-1}}x(t) + f(t)$ in X_{-2} for almost all $t \geq s$, and, since both sides are continuous in X_{-2}, we must actually have equality for *all* $t \geq s$. Choose some α in the resolvent set of $A_{|X_{-1}}$ (or equivalently, from the resolvent set of A) and subtract $\alpha x(t)$ from both sides of this identity to get (as an identity in X_{-2})

$$\alpha x(t) - \dot{x}(t) = (\alpha - A_{|X_{-1}})x(t) - f(t),$$

that is

$$x(t) = (\alpha - A_{|X_{-1}})^{-1}\bigl(f(t) + \alpha x(t) - \dot{x}(t)\bigr).$$

As $(\alpha - A_{|X_{-1}})^{-1} \in \mathcal{B}(X_{-1};X)$, and x, $\dot{x}$ and f belong to $C([s,\infty);X_{-1})$, the latter equation shows that $x \in C([s,\infty);X)$, and that $\dot{x}(t) = A_{|X}x(t) + f(t)$ in X_{-1} for all $t \geq s$. Thus, x is a strong solution of (3.8.1) in X.

(vi) Since $\pi_{[\alpha,\beta)}f_1 = \pi_{[\alpha,\infty)}f_1 - \pi_{[\beta,\infty)}f_1$, we can without loss of generality suppose that $\beta = \infty$ (cf. (vii)).

Clearly, the restriction of x to $[s,\alpha)$ is the zero function, so in order to prove the theorem it suffices to show that $\pi_{[\alpha,\infty)}x \in C([\alpha,\infty);X) \cap W^{1,1}_{\mathrm{loc}}([\alpha;\infty);X_{-1})$ and that $x(\alpha)=0$, because this implies that $x \in C([\alpha,\infty);X) \cap W^{1,1}_{\mathrm{loc}}([\alpha;\infty);X_{-1})$. But this follows from (v) with s replaced by α (and $x_\alpha = 0$).[6]

(vii) This follows from the linearity of (3.8.1) and (3.8.4). □

Sometimes we need more smoothness of a solution than we get from Theorem 3.8.2.

Theorem 3.8.3 *Let* $s \in \mathbb{R}$, $x_s \in X$, $f \in W^{2,1}_{\mathrm{loc}}([s,\infty);X_{-1})$, *and* $A_{|X}x_s + f(s) \in X$. *Then the strong solution* x *of* (3.8.1) *satisfies* $x \in C^2([s,\infty);X_{-1}) \cap C^1([s,\infty);X)$, $\dot{x} = z$ *is the strong solution of* (3.8.5) *in* X, *and* $\ddot{x} = y$ *is the strong solution of*

$$\begin{aligned} \dot{y}(t) &= Ay(t) + \ddot{f}(t), \qquad t \ge s, \\ y(s) &= A\dot{x}(s) + \dot{f}(s) \end{aligned} \tag{3.8.6}$$

in X_{-1}. *In particular,* $\dot{x} = A_{|X}x + f \in C^1([s,\infty);X_{-1}) \cap C([s,\infty);X)$ *and the identities* $\dot{x}(t) = A_{|X_{-1}}x(t) + f(t)$ *and* $\ddot{x}(t) = A_{|X_{-1}}\dot{x}(t) + \dot{f}(t)$ *are hold* X_{-1} *for all* $t \ge s$.

Proof By Theorem 3.8.2(v), $x \in C([s,\infty);X) \cap C^1([s,\infty);X_{-1})$, and, of course,

$$\dot{x}(t) = A_{|X}x(t) + f(t), \qquad t \ge s.$$

Arguing as in the proof of Theorem 3.8.2(v) (using the density of C^2 in $W^{2,1}$) we can use the extra differentiability assumption on u to show that $x \in C^2([s,\infty);X_{-2})$, and that

$$\ddot{x}(t) = A_{|X_{-1}}\dot{x}(t) + \dot{f}(t), \qquad t \ge s.$$

Let $z = \dot{x}$. Then $z(s) = A_{|X}x_s + f(s) \in X$, and z is the strong solution of the equation (3.8.5) in X_{-1}. However, by Theorem 3.8.2(v), this solution is actually a strong solution in X, i.e., $z \in C([s,\infty);X)$, and it has some additional smoothness, namely $z \in C^1([s,\infty);X_{-1})$. Since $z = \dot{x}$, this means that $x \in C^2([s,\infty);X_{-1}) \cap C^1([s,\infty);X)$, as claimed. □

Above we have only looked at the *local smoothness* of a strong solution of (3.8.1) (or more generally, of the function x defined by the variation of constants formula (3.8.2)). There are also some corresponding *global growth bounds* on the solution and its derivatives.

[6] Although x is continuous, there will be a jump discontinuity in $\dot{x}$ at the cutoff point. Thus, we will not in general have $x \in C^1([s,\infty);X_{-1})$ in this case, but we will still have $x \in Reg^1_{\mathrm{loc}}([s,\infty);X_{-1})$ and $\dot{x} - f \in C([s,\infty);X_{-1})$.

Theorem 3.8.4 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ with growth bound $\omega_{\mathfrak{A}}$. Let $\omega > \omega_{\mathfrak{A}}$, and let $1 \le p < \infty$. Under the following additional assumptions on the function f in Theorems 3.8.2 and 3.8.3 we get the following additional conclusions about the strong solution x of* (3.8.1) *(and all the listed derivatives exist in the given sense):*

(i) *If $f \in L^p_\omega([s,\infty); X)$, then*

$$\begin{aligned} x &\in BC_{0,\omega}([s,\infty); X) \cap L^p_\omega([s,\infty); X), \\ \dot{x} &\in L^p_\omega([s,\infty); X_{-1}). \end{aligned}$$

(ii) *If $f \in W^{1,p}([s,\infty); X_{-1})$, then*

$$\begin{aligned} x &\in BC_{0,\omega}([s,\infty); X) \cap L^p_\omega([s,\infty); X), \\ \dot{x} &\in BC_{0,\omega}([s,\infty); X_{-1}) \cap L^p_\omega([s,\infty); X_{-1}), \\ \ddot{x} &\in L^p_\omega([s,\infty); X_{-2}). \end{aligned}$$

(iii) *If $f \in W^{2,p}([s,\infty); X_{-1})$, then*

$$\begin{aligned} x &\in BC_{0,\omega}([s,\infty); X) \cap L^p_\omega([s,\infty); X), \\ \dot{x} &\in BC_{0,\omega}([s,\infty); X) \cap L^p_\omega([s,\infty); X), \\ \ddot{x} &\in BC_{0,\omega}([s,\infty); X_{-1}) \cap L^p_\omega([s,\infty); X_{-1}), \\ \dddot{x} &\in L^p_\omega([s,\infty); X_{-2}). \end{aligned}$$

Proof (i) Let Σ be the L^p-well-posed linear system on (X, X, X) described in Proposition 2.3.1 with $B = 1$, $C = 1$, and $D = 0$. Then, according to Theorem 3.8.2(iv), the strong solution x of (3.8.1) can be interpreted as the state trajectory of this system, and furthermore, its output y satisfies $y(t) = x(t)$ for all $t \ge s$. By Theorem 2.5.4, $x \in BC_{0,\omega}([s,\infty); X)$ and $x = y \in L^p_\omega([s,\infty); X)$. This implies that $\dot{x} = A_{|X}x + f \in L^p_\omega([s,\infty); X_{-1})$.

(ii) We again consider the same system as above, but this time on (X_{-1}, X_{-1}, X_{-1}). As above we first conclude that $x \in BC_{0,\omega}([s,\infty); X_{-1}) \cap L^p_\omega([s,\infty); X_{-1})$. We can also apply the same argument with x replaced by $\dot{x}$ (recall that, by Theorem 3.8.2(v), $\dot{x}$ is the strong solution of (3.8.5) in X_{-1}) to get $\dot{x} \in BC_{0,\omega}([s,\infty); X_{-1}) \cap L^p_\omega([s,\infty); X_{-1})$ and $\ddot{x} = A_{|X_{-1}}\dot{x} + \dot{f} \in L^p_\omega([s,\infty); X_{-2})$. Finally, we choose some $\alpha \in \rho(A) = \rho(A_{|X})$ and write the equation $\dot{x} = A_{|X}x + f$ in the form $(\alpha - A_{|X})^{-1}(\alpha x - \dot{x} + f)$ to conclude that $\dot{x} \in BC_{0,\omega}([s,\infty); X) \cap L^p_\omega([s,\infty); X)$.

(iii) Apply (ii) both to the function x itself and to the function $\dot{x}$. □

Another instance where we need a global growth bound on the solution, this time on $\mathbb{R}^-$, is when we want to study the existence and uniqueness of strong solutions of the equation $\dot{x}(t) = Ax(t) + f(t)$ on all of $\mathbb{R}$.

Definition 3.8.5 Let $n = 0, \pm 1, \pm 2, \dots$, and $f \in L^1_{\mathrm{loc}}(\mathbb{R}; X_{n-1})$. A function x is a strong solution of the equation

$$\dot{x}(t) = Ax(t) + f(t), \qquad t \in \mathbb{R}, \tag{3.8.7}$$

in X_n (on all of $\mathbb{R}$) if $x \in C(\mathbb{R}; X_n) \cap W^{1,1}_{\mathrm{loc}}(\mathbb{R}; X_{n-1})$, and $\dot{x}(t) = A_{|X_n} x(t) + f(t)$ in X_{n-1} for almost all $t \in \mathbb{R}$. By a strong solution of (3.8.7) (without any reference to a space X_n) we mean a strong solution of (3.8.7) in X $(= X_0)$.

Without any further conditions we cannot expect a strong solution of (3.8.7) to be unique. For example, if A generates a C_0 group on X, then for every $x_0 \in X$, the function $x(t) = \mathfrak{A}^t x_0$, $t \in \mathbb{R}$, is a strong solution of (3.8.7). We can rule out this case by, e.g., imposing a growth restriction on x at $-\infty$.

Lemma 3.8.6 *Let $\omega \in \mathbb{R}$, and suppose that the semigroup $\mathfrak{A}$ generated by A is ω-bounded (see Definition 2.5.6). Then, for each $f \in L^1_{\mathrm{loc}}(\mathbb{R}; X_{-1})$, the equation* (3.8.7) *can have at most one strong solution x satisfying* $\lim_{t\to-\infty} \mathrm{e}^{-\omega t} x(t) = 0$.

If such a solution exists, then we refer to it as the *strong solution of* (3.8.7) *which vanishes at $-\infty$.*

Proof The difference of two strong solutions of (3.8.7) is a strong solution of the equation $\dot{x}(t) = Ax(t)$ on $\mathbb{R}$, so it suffices to show that the only strong solution of (3.8.7) which satisfies $\lim_{t\to-\infty} \mathrm{e}^{-\omega t} x(t) = 0$ is the zero solution. Since it is a strong solution on $\mathbb{R}$, it is also a strong solution on $[s, \infty)$ with initial state $x(s)$ for every $s \in \mathbb{R}$, hence by Theorem 3.8.2(iv), $x(t) = \mathfrak{A}^{t-s} x(s)$ for every $t \geq s$. By the ω-boundedness of $\mathfrak{A}$, there is a constant M such that $|x(t)| \leq M \mathrm{e}^{\omega(t-s)} |x(s)|$, or equivalently, $\mathrm{e}^{-\omega t} |x(t)| \leq M \mathrm{e}^{-\omega s)} |x(s)|$. Let $s \to -\infty$ to conclude that $x(t) = 0$ for all $t \in \mathbb{R}$. □

Theorem 3.8.7 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ with growth bound $\omega_{\mathfrak{A}}$. Let $\omega > \omega_{\mathfrak{A}}$, and let $1 \leq p < \infty$. In all the cases (i)–(iii) listed below the equation* (3.8.7) *has a unique strong solution x satisfying* $\lim_{t\to-\infty} \mathrm{e}^{-\omega t} x(t) = 0$*, namely the function*

$$x(t) = \int_{-\infty}^{t} \mathfrak{A}^{t-v} f(v)\, dv, \tag{3.8.8}$$

and this solution has the additional properties listed below.

(i) $f \in L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X)$. *In this case*

$$x \in BC_{0,\omega,\mathrm{loc}}(\mathbb{R}; X) \cap L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X),$$
$$\dot{x} \in L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X_{-1}).$$

(ii) $f \in W^{1,p}_{\omega,\mathrm{loc}}(\mathbb{R}; X_{-1})$. *In this case*

$$x \in BC_{0,\omega,\mathrm{loc}}(\mathbb{R}; X) \cap L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X),$$
$$\dot{x} \in BC_{0,\omega,\mathrm{loc}}(\mathbb{R}; X_{-1}) \cap L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X_{-1}),$$
$$\ddot{x} \in L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X_{-2}).$$

(iii) $f \in W^{2,p}_{\omega,\mathrm{loc}}(\mathbb{R}; X_{-1})$. *In this case*

$$x \in BC_{0,\omega,\mathrm{loc}}(\mathbb{R}; X) \cap L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X),$$
$$\dot{x} \in BC_{0,\omega,\mathrm{loc}}(\mathbb{R}; X) \cap L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X),$$
$$\ddot{x} \in BC_{0,\omega,\mathrm{loc}}(\mathbb{R}; X_{-1}) \cap L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X_{-1}),$$
$$\dddot{x} \in L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X_{-2}).$$

Proof (i) (This proof is very similar to the proof of Theorem 3.8.4.) Let Σ be the L^p-well-posed linear system on (X, X, X) described in Proposition 2.3.1 with $B = 1$, $C = 1$, and $D = 0$. It follows from Theorems 2.5.7 and 3.8.2(iv) and Example 2.5.10 that the function x defined by (3.8.8) is a strong solution (3.8.7) satisfying $\lim_{t\to-\infty} \mathrm{e}^{-\omega t} x(t) = 0$, hence *the* strong solution satisfying this growth bound. Moreover, by Theorem 2.5.7 and Example 2.5.10, $x \in BC_{0,\omega,\mathrm{loc}}(\mathbb{R}; X)$ and $x = y \in L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X)$. Since $\dot{x} = A_{|X} x + f$, this implies that $\dot{x} \in L^p_{\omega,\mathrm{loc}}(\mathbb{R}; X_{-1})$.

(ii)–(iii) The proofs of (ii)–(iii) are analogous to the proofs of parts (ii)–(iii) of Theorem 3.8.4, and we leave them to the reader. □

Remark 3.8.8 Theorem 3.8.4 remains valid if we replace L^p_ω by $L^\infty_{0,\omega}$ or $Reg_{0,\omega}$ throughout. Theorem 3.8.7 remains valid if we delete the subindex 'loc', or if we replace $L^p_{\omega,\mathrm{loc}}$ by $L^\infty_{0,\omega,\mathrm{loc}}$ or $Reg_{0,\omega,\mathrm{loc}}$ throughout, or if we do both of these operations at the same time. The proofs remain the same.

3.9 Symbolic calculus and fractional powers

In this section we shall develop a basic symbolic calculus for the generators of C_0 semigroups.[7] We shall here consider only two classes of mappings of generators. The first class it the one where the generator A is mapped conformally into $f(A)$ where f is a complex-valued function which is analytic at the spectrum of A (including the point at infinity if A is unbounded). The other class of mapping is the one which gives us the fractional powers of $\gamma - A$ where $\gamma \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$. In

[7] With some trivial modifications this functional calculus can be applied to any closed operator with a nonempty resolvent set.

Section 3.10 we shall use a similar calculus to construct the semigroup generated by A in a special (analytic) case.

Let us begin with the simplest case where A is bounded. Let Γ be a piecewise continuously differentiable Jordan curve which encircles $\sigma(A)$ counterclockwise, i.e., the index of $\sigma(A)$ with respect to Γ is one. If f is analytic on Γ and inside Γ, then we define $f(A)$ by

$$f(A) = \frac{1}{2\pi j} \oint_{\Gamma} (\lambda - A)^{-1} f(\lambda)\, d\lambda. \tag{3.9.1}$$

This integral converges in the operator norm topology, e.g., as a Riemann integral (but it can, of course, also be interpreted in the strong sense, where we apply each side to a vector $x \in X$). This definition of $f(A)$ given is standard, and it is found in most books on functional analysis (see, e.g., Rudin (1973, p. 243)).

Let us check that the definition (3.9.1) of $f(A)$ coincides with the standard definition in the case where $f(z) = \sum_{k=0}^{n} a_k z^k$ is a polynomial. In this case we expect to have $f(A) = \sum_{k=0}^{n} a_k A^k$. By the linearity of the integral in (3.9.1), to prove this it suffices to verify the special case where $f(z) = z^n$ for some $n = 0, 1, 2, \ldots$. In this case we get

$$\begin{aligned}
\frac{1}{2\pi j} \oint_{\Gamma} \lambda^n (\lambda - A)^{-1}\, d\lambda &= \frac{1}{2\pi j} \oint_{\Gamma} (\lambda - A + A)^n (\lambda - A)^{-1}\, d\lambda \\
&= \sum_{k=0}^{n} \binom{n}{k} A^k \frac{1}{2\pi j} \oint_{\Gamma} (\lambda - A)^{n-k-1}\, d\lambda \\
&= A^n,
\end{aligned}$$

where the last step uses Lemma 3.9.2. Thus (3.9.1) is consistent with the standard definition of $f(A)$ in terms of powers of A when f is a polynomial.

If A is unbounded, then (3.9.1) must be slightly modified. In the following discussion, we denote the compactified complex plane $\mathbb{C} \cup \{\infty\}$ by $\overline{\mathbb{C}}$, and we let $\overline{\sigma}(A)$ be the (extended) spectrum of A in $\overline{\mathbb{C}}$, i.e., $\overline{\sigma}(A) = \sigma(A)$ if A is bounded, and $\overline{\sigma}(A) = \sigma(A) \cup \{\infty\}$ if A is unbounded.

Let A be the generator of a C_0 semigroup $\mathfrak{A}$ with growth bound $\omega_{\mathfrak{A}}$. Then we know from Theorem 3.2.9(ii) that $\overline{\sigma}(A) \subset \overline{\mathbb{C}}^{-}_{\omega_{\mathfrak{A}}} \cup \{\infty\}$ (where we can remove the point at infinity of A is bounded). Let f be a complex-valued function which is analytic on $\overline{\mathbb{C}}^{-}_{\omega_{\mathfrak{A}}} \cup \{\infty\}$ (f need not be analytic at infinity if A is bounded). We denote the set of points $\lambda \in \overline{\mathbb{C}}$ in which f is not analytic by $\overline{\sigma}(f)$ (this includes the point at infinity if f is not analytic there).

If A and f satisfies the conditions listed in the preceding paragraph, then it is possible to choose a piecewise continuously differentiable Jordan curve Γ in the complex plane which separates $\overline{\sigma}(A)$ from $\overline{\sigma}(f)$, with $\overline{\sigma}(A)$ 'to the left' of Γ and $\overline{\sigma}(f)$ 'to the right' of Γ. If A is bounded, then we can choose

Γ to be a curve encircling $\sigma(A)$ counter-clockwise with $\overline{\sigma}(f)$ on the outside, and if f is analytic at infinity, then we can choose Γ to be a curve encircling $\sigma(f)$ clockwise with $\overline{\sigma}(A)$ on the outside. If both of these conditions hold, then both choices are possible. Unfortunately, they do not produce exactly the same result, so before we try this approach we have to modify (3.9.1) slightly.

Before proceeding further, let us recall two different versions of the *Cauchy formula* for the derivatives of a function.

Lemma 3.9.1 *Let U be a Banach space, and let Γ be a positively oriented piecewise continuously differentiable Jordan curve in $\mathbb{C}$ (i.e., the index of the inside is one).*

(i) *If f is a U-valued function which is analytic on Γ and inside Γ, then, for every λ_0 inside Γ,*

$$\frac{1}{2\pi j}\oint_\Gamma \frac{f(\lambda)}{(\lambda-\lambda_0)^{n+1}}\,d\lambda = \begin{cases} 0, & n<0, \\ 1/(n!) f^{(n)}(\lambda_0), & n \ge 0. \end{cases}$$

(ii) *If instead f is analytic on Γ and outside Γ (including the point at infinity), then for every λ_0 inside Γ,*

$$\frac{1}{2\pi j}\oint_\Gamma (\lambda-\lambda_0)^{n-1} f(\lambda)\,d\lambda = \begin{cases} 0, & n<0, \\ 1/(n!)\,\frac{\mathrm{d}^n}{\mathrm{d}z^n}\, f(\lambda_0+1/z)_{|z=0}, & n \ge 0. \end{cases}$$

Proof (i) In the scalar case this is the standard Cauchy formula for the derivative found in all textbooks (if $n \in \mathbb{Z}_-$ then the integrand is analytic inside Γ, so the result is zero). The operator-valued case can be reduced to the scalar-values case: if f is $\mathcal{B}(X;Y)$-valued, then we choose arbitrary $x \in X$ and $y^* \in Y^*$ and apply the scalar case to $y^* f x$.

(ii) We make a change of integration variable from λ to $\zeta = 1/(\lambda-\lambda_0)$, $(\lambda-\lambda_0)^{-1}\,d\lambda = -z^{-1}\,dz$. If Γ' is the image of Γ under the mapping $\lambda \mapsto 1/(\lambda-\lambda_0)$, then Γ' is negatively oriented (the outside of Γ is mapped onto the inside of Γ'), and it encircles the origin. Part (i) gives (if we take the negative orientation of Γ' into account)

$$\begin{aligned}\frac{1}{2\pi j}\oint_\Gamma \frac{f(\lambda)}{(\lambda-\lambda_0)^{-n+1}}\,d\lambda &= -\frac{1}{2\pi j}\oint_{\Gamma'} \frac{f(\lambda_0+1/z)}{z^{n+1}}\,dz \\ &= \begin{cases} 0, & n<0, \\ 1/(n!)\,\frac{\mathrm{d}^n}{\mathrm{d}z^n}\, f(\lambda_0+1/z)_{|z=0}, & n \ge 0. \end{cases}\end{aligned}$$

□

Lemma 3.9.2 *Let $A \in \mathcal{B}(X)$, and let Γ be a positively oriented piecewise continuously differentiable Jordan curve which encircles $\rho(A)$. Then,*

$$\frac{1}{2\pi j}\oint_\Gamma (\lambda - A)^{-k}\,d\lambda = \begin{cases} 1, & k = 1, \\ 0, & k \in \mathbb{Z},\ k \neq 1. \end{cases}$$

Proof If $k \in \mathbb{Z}_-$, then the integrand is analytic inside Γ, and the result is zero. If $k \in \mathbb{Z}_+$, then the integrand is analytic outside Γ, including the point at infinity, and the result follows from Lemma 3.9.1(ii) with $n = 1$ and $f(\lambda) = (\lambda - A)^{-k}$. □

By Lemma 3.9.2, if A is bounded and if f is analytic at infinity, then (3.9.1) is equivalent to

$$f(A) = f(\infty) + \frac{1}{2\pi j}\oint_\Gamma (\lambda - A)^{-1}(f(\lambda) - f(\infty))\,d\lambda. \tag{3.9.2}$$

The function inside the integral has a second order zero at infinity, so if we replace Γ by a curve encircling both $\sigma(A)$ and $\sigma(f)$, then it follows from Lemma 3.9.1(ii) (with $n = 1$ and $f(\lambda)$ replaced by $(\lambda - A)^{-1}(f(\lambda) - f(\infty))$ that the resulting integral is zero. Thus, in (3.9.3) we may replace the positively oriented curve Γ which encircles $\sigma(A)$ with $\sigma(f)$ on the outside by a negatively oriented curve which encircles $\sigma(f)$ with $\sigma(A)$ on the outside. If we do so, then $\oint_\Gamma (\lambda - A)^{-1}\,d\lambda = 0$, and (3.9.2) can alternatively be written in the form

$$f(A) = f(\infty) + \frac{1}{2\pi j}\oint_\Gamma (\lambda - A)^{-1} f(\lambda)\,d\lambda. \tag{3.9.3}$$

Here it does not matter if A is bounded or unbounded, as long as Γ and the inside of Γ belong to $\rho(A)$, and f is analytic on Γ and the outside of Γ, including the point at infinity.

From (3.9.3) we immediately conclude the following:

Lemma 3.9.3 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ with growth rate $\omega_{\mathfrak{A}}$, let f be analytic on $\overline{\mathbb{C}^-_{\omega_{\mathfrak{A}}}} \cup \{\infty\}$, and define $f(A)$ as explained above. Then $f(A) - f(\infty) \in \mathcal{B}(X; X_1)$.*

Proof This follows from (3.9.3): for an arbitrary $\alpha \in \rho(A)$ we have

$$\begin{aligned} f(A) - f(\infty) &= (\alpha - A)^{-1}\frac{1}{2\pi j}\oint_\Gamma (\alpha - A)(\lambda - A)^{-1} f(\lambda)\,d\lambda \\ &= (\alpha - A)^{-1}\frac{1}{2\pi j}\oint_\Gamma \big[(\alpha - \lambda)(\lambda - A)^{-1} - 1\big] f(\lambda)\,d\lambda, \end{aligned}$$

where the integral defines an operator in $\mathcal{B}(X)$. □

As we already mentioned above, the definition of $f(A)$ given in (3.9.1) in the case where A is bounded is standard, but the definition of $f(A)$ in (3.9.3) with unbounded A is less common. However, (3.9.3) can be reduced to (3.9.1) by, e.g., a linear fractional transformation. For example, we can take some $\alpha \in \mathbb{C}^+_{\alpha_{\mathfrak{A}}} \cap \mathbb{C}^+$, and define $\varphi(\lambda) = 1/(\alpha - \lambda)$. The inverse transformation is $z \mapsto \varphi^{-1}(z) = \alpha - 1/z$. Note that $\alpha \in \rho(A)$, that α is mapped into ∞, and that ∞ is mapped into zero. Let Γ' be the image of Γ under this mapping. If Γ is negatively oriented, then the orientation of Γ' is positive and it encircles the origin (assuming that α lies inside Γ). By changing the integration variable in (3.9.3) we get (note that $d\lambda = z^{-2}\,dz$ and that $1/(2\pi j)\oint_{\Gamma'} z^{-1} f(\alpha - 1/z)\,dz = f(\infty)$)

$$\begin{aligned} f(A) &= f(\infty) + \frac{1}{2\pi j}\oint_{\Gamma'} (\alpha - 1/z - A)^{-1} z^{-2} f(\alpha - 1/z)\,dz \\ &= \frac{1}{2\pi j}\oint_{\Gamma'} \left[1 + (\alpha z - 1 - zA)^{-1}\right] z^{-1} f(\alpha - 1/z)\,dz \\ &= \frac{1}{2\pi j}\oint_{\Gamma'} (\alpha - A)\big(\alpha z - 1 - zA\big)^{-1} f(\alpha - 1/z)\,dz. \end{aligned}$$

Let $B_\alpha = (\alpha - A)^{-1}$ (thus, formally $B_\alpha = \varphi(A)$). Then $B_\alpha \in \mathcal{B}(X)$, and a short algebraic computation shows that

$$(z - B_\alpha)^{-1} = (\alpha - A)\big(\alpha z - 1 - zA\big)^{-1}.$$

Substituting this into the expression for $f(A)$ given above we get

$$f(A) = \frac{1}{2\pi j}\oint_{\Gamma'} (z - B_\alpha)^{-1} f(\alpha - 1/z)\,dz, \qquad B_\alpha = (\alpha - A)^{-1}. \tag{3.9.4}$$

Here Γ' is a positively oriented piecewise continuously differentiable Jordan curve which encircles $\sigma(B_\alpha)$, and the function $z \mapsto f(\alpha - 1/z)$ is analytic on Γ' and inside Γ'. Since we have obtained this from (3.9.3) (which does not depend on α) through a change of integration variable, the right-hand side of (3.9.4) does not depend on α, and it can be used as an alternative definition of $f(A)$.

If f is a rational function whose poles are located in $\mathbb{C}^+_{\omega_{\mathfrak{A}}}$ and which is analytic at infinity, then there is still another way of defining $f(A)$. Each such function can be written as a constant plus a linear combination of terms of the type $(\alpha_i - \lambda)^{-k_i}$, where each $\alpha_i \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$ and $k_i > 0$. It is then natural to define $f(A)$ to be the corresponding linear combination of $(\alpha_i - A)^{-k_i}$. Let us check that this definition is consistent with the one given earlier. To do this it suffices to show that, for all $\alpha \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$ and all $k = 1, 2, 3\ldots$,

$$(\alpha - A)^{-k} = \frac{1}{2\pi j}\oint_{\Gamma} (\lambda - A)^{-1}(\alpha - \lambda)^{-k}\,d\lambda, \tag{3.9.5}$$

where Γ is a negatively oriented piecewise continuously differentiable Jordan curve which encircles α with $\overline{\sigma}(A)$ on the outside. We begin with the case $k = 1$.

Then Lemmas 3.2.8 and 3.9.1 give

$$\begin{aligned}
&\frac{1}{2\pi j}\oint_\Gamma (\lambda - A)^{-1}(\alpha - \lambda)^{-1}\,d\lambda - (\alpha - A)^{-1} \\
&\quad = \frac{1}{2\pi j}\oint_\Gamma [(\lambda - A)^{-1} - (\alpha - A)^{-1}](\alpha - \lambda)^{-1}\,d\lambda \\
&\quad = \frac{1}{2\pi j}\oint_\Gamma (\lambda - A)^{-1}(\alpha - A)^{-1}\,d\lambda \\
&\quad = (\alpha - A)^{-1}\frac{1}{2\pi j}\oint_\Gamma (\lambda - A)^{-1}\,d\lambda = 0.
\end{aligned}$$

The case $k \geq 2$ follows from the case $k = 1$ if we differentiate the special case $k = 1$ of (3.9.5) $k - 1$ times with respect to α.

We shall next look at the related problem of how to define *fractional powers* of $(\gamma - A)$, where A is the generator of a C_0 semigroup $\mathfrak{A}$ and $\gamma > \omega_{\mathfrak{A}}$. This can be done in several different ways, see Pazy [1983]. Usually one starts with the *negative* fractional powers of $(\gamma - A)$, and then inverts these to get the positive fractional powers. One method, explained, e.g. in Pazy [1983], is to imitate (3.9.1) with $f(\lambda) = (\gamma - \lambda)^{-\alpha}$, and to let Γ be a path from $\infty e^{-j\epsilon}$ to $\infty e^{j\epsilon}$, where $0 < \epsilon < \pi/2$, passing between $\sigma(A)$ and the interval $[\gamma, \infty)$.[8] Here we shall used a different approach and instead extend the formula for $(\gamma - A)^{-n}$ given in Theorem 3.2.9(i) to fractional values of n.

Definition 3.9.4 Let A be the generator of a C_0 semigroup $\mathfrak{A}$ with growth bound $\omega_{\mathfrak{A}}$. For each $\gamma \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$ and $\alpha \geq 0$ we define $(\gamma - A)^{-\alpha}$ by

$$\begin{aligned}
(\gamma - A)^0 &= 1, \\
(\gamma - A)^{-\alpha}x &= \frac{1}{\Gamma(\alpha)}\int_0^\infty t^{\alpha-1}e^{-\gamma t}\mathfrak{A}^t x\,dt, \qquad \alpha > 0, \qquad x \in X.
\end{aligned}$$

Lemma 3.9.5 *The operators $(\gamma - A)^{-\alpha}$ introduced in Definition 3.9.4 are bounded linear operators on X, and $\alpha \mapsto (\gamma - A)^{-\alpha}$ is a semigroup, i.e.,*

$$(\gamma - A)^{-(\alpha+\beta)} = (\gamma - A)^{-\alpha}(\gamma - A)^{-\beta}$$

for all α, $\beta > 0$. Moreover, $(\gamma - A)^{-\alpha}$ is injective for all $\alpha \geq 0$.

Proof By assumption, the growth bound of $\mathfrak{A}$ is less than γ, hence the integral used in the definition of $(\gamma - A)^{-\alpha}$ converges absolutely, and it defines an operator in $\mathcal{B}(X)$.

To simplify the notation in our verification of the semigroup property we take $\gamma = 0$ (i.e., we denote $(A - \gamma)$ by A and $e^{-\gamma t}\mathfrak{A}^t$ by $\mathfrak{A}^t$). We take $x \in X$

[8] This method is quite general, and it can be used even in some cases where A is not a generator of a C_0 semigroup.

and make two changes of integration variable as follows:

$$\begin{aligned}(-A)^{-\alpha}(-A)^{-\beta}x &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\int_0^\infty\int_0^\infty s^{\alpha-1}t^{\beta-1}\mathfrak{A}^{s+t}x\,ds\,dt\\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\int_0^\infty\int_0^\infty s^{\alpha-1}(v-s)^{\beta-1}\mathfrak{A}^{v}x\,dv\,ds\\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\int_0^\infty\Big(\int_0^v s^{\alpha-1}(v-s)^{\beta-1}\,ds\Big)\mathfrak{A}^{v}x\,dv\\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 s^{\alpha-1}(1-s)^{\beta-1}\,ds\int_0^\infty v^{\alpha+\beta-1}\mathfrak{A}^{v}x\,dv\\ &= (-A)^{-(\alpha+\beta)}x;\end{aligned}$$

here the last equality follows from Definition 3.9.4 and the fact that the Beta function satisfies (for all $\alpha, \beta > 0$)

$$B(\alpha,\beta) = \int_0^1 s^{\alpha-1}(1-s)^{\beta-1}\,ds = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}.$$

To show that $(\gamma - A)^{-\alpha}$ is injective we can use the semigroup property in the following way. We choose β so that $\alpha + \beta = n$ is an integer. The operator $(\gamma - A)^{-n} = (\gamma - A)^{-\beta}(\gamma - A)^{-\alpha}$ in injective since $\gamma \in \rho(A)$ (recall that we take $\gamma > \gamma_{\mathfrak{A}}$), hence $(\gamma - A)^{-\alpha}$ in injective. □

The semigroup $\alpha \mapsto (\gamma - A)^{-\alpha}$ is actually a C_0 semigroup (i.e., it is strongly continuous). See Pazy [1983, Corollary 6.5, p. 72].

Since $(\gamma - A)^{-\alpha}$ is injective, it has an inverse defined on its range:

Definition 3.9.6 Let A be the generator of a C_0 semigroup $\mathfrak{A}$ with growth bound $\omega_{\mathfrak{A}}$. For each $\gamma \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$ and $\alpha \geq 0$ we define $(\gamma - A)^{\alpha}$ to be the inverse of the operator $(\gamma - A)^{-\alpha}$ defined in Definition 3.9.4, with domain $\mathcal{D}((\gamma - A)^{\alpha}) = \mathcal{R}\left((\gamma - A)^{-\alpha}\right)$.

Lemma 3.9.7 *With the notation of Definitions 3.9.4 and 3.9.6, let $\gamma \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$ and $\delta \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$. Then the fractional powers of $(\gamma - A)$ and $(\delta - A)$ have the following properties:*

(i) $(\gamma - A)^{\alpha} \in \mathcal{B}(X)$ *if* $\alpha \leq 0$*, and* $(\gamma - A)^{\alpha}$ *is closed if* $\alpha > 0$*;*
(ii) $(\delta - A)^{\alpha}(\gamma - A)^{\beta} = (\gamma - A)^{\beta}(\delta - A)^{\alpha}$ *if* $\alpha \leq 0$ *and* $\beta \leq 0$*;*
(iii) $\mathcal{D}((\gamma - A)^{\alpha}) \subset \mathcal{D}\left((\gamma - A)^{\beta}\right)$ *if* $\alpha \geq \beta$*;*
(iv) $\mathcal{D}((\gamma - A)^{\alpha})$ *is dense in X for all* $\alpha > 0$ *(and equal to X for all* $\alpha \leq 0$*);*
(v) $\mathcal{D}((\gamma - A)^{\alpha}) = \mathcal{D}((\delta - A)^{\alpha})$ *and* $(\delta - A)^{\alpha}(\gamma - A)^{-\alpha} \in \mathcal{B}(X)$ *if* $\alpha \geq 0$*;*

Proof (i) The case $\alpha \leq 0$ is contained in Lemma 3.9.5, and the inverse of a bounded (hence closed) operator is closed.

(ii) Use Fubini's theorem in Definition 3.9.4.

(iii) This is trivial if $\beta \le 0$ or $\alpha = \beta$. Otherwise, by Lemma 3.9.5,

$$(\gamma - A)^{-\alpha} = (\gamma - A)^{-\beta}(\gamma - A)^{-(\alpha-\beta)},$$

hence $\mathcal{R}\left((\gamma - A)^{-\alpha}\right) \subset \mathcal{R}\left((\gamma - A)^{-\beta}\right)$, or equivalently, $\mathcal{D}((\gamma - A)^{\alpha}) \subset \mathcal{D}\left((\gamma - A)^{\beta}\right)$.

(iv) This follows from (iii), since $\mathcal{D}((\gamma - A)^{\alpha})$ contains $\mathcal{D}((\gamma - A)^{n})$ for some positive integer n, and by Theorem 3.2.1(vi), $\mathcal{D}((\gamma - A)^{n}) = \mathcal{D}(A^{n})$ is dense in X.

(v) The boundedness of the operator $(\delta - A)^{\alpha}(\gamma - A)^{-\alpha}$ follows from the closed graph theorem as soon as we have shown that $\mathcal{D}((\gamma - A)^{\alpha}) = \mathcal{D}((\delta - A)^{\alpha})$, or equivalently, that $\mathcal{R}\left((\gamma - A)^{-\alpha}\right) = \mathcal{R}\left((\delta - A)^{-\alpha}\right)$. This is true for integer values of α, so by Theorem 3.9.5, it suffices to consider the case where $0 < \alpha < 1$. Moreover, by (iii), it suffices to show that

$$\mathcal{R}\left((\gamma - A)^{-\alpha} - (\delta - A)^{-\alpha}\right) \subset X_1.$$

By Definition 3.9.4, for all $x \in X$,

$$(\gamma - A)^{-\alpha}x - (\delta - A)^{-\alpha}x = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1}[1 - \mathrm{e}^{-(\delta-\gamma)t}]\mathrm{e}^{-\gamma t}\mathfrak{A}^t x\, dt.$$

Therefore, for $x \in X_1$ we have $[(\gamma - A)^{-\alpha} - (\delta - A)^{-\alpha}]x \in X_1$ and (integrate by parts)

$$[(\gamma - A)^{-\alpha} - (\delta - A)^{-\alpha}]x = \frac{1}{\Gamma(\alpha)}(\gamma - A)^{-1} \int_0^\infty \dot{h}(t)\mathrm{e}^{-\gamma t}\mathfrak{A}^t x\, dt,$$

where $h(t) = -t^{\alpha-1}[1 - \mathrm{e}^{-(\delta-\gamma)t}]$. Without loss of generality, suppose that $\delta > \gamma$. Then $\dot{h} \in L^1([0, 1]) \cap L^\infty([1, \infty)$ and $t \mapsto \mathrm{e}^{-\gamma t}\|\mathfrak{A}^t\| \in L^\infty([0, 1]) \cap L^1([1, \infty)$, so the integral converges absolutely for all $x \in X$. This implies that $\mathcal{R}\left((\gamma - A)^{-\alpha} - (\delta - A)^{-\alpha}\right) \subset X_1$, as claimed. □

With the fractional powers of $(\gamma - A)$ to our disposal, we can construct a continuous scale of Banach spaces X_α, $\alpha \in \mathbb{R}$, in the same way as we constructed the spaces X_n with integral indices n in Section 3.6. For $\alpha > 0$ we let X_α be the range of $(\gamma - A)^{-\alpha}$ (i.e., the image of X under $(\gamma - A)^{-\alpha}$), with norm

$$|x|_\alpha = |x|_{X_\alpha} = \left|(\gamma - A)^{\alpha}x\right|_X.$$

For $\alpha < 0$ we let X_α be the completion of X with the weaker norm

$$|x|_{-\alpha} = \left|(\gamma - A)^{-\alpha}x\right|_X, \qquad \alpha > 0.$$

All the earlier conclusions listed in Section 3.6 remain valid. In particular, for all $\gamma \in \mathbb{C}^+$, all $\alpha, \beta, \delta \in \mathbb{R}$, and all $t \geq 0$,

$$\begin{aligned}
&(\gamma - A)^{\beta}_{|X_{\alpha+\beta}} \text{ is an isometry of } X_{\alpha+\beta} \text{ onto } X_\alpha,\\
&(\gamma - A)^{\alpha+\beta}_{|X_\delta} = (\gamma - A)^{\alpha}_{|X_{\delta-\beta}}(\gamma - A)^{\beta}_{|X_\delta},\\
&(\gamma - A)^{\alpha}_{|X_\delta}, \mathfrak{A}^t_{|X_\delta} = \mathfrak{A}^t_{|X_{\delta-\alpha}}(\gamma - A)^{\alpha}_{|X_\delta}.
\end{aligned}$$

Different choices of γ give identical spaces with equivalent norms, and $(\gamma - A)^\alpha$ commutes with $(\delta - A)^\beta$ for all $\alpha, \beta \in \mathbb{R}$, and all $\gamma, \delta \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$,

The spaces X_α can be interpreted as *interpolation spaces* between the spaces X_n with integral indices; see Lunardi [1995, Chapters 1,2]. The following lemma is related to this fact:

Lemma 3.9.8 *Define the fractional space X_α as above. Then, there is a constant $C > 0$ such that for all $0 < \alpha < 1$, all $x \in X_1 = \mathcal{D}(A)$, and all $\rho > 0$,*

$$\begin{aligned}
|x|_{X_\alpha} &\leq C\big(\rho^\alpha |x|_X + \rho^{\alpha-1}|x|_{X_1}\big),\\
|x|_{X_\alpha} &\leq 2C|x|_X^{1-\alpha}|x|_{X_1}^{\alpha}.
\end{aligned} \tag{3.9.6}$$

The proof of this lemma is based on the the following representation of $(\gamma - A)^{-\alpha}$, valid for $0 < \alpha < 1$:

Lemma 3.9.9 *For $0 < \alpha < 1$ the operator $(\gamma - A)^{-\alpha}$ defined in Definition 3.9.4 has the representation*

$$(\gamma - A)^{-\alpha} = \frac{\sin \pi\alpha}{\pi} \int_0^\infty s^{-\alpha}\big(s + \gamma - A\big)^{-1}\, ds,$$

where the integral converges absolutely in operator norm.

Proof The absolute convergence in operator norm follows from the Hille–Yosida Theorem 3.4.1 and the assumption that $\gamma \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$. By using Theorem 3.2.9(i), Fubini's theorem, a change of integration variable $s = v/t$, and the fact that the Gamma-function

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}\, dt \tag{3.9.7}$$

satisfies (for $0 < \alpha < 1$) the reflection formula $\Gamma(\alpha)\Gamma(1-\alpha) = \frac{\sin \pi\alpha}{\pi}$, we get for all $x \in X$,

$$\begin{aligned}
&\frac{\sin \pi\alpha}{\pi} \int_0^\infty s^{-\alpha}\big(s+\gamma-A\big)^{-1} x\, ds \\
&\quad = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^\infty s^{-\alpha} \int_0^\infty \mathrm{e}^{-(s+\gamma)t} \mathfrak{A}^t x\, dt\, ds \\
&\quad = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^\infty \Big(\int_0^\infty s^{-\alpha} \mathrm{e}^{-st}\, ds\Big) \mathrm{e}^{-\gamma t} \mathfrak{A}^t x\, dt \\
&\quad = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \Big(\int_0^\infty v^{-\alpha} \mathrm{e}^{v}\, dv\Big) \int_0^\infty t^{\alpha-1} \mathrm{e}^{-\gamma t} \mathfrak{A}^t x\, dt \\
&\quad = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} \mathrm{e}^{-\gamma t} \mathfrak{A}^t x\, dt \\
&\quad = (\gamma - A)^{-\alpha} x.
\end{aligned}$$

□

Proof of Lemma 3.9.8. Let $\alpha \in (0,1)$, $\rho > 0$, $x \in X_1$, and recall that $|x|_{X_\alpha} = |(\gamma - A)^\alpha x|_X$ and that $|x|_{X_1}$ is (equivalent to) $|(\gamma - A)x|_X$. Since $0 < \alpha < 1$, we have $0 < 1 - \alpha < 1$, hence by Lemma 3.9.9, Theorem 3.4.1, and the assumption that $\gamma \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$ (observe that $\sin \pi(1-\alpha) = \sin \pi\alpha$ and $\sin \pi\alpha \le \pi\alpha$)

$$\begin{aligned}
|x|_{X_\alpha} &= |(\gamma - A)^{-(1-\alpha)}(\gamma - A)x|_X \\
&\le \frac{\sin \pi(1-\alpha)}{\pi} \int_0^\infty s^{\alpha-1} \big|\big((s+\gamma) - A\big)^{-1}(\gamma - A)x\big|\, ds \\
&\le \frac{\sin \pi\alpha}{\pi} \int_0^\rho s^{\alpha-1} \big\|\big(s+\gamma-A\big)^{-1}(\gamma - A)\big\| |x|_X\, ds \\
&\quad + \frac{\sin \pi(1-\alpha)}{\pi} \int_\rho^\infty s^{\alpha-1} \big\|\big(s+\gamma-A\big)^{-1}\big\| |x|_{X_1}\, ds \\
&\le |x|_X \alpha \int_0^\rho s^{\alpha-1} \big\|1 - s\big(s+\gamma-A\big)^{-1}\big\|\, ds \\
&\quad + |x|_{X_1}(1-\alpha) \int_\rho^\infty s^{\alpha-1} \big\|\big(s+\gamma-A\big)^{-1}\big\|\, ds \\
&\le (1+C)|x|_X \alpha \int_0^\rho s^{\alpha-1}\, ds + C|x|_{X_1}(1-\alpha) \int_\rho^\infty s^{\alpha-2}\, ds \\
&= (1+C)|x|_X \rho^\alpha + C|x|_{X_1} \rho^{\alpha-1}.
\end{aligned}$$

This proves the first inequality (with $C+1$ instead of C) in (3.9.6). To get the second inequality in (3.9.6) for nonzero x we simply take $\rho = |x|_{X_1}/|x|_X$. □

3.10 Analytic semigroups and sectorial operators

Semigroups obtained from the Cauchy problem for partial differential equations of parabolic type have some extra smoothness properties. They are of the following type.

Definition 3.10.1 Let X be a Banach space, let $0 < \delta \le \pi/2$, and let Δ_δ be the open sector $\Delta_\delta = \{t \in \mathbb{C} \mid t \neq 0,\ |\arg t| < \delta\}$ (see Figure 3.1). The family of operators $\mathfrak{A}^t \in \mathcal{B}(X)$, $t \in \Delta_\delta$, is an *analytic semigroup (with uniformly bounded growth bound ω) in* Δ if the following conditions holds:

(i) $t \mapsto \mathfrak{A}^t$ is analytic in Δ_δ;
(ii) $\mathfrak{A}^0 = 1$ and $\mathfrak{A}^s\mathfrak{A}^t = \mathfrak{A}^{s+t}$ for all $s, t \in \Delta_\delta$;
(iii) There exist constants $M \ge 1$ and $\omega \in \mathbb{R}$ such that

$$\|\mathfrak{A}^t\| \le M e^{\omega t}, \qquad t \in \Delta_\delta.$$

(iv) For all $x \in X$, $\lim_{\substack{t \to 0 \\ t \in \Delta_\delta}} \mathfrak{A}^t x = x$.

A semigroup $\mathfrak{A}$ is *analytic* if it is analytic in some sector Δ of the type described above.

We warn the reader that the sector Δ_δ in Definition 3.10.1 need not be maximal: If $\mathfrak{A}$ is analytic on some sector Δ_δ, then it can often be extended to an analytic semigroup on a larger sector $\Delta_{\delta'}$ with $\delta' > \delta$. If we take the union of all the sectors Δ_δ where $\mathfrak{A}^t$ is analytic (with a uniformly bounded growth bound), then the constants M and ω in (iii) typically deteriorate as we approach the sector boundary.

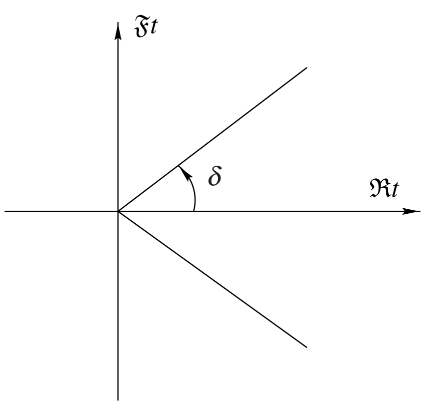

Figure 3.1 The sector Δ_δ

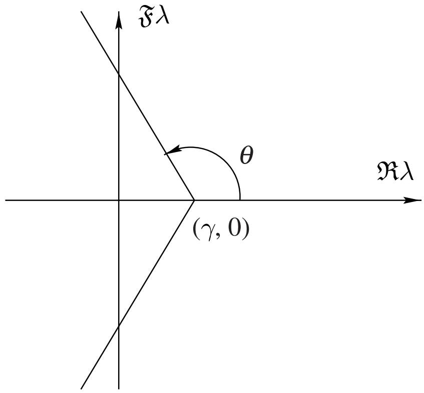

Figure 3.2 The sector $\Sigma_{\theta,\gamma}$

As we shall see in a moment, the generators of the analytic semigroups in Definition 3.10.1 are *sectorial operators*, which are defined as follows.

Definition 3.10.2 For each $\gamma \in \mathbb{R}$ and $\pi/2 < \theta < \pi$, let $\Sigma_{\theta,\gamma}$ be the open sector (see Figure 3.2)

$$\Sigma_{\theta,\gamma} = \{\lambda \in \mathbb{C} \mid \lambda \neq \gamma,\ |\arg(\lambda - \gamma)| < \theta\}.$$

A (closed) densely defined linear operator $X \supset \mathcal{D}(A) \to X$ is *sectorial on* $\Sigma_{\theta,\gamma}$ (with a uniform bound) if the resolvent set of A contains $\Sigma_{\theta,\gamma}$, and if

$$\|(\lambda - A)^{-1}\| \leq \frac{C}{|\lambda - \gamma|}, \qquad \lambda \in \Sigma_{\theta,\gamma}, \tag{3.10.1}$$

for some $C \geq 1$. The operator A is *sectorial* if it is sectorial on some sector $\Sigma_{\theta,\gamma}$ (with $\pi/2 < \theta < \pi$).

Again we warn the reader that the constant θ in Definition 3.10.2 is not maximal: it can always be replaced by a larger constant θ':

Lemma 3.10.3 *If A is sectorial on some sector $\Sigma_{\theta,\gamma}$ with $\pi/2 < \theta < \pi$ then it is also sectorial on some bigger sector $\Sigma_{\theta',\gamma}$ with $\theta' > \theta$. More precisely, this is true for every θ' satisfying $\theta < \theta' < \pi$ and* $\sin(\theta' - \theta) < 1/C$*, where C is the constant in* (3.10.1).

Proof Without loss of generality we can take $\gamma = 0$ (i.e., we replace $A - \gamma I$ by A).

By Lemma 3.2.8(ii), the rays $\{\lambda \neq 0 \mid \arg\lambda = \pm\theta\}$ bounding the sector $\Sigma_{\theta,0}$ belong to the resolvent set of A, and (by continuity) (3.10.1) holds for these λ, too. Let $0 < k < 1$. By (3.10.1) and Lemma 3.2.8(ii), if we take λ close enough to $\alpha \in \Sigma_{\theta,0}$ so that $|\lambda - \alpha| \le \frac{k|\alpha|}{C}$, then $\lambda \in \rho(A)$ and

$$\left\|(\lambda - A)^{-1}\right\| \le \frac{C}{(1-k)|\alpha|}.$$

Choose θ' to satisfy $\theta < \theta' < \pi$ and $\sin(\theta' - \theta) = k/C$. By letting α vary over the rays $\{\lambda \neq 0 \mid \arg \lambda = \pm\theta\}$ and taking $\lambda - \alpha$ orthogonal to λ (and $\Re\lambda < \Re\alpha$) we reach all λ in $\Sigma_{\theta',0} \backslash \Sigma_{\theta,0}$. Moreover, with this choice of λ and α, we have $|\lambda| < |\alpha|$, hence

$$\left\|(\lambda - A)^{-1}\right\| \le \frac{C}{(1-k)\lambda}.$$

Thus, A is sectorial on $\Sigma_{\theta',0}$ with C in (3.10.1) replaced by $C/(1-k)$. □

Lemma 3.10.3 (and its proof) implies that the constant C in (3.10.1) must satisfy $C \ge 1$, because if $C < 1$, then the proof of Lemma 3.10.3 shows that $(\lambda - A)^{-1}$ is a bounded entire function vanishing at infinity, hence identically zero. The optimal constant for $A = \gamma$ is $C = 1$. A similar argument shows that (3.10.1) holds with $\theta = \pi$ if and only if $A = \gamma$.

It is sometimes possible to increase the sector $\Sigma_{\theta,\gamma}$ in which A is sectorial in a different way: we keep θ fixed, but replace γ by $\gamma' < \gamma$. This is possible under the following assumptions:

Lemma 3.10.4 *Let A be a closed linear operator on X, and let $\gamma \in \mathbb{R}$ and $\pi/2 < \theta < \pi$. Then the following conditions are equivalent:*

(i) *A is sectorial on $\Sigma_{\theta,\gamma}$ and $\gamma \in \rho(A)$;*
(ii) *A is sectorial on some sector $\Sigma_{\theta,\gamma'}$ with $\gamma' < \gamma$.*

Proof (i) ⇒ (ii): It follows from Lemma 3.10.3 and the assumption $\gamma \in \rho(A)$ that the distance from $\sigma(A)$ to the boundary of the sector Σ_θ, γ is strictly positive, hence $\rho(A) \supset \Sigma_{\theta,\gamma'}$ for some $\gamma' < \gamma$. The norm of the resolvent $(\lambda - A)^{-1}$ is uniformly bounded for all $\lambda \in \Sigma_{\theta,\gamma'}$ satisfying $|\lambda - \gamma| \le |\gamma - \gamma'|$, and for $|\lambda - \gamma| > |\gamma - \gamma'|$ we can estimate

$$\|(\lambda - A)^{-1}\| \le \frac{C}{|\lambda - \gamma|} = \frac{C}{|\lambda - \gamma'|}\frac{|\lambda - \gamma'|}{|\lambda - \gamma|} < \frac{2C}{|\lambda - \gamma'|}.$$

(ii) ⇒ (i): This proof is similar to the one above (but slightly simpler). □

Our next result is a preliminary version of Theorem 3.10.6, and we refer the reader to that theorem for a more powerful result.

Theorem 3.10.5 *Let $A \colon X \to X$ be a (densely defined) sectorial operator on the sector $\Sigma_{\theta,\gamma}$. Then A is the generator of a C_0 semigroup $\mathfrak{A}$ satisfying $\|\mathfrak{A}^t\| \le$*

$Me^{\gamma t}$ *for some* $M \geq 1$ *which is continuous in the uniform operator norm on* $(0, \infty)$. *This semigroup has the representation*

$$\mathfrak{A}^t = \frac{1}{2\pi j} \int_\Gamma e^{\lambda t} (\lambda - A)^{-1} \, d\lambda, \tag{3.10.2}$$

where Γ *is a smooth curve in* $\Sigma_{\theta,\gamma}$ *running from* $\infty e^{-j\vartheta}$ *to* $\infty e^{j\vartheta}$, *where* $\pi/2 < \vartheta < \theta$. *For each* $t > 0$ *the integral converges in the uniform operator topology.*

Proof Throughout this proof we take, without loss of generality, $\gamma = 0$. (If $\gamma \neq 0$, then we replace $A - \gamma$ by A and $e^{-\gamma t}\mathfrak{A}^t$ by $\mathfrak{A}^t$; see Examples 2.3.5 and 3.2.6.)

The absolute convergence in operator norm of the integral in (3.10.2) is a consequence of (3.10.1) and the fact that $|e^{e^{j\theta} r}| = e^{\Re e^{j\theta} r} = e^{r \cos\theta}$ for all $r \in \mathbb{R}$ (observe that $\cos\theta < 0$ for $\pi/2 < |\theta| < \pi$). The continuity in the uniform operator topology on $(0, \infty)$ follows from a straightforward estimate (and the Lebesgue dominated convergence theorem). Since the integrand is analytic, we can deform the path of integration without changing the value of the integral to $\Gamma_t = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ (see Figure 3.3), where

$$\begin{aligned}
\Gamma_1 &= \{r e^{-j\vartheta} \mid \infty > r \geq 1/t\}, \\
\Gamma_2 &= \{t^{-1} e^{j\theta} \mid -\vartheta \leq \theta \leq \vartheta\}, \\
\Gamma_3 &= \{r e^{j\vartheta} \mid 1/t \leq r < \infty\},
\end{aligned}$$

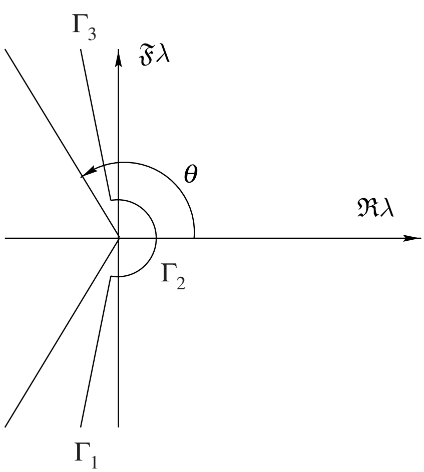

Figure 3.3 The path in the proof of Theorem 3.10.5

Then, by (3.10.1)

$$\begin{aligned}\left\| \frac{1}{2\pi j} \int_{\Gamma_1} e^{\lambda t} (\lambda - A)^{-1} \, d\lambda \right\| &\le \frac{1}{2\pi} \int_{\Gamma_1} \left\| e^{\lambda t} (\lambda - A)^{-1} \right\| d\lambda \\ &\le \frac{C}{2\pi} \int_{1/t}^{\infty} e^{rt\cos\vartheta} r^{-1} \, dr \\ &= \frac{C}{2\pi} \int_{1}^{\infty} e^{s\cos\vartheta} s^{-1} \, ds \\ &= M_1.\end{aligned}$$

The same estimate is valid for the integral over Γ_3. On Γ_2 we estimate

$$\begin{aligned}\left\| \frac{1}{2\pi j} \int_{\Gamma_2} e^{\lambda t} (\lambda - A)^{-1} \, d\lambda \right\| &\le \frac{1}{2\pi} \int_{\Gamma_2} \left\| e^{\lambda t} (\lambda - A)^{-1} \right\| d\lambda \\ &\le \frac{C}{2\pi} \int_{-\vartheta}^{\vartheta} e^{\cos\theta} \, d\theta \\ &= M_2.\end{aligned}$$

Together these estimates prove that $\|\mathfrak{A}^t\| \le M$ with $M = 2M_1 + M_2$.

Next we claim that for all $\alpha > 0$,

$$(\alpha - A)^{-1} = \int_0^{\infty} e^{-\alpha t} \mathfrak{A}^t \, dt. \tag{3.10.3}$$

To prove this we choose the curve Γ as we did above, but replace Γ_2 by $\Gamma_2 = \{\epsilon e^{-j\theta} \mid \vartheta \le \theta \le \vartheta\}$, where $0 < \epsilon < \alpha$ is fixed, and adjust Γ_1 and Γ_3 accordingly. We repeat the calculations presented above for this choice of Γ, and find that

$$\frac{1}{2\pi} \int_{\Gamma} \|e^{\lambda t} (\lambda - A)^{-1}\| \, d\lambda \le M(|\log t| + e^{\epsilon t})$$

for some finite M (the logarithm comes from Γ_1 and Γ_3, and the exponent from Γ_2). Thus, with this choice of Γ, if we multiply (3.10.2) by $e^{-\alpha t}$ and integrates over $(0, \infty)$, then the resulting double integral also converges absolutely. By Fubini's theorem, we may change the order of integration to get

$$\int_0^{\infty} e^{-\alpha t} \mathfrak{A}^t \, dt = \frac{1}{2\pi j} \int_{\Gamma} (\alpha - \lambda)^{-1} (\lambda - A)^{-1} \, d\lambda.$$

By (3.10.1), this integral is the limit as $n \to \infty$ of the integral over $\widetilde{\Gamma}_n$, where $\widetilde{\Gamma}_n$ is the closed curve that we get by restricting the variable r in the definition of Γ_1 and Γ_3 to the interval $\epsilon \le r \le n$, and joining the points $ne^{e^{-i\varphi}}$ and $ne^{e^{i\varphi}}$ with the arc $\Gamma_4 = \{ne^{-j\theta} \mid \vartheta \ge \theta \ge -\vartheta\}$. However, by the residue theorem, this integral is equal to $(\alpha - A)^{-1}$. This proves (3.10.3).

Equation (3.10.3) together with the bound $\|\mathfrak{A}^t\| \le M$ enables us to repeat the argument that we used in the proof of Theorem 3.2.9(i)–(ii) (starting with formula (3.2.5)) to show that

$$\|(\alpha - A)^{-n}\| \le M\alpha^{-n}$$

for all $n = 1, 2, 3, \ldots$ and $\alpha > 0$. According to Theorem 3.4.1, this implies that A generates a C_0 semigroup $\mathfrak{A}_1$. Thus, to complete the proof it suffices to show that $\mathfrak{A}_1^t = \mathfrak{A}^t$ for all $t > 0$. However, by (3.10.3) and Theorem 3.2.9(i), for every $x^* \in X^*$, $x \in X$, and $\alpha > 0$,

$$\int_0^\infty e^{-\alpha t} x^*(\mathfrak{A}^t x - \mathfrak{A}_1^t x)\, dt = 0.$$

Thus, since the Laplace transform of a (scalar continuous) function determines the function uniquely (see Section 3.12), we must have $x^*\mathfrak{A}^t x = \mathfrak{A}_1^t x$ for all $x^* \in X^*$, $x \in X$, and $t \ge 0$. Thus, $\mathfrak{A}_1^t = \mathfrak{A}^t$ for all $t > 0$. In particular, $\mathfrak{A}$ is strongly continuous. □

As the following theorem shows, the class of semigroups generated by sectorial operators can be characterized in several different ways.

Theorem 3.10.6 *Let A be a closed operator on the Banach space X, and let $\gamma \in \mathbb{R}$. Then the following conditions are equivalent:*

(i) *A is the generator of an analytic semigroup $\mathfrak{A}^t$ with uniformly bounded growth bound γ on a sector $\Delta_\delta = \{t \in \mathbb{C} \mid |\arg t| < \delta\}$ (with $\delta > 0$; see Definition 3.10.1);*

(ii) *Every $\lambda \in \mathbb{C}_\gamma^+$ belongs to the resolvent set of A, and there is a constant C such that*

$$\|(\lambda - A)^{-1}\| \le \frac{C}{|\lambda - \gamma|}, \qquad \Re\lambda > \gamma;$$

(iii) *A is sectorial on some sector $\Sigma_{\theta,\gamma}$ (with $\pi/2 < \theta < \pi$; see Definition 3.10.2);*

(iv) *A is the generator of a semigroup $\mathfrak{A}$ which is differentiable on $(0, \infty)$, and there exist finite constants M_0 and M_1 such that*

$$\|\mathfrak{A}^t\| \le M_0 e^{\gamma t}, \qquad \|(A - \gamma)\mathfrak{A}^t\| \le \frac{M_1 e^{\gamma t}}{t}, \qquad t > 0.$$

Proof Throughout this proof we take, without loss of generality, $\gamma = 0$. (If $\gamma \ne 0$, then we replace $A - \gamma$ by A and $e^{-\gamma t}\mathfrak{A}^t$ by $\mathfrak{A}^t$; see Examples 2.3.5 and 3.2.6.)

(i) ⇒ (ii): Let $0 < \varphi < \delta$. Let A_φ be the generator of the semigroup $t \mapsto \mathfrak{A}^{e^{j\varphi}t}$. Then, by Theorem 3.2.9(i), every (real) $\alpha > 0$ belongs to the resolvent set of A_φ, and

$$(\alpha - A_\varphi)^{-1} x = \int_0^\infty e^{-\alpha s} \mathfrak{A}^{e^{j\varphi}s} x\, ds.$$

By the estimate that we have on $\mathfrak{A}$ in (i), we can make a change of integration variable from s to $t = e^{j\varphi}s$, $s = e^{-j\varphi}t$ to get

$$\begin{aligned}(\alpha - A_\varphi)^{-1}x &= e^{-j\varphi}\int_0^\infty e^{-\alpha e^{-j\varphi}t}\mathfrak{A}^t x\,ds\\ &= e^{-j\varphi}(\alpha e^{-j\varphi} - A)^{-1}\\ &= (\alpha - e^{j\varphi}A)^{-1},\end{aligned}$$

where the second equality follows from Theorem 3.2.9(i). Thus, $A_\varphi = e^{j\varphi}A$. The estimate in (ii) then follows from (i) and Theorem 3.2.9(ii).

(ii) $\Rightarrow$ (iii): See Lemma 3.10.3.

(iii) $\Rightarrow$ (iv): We know from Theorem 3.10.5 that A generates a C_0 semigroup $\mathfrak{A}$, which has the integral representation (3.10.2). Thus, for all $h > 0$,

$$\frac{1}{h}\big(\mathfrak{A}^t - 1\big) = \frac{1}{2\pi j}\int_\Gamma h^{-1}(e^{\lambda h} - 1)e^{\lambda t}(\lambda - A)^{-1}\,d\lambda.$$

The same type of estimates that we developed in the proof of Theorem 3.10.5 (where we split Γ into $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$) together with Lemma 3.2.10 and the Lebesgue dominated convergence theorem show that this has a limit as $h \downarrow 0$. Thus, $\mathfrak{A}^t$ is differentiable (in operator norm), and

$$A\mathfrak{A}^t = \frac{d}{dt}(\mathfrak{A}^t) = \frac{1}{2\pi j}\int_\Gamma \lambda e^{\lambda t}(\lambda - A)^{-1}\,d\lambda. \tag{3.10.4}$$

As in the proof of Theorem 3.10.5 we again split Γ into $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ and estimate. We leave it to the reader to check that the final estimate that we get in this way is of the type

$$\|A\mathfrak{A}^t\| \le M/t,$$

with the same constant $M = 2M_1 + M_2$ that we get in the proof of Theorem 3.10.5.

(iv) $\Rightarrow$ (i): We first claim that, for each $t > 0$, $\mathfrak{A}^t$ maps X into $\bigcap_{n=1}^\infty \mathcal{D}(A^n)$, and prove this as follows. By the differentiability assumption on $\mathfrak{A}^t$, for each $x \in X$, the function $x(t) = \mathfrak{A}^t x$ is differentiable for $t > 0$. By the semigroup property,

$$\dot{x}(t) = \lim_{h\downarrow 0}\frac{1}{h}(\mathfrak{A}^{t+h} - \mathfrak{A}^t)x, = \lim_{h\downarrow 0}\frac{1}{h}(\mathfrak{A}^h - 1)\mathfrak{A}^t x,$$

hence $x(t) = \mathfrak{A}^t x \in \mathcal{D}(A)$, and $\frac{d}{dt}(\mathfrak{A}^t) = A\mathfrak{A}^t$. The same argument shows that

$\mathfrak{A}^t$ maps $\mathcal{D}\left(A^k\right)$ into $\mathcal{D}\left(A^{k+1}\right)$ for all $n = 0, 1, 2, \dots$ (because $A^k\mathfrak{A}^t x = \mathfrak{A}^t A^k x$ for all $x \in \mathcal{D}\left(A^k\right)$; see Theorem 3.2.1(iii)). By the semigroup property, $\mathfrak{A}^t = (\mathfrak{A}^{t/n})^n$, and this together with our earlier observation that $\mathfrak{A}^{t/n}$ maps $\mathcal{D}\left(A^k\right)$ into $\mathcal{D}\left(A^{k+1}\right)$ implies that $\mathfrak{A}^t$ maps X into $\mathcal{D}(A^n)$ for each $t > 0$ and $n = 1, 2, 3, \dots$. Thus, $\mathfrak{A}^t$ maps X into $\bigcap_{n=1}^{\infty} \mathcal{D}(A^n)$. This together with Theorem 3.2.1(vi) implies that $x(t) = \mathfrak{A}^t x$ is infinitely many times differentiable on $(0, \infty)$, and that

$$x^{(n)}(t) = (\mathfrak{A}^t)^{(n)} = A^n \mathfrak{A}^t x, \qquad t > 0.$$

Since A commutes with $\mathfrak{A}^t$ on $\bigcap_{n=1}^{\infty} \mathcal{D}(A^n)$, and since $\mathfrak{A}^t = (\mathfrak{A}^{t/n})^n$, this implies that

$$(\mathfrak{A}^t)^{(n)} = A^n \mathfrak{A}^t = (A\mathfrak{A}^{t/n})^n, \qquad t > 0. \tag{3.10.5}$$

Observe that the operators $A^n\mathfrak{A}^t$ are closed with $\mathcal{D}\left(A^n\mathfrak{A}^t\right) = X$, hence, by the closed graph theorem, they are bounded.

So far we have only used the differentiability of $\mathfrak{A}^t$ and not the norm estimate in (iv). This norm estimate, combined with (3.10.5) and the fact that $\mathrm{e}^n > n^n/n!$ (this is one of the terms in the power series expansion of e^n) gives

$$\left\| \frac{(\mathfrak{A}^t)^{(n)}}{n!} \right\| = \left\| \frac{A^n\mathfrak{A}^t}{n!} \right\| \le \left(\frac{M_1 \mathrm{e}}{t}\right)^n, \quad t > 0, \quad n = 1, 2, 3, \dots. \tag{3.10.6}$$

For each $t > 0$ we can define a formal power series

$$\widetilde{\mathfrak{A}}^z = \sum_{n=0}^{\infty} \frac{(\mathfrak{A}^t)^{(n)}}{n!}(z - t)^n. \tag{3.10.7}$$

The estimate (3.10.6) shows that this power series converges in operator norm for $|z - t| < t/(M_1\mathrm{e})$. Being the limit of a power series, the function $\widetilde{\mathfrak{A}}^z$ is analytic in the circle $|z - t| < t/(M_1\mathrm{e})$.

We claim $\widetilde{\mathfrak{A}}^z = \mathfrak{A}^z$ for real z satisfying $-t/(1 + M_1\mathrm{e}) < z - t < t/(M_1\mathrm{e})$. This is obvious for $z = t$. To see that it is true for other values of z we fix $x \in X$ and use a Taylor expansion of order $N - 1$ with Lagrangian remainder term for $\mathfrak{A}^z$, i.e.,

$$\mathfrak{A}^z x = \sum_{n=0}^{N-1} \frac{(\mathfrak{A}^t)^{(n)}x}{n!}(z - t)^n + \frac{(\mathfrak{A}^\xi)^{(N)}x}{N!}(z - t)^N,$$

where $z < \xi < t$ or $t < \xi < z$, depending on whether $z < t$ or $z > t$. By (3.10.6), the remainder term can be estimated by $(M_1\mathrm{e})^N|z - t|^N(\min\{t, z\})^{-N}$, which tends to zero in the given interval. Thus, $\widetilde{\mathfrak{A}}^z = \mathfrak{A}^z$ for real z satisfying $-t/(1 + M_1\mathrm{e}) < z - t < t/(M_1\mathrm{e})$. Letting $t > 0$ vary we conclude that $\widetilde{\mathfrak{A}}^t$ is an analytic extension of $\mathfrak{A}^t$ to the sector Δ_δ, where $0 < \delta \le \pi/2$ satisfies $\sin\delta = 1/(M_1\mathrm{e})$ (we choose z perpendicular to $z - t$, and assume, without loss of generality, that $M_1 e \geq 1$). The analyticity of $\widetilde{\mathfrak{A}}$ implies that $\widetilde{\mathfrak{A}}$ inherits the semigroup property $\widetilde{\mathfrak{A}}^{s+t} = \widetilde{\mathfrak{A}}^s \widetilde{\mathfrak{A}}^t$ from $\mathfrak{A}$ (two analytic functions which coincide on the real axis coincide everywhere). In each smaller subsector Δ_φ with $\varphi < \delta$ we can choose z and t (with z still perpendicular to $z - t$) so that $(|z - t| M_1 e)/t \leq k < 1$ (uniformly in z and t), and this implies that

$$\|\widetilde{\mathfrak{A}}^z\| \leq \|\mathfrak{A}^t\| + \sum_{n=1}^{\infty} \frac{\|(\mathfrak{A}^t)^{(n)}\|}{n!} |z - t|^n \leq M_0 + k/(1 - k).$$

Thus, the extended semigroup is uniformly bounded in each proper subsector Σ_φ of Σ_δ. If we rename φ to δ, then we have shown that $\widetilde{\mathfrak{A}}$ satisfies conditions (i)–(iii) in Definition 3.10.1.

The only thing that remains to be proved is the strong continuity, i.e., condition (iv) in Definition 3.10.1. Because of the density of $\mathcal{D}(A)$ in X and the uniform bound that we have on $\|\widetilde{\mathfrak{A}}^z\|$ for all $z \in \Delta_\delta$, it suffices to show that, for all $x \in \mathcal{D}(A)$,

$$\lim_{\substack{t \to 0 \\ t \in \Delta_\delta}} \mathfrak{A}^t x = x. \tag{3.10.8}$$

By using (3.10.6) and (3.10.7) we can estimate, for all $x \in \mathcal{D}(A)$,

$$\begin{aligned}
|\widetilde{\mathfrak{A}}^z x - x| &\leq |\mathfrak{A}^t x - x| + |\widetilde{\mathfrak{A}}^z x - \mathfrak{A}^t x| \\
&\leq |\mathfrak{A}^t x - x| + \sum_{n=1}^{\infty} \frac{|A^n \mathfrak{A}^t x|}{n!} |z - t|^n \\
&\leq |\mathfrak{A}^t x - x| + |z - t| \sum_{n=1}^{\infty} \frac{\|A^{n-1} \mathfrak{A}^t\| |Ax|}{n(n-1)!} |z - t|^{n-1} \\
&\leq |\mathfrak{A}^t x - x| + |z - t| |Ax| \sum_{n=1}^{\infty} \Big(\frac{|z - t| M_1 e}{t} \Big)^{n-1} \\
&\leq |\mathfrak{A}^t x - x| + \frac{|z - t| |Ax|}{1 - k},
\end{aligned}$$

where the last inequality is true provided we choose z and t to satisfy $(|z - t| M_1 e)/t \leq k < 1$. As before, we take z perpendicular to $z - t$. Let $\varphi = \arg z$. Then $|\varphi| < \delta$, $|z - t|/t = |\sin \varphi| < \sin \delta \leq 1$, and

$$t \cos \delta < t \cos \varphi = |z| \leq t.$$

Thus, with this choice of z and t, the condition $z \to 0$, $z \in \Delta_\delta$, is equivalent to $t \downarrow 0$, and

$$|\widetilde{\mathfrak{A}}^z x - x| \leq |\mathfrak{A}^t x - x| + \frac{t |Ax|}{1 - k}.$$

Because of the strong continuity of $\mathfrak{A}^t$, this tends to zero as $t \downarrow 0$. Thus, (3.10.8) holds, and the proof is complete. □

Let us record the following facts, which were established as a part of the proof of Theorem 3.10.6:

Corollary 3.10.7 *If the equivalent conditions listed in Theorem 3.10.6 hold, then, for each $\varphi \in (0, \delta)$, the generator of the semigroup $\mathfrak{A}^{e^{j\varphi}t}$, $t \geq 0$, is $e^{j\varphi}A$, where A is the generator of $\mathfrak{A}$.*

This was established as a part of the proof that (i) ⇒ (ii).

Corollary 3.10.8 *Condition (iv) in Theorem 3.10.6 implies that*

$$\|(A-\gamma)^n\mathfrak{A}^t\| \leq \frac{(nM_1)^n e^{\gamma t}}{t^n}, \quad t > 0, \quad n = 1, 2, 3, \ldots.$$

This follows from (3.10.5) (with A replaced by $A-\gamma$ and $\mathfrak{A}^t$ replaced by $e^{-\gamma t}\mathfrak{A}^t$).

In our applications to well-posed linear systems we are especially interested in the following extension of the estimates in Definition 3.10.2 and Corollary 3.10.8 to fractional powers of $(\gamma - A)$:

Lemma 3.10.9 *Let A be sectorial on some sector $\Sigma_{\theta,\gamma'}$, let $\gamma > \gamma'$, and let $\mathfrak{A}$ be the analytic semigroup generated by A. Then there exist constants M and C such that, for all $0 \leq \alpha \leq 1$,*

$$\begin{aligned}\|(\gamma - A)^\alpha \mathfrak{A}^t\| &\leq M(1+t^{-\alpha})e^{\gamma' t}, \quad t > 0,\\ \|(\gamma - A)^\alpha(\lambda - A)^{-1}\| &\leq C|\lambda-\gamma'|^{-1}\big(1+|\lambda-\gamma'|^\alpha\big), \quad \lambda \in \Sigma_{\theta,\gamma'}.\end{aligned}$$

Proof For $\alpha = 0$ and $\alpha = 1$ these estimates follow from Definition 3.10.2 and Theorem 3.10.6(iv). For intermediate values of α we interpolate between these two extreme values by using Lemma 3.9.8 as follows. By that lemma and Theorem 3.10.6(iv), we have for all $x \in X$ and $t > 0$ (recall that $\mathfrak{A}^t x \in \mathcal{D}(A)$ for $t > 0$)

$$\begin{aligned}\|(\gamma - A)^\alpha \mathfrak{A}^t x\| &\leq 2C|\mathfrak{A}^t x|_X^{1-\alpha}|(\gamma' - A)\mathfrak{A}^t x + (\gamma-\gamma')\mathfrak{A}^t x|_X^\alpha\\ &\leq 2CM_0^{1-\alpha}(M_1/t + |\gamma-\gamma'|M_0)^\alpha e^{\gamma' t}|x|_X\\ &\leq C_1(1+t^{-\alpha})e^{\gamma' t}|x|_X.\end{aligned}$$

To get the second inequality we argue essentially in the same way but replace Theorem 3.10.6(iv) by Definition 3.10.2:

$$\begin{aligned}&\|(\gamma - A)^{\alpha}(\lambda - A)^{-1}x\|\\&\quad \le 2C|(\lambda - A)^{-1}x|_X^{1-\alpha}|(\gamma - A)(\lambda - A)^{-1}x|_X^{\alpha}\\&\quad \le C_1|\lambda - \gamma'|^{\alpha-1}\big(1 + |\lambda - \gamma|/|\lambda - \gamma'|\big)^{\alpha}|x|_X\\&\quad \le C_2|\lambda - \gamma'|^{-1}\big(1 + |\lambda - \gamma'|^{\alpha}\big)|x|_X.\end{aligned}$$

□

Lemma 3.10.9 enables us to add the following conclusion to Theorem 3.8.2 in the case of an analytic semigroup:

Theorem 3.10.10 *Let $\mathfrak{A}$ be an analytic semigroup on X, and define the spaces X_α, $\alpha \in \mathbb{R}$ as above. Let $s \in \mathbb{R}$, $x_s \in X$, $1 < p \le \infty$, and $f \in L^p_{\mathrm{loc}}([s,\infty); X_{-\alpha})$, with*

$$\alpha < 1 - 1/p.$$

Then the function x given by (3.8.2) *is a strong solution of* (3.8.1) *in X.*

Proof This follows from (3.8.2), Lemma 3.10.9, and Hölder's inequality. □

Let us end this section with a perturbation result. The feedback transform studied in Chapter 7 leads to a perturbation of the original semigroup of the system, so that the generator A of this semigroup is replaced by $A + T$ for some operator T. In the analytic case we are able to allow a fairly large class of perturbations T without destroying the analyticity of the perturbed semigroup.

Theorem 3.10.11 *Let A be the generator of an analytic semigroup on the Banach space X, and define the fractional spaces X_α, $\alpha \in \mathbb{R}$, as in Section 3.9. If $T \in \mathcal{B}(X_\alpha; X_\beta)$ for some $\alpha, \beta \in \mathbb{R}$ with $\alpha - \beta < 1$, then the operator $(A + T)_{|X_\alpha}$ generates an analytic semigroup $\mathfrak{A}^T_{|X_{\alpha-1}}$ on $X_{\alpha-1}$. For $\gamma \in [\alpha - 1, \beta + 1]$, the spaces X_γ are invariant under $\mathfrak{A}^T_{|X_{\alpha-1}}$, and the restriction $\mathfrak{A}^T_{|X_\gamma}$ of $\mathfrak{A}^T_{|X_{\alpha-1}}$ to X_γ is an analytic semigroup on X_γ. The generator of $\mathfrak{A}^T_{|X_\gamma}$ is $(A + T)_{|X_{\gamma+1}}$ if $\gamma \in [\alpha - 1, \beta]$, and it is the part of $A + T$ in X_γ if $\gamma \in (\beta, \beta + 1]$.*[9] *Moreover, if $0 \in [\alpha - 1, \beta + 1]$ (so that $\mathfrak{A}^T_{|X}$ is an analytic semigroup on X) and if we let X^T_α, $\alpha \in \mathbb{R}$, be the analogues of the spaces X_α with A replaced by $A + T$, then $X^T_\gamma = X_\gamma$ for all $\gamma \in [\alpha - 1, \beta + 1]$.*

Proof We begin by studying the special case where $\alpha = 1$ and $0 < \beta \le 1$. By Theorem 3.10.6, every λ in some half-plane $\mathbb{C}^+_\mu$ belongs to the resolvent set of

[9] See Definition 3.14.12 and Theorem 3.14.14.

A, and for all $\lambda \in \mathbb{C}_\mu^+$,

$$\|(\lambda - A)^{-1}\| \le \frac{C}{|\lambda - \mu|}. \tag{3.10.9}$$

We claim that $A + T$ has the same property on some other half-plane $\mathbb{C}_\nu^+$. For all $\lambda \in \mathbb{C}_\mu^+$ we may write $\lambda - A - T$ in the form

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

Here $\lambda - A$ maps X_1 one-to-one onto X, so to show that $\lambda - A - T$ is invertible it suffices to show that $1 - (\lambda - A)^{-1}T$ is invertible in $\mathcal{B}(X_1)$. Fix some $\delta \in \rho(A)$. Then $(\delta - A)^{-\beta}T \in \mathcal{B}(X_1)$, and

$$\begin{aligned}\|(\lambda - A)^{-1}T\|_{\mathcal{B}(X_1)} &= \|(\delta - A)^{\beta}(\lambda - A)^{-1}(\delta - A)^{-\beta}T\|_{\mathcal{B}(X_1)} \\ &\le \|(\delta - A)^{\beta}(\lambda - A)^{-1}\|_{\mathcal{B}(X_1)}\|(\delta - A)^{-\beta}T\|_{\mathcal{B}(X_1)}.\end{aligned}$$

By Lemma 3.10.9 (with X replaced by X_1), we can make the right-hand side less than $\frac{1}{2}$ by choosing $|\lambda|$ large enough, and then it follows from the contraction mapping principle, $1 - (\lambda - A)^{-1}T$ is invertible in $\mathcal{B}(X_1)$ and $\|(1 - (\lambda - A)^{-1}T))^{-1}\|_{\mathcal{B}(X_1)} \le 2$. This proves that, for some sufficiently large $\mu_1 \ge \mu$, every $\mathbb{C}_{\mu_1}^+ \subset \rho(A + T)$, and that for all $\lambda \in \mathbb{C}_{\mu_1}^+$, $\lambda - A - T$ maps X_1 one-to-one onto X and

$$\|(\lambda - A - T)^{-1}\|_{\mathcal{B}(X;X_1)} \le 2\|(\lambda - A)^{-1}\|_{\mathcal{B}(X;X_1)}.$$

Fix some $\nu \in \mathbb{C}_{\mu_1}^+$. Then both $\nu - A$ and $\nu - A - T$ are boundedly invertible maps of X_1 onto X, so the above inequality implies the existence of some $C > 0$ such that, for all $\lambda \in \mathbb{C}_{\mu_1}^+$,

$$\|(\nu - A - T)(\lambda - A - T)^{-1}\|_{\mathcal{B}(X)} \le C\|(\nu - A)(\lambda - A)^{-1}\|_{\mathcal{B}(X)}.$$

Equivalently,

$$\|(\nu - \lambda)(\lambda - A - T)^{-1} + 1\|_{\mathcal{B}(X)} \le C\|(\nu - \lambda)(\lambda - A)^{-1} + 1\|_{\mathcal{B}(X)},$$

hence, by the triangle inequality,

$$\|(\lambda - A - T)^{-1}\|_{\mathcal{B}(X)} \le C\|(\lambda - A)^{-1}\|_{\mathcal{B}(X)} + \frac{C + 1}{|\lambda - \nu|}. \tag{3.10.10}$$

This, together with (3.10.9) and Theorem 3.10.6 implies that $A + T$ generates an analytic semigroup.

Still assuming that $\alpha = 1$ and $0 < \beta \le 1$, let us show that $X_\gamma^T = X_\gamma$ for all $\gamma \in (0, 1)$ (we already know that this is true for $\gamma = 0$ and $\gamma = 1$). This is equivalent to the claim that $(\nu' - A - T)^{-\gamma}$ has the same range as $(\nu' - A)^{-\gamma}$, where, for example, $\nu' > \nu$, and ν is the same constant as in the proof above.

A short algebraic computation shows that, for $s \ge 0$,

$$(s + \nu' - A)^{-1}(1 + T(s + \nu' - A - T)^{-1}) = (s - A)^{-1},$$

and therefore, by Lemma 3.9.9,

$$\begin{aligned}(\nu' - A - T)^{-\gamma} &- (\nu' - A)^{-\gamma} \\ &= c \int_0^\infty s^{-\gamma}\big(s + \nu' - A\big)^{-1} T\big(s + \nu' - A - T\big)^{-1}\, ds,\end{aligned} \tag{3.10.11}$$

where $c = \frac{\sin \pi\alpha}{\pi}$. Since $T \in \mathcal{B}(X_1; X_\beta)$, we get by a computation similar to the one leading to (3.10.10) that there exist constants $C_1, C_2 > 0$ such that for all $s \ge 0$ (see also (3.10.9)),

$$\big\| T\big(s + \nu' - A - T\big)^{-1} \big\|_{\mathcal{B}(X;X_\beta)} \le C_1 \big\| s(s + \nu' - A)^{-1} + 1 \big\|_{\mathcal{B}(X)} \le C_2.$$

If $\gamma \le \beta$, then we can estimate

$$\big\| \big(s + \nu' - A\big)^{-1} \big\|_{\mathcal{B}(X_\beta;X_\gamma)} \le C_3 \big\| \big(s + \nu' - A\big)^{-1} \big\|_{\mathcal{B}(X_\beta)} \le C_4 (s + \nu' - \nu)^{-1},$$

and if $\gamma > \beta$ then we get from Lemma 3.10.9 (with $\alpha = \gamma - \beta$),

$$\big\| \big(s + \nu - A\big)^{-1} \big\|_{\mathcal{B}(X_\beta;X_\gamma)} \le C((s + \nu' - \nu)^{-1} + (s + \nu' - \nu)^{\gamma-\beta-1}).$$

In either case we find that the operator-norm in $\mathcal{B}(X; X_\gamma)$ of the integrand in (3.10.11) is bounded by a constant times $s^{-\gamma}\big((s + \nu' - \nu)^{-1} + (s + \nu' - \nu)^{\gamma-\beta-1}\big)$, and this implies that the integral converges in $\mathcal{B}(X; X_\gamma)$. Since also $(\nu' - A)^{-\gamma} \in \mathcal{B}(X; X_\gamma)$, this implies that $\mathcal{R}\left((\nu' - A - T)^{-\gamma}\right) \subset X_\gamma$, or in other words, $X_\gamma^T \subset X_\gamma$. To prove the opposite inclusion it suffices to show apply the same argument with A replaced by $A + T$ and T replaced by $-T$ (i.e., we interchange the roles of A and $A + T$.

Let us recall what we have proved so far: If $\alpha = 1$ and $0 < \beta \le 1$, then $A + T$ generates an analytic semigroup on X, and $X_\gamma^T = X_\gamma$ for all $\gamma \in [0, 1]$. The restriction $\beta \le 1$ is irrelevant, because if $\beta \ge 1$, then $T \in \mathcal{B}(X_1)$, and we can repeat the same argument with β replaced by one. Thus, the conclusion that we have established so far is valid for all $\beta > 0$ when $\alpha = 1$.

Next we look at another special case, namely the one where $\beta = 0$ and $0 \le \alpha < 1$. The proof is very similar to the one above, so let us only indicate the differences, and leave the details to the reader. The estimate (3.10.9) is still valid. This time we write $\lambda - A - T$ in the form

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

where $\lambda - A$ still maps X_1 one-to-one onto X. To prove that $\lambda - A - T$ maps X_1 one-to-one onto X it suffices to show that $1 - T(\lambda - A)^{-1}$ is invertible in $\mathcal{B}(X)$ for sufficiently large $|\lambda|$, and this is done by using the same type of estimates as we saw above. In particular, we find that $\|T(\lambda - A)^{-1}\|_{\mathcal{B}(X)} \le \frac{1}{2}$

and that (3.10.10) holds for $|\lambda|$ large enough. Thus, $A + T$ generates an analytic semigroup on X.

If $\alpha < 0$ and $\beta = 0$, then $T \in \mathcal{B}(X)$, and we may simply replace α by zero in the preceding argument. Thus, $A + T$ generates an analytic semigroup on X whenever $\alpha < 1$ and $\beta = 0$.

Let us finally return to the general case where $\alpha, \beta \in \mathbb{R}$ with $\alpha - \beta < 1$. To deal with this case we replace X by the new state space $X' = X_{\beta-\epsilon}$, where ϵ is an arbitrary number satisfying $0 < \epsilon \le \beta - \alpha + 1$. With respect to this new state space the index α is replaced by $\alpha' = \alpha - \beta + \epsilon \le 1$, β is replaced by $\beta' = \epsilon > 0$, and $T \in \mathcal{B}(X'_{\alpha'}; X'_{\beta'})$. This implies that $T \in \mathcal{B}(X'_1; X'_{\beta'})$ (since $\alpha' \le 1$), so we can apply the first special case where we had $\alpha = 1$ and $\beta > 0$ (replacing X by X'). We conclude that $(A + T)_{|X'_1}$ generates an analytic semigroup $\mathfrak{A}^T_{|X'}$ on $X' = X_{\beta-\epsilon}$ and that $X'^{\,T}_{\gamma'} = X'_{\gamma'}$ for all $\gamma' \in [0, 1]$, where $X'^{\,T}_{\gamma'}$ are the fractional spaces constructed by means of the semigroup $\mathfrak{A}^T_{|X'}$ as described in Section 3.9. Moreover, by restricting $\mathfrak{A}^T_{|X'}$ to $X'^{\,T}_{\gamma'} = X'_{\gamma'}$ we get an analytic semigroup $\mathfrak{A}^T_{|X'_{\gamma'}}$ on $X'_{\gamma'}$ for all $\gamma' \in [0, 1]$. The above ϵ could be any number in $(0, \beta - \alpha + 1]$. In particular, taking $\epsilon = \beta - \alpha + 1$ we get a semigroup $\mathfrak{A}^T_{|X_{\alpha-1}}$ generated by $(A + T)_{|X_\alpha}$ on $X_{\alpha-1}$.

One drawback with the above construction is that the fractional spaces $X'^{\,T}_{\gamma'}$ that we get depend on X', hence on the choice of the parameter ϵ. However, as we saw in Section 3.9, we get the same spaces if we replace X' by any of the spaces $X'^{\,T}_{\gamma'}$, and adjust the index accordingly. Since $X'^{\,T}_{\gamma'} = X'_{\gamma'}$ for all $\gamma' \in [0, 1]$, this means that we may fix one particular value of the parameter ϵ (for example, $\epsilon = \beta - \alpha + 1$ which gives $X' = X_{\alpha-1}$), and base the definition of the fractional spaces induced by $A + T$ on this fixed value of ϵ. After that, by letting ϵ vary in $(0, \beta - \alpha + 1]$ we find that for all $\gamma \in [\alpha - 1, \beta)$, the operator $(A + T)_{|X_{\gamma+1}}$ generates the analytic semigroup $\mathfrak{A}^T_{|X_\gamma}$ on X_γ, where $\mathfrak{A}^T_{|X_\gamma}$ is the restriction of $\mathfrak{A}^T_{|X_{\alpha-1}}$ to X_γ. Moreover, for $\gamma \in [\beta, \beta + 1)$, the restriction $\mathfrak{A}^T_{|X_\gamma}$ of $\mathfrak{A}^T_{|X_{\alpha-1}}$ to X_γ is an analytic semigroup on X_γ. By Theorem 3.14.14, the generator of this semigroup is the part of $A + T$ in X_γ. If $0 \in [\alpha - 1, \beta + 1)$, then we may base our definition of the fractional spaces induced by $A + T$ on the original state space X.

We have now proved most of Theorem 3.10.11. The only open question is whether we can also take $\gamma = \beta$ (instead of $\gamma < \beta$). To deal with this case we replace the state space X by $X' = X_\beta$. Then α is replaced by $\alpha' = \alpha - \beta < 1$ and β is replaced by $\beta' = 0$. By the second of the two special cases that we studied above, $(A + T)_{|X'_1}$ generates an analytic semigroup $\mathfrak{A}^T_{|X_\beta}$ on $X' = X_\beta$ and $X'^{\,T}_1 = X'_1 = X_{\beta+1}$. Thus, $(A + T)_{|X_{\beta+1}}$ generates an analytic semigroup on X_β and $X^T_{\beta+1} = X_{\beta+1}$. Furthermore, the restriction of $\mathfrak{A}^T_{|X_\beta}$ to $X_{\beta+1}$ is an analytic semigroup on this space. $\square$

3.11 Spectrum determined growth

One of the most important properties of a semigroup $\mathfrak{A}$ is its growth bound $\omega_{\mathfrak{A}}$. We defined $\omega_{\mathfrak{A}}$ in Definition 2.5.6 in terms of the behavior of $\|\mathfrak{A}^t\|$ as $t \to \infty$. In practice it is often more convenient to use characterizations of $\omega_{\mathfrak{A}}$ which refer to the generator A of $\mathfrak{A}$ and not directly to $\mathfrak{A}$.

There is one obvious condition that the generator A of $\mathfrak{A}$ must satisfy: by Theorem 3.4.1 the half-plane $\mathbb{C}^+_{\omega_{\mathfrak{A}}}$ belongs to the resolvent set of A. Thus, it is always true that

$$\omega_{\mathfrak{A}} \geq \sup\{\Re\lambda \mid \lambda \in \sigma(A)\}. \tag{3.11.1}$$

If the converse inequality also holds, then we say that $\mathfrak{A}$ has the *spectrum determined growth property*:

Definition 3.11.1 The C_0 semigroup $\mathfrak{A}$ has the spectrum determined growth property if

$$\omega_{\mathfrak{A}} = \sup\{\Re\lambda \mid \lambda \in \sigma(A)\}. \tag{3.11.2}$$

Example 3.11.2 *Suppose that the semigroup $\mathfrak{A}$ on X has the spectrum determined growth property. Then so do the following semigroups derived from $\mathfrak{A}$:*

(i) *the exponentially shifted semigroup $t \mapsto \mathrm{e}^{\alpha t}\mathfrak{A}^t$ for every $\alpha \in \mathbb{C}$ (see Example 2.3.5);*
(ii) *the time compressed semigroup $t \mapsto \mathfrak{A}^{\lambda t}$ for every $\lambda > 0$ (see Example 2.3.6;*
(iii) *the similarity transformed semigroup $t \mapsto E^{-1}\mathfrak{A}^t E$ for every invertible $E \in \mathcal{B}(X_1; X)$ (see Example 2.3.7).*

The easy proof is left to the reader.

Some other classes of semigroups which have the spectrum determined growth property are the following:

Example 3.11.3 *The following semigroups have the spectrum determined growth property:*

(i) *the left shift semigroups τ, τ_+, τ_-, $\tau_{[0,T)}$, and $\tau_{\mathbb{T}_T}$ in Examples 2.3.2 and 2.5.3;*
(ii) *diagonal semigroups (see Example 3.3.3 and Definition 3.3.4);*
(iii) *analytic semigroups (see Definition 3.10.1).*

Proof (i) See Examples 2.5.3 and 3.3.1.

(ii) See Examples 3.3.3 and 3.3.5.

(iii) We know that (3.11.1) holds for all semigroups, so it suffices to prove the opposite inequality. Choose $\omega \in \mathbb{R}$ so that $\Re\lambda < \omega - \epsilon$ for some $\epsilon > 0$ and all $\omega \in \sigma(A)$. The fact that A is sectorial on some sector $\Sigma_{\theta,\omega'}$ for some $\pi/2 < \theta < \pi$ and some $\omega' \in \mathbb{R}$ then implies that condition (ii) in Theorem 3.10.6 holds. By the same theorem, the growth bound of $\mathfrak{A}$ is at most ω. Since $\epsilon > 0$ is arbitrary we get $\omega_{\mathfrak{A}} \le \sup\{\Re\lambda \mid \lambda \in \sigma(A)\}$. □

Example 3.11.3(iii) is a special case of the following theorem:

Theorem 3.11.4 *A semigroup $\mathfrak{A}^t$ on a Banach space which is continuous in the operator norm on an interval $[t_0, \infty)$, where $t_0 \ge 0$, has the spectral determined growth property.*

Proof Let $\omega_A = \sup\{\Re\lambda \mid \lambda \in \sigma(A)\}$. By Theorem 2.5.4(i), it suffices to show that, for some $t > 0$, the spectral radius of $\mathfrak{A}^t$ is equal to $e^{\omega_A t}$. However, this follows from the operator norm continuity of $\mathfrak{A}^t$ which implies that $e^{t\sigma(A)} \subset \sigma(\mathfrak{A}^t) \subset \{0\} \cup e^{t\sigma(A)}$; see Davies (1980, Theorems 2.16 and 2.19). □

Corollary 3.11.5 *The semigroup $\mathfrak{A}^t$ on the Banach space X has the spectrum determined growth property (at least) in the following cases:*

(i) *$t \mapsto \mathfrak{A}^t x$ is differentiable on $[t_0, \infty)$ for some $t_0 \ge 0$ and all $x \in X$, or equivalently, $\mathcal{R}\left(\mathfrak{A}^{t_0}\right) \subset \mathcal{D}(A)$ for some $t_0 \ge 0$.*
(ii) *$\mathfrak{A}^t$ is compact for all $t \ge t_0$, where $t_0 \ge 0$, or equivalently, $\mathfrak{A}^{t_0}$ is compact for some $t_0 \ge 0$.*

Proof This follows from Theorem 3.11.4 and the fact that in both cases, $t \mapsto \mathfrak{A}^t$ is continuous in the operator norm on $[t_0, \infty)$. See, for example, Pazy (1983, Theorem 3.2 and Lemma 4.2). □

One way of looking at the spectrum determined growth property is to interpret it as a condition that the generator A must not have a spectrum 'at infinity' in the right half-plane $\mathbb{C}^+_{\omega_{\mathfrak{A}}}$. This idea can be made more precise. Theorem 3.4.1 not only implies that the half-plane $\mathbb{C}^+_{\omega_{\mathfrak{A}}}$ belongs to the resolvent set of A, but in fact,

$$\sup_{\Re\lambda \ge \omega_{\mathfrak{A}}+\epsilon} \|(\lambda - A)^{-1}\| < \infty, \qquad \epsilon > 0.$$

As the following theorem shows, this condition can be used to determine $\omega_{\mathfrak{A}}$ whenever the state space is a Hilbert space.

Theorem 3.11.6 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ on a Hilbert space X. Then*

$$\omega_{\mathfrak{A}} = \inf\left\{\omega \in \mathbb{R} \;\middle|\; \sup_{\Re\lambda \ge \omega} \|(\lambda - A)^{-1}\| < \infty\right\}.$$

The proof of this theorem will be given in Section 10.3 (there this theorem is reformulated as Theorem 10.3.7).

This theorem has the following interesting consequence:

Lemma 3.11.7 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ on a Hilbert space X, and suppose that, for some $\omega \in \mathbb{R}$, some $M > 0$, and some $0 < \epsilon \le 1$, the half-plane $\mathbb{C}^+_\omega$ belongs to the resolvent set of A, and*

$$\|(\lambda - A)^{-1}\| \le M(\Re\lambda - \omega)^{\epsilon - 1}, \qquad \Re\lambda > \omega. \tag{3.11.3}$$

Then the growth bound $\omega_{\mathfrak{A}}$ of $\mathfrak{A}$ satisfies

$$\omega_{\mathfrak{A}} \le \omega - \delta,$$

where $\delta > 0$ is given by

$$\delta = \begin{cases} 1/M, & \epsilon = 1, \\ \epsilon(1-\epsilon)^{1/\epsilon - 1} M^{-1/\epsilon}, & 0 < \epsilon < 1. \end{cases}$$

In particular, we conclude that (3.11.3) cannot possibly hold with $\omega = \omega_{\mathfrak{A}}$ (because that would imply $\omega_{\mathfrak{A}} < \omega_{\mathfrak{A}}$). (It cannot hold for $\epsilon > 1$, either.)

Proof The proof is based on Lemma 3.2.8(ii) and Theorem 3.11.6. The case $\epsilon = 1$ follows directly from these two results (the line $\Re\lambda = \omega$ must belong to the resolvent set in this case, and, by continuity, (3.11.3) holds for $\Re\lambda = \omega$, too).

If $0 < \epsilon < 1$, then we take α in Lemma 3.2.8(ii) to lie on the line

$$\Re\alpha = \omega + (1-\epsilon)^{1/\epsilon} M^{-1/\epsilon}$$

(this is the choice that will maximize the constant δ), and take λ to have $\Re\lambda \le \Re\alpha$ and $\Im\lambda = \Im\alpha$. Then, by Lemma 3.2.8(ii), λ will belong to the resolvent set of A if

$$\Re(\alpha - \lambda) = \alpha - \lambda < 1/\|(\alpha - A)^{-1}\|,$$

which by (3.11.3) is true whenever $\Re(\alpha - \lambda) < 1/\kappa$, where

$$\kappa = M^{-1}(\Re\alpha - \omega)^{\epsilon - 1} = (1-\epsilon)^{1/\epsilon - 1} M^{-1/\epsilon}.$$

This can equivalently be rewritten as

$$\Re\alpha - 1/\kappa = \omega - \delta < \Re\lambda \le \omega + (1-\epsilon)^{1/\epsilon} M^{-1/\epsilon}.$$

Moreover, by the same lemma, $\|(\lambda - A)^{-1}\|$ is bounded by

$$\begin{aligned} \|(\lambda - A)^{-1}\| &\le \frac{\|(\lambda - A)^{-1}\|}{1 - \|(\lambda - A)^{-1}\|(\alpha - \lambda)} \\ &\le \frac{\kappa}{1 - \kappa(\alpha - \lambda)} = \frac{1}{\Re\lambda - (\omega - \delta)}. \end{aligned}$$

The conclusion now follows from Theorem 3.11.6. □

There is a related result which connects the growth bound of a semigroup to some L^p-estimates in the time domain:

Theorem 3.11.8 *Let $\mathfrak{A}$ be a C_0 semigroup on the Banach space X and let $\omega \in \mathbb{R}$. Then the following conditions are equivalent:*

(i) $\omega_{\mathfrak{A}} < \omega$;

(ii) $e^{-\omega t}\|\mathfrak{A}^t\| \to 0$ *as* $t \to \infty$;

(iii) $\|\mathfrak{A}^t\| < e^{\omega t}$ *for some* $t > 0$;

(iv) *For all $x_0 \in X$, the function $x(t) = \mathfrak{A}^t x_0$, $t \geq 0$, belongs to $L^p_\omega(\mathbb{R}^+; X)$ for all $p \in [1, \infty]$;*

(v) *For some $p \in [1, \infty)$ and all $x_0 \in X$, the function $x(t) = \mathfrak{A}^t x_0$, $t \geq 0$, belongs to $L^p_\omega(\mathbb{R}^+; X)$;*

(vi) *For all $q \in [1, \infty]$, there is a finite constant M_q such that, for all $u \in C_c(\mathbb{R}^-; X)$,*

$$\left| \int_{-\infty}^{0} \mathfrak{A}^{-s} u(s)\, ds \right|_X \leq M_q \|u\|_{L^q_\omega(\mathbb{R}^-;X)}.$$

(vii) *For some $q \in (1, \infty]$, some finite constant M_q, and all $u \in C^1_c(\mathbb{R}^-; X_1)$,*

$$\left| \int_{-\infty}^{0} \mathfrak{A}^{-s} u(s)\, ds \right|_X \leq M_q \|u\|_{L^q_\omega(\mathbb{R}^-;X)}.$$

(viii) *For all $p \in [1, \infty]$, there is a finite constant M_p such that, for all $u \in C_c(\mathbb{R}; X)$,*

$$\left\| t \mapsto \int_{-\infty}^{t} \mathfrak{A}^{t-s} u(s)\, ds \right\|_{L^p_\omega(\mathbb{R};X)} \leq M_p \|u\|_{L^p_\omega(\mathbb{R};X)}.$$

(ix) *For some $p \in [1, \infty]$, some finite constant M_p, and all $u \in C^1_c(\mathbb{R}^+; X)$,*

$$\left\| t \mapsto \int_{0}^{t} \mathfrak{A}^{t-s} u(s)\, ds \right\|_{L^p_\omega(\mathbb{R}^+;X)} \leq M_p \|u\|_{L^p_\omega(\mathbb{R}^+;X)}.$$

In particular, if $\omega = 0$, then all the preceding conditions are equivalent to the exponential stability of $\mathfrak{A}$.

Proof Without loss of generality, we take $\omega = 0$ (see Examples 2.3.5 and 3.2.6).

(i) ⇒ (ii) and (i) ⇒ (iv): This follows from Theorem 2.5.4(i).

(i) ⇒ (vi): Use Theorem 2.5.4(i) and Hölder's inequality.

(i) ⇒ (viii): Apply Theorem 2.5.4(ii) to the system in Proposition 2.3.1 with $B = C = 1$ and $D = 0$.

(ii) ⇒ (iii), (iv) ⇒ (v), (vi) ⇒ (vii), and (viii) ⇒ (ix): These implications are obvious.

(iii) $\Rightarrow$ (i): Clearly, the spectral radius of $\mathfrak{A}^t$ is less than one, and Theorem 2.5.4(i) implies that $\omega_{\mathfrak{A}} < 0$.

(v) $\Rightarrow$ (ii): The operator $x_0 \mapsto x$ is continuous $X \to C(\overline{\mathbb{R}}^+; X)$, hence it is closed $X \to L^p(\mathbb{R}^+; X)$. By the closed graph theorem, there is a constant $M_p > 0$ such that

$$\left(\int_0^\infty |\mathfrak{A}^s x_0|_X^p \, ds\right)^{1/p} \le M_p |x_0|_X, \qquad x_0 \in X. \tag{3.11.4}$$

By Theorem 2.5.4(i), there are constants $\alpha > 0$ and $M > 0$ such that $\|\mathfrak{A}^t\| \le M e^{\alpha t}$. Therefore, for all $x_0 \in X$ and $t > 0$,

$$\begin{aligned}
\frac{1 - e^{-p\alpha t}}{p\alpha} |\mathfrak{A}^t x_0|_X^p &= \int_0^t e^{-p\alpha s} \, ds |\mathfrak{A}^t x_0|_X^p = \int_0^t e^{-p\alpha s} |\mathfrak{A}^s \mathfrak{A}^{t-s} x_0|_X^p \, ds \\
&\le \int_0^t \|e^{-\alpha s} \mathfrak{A}^s\|^p |\mathfrak{A}^{t-s} x_0|_X^p \, ds \le M^p \int_0^t |\mathfrak{A}^{t-s} x_0|_X^p \, ds \\
&\le M^p M_p^p |x_0|_x^p.
\end{aligned}$$

This implies that there is a finite constant M_∞ such that $\|\mathfrak{A}^t\| \le M_\infty$ for $t \ge 0$. We can now repeat the same computation with $\alpha = 0$ to get

$$\begin{aligned}
t \|\mathfrak{A}^t x_0\|_X^p &= \int_0^t |\mathfrak{A}^t x_0|_X^p \, ds \le \int_0^t \|\mathfrak{A}^s\|^p |\mathfrak{A}^{t-s} x_0|_X^p \, ds \\
&\le M_\infty^p \int_0^t |\mathfrak{A}^{t-s} x_0|_X^p \, ds \le M_\infty^p M_p^p |x_0|_x^p,
\end{aligned}$$

which implies that

$$\|\mathfrak{A}^t\| \le (M_p M_\infty) t^{-1/p}.$$

(vii) $\Rightarrow$ (ii): Let $t > 0$. If $q < \infty$ then we can use the density of $C^1([-t, 0); X_1)$ in $L^q([-t, 0); X)$, and if $q = \infty$ then we can use a simple approximation argument and the Lebesgue dominated convergence theorem to show that, for all $t > 0$ and all $u \in C([0, t); X)$,

$$\left| \int_0^t \mathfrak{A}^s u(s) \, ds \right|_X = \left| \int_{-t}^0 \mathfrak{A}^{-s} u(-s) \, ds \right|_X \le M_q \|u\|_{L^q([0,t);X)}.$$

In particular, taking $u(s) = e^{\alpha s} \mathfrak{A}^{t-s} x_0$, where $\alpha > 0$ and $x_0 \in X$ we get

$$\begin{aligned}
\frac{e^{\alpha t} - 1}{\alpha} |\mathfrak{A}^t x_0|_X &= \left| \int_0^t e^{\alpha s} \mathfrak{A}^t x_0 \, ds \right|_X = \left| \int_0^t \mathfrak{A}^s e^{\alpha s} \mathfrak{A}^{t-s} x_0 \, ds \right|_X \\
&\le e^{\alpha t} M_q K_q |x_0|_X,
\end{aligned}$$

where K_q is the (finite) L^q-norm over $\mathbb{R}$ of the function $s \mapsto e^{-\alpha s} \|\mathfrak{A}^s\|$. This implies that $\|\mathfrak{A}^t\| \le M_1$ for some finite M_1 and all $t \ge 0$. We then repeat the same computation with $\alpha = 0$ to get

$$t|\mathfrak{A}^t x_0|_X = \left|\int_0^t \mathfrak{A}^t x_0\,ds\right|_X = \left|\int_0^t \mathfrak{A}^s\mathfrak{A}^{t-s} x_0\,ds\right|_X$$
$$\le M_q \left\|(s \mapsto \|\mathfrak{A}^s\|)\right\|_{L^q([0,t))}|x_0|_X \le M_q M_1 t^{1/q}|x_0|_X,$$

i.e., $\|\mathfrak{A}^t\| \le M_q M_1 t^{1/q-1}$.

(ix) $\Rightarrow$ (i): First we consider the case $p = \infty$. In (ix) we can replace $C_c^1(\mathbb{R}^+; X)$ by $C_c^1(\mathbb{R}; X)$ if we at the same time replace $L^\infty(\mathbb{R}^+; X)$ by $L^\infty(\mathbb{R}; X)$ and $\int_0^t \mathfrak{A}^{t-s}u(s)\,ds$ by $\int_{-\infty}^t \mathfrak{A}^{t-s}u(s)\,ds$ since the convolution operator $t \mapsto \int_{-\infty}^t \mathfrak{A}^{t-s}u(s)\,ds$ is time-invariant (shift $u \in C_c^1(\mathbb{R}; X)$ to the right until it is supported on $\mathbb{R}^+$, apply the convolution operator, and then shift the result back). The integral $\int_{-\infty}^t \mathfrak{A}^{t-s}u(s)\,ds$ is continuous in t, and by evaluating this integral at zero we conclude that (vii) holds with $q = \infty$. As we saw above, this implies (i).

In the case $p < \infty$ we argue as follows. As $C_c^1(\mathbb{R}^+; X)$ is dense in $L^p(\mathbb{R}^+; X)$, we can weaken the condition $u \in C_c^1(\mathbb{R}^+; X)$ to $u \in L^p(\mathbb{R}^+; X)$, and the same estimate still holds. Let $x_0 \in X$, and define

$$u(t) = \begin{cases} \mathfrak{A}^t x_0, & 0 \le t < 1, \\ 0, & t \ge 1, \end{cases}$$

and $y(t) = \int_0^t \mathfrak{A}^{t-s}u(s)\,ds$. Then

$$y(t) = \begin{cases} t\mathfrak{A}^t x_0, & 0 \le t < 1, \\ \mathfrak{A}^t x_0, & t \ge 1, \end{cases}$$

and by (ix), $y \in L^p(\mathbb{R}^+; X)$. In particular, $t \mapsto \mathfrak{A}^t x_0 \in L^p(\mathbb{R}^+; X)$ for every $x_0 \in X$. Thus (v) is satisfied, hence so is (i). □

3.12 The Laplace transform and the frequency domain

Some of the shift semigroups in Examples 2.3.2 and 2.5.3 and their generators and resolvents described in Examples 3.3.1 and 3.3.2 have simple frequency domain descriptions, which will be presented next. When we say ‘frequency domain description’ we mean a description given in terms of Laplace transforms of the original functions.

Definition 3.12.1 The *(right-sided) Laplace transform* of a function $u \in L^1_{\mathrm{loc}}(\mathbb{R}^+; U)$ is given by

$$\hat{u}(z) = \int_0^\infty \mathrm{e}^{-zt}u(t)\,dt$$

for all $z \in \mathbb{C}$ for which the integral converges (absolutely). The *left-sided Laplace transform* of a function $u \in L^1_{\mathrm{loc}}(\mathbb{R}^-; U)$ is given by

$$\hat{u}(z) = \int_{-\infty}^{0} \mathrm{e}^{-zt} u(t)\, dt$$

for all $z \in \mathbb{C}$ for which the integral converges (absolutely). The *bilateral Laplace transform* of a function $u \in L^1_{\mathrm{loc}}(\mathbb{R}; U)$ is given by

$$\hat{u}(z) = \int_{-\infty}^{\infty} \mathrm{e}^{-zt} u(t)\, dt$$

for all $z \in \mathbb{C}$ for which the integral converges (absolutely). The *finite Laplace transform over the interval* $[0, T]$ of a function $u \in L^1([0, T]; U)$ is given by

$$\hat{u}(z) = \int_{0}^{T} \mathrm{e}^{-zt} u(t)\, dt$$

for all $z \in \mathbb{C}$.

The finite Laplace transform is always an entire function. It is easy to see that the domains of definition of the other Laplace transforms, if nonempty, are vertical strips $\{z \in \mathbb{C} \mid \Re z \in J\}$, where J is an interval in $\mathbb{R}$ (open, closed, or semi-closed, bounded or unbounded). The one-sided, left-sided, and finite Laplace transforms can be interpreted as special cases of the bilateral Laplace transform, namely, the case where u vanishes on the complements of $\mathbb{R}^+$, $\mathbb{R}^-$, or $[0, T]$. For the one-sided Laplace transform, either J is empty or the right end-point of J is $+\infty$, and for the left-sided Laplace transform, either J is empty or the left end-point of J is $-\infty$. In the case of the bilateral Laplace transform J may be bounded or unbounded, and it may consist of one point only. In the interior of their domains all the different Laplace transforms are analytic. For example, if $u \in L^1_c(\mathbb{R}; U)$, then the bilateral Laplace transform of u is entire (as in the case of the finite Laplace transform).

All the Laplace transforms listed above determine the original function u uniquely (on its given domain), whenever the domain of the Laplace transform is nonempty. To prove this it suffices to consider the case of the bilateral transform (since the others can be reduced to this one). If $\hat{u}(z)$ is defined for some $z \in \mathbb{C}$, then $u \in L^1_\alpha(\mathbb{C}; U)$ where $\alpha = \Re z$. Under some further conditions we can get the following explicit formula for u in terms of $\hat{u}$.

Proposition 3.12.2 *Let $u \in L^1_\alpha(\mathbb{R}; U)$, where U is a Banach space, and suppose that the bilateral Laplace transform $\hat{u}$ of u satisfies $\hat{u} \in L^1(\alpha + j\mathbb{R}; U)$. Define*

$$v(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{(\alpha + j\omega)t} \hat{u}(\alpha + j\omega)\, d\omega, \qquad t \in \mathbb{R}.$$

Then $u = v$ almost everywhere, $v \in BC_{\alpha,0}(\mathbb{R}; U)$, and $\hat{u} \in BC_0(\alpha + j\mathbb{R}; U)$.

Proof In the scalar case with $\alpha = 0$ this is shown in Rudin [1987, Theorem 9.12, p. 185]. We can introduce a general $\alpha \in \mathbb{R}$ by applying the same result with $u(t)$ replaced by $e^{-\alpha t}u(t)$. The proof of the vector-valued case is identical to the proof of the scalar case given in Rudin (1987). □

It follows from Proposition 3.12.2 that the Laplace transform is injective: If two functions u and v have the same Laplace transform, defined on the same vertical line $\Re z = \alpha$, then they must be equal almost everywhere (apply Proposition 3.12.2 to the difference $u - v$).

In the case of the left-sided or right-sided Laplace transform, the most restrictive condition in Proposition 3.12.2 is the requirement that $\hat{u} \in L^1(\alpha + j\mathbb{R}; U)$. If u is continuous with $u(0) \neq 0$, then Proposition 3.12.2 cannot be applied to the one-sided transforms, due to the fact that $\pi_+ u$ and/or $\pi_- u$ will have a jump discontinuity at zero. The following proposition has been designed to take care of this problem.

Proposition 3.12.3 *Let U be a Banach space.*

(i) *Suppose that $u \in L^1_\alpha(\mathbb{R}^+; U)$ and that there exists some $u_0 \in \mathbb{C}$ such that the function $\omega \mapsto \hat{u}(\alpha + j\omega) - (1 + j\omega)^{-1}u_0$ belongs to $L^1(\mathbb{R}; U)$, where $\hat{u}$ is the (right-sided) Laplace transform of u. Let $\beta \in \mathbb{C}^-_\alpha$, and define for all $t \in \mathbb{R}^+$,*

$$v(t) = e^{\beta t}u_0 + \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{(\alpha + j\omega)t}\big[\hat{u}(\alpha + j\omega) - (\alpha + j\omega - \beta)^{-1}u_0\big]\,d\omega.$$

Then v is independent of β, $v \in BC_{\alpha,0}(\mathbb{R}^+; U)$, $v(0+) = u_0$, and $u = v$ almost everywhere on $\mathbb{R}^+$.

(ii) *Suppose that $u \in L^1_\alpha(\mathbb{R}^-; U)$ and that there exists some $u_0 \in \mathbb{C}$ such that the function $\omega \mapsto \hat{u}(\alpha + j\omega) - (1 + j\omega)^{-1}u_0$ belongs to $L^1(\mathbb{R}; U)$, where $\hat{u}$ is the left-sided Laplace transform of u. Let $\beta \in \mathbb{C}^+_\alpha$, and define for all $t \in \mathbb{R}^-$,*

$$v(t) = -e^{\beta t}u_0 + \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{(\alpha + j\omega)t}\big[\hat{u}(\alpha + j\omega) - (\alpha + j\omega - \beta)^{-1}u_0\big]\,d\omega.$$

Then v is independent of β, $v \in BC_{\alpha,0}(\mathbb{R}^-; U)$, $v(0-) = -u_0$, and $u = v$ almost everywhere on $\mathbb{R}^+$.

Proof The proof is essentially the same in both cases. First we observe that the function $\omega \mapsto \hat{u}(\alpha + j\omega) + (\alpha + j\omega - \beta)^{-1}u_0$ belongs to $L^1_\alpha(\mathbb{R}; U)$ since the difference

$$\frac{1}{\alpha + j\omega - \beta} - \frac{1}{1 + j\omega} = \frac{1 - \alpha + \beta}{(\alpha + j\omega - \beta)(1 + j\omega)}$$

belongs to $L^1(\mathbb{R})$. We get the conclusion of (i) by applying Proposition 3.12.2 to the function $t \mapsto u(t) - e^{\beta t}u_0$ (whose right-sided Laplace transform is $\hat{u}(\lambda) - (\lambda - \beta)^{-1}u_0$ for $\Re\lambda \geq \alpha$), and we get the conclusion of (ii) by applying Proposition 3.12.2 to the function $t \mapsto u(t) + e^{\beta t}u_0$ (whose left-sided Laplace transform is $\hat{u}(\lambda) - (\lambda - \beta)^{-1}u_0$ for $\Re\lambda \leq \alpha$). □

In the proof of Proposition 3.12.3 given above we observe that, although the Laplace transform is injective in the sense that we explained earlier, it is possible to have two functions, one supported on $\mathbb{R}^+$ and another supported on $\mathbb{R}^-$, which have 'the same' one-sided Laplace transforms in the sense that the two Laplace transforms are analytic continuations of each other. As our following theorem shows, this situation is not at all unusual.

Theorem 3.12.4 *Let U be a Banach space, and let f be a U-valued function which is analytic at infinity with $f(\infty) = 0$. Let Γ be a positively oriented piecewise continuously differentiable Jordan curve such that f is analytic on Γ and outside of Γ, and define*

$$u(t) = \frac{1}{2\pi j}\oint_\Gamma e^{\lambda t} f(\lambda)\, d\lambda.$$

Then u is entire, $u(0) = \lim_{\lambda\to\infty} \lambda f(\lambda)$, $u(t) = O(e^{\beta t})$ as $t \to \infty$ and $u(t) = O(e^{-\alpha t})$ as $t \to -\infty$, where $\beta = \sup\{\Re\lambda \mid \lambda \in \Gamma\}$ and $\alpha = \inf\{\Re\lambda \mid \lambda \in \Gamma\}$. Moreover, the restriction of f to $\mathbb{C}^+_\omega$ is the (right-sided) Laplace transform of $\pi_+ u$, and the restriction of f to $\mathbb{C}^-_\alpha$ is the (left-sided) Laplace transform of $-\pi_- u$.

Proof That u is entire and satisfies the two growth bounds follows immediately from the integral representation. The expression for $u(0)$ follows from Lemma 3.9.1(ii). Thus, it only remains to prove that f, suitably restricted, is the right-sided Laplace transform of $\pi_+ u$ and the left-sided Laplace transform of $-\pi_- u$.

We begin with the case where $f(\lambda) = (\lambda - \gamma)^{-1}u_0$, where γ is encircled by Γ (in particular, $\alpha < \Re\gamma < \beta$). Then a direct inspection shows that the theorem is true with $u(t) = e^{\gamma t}u_0$. Since f is analytic at infinity with $f(\infty) = 0$, it has an expansion of the type $f(\lambda) = u_0/\lambda + O(\lambda^{-2})$ as $\lambda \to \infty$. After subtracting $(\lambda - \gamma)^{-1}u_0$ from f, the new function f satisfies $f(\lambda) = O(\lambda^{-2})$ as $\lambda \to \infty$. Thus, we may, without loss of generality, assume that $f(\lambda) = O(\lambda^{-2})$ as $\lambda \to \infty$. Then $u(0) = 0$. (By subtracting off further terms it is even possible to assume that $f(\lambda) = O(\lambda^{-k})$ as $\lambda \to \infty$ for any finite k.)

Choose some $\beta' > \beta$ and $\alpha' < \alpha$, and define, for all $t \in \mathbb{R}$,

$$v(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{(\beta'+j\omega)t} f(\beta' + j\omega)\, d\omega,$$
$$w(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{(\alpha'+j\omega)t} f(\alpha' + j\omega)\, d\omega.$$

We claim that $v(t) = 0$ for $t \leq 0$, that $w(t) = 0$ for $t \geq 0$, and that $v(t) - w(t) = u(t)$ for all $t \in \mathbb{R}$. Suppose that this is true. Then $v = \pi_+ u$, $w = -\pi_- u$, v is the inverse right-sided Laplace transform of the restriction of f to the line $\Re\lambda = \beta'$, and w is the inverse left-sided Laplace transform of the restriction of f to the line $\Re\lambda = \alpha'$. Since the inverse Laplace transform is injective (the proof of this is identical to the proof of the fact that the bilateral Laplace transform is injective), the right-sided Laplace transform of $\pi_+ u$ must coincide with u on $\mathbb{C}^+_{\beta'}$ and the left-sided Laplace transform of $-\pi_- u$ must coincide with u on $\mathbb{C}^-_{\alpha'}$. Thus, it only remains to verify the claim that $v(t) = 0$ for $t \leq 0$, that $w(t) = 0$ for $t \geq 0$, and that $v(t) - w(t) = u(t)$ for all $t \in \mathbb{R}$.

We begin with the claim that $v(t) = 0$ for all $t \leq 0$. For each $R > 0$, by Cauchy's theorem, we have

$$\frac{1}{2\pi j} \oint_{\Gamma_R} \mathrm{e}^{\lambda t} f(\lambda)\, d\lambda = 0,$$

where Γ_R is the closed path running from $\beta' - jR$ to $\beta' + jR$ along the line $\Re\lambda = \beta'$, and then back to $\beta' - jR$ along a semi-circle in $\overline{\mathbb{C}}^+_{\beta'}$ centered at β'. If $t \leq 0$, then $\mathrm{e}^{\lambda t}$ is bounded on $\mathbb{C}^+_{\beta}$, and we can use the fact that $f(\lambda) = O(\lambda^{-2})$ as $\lambda \to \infty$ to let $R \to \infty$ (the length of the semi-circle is πR, hence this part of the integral tends to zero), and get

$$v(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{(\beta'+j\omega)t} f(\beta' + j\omega)\, d\omega = \lim_{R\to\infty} \frac{1}{2\pi j} \oint_{\Gamma_R} \mathrm{e}^{\lambda t} f(\lambda)\, d\lambda = 0.$$

This proves that $v(t) = 0$ for $t \leq 0$. An analogous proof shows that $w(t) = 0$ for $t \geq 0$.

Finally, let us show that $v(t) - w(t) = u(t)$ for all $t \in \mathbb{R}$. By analyticity, for sufficiently large values of R, we can deform the original path Γ given in the theorem to a rectangular path Γ_R, running from $\beta' - jR$ to $\beta' + jR$ to $\alpha' + jR$ to $\alpha' - jR$, and back to $\beta' - jR$, and get for all $t \in \mathbb{R}$,

$$u(t) = \frac{1}{2\pi j} \oint_{\Gamma_R} \mathrm{e}^{\lambda t} f(\lambda)\, d\lambda.$$

Letting again $R \to \infty$ we get

$$\begin{aligned} u(t) &= \lim_{R\to\infty} \frac{1}{2\pi j} \oint_{\Gamma_R} e^{\lambda t} f(\lambda)\, d\lambda \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(\beta' + j\omega)t} f(\beta' + j\omega)\, d\omega \\ &\qquad - \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(\alpha' + j\omega)t} f(\alpha' + j\omega)\, d\omega \\ &= v(t) - w(t). \end{aligned}$$

□

Corollary 3.12.5 *Let U be a Banach space.*

(i) *Suppose that $u \in L^1_\alpha(\mathbb{R}^+; U)$ and that $\hat{u}$ is analytic at infinity. For all $t \in \mathbb{R}^+$, define*

$$v(t) = \frac{1}{2\pi j} \oint_{\Gamma} e^{\lambda t} \hat{u}(\lambda)\, d\lambda,$$

where Γ is a positively oriented piecewise continuously differentiable Jordan curve which encircles $\sigma(\hat{u})$. Then v is the restriction to $\mathbb{R}^+$ of an entire function, and $u = v$ almost everywhere on $\mathbb{R}^+$.

(ii) *Suppose that $u \in L^1_\alpha(\mathbb{R}^-; U)$ and that $\hat{u}$ is analytic at infinity. For all $t \in \mathbb{R}^-$, define*

$$v(t) = \frac{1}{2\pi j} \oint_{\Gamma} e^{\lambda t} \hat{u}(\lambda)\, d\lambda,$$

where Γ is a negatively oriented piecewise continuously differentiable Jordan curve which encircles $\sigma(\hat{u})$. Then v is the restriction to $\mathbb{R}^-$ of an entire function, and $u = v$ almost everywhere on $\mathbb{R}^-$.

Proof This follows from Theorem 3.12.4. □

By combining the symbolic calculus developed in Section 3.9 with Corollary 3.12.5 we get the following result.

Theorem 3.12.6 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ on X with growth bound $\omega_{\mathfrak{A}}$. Let $u \in L^1_\omega(\mathbb{R}^-; \mathbb{C})$ for some $\omega > \omega_{\mathfrak{A}}$, and suppose that the left-sided Laplace transform $\hat{u}$ of u is analytic at infinity. Then $\hat{u}(A) \in \mathcal{B}(X; X_1)$ and*

$$\hat{u}(A)x = \int_{-\infty}^{0} u(t)\mathfrak{A}^{-t} x\, dt, \qquad x \in X.$$

Proof The assumptions imply that $\hat{u}$ is analytic on $\mathbb{C}^-_\omega \cup \infty$ with $\hat{u}(\infty) = 0$. In particular, by Lemma 3.9.3, $\hat{u}(A) \in \mathcal{B}(X; X_1)$.

Let Γ be a negatively oriented piecewise continuously differentiable Jordan curve which encircles $\sigma(\hat{u})$ with $\overline{\sigma}(A)$ on the outside. Then Theorem 3.2.9, Corollary 3.12.5, Fubini's theorem, and (3.9.3) give, for all $x \in X$,

$$\begin{aligned}
\int_{-\infty}^{0} u(t)\mathfrak{A}^{-t}x\,dt &= \frac{1}{2\pi j}\int_{-\infty}^{0}\oint_{\Gamma} e^{\lambda t}\hat{u}(\lambda)\,d\lambda\,\mathfrak{A}^{-t}x\,dt \\
&= \frac{1}{2\pi j}\oint_{\Gamma}\hat{u}(\lambda)\int_{-\infty}^{0} e^{\lambda t}\mathfrak{A}^{-t}x\,dt\,d\lambda \\
&= \frac{1}{2\pi j}\oint_{\Gamma}\hat{u}(\lambda)\int_{0}^{\infty} e^{-\lambda t}\mathfrak{A}^{t}x\,dt\,d\lambda \\
&= \frac{1}{2\pi j}\oint_{\Gamma}\hat{u}(\lambda)(\lambda - A)^{-1}x\,d\lambda \\
&= \hat{u}(A)x.
\end{aligned}$$

□

Corollary 3.12.7 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ on X with growth bound $\omega_{\mathfrak{A}}$. Let $u \in L^1_{-\omega}(\mathbb{R}^+;\mathbb{C})$ for some $\omega > \omega_{\mathfrak{A}}$, and suppose that the right-sided Laplace transform $\hat{u}$ of u is analytic at infinity. Then $\hat{u}(-A) \in \mathcal{B}(X;X_1)$ and*

$$\hat{u}(-A)x = \int_0^\infty u(t)\mathfrak{A}^t x\,dt, \qquad x \in X. \tag{3.12.1}$$

Proof This follows from Theorem 3.12.6 through a change of integration variables (i.e., we apply Theorem 3.12.6 to the function $t \mapsto u(-t)$). □

It is possible to use (3.12.1) as a *definition* of $\hat{u}(-A)$, and in this way it is possible to extend the functional calculus presented in the first half of Section 3.9 to a larger class of functions of A, namely those that correspond to functions $u \in L^1_{\text{loc}}(\mathbb{R}^+)$ satisfying

$$\int_0^\infty |u(t)|\|\mathfrak{A}^t\|\,dt < \infty \tag{3.12.2}$$

(thus, u should be integrable with respect to the weight function $\|\mathfrak{A}\|$). One such example was the definition of the fractional powers of $\gamma - A$ given in the second half of Section 3.9.

Theorem 3.12.8 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ on X with growth bound $\omega_{\mathfrak{A}}$. Let $u, v \in L^1_{\omega}(\mathbb{R}^-;\mathbb{C})$ for some $\omega > \omega_{\mathfrak{A}}$, and suppose that the left-sided Laplace transforms of u and v are analytic at infinity. Define $w(t) = \int_t^0 u(t-s)v(s)\,ds$ for all those $t \in \mathbb{R}^-$ for which the integral converges. Then $w \in L^1_{\omega}(\mathbb{R}^-;\mathbb{C})$ (in particular, w is defined for almost all t), $\widehat{w}(\lambda) = \hat{u}(\lambda)\hat{v}(\lambda)$*

for $\lambda \in \mathbb{C}^-_\omega$, *and*

$$\widehat{w}(A) = \hat{u}(A)\hat{v}(A).$$

Proof The function $(s,t) \mapsto u(t-s)v(s)$ is measurable (see, e.g., Hewitt and Stromberg [1965, pp. 396–397]). By Fubini's theorem and a change of integration variable,

$$\begin{aligned}\int_{-\infty}^0 \mathrm{e}^{\omega t}|w(t)|\,dt &\le \int_{-\infty}^0 \int_t^0 |\mathrm{e}^{\omega(t-s)}u(t-s)||\mathrm{e}^{\omega s}v(s)|\,ds\,dt \\ &= \int_{-\infty}^0 \int_{-\infty}^s |\mathrm{e}^{\omega(t-s)}u(t-s)|\,dt\,|\mathrm{e}^{\omega s}v(s)|\;ds \\ &= \int_{-\infty}^0 |\mathrm{e}^{\omega t}u(t)|\,dt \int_{-\infty}^0 |\mathrm{e}^{\omega s}v(s)|\,ds < \infty.\end{aligned}$$

This proves that $w \in L^1_\omega(\mathbb{R}^-;\mathbb{C})$. Using Fubini's theorem once more we get for all $\lambda \in \mathbb{C}^-_\omega$,

$$\begin{aligned}\widehat{w}(\lambda) &= \int_{-\infty}^0 \mathrm{e}^{-\lambda t}w(t)\,dt \\ &= \int_{-\infty}^0 \int_t^0 \mathrm{e}^{-\lambda(t-s)}u(t-s)\mathrm{e}^{-\lambda s}v(s)\,ds\,dt \\ &= \int_{-\infty}^0 \int_{-\infty}^s \mathrm{e}^{-\lambda(t-s)}u(t-s)\,dt\;\mathrm{e}^{-\lambda s}v(s)\;ds \\ &= \int_{-\infty}^0 \mathrm{e}^{-\lambda t}u(t)\,dt \int_{-\infty}^0 \mathrm{e}^{-\lambda s}v(s)\,ds \\ &= \hat{u}(\lambda)\hat{v}(\lambda).\end{aligned}$$

In particular, $\widehat{w}$ is analytic at infinity. To show that $\widehat{u}(A) = \hat{u}(A)\hat{v}(A)$ it suffices to repeat the computation above with $\mathrm{e}^{-\lambda t}$ replaced by $\mathfrak{A}^{-t}$ (and use Theorem 3.12.8). □

Theorem 3.12.9 *Let A be the generator of a* C_0 *semigroup* $\mathfrak{A}$ *on X with growth bound* $\omega_{\mathfrak{A}}$. *Let* $u \in L^1_\omega(\mathbb{R}^-;\mathbb{C})$ *and* $v \in L^1_\omega(\mathbb{R}^-;X)$ *for some* $\omega > \omega_{\mathfrak{A}}$, *and suppose that the left-sided Laplace transforms of u and v are analytic at infinity. Define* $w(t) = \int_t^0 u(t-s)v(s)\,ds$ *for all those* $t \in \mathbb{R}^-$ *for which the integral converges. Then* $w \in L^1_\omega(\mathbb{R}^-;X)$ *(in particular, w is defined for almost all t),* $\widehat{w}(\lambda) = \hat{u}(\lambda)\hat{v}(\lambda)$ *for* $\lambda \in \mathbb{C}^-_\omega$, *and*

$$\int_{-\infty}^0 \mathfrak{A}^{-t}w(t)\,dt = \hat{u}(A)\int_{-\infty}^0 \mathfrak{A}^{-t}v(t)\,dt.$$

Proof The proof of this theorem is the same as the proof of Theorem 3.12.8. □

Lemma 3.12.10 *Let $\beta \in \mathbb{R}$ and $\alpha > 0$, and define*

$$f(t) = \frac{1}{\Gamma(\alpha)} e^{\beta t} t^{\alpha-1}, \qquad t \in \mathbb{R}^+.$$

Then the (right-sided) Laplace transform of f is given by

$$\hat{f}(\lambda) = (\lambda - \beta)^{-\alpha}, \qquad \lambda \in \mathbb{C}^+_\beta.$$

We leave the proof to the reader (one possibility is to make a (complex) change of variable from t to $u = t/(\lambda - \beta)$ in (3.9.7)).

If U is a Hilbert space and $p = 2$, then it is possible to say much more about the behavior of the various Laplace transforms. Some results of this type are described in Section 10.3.

3.13 Shift semigroups in the frequency domain

The following proposition describes how the shift semigroups introduced in Examples 2.3.2 and 2.5.3 behave in the terms of Laplace transforms.

Proposition 3.13.1 *Let U be a Banach space, let $1 \le p \le \infty$, let $\omega \in \mathbb{R}$, and let $T > 0$.*

(i) *Let τ^t be the bilateral shift $(\tau^t u)(s) = u(s + t)$ for all $s, t \in \mathbb{R}$ and all $u \in L^1_{loc}(\mathbb{R}; U)$. Then, for all $t \in \mathbb{R}$ and all $u \in L^1_{loc}(\mathbb{R}; U)$, the bilateral Laplace transforms $\widehat{\tau^t u}$ and $\hat{u}$ of $\tau^t u$, respectively u, have the same domain (possibly empty), and for all z in this common domain,*

$$\widehat{\tau^t u}(z) = e^{zt} \hat{u}(z).$$

(ii) *Let $\tau^t_+ = \pi_+ \tau^t$ for all $t \ge 0$, where τ^t is the bilateral shift defined in (i). Then, for all $t \ge 0$ and all $u \in L^1_{loc}(\mathbb{R}^+; U)$, the (one-sided) Laplace transforms $\widehat{\tau^t_+ u}$ and $\hat{u}$ of $\tau^t_+ u$, respectively u, have the same domain (possibly empty), and for all z in this common domain,*

$$\widehat{\tau^t_+ u}(z) = e^{zt} \int_t^\infty e^{-zs} u(s)\, ds.$$

In particular, if $u \in L^p_\omega(\mathbb{R}^+; U)$, then $\widehat{\tau^t u}$ and $\hat{u}$ are defined at least for all $z \in \mathbb{C}^+_\omega$ and the above formula holds.

(iii) *Let $\tau^t_- =< \tau^t \pi_-$ for all $t \ge 0$, where τ^t is the bilateral shift defined in (i). Then, for all $t \ge 0$ and all $u \in L^1_{loc}(\mathbb{R}^-; U)$, the left-sided Laplace transforms $\widehat{\tau^t_- u}$ and $\hat{u}$ of $\tau^t_- u$, respectively u, have the same domain (possibly empty), and for all z in this common domain,*

$$\widehat{\tau^t_- u}(z) = e^{zt} \hat{u}(z).$$

In particular, if $u \in L^p_\omega(\mathbb{R}^-;U)$, then $\widehat{\tau^t u}$ and $\hat{u}$ are defined at least for all $z \in \mathbb{C}^-_\omega$ and the above formula holds.

(iv) *Let $\tau^t_{[0,T)}$, $t \geq 0$, be the finite left shift introduced in Example 2.3.2(iv). Then, for all $u \in L^1([0,T];U)$, the finite Laplace transform of $\tau^t_{[0,T)}u$ over $[0,T]$ is given by (for all $z \in \mathbb{C}$)*

$$\widehat{\tau^t_{[0,T)}u}(z) = \begin{cases} e^{zt}\int_t^T e^{-zs}u(s)\,ds, & 0 \leq t \leq T, \\ 0, & \textit{otherwise}. \end{cases}$$

(v) *Let $\tau^t_{\mathbb{T}_T}$, $t \in \mathbb{R}$, be the circular left shift introduced in Example 2.3.2(v). Then, for all $u \in L^1(\mathbb{T}_T;U)$, the finite Laplace transform of $\tau^t_{\mathbb{T}_T}u$ over $[0,T]$ is given by (for all $z \in \mathbb{C}$)*

$$\widehat{\tau^t_{\mathbb{T}_T}u}(z) = e^{zt}\int_t^{t+T} e^{-zs}u(s)\,ds.$$

In particular, this Laplace transform is a T-periodic function of t.

Proof The proof is the same in all cases: it suffices to make a change of variable $v = s + t$ in the integral defining the Laplace transform of the shifted u (and adjust the bounds of integration appropriately):

$$\int e^{-zs}u(s+t)\,ds = e^{zt}\int e^{-zv}u(v)\,dv.$$

□

From Proposition 3.13.1 we observe that especially the bilateral shift τ^t and the outgoing shift τ^t_- have simple Laplace transform descriptions, whereas the descriptions of the other shifts are less transparent. Fortunately, all the generators and the resolvents of the various shifts have simple descriptions.

Proposition 3.13.2 *The generators of the (semi)groups τ^t, τ^t_+, τ^t_-, $\tau^t_{[0,T)}$, and $\tau^t_{\mathbb{T}_T}$ in Examples 2.3.2 and 2.5.3 (see Example 3.2.3), have the following descriptions in terms of Laplace transforms:*

(i) *For all u in the domain of the generator $\frac{d}{ds}$ of the bilateral left shift group τ^t on $L^1_\omega(\mathbb{R};U)$, the bilateral Laplace transforms of u and $\dot{u} = \frac{d}{ds}u$ are defined (at least) on the vertical line $\Re z = \omega$, and*

$$\widehat{\dot{u}}(z) = z\hat{u}(z), \qquad \Re z = \omega.$$

(ii) *For all u in the domain of the generator $\frac{d}{ds}_+$ of the incoming left shift semigroup τ^t_+ on $L^p_\omega(\mathbb{R}^+;U)$ or on $BUC_\omega(\mathbb{R}^+;U)$, the Laplace transforms of u and $\dot{u} = \frac{d}{ds}_+u$ are defined (at least) on the half-plane $\mathbb{C}^+_\omega$, and*

$$\widehat{\dot{u}}(z) = z\hat{u}(z) - u(0), \qquad \Re z > \omega.$$

(iii) *For all u in the domain of the generator $\frac{\mathrm{d}}{\mathrm{d}s}_-$ of the outgoing left shift semigroup τ_-^t on $L^p_\omega(\mathbb{R}^-;U)$ or on $\{u \in BUC_\omega(\overline{\mathbb{R}}^-;U) \mid u(0)=0\}$, the left-sided Laplace transforms of u and $\dot u = \frac{\mathrm{d}}{\mathrm{d}s}_+ u$ are defined (at least) on the half-plane $\mathbb{C}^-_\omega$, and*

$$\widehat{\dot u}(z) = z\hat u(z), \qquad \Re z < \omega.$$

(iv) *For all u in the domain of the generator $\frac{\mathrm{d}}{\mathrm{d}s}_{[0,T)}$ of the finite left shift semigroup $\tau^t_{[0,T)}$ on $L^p([0,T);U)$ or on $\{u \in C([0,T];U) \mid u(T)=0\}$, the finite Laplace transforms over $[0,T]$ of u and $\dot u = \frac{\mathrm{d}}{\mathrm{d}s}_{[0,T)} u$ are related as follows:*

$$\widehat{\dot u}(z) = z\hat u(z) - u(0), \qquad z \in \mathbb{C}.$$

(v) *For all u in the domain of the generator $\frac{\mathrm{d}}{\mathrm{d}s}_{\mathbb{T}_T}$ of the circular left shift group $\tau^t_{\mathbb{T}_T}$ on $L^p(\mathbb{T}_T;U)$ or on $C(\mathbb{T}_T;U)$, the finite Laplace transforms over $[0,T]$ of u and $\dot u = \frac{\mathrm{d}}{\mathrm{d}s}_{\mathbb{T}_T} u$ are related as follows:*

$$\widehat{\dot u}(z) = z\hat u(z) + (\mathrm{e}^{-zT} - 1)u(0), \qquad z \in \mathbb{C}.$$

Proof All the proofs are identical: it suffices to integrate by parts in the integral defining the particular Laplace transform, and to take into account possible nonzero boundary terms (i.e., the terms $-u(0)$ and $(\mathrm{e}^{-zT}-1)u(0)$ in (ii), (iv) and (v)). □

Proposition 3.13.3 *The resolvents of the generators $\frac{\mathrm{d}}{\mathrm{d}s}$, $\frac{\mathrm{d}}{\mathrm{d}s}_+$, $\frac{\mathrm{d}}{\mathrm{d}s}_-$, $\frac{\mathrm{d}}{\mathrm{d}s}_{[0,T)}$, and $\frac{\mathrm{d}}{\mathrm{d}s}_{\mathbb{T}_T}$ in Examples 3.2.3 and 2.5.3 (see Example 3.3.2), have the following descriptions in terms of Laplace transforms:*

(i) *For all $f \in L^1_\omega(\mathbb{R};U)$ and all λ with $\Re\lambda \neq \omega$, the bilateral Laplace transforms of f and $u = \left(\lambda - \frac{\mathrm{d}}{\mathrm{d}s}\right)^{-1} f$ are defined (at least) on the vertical line $\Re z = \omega$, and*

$$\hat u(z) = (\lambda - z)^{-1}\hat f(z), \qquad \Re z = \omega.$$

(ii) *For all $f \in L^p_\omega(\mathbb{R}^+;U)$ or $f \in BUC_\omega(\mathbb{R}^+;U)$ and all $\lambda \in \mathbb{C}^+_\omega$, the Laplace transforms of f and $u = \left(\lambda - \frac{\mathrm{d}}{\mathrm{d}s}_+\right)^{-1} f$ are defined (at least) on the half-plane $\mathbb{C}^+_\omega$, and*

$$\hat u(z) = (\lambda - z)^{-1}(\hat f(z) - \hat f(\lambda)), \qquad \Re z > \omega.$$

(iii) *For all $f \in L^p_\omega(\mathbb{R}^-;U)$ or $f \in BUC_\omega(\mathbb{R}^-;U)$ with $f(0)=0$ and all λ with $\Re\lambda > \omega$, the left-sided Laplace transforms of f and $u = \left(\lambda - \frac{\mathrm{d}}{\mathrm{d}s}_-\right)^{-1} f$ are defined (at least) on the half-plane $\mathbb{C}^-_\omega$, and*

$$\hat u(z) = (\lambda - z)^{-1}\hat f(z), \qquad \Re z < \omega.$$

(iv) *For all* $f \in L^p_\omega([0,T];U)$ *or* $f \in C_\omega([0,T];U)$ *with* $f(T) = 0$ *and all* $\lambda \in \mathbb{C}$*, the finite Laplace transforms over* $[0,T]$ *of* f *and* $u = \big(\lambda - \frac{\mathrm{d}}{\mathrm{d}s}_{[0,T)}\big)^{-1} f$ *are related as follows:*

$$\hat{u}(z) = (\lambda - z)^{-1}(\hat{f}(z) - \hat{f}(\lambda)), \qquad z \neq \lambda.$$

(v) *For all* $f \in L^p(\mathbb{T}_T;U)$ *or* $f \in C(\mathbb{T}_T;U)$ *and all* $\lambda \notin \{2\pi jm/T \mid m = 0, \pm 1, \pm 2, \ldots\}$*, the finite Laplace transforms over* $[0,T]$ *of* f *and* $u = \big(\lambda - \frac{\mathrm{d}}{\mathrm{d}s}_{\mathbb{T}_T}\big)^{-1} f$ *are related as follows:*

$$\hat{u}(z) = (\lambda - z)^{-1}(\hat{f}(z) - (1 - \mathrm{e}^{-zT})(1 - \mathrm{e}^{-\lambda T})^{-1}\hat{f}(\lambda)), \qquad z \neq \lambda.$$

Proof Throughout this proof the convergence of the indicated Laplace transforms are obvious, so let us only concentrate on the formulas relating $\hat{u}(z)$ to $\hat{f}(z)$.

(i) This follows from Proposition 3.13.2(i), which given $\hat{f}(z) = \lambda\hat{u}(z) - \widehat{\dot{u}}(z) = (\lambda - z)\hat{u}(z)$.

(ii) By Proposition 3.13.2(ii), $\hat{f}(z) = \lambda\hat{u}(z) - \widehat{\dot{u}}(z) = (\lambda - z)\hat{u}(z) + u(0)$. Taking $z = \lambda$ we get $u(0) = \hat{f}(\lambda)$, and so

$$(\lambda - z)\hat{u}(z) = \hat{f}(z) - \hat{f}(\lambda).$$

(iii)–(iv) These proofs are identical to the proofs of (i) and (ii), respectively.

(v) By Proposition 3.13.2(v),

$$\hat{f}(z) = \lambda\hat{u}(z) - \widehat{\dot{u}}(z) = (\lambda - z)\hat{u}(z) + (1 - \mathrm{e}^{-zT})u(0).$$

Taking $z = \lambda$ we get $u(0) = (1 - \mathrm{e}^{-\lambda T})^{-1}\hat{f}(\lambda)$, and so

$$(\lambda - z)\hat{u}(z) = \hat{f}(z) - (1 - \mathrm{e}^{-zT})(1 - \mathrm{e}^{-\lambda T})^{-1}\hat{f}(\lambda).$$

□

3.14 Invariant subspaces and spectral projections

We begin by introducing some terminology.

Definition 3.14.1 We say that the Banach space X is the *direct sum* of Y and Z and write either $X = Y \dotplus Z$ or $X = \left[\begin{smallmatrix} Y \\ Z \end{smallmatrix}\right]$ if Y and Z are closed subspaces of X and every $x \in X$ has a unique representation of the form $x = y + z$ where $y \in Y$ and $z \in Z$. A subspace Y of X is *complemented* if X is the direct sum of Y and some other subspace Z.

Instead of writing $x = y + z$ (corresponding to the notation $X = Y \dotplus Z$) as we did above we shall also use the alternative notation $x = \left[\begin{smallmatrix} y \\ z \end{smallmatrix}\right]$ (corresponding to the notation $X = \left[\begin{smallmatrix} Y \\ Z \end{smallmatrix}\right]$), and we identify $y \in Y$ with $\left[\begin{smallmatrix} y \\ 0 \end{smallmatrix}\right] \in X$. Note that every closed subspace Y of a Hilbert space X is complemented: we may take the complement of $Y \subset X$ to be $Z = Y^{\perp}$.

Lemma 3.14.2 *Let X be a Banach space.*

(i) *If π is an arbitrary projection on X (i.e., $\pi \in \mathcal{B}(X)$ and $\pi = \pi^2$), then $X = \mathcal{N}(\pi) \dotplus \mathcal{R}(\pi)$. (In particular, the range of a projection operator is closed.)*

(ii) *Conversely, to each splitting of X into a direct sum $X = Y \dotplus Z$ there is a unique projection π such that $Y = \mathcal{N}(\pi)$ and $Z = \mathcal{R}(\pi)$. We call π the* projection of X onto Z along Y.

Proof (i) Clearly, every $x \in X$ can be split into $x = \pi x + (1-\pi)x$ where $\pi x \in \mathcal{R}(\pi)$ and $(1-\pi)x \in \mathcal{N}(\pi)$ (since $\pi(1-\pi)x = \pi x - \pi^2 x = 0$). Obviously, if $x \in \mathcal{N}(1-\pi)$, then $x = \pi x$, so $x \in \mathcal{R}(\pi)$, and conversely, if $x \in \mathcal{R}(\pi)$, then $x = \pi y$ for some $y \in X$, so $\pi x = \pi^2 y = \pi y = x$, i.e., $x \in \mathcal{N}(1-\pi)$. This means that $\mathcal{R}(\pi) = \mathcal{N}(1-\pi)$, hence $\mathcal{R}(\pi)$ is closed.

(ii) For each $x \in X$, split x (uniquely) into $x = y + z$ where $y \in Y$ and $z \in Z$. Define $\pi x = y$. Then $\pi x = x$ iff $x \in Y$ and $\pi x = 0$ iff $x \in Z$. It is easy to see that the operator π defined in this way is linear and closed, and that it satisfies $\pi = \pi^2$. By the closed graph theorem, $\pi \in \mathcal{B}(X)$. □

Definition 3.14.3 Let X be a Banach space, and let $A\colon X \supset \mathcal{D}(A) \to X$ be a linear operator.

(i) A subspace Y of X is an *invariant* subspace of A if $Ax \in Y$ for every $x \in \mathcal{D}(A) \cap Y$.

(ii) A pair of subspaces Y and Z of X are *reducing subspaces* of A if $X = Y \dotplus Z$, every $x \in \mathcal{D}(A)$ is of the form $x = y + z$ where $y \in \mathcal{D}(A) \cap Y$ and $z \in \mathcal{D}(A) \cap Z$, and both Y and Z are invariant subspaces of A,

(iii) By an invariant subspace of a C_0 semigroup $\mathfrak{A}$ on X we mean a subspace Y which is an invariant subspace of $\mathfrak{A}^t$ for every $t \geq 0$.

(iv) By a pair of reducing subspaces of a C_0 semigroup $\mathfrak{A}$ on X we mean a pair of subspaces which are reducing subspaces of $\mathfrak{A}^t$ for every $t \geq 0$.

Theorem 3.14.4 *Let $\mathfrak{A}$ be a C_0 semigroup on a Banach space X with generator A and growth bound $\omega_{\mathfrak{A}}$, and let Y be a closed subspace of X. Denote the component of $\rho(A)$ which contains an interval $[\omega, \infty)$ by $\rho_\infty(A)$ (by Theorem 3.2.9(i) such an interval always exists). Then the following conditions are equivalent.*

(i) *Y is an invariant subspace of $\mathfrak{A}$.*
(ii) *Y is an invariant subspace of $(\lambda - A)^{-1}$ for some $\lambda \in \rho_\infty(A)$.*
(iii) *Y is an invariant subspace of $(\lambda - A)^{-1}$ for all $\lambda \in \rho_\infty(A)$.*
(iv) *Y is an invariant subspace of A and $\rho(A_{|\mathcal{D}(A)\cap Y}) \cap \rho_\infty(A) \neq \emptyset$.*

If these equivalent conditions hold, then it is also true that

(v) *$\mathfrak{A}_{|Y}$ is a C_0 semigroup on Y whose generator is $A_{|\mathcal{D}(A)\cap Y}$.*

Here it is important that Y is *closed*. For example, $\mathcal{D}(A)$ is an invariant subspace of $\mathfrak{A}$ but not of A.

Proof (i) $\Rightarrow$ (ii): If (i) holds, then by Theorem 3.2.9(i), for all $\lambda \in \omega_{\mathfrak{A}}^+$ and all $x \in Y$,

$$(\lambda - A)^{-1}x = \int_0^\infty \mathrm{e}^{-\lambda s}\mathfrak{A}^s x \in Y.$$

(ii) $\Rightarrow$ (iii): Take an arbitrary $x^* \in Y^\perp$, where

$$Y^\perp = \{x^* \in X^* \mid \langle x, x^*\rangle_{(X,X^*)} = 0 \text{ for all } x \in Y\}.$$

Fix $x \in Y$, and take some $\lambda_0 \in \rho_\infty(A)$ such that Y is invariant under $(\lambda_0 - A)^{-1}$. Then $(\lambda_0 - A)^{-n}x \in Y$ for all $n = 1, 2, 3, \ldots$, hence $\langle(\lambda_0 - A)^{-n}x, x^*\rangle_{(X,X^*)} = 0$ for all $n = 1, 2, 3, \ldots$. Define $f(\lambda) = \langle(\lambda - A)^{-n}x, x^*\rangle_{(X,X^*)}$. Then f is analytic in $\rho_\infty(A)$, and all its derivatives vanish at the point λ_0 (see (3.2.6)). Therefore $f(\lambda) = \langle(\lambda - A)^{-n}x, x^*\rangle_{(X,X^*)} = 0$ for all $\lambda \in \rho_\infty(A)$. Taking the intersection over all $x^* \in Y^\perp$ we find that $(\lambda - A)^{-n}x \in Y$.

(iii) $\Rightarrow$ (i): If (iii) holds, then by Theorem 3.7.5, for all $t \geq 0$ and all $x \in Y$,

$$\mathfrak{A}^t x = \lim_{n\to\infty}\Big(1 - \frac{t}{n}A\Big)^{-n} x \in Y.$$

(i) $\Rightarrow$ (iv): For all $x \in \mathcal{D}(A) \cap Y$,

$$Ax = \lim_{h\downarrow 0}\frac{1}{h}(\mathfrak{A}^h - 1)x \in Y.$$

That $\rho(A_{|\mathcal{D}(A)\cap Y}) \cap \rho_\infty(A) \neq \emptyset$ follows from (v) and Theorem 3.2.9(i).

(i) $\Rightarrow$ (v): Trivially, $\mathfrak{A}_{|Y}$ is a C_0 semigroup on Y (since Y is invariant, and we use the same norm in Y as in X). Let us denote its generator by $\widetilde{A}$. We have just seen that $x \in \mathcal{D}(\widetilde{A})$ whenever $x \in \mathcal{D}(A) \cap Y$, and that $\widetilde{A}x = Ax$ for all $x \in \mathcal{D}(A) \cap Y$. Conversely, if $x \in \mathcal{D}(\widetilde{A})$, then $x \in Y$ and $\frac{1}{h}(\mathfrak{A}^h - 1)x$ has a limit in Y as $h \downarrow 0$, so $x \in \mathcal{D}(A) \cap Y$.

(iv) $\Rightarrow$ (ii): Assume (iv). Let $\lambda \in \rho(A_{|\mathcal{D}(A)\cap Y}) \cap \rho_\infty(A)$. Then $\lambda - A$ maps $\mathcal{D}(A) \cap Y$ one-to-one onto Y (since $\lambda \in \rho(A_{|\mathcal{D}(A)\cap Y})$ and $(\lambda - A_{|\mathcal{D}(A)\cap Y})y = (\lambda - A)y$ for every $y \in \mathcal{D}(A) \cap Y$). This implies that $(\lambda - A)^{-1}$ maps Y one-to-one onto $\mathcal{D}(A) \cap Y$. In particular, Y is invariant under $(\lambda - A)^{-1}$. □

Invariance is preserved under the symbolic calculus described at the beginning of Section 3.9.

Lemma 3.14.5 *Let $A \in \mathcal{B}(X)$, let Γ be a piecewise continuously differentiable Jordan curve which encircles $\sigma(A)$ counter-clockwise, and let f be analytic on Γ and inside Γ. Define $f(A)$ by* (3.9.1). *Then the following claims are true.*

(i) *If Y is a closed invariant subspace of A, then Y is also invariant under $f(A)$.*
(ii) *If Y and Z are a pair of reducing subspaces of A, then they are also reducing for $f(A)$.*

Proof This follows from (3.9.1), the fact that Y and Z are closed, and Theorem 3.14.4. □

In the decomposition of systems into smaller parts we shall encounter closed invariant subspaces contained in the domain of the generator.

Theorem 3.14.6 *Let $A\colon X \supset \mathcal{D}(A) \to X$ be a closed linear operator, and let Y be an invariant subspace of A which is contained in $\mathcal{D}(A)$ and closed in X. Then $Y \subset \cap_{n=1}^{\infty}\mathcal{D}(A^n)$, $A_{|Y} \in \mathcal{B}(Y)$, and $(A^n)_{|Y} = (A_{|Y})^n$ for all $n =$ $1, 2, 3, \ldots$. If A is the generator of a C_0 semigroup $\mathfrak{A}$ on X, then Y is invariant under $\mathfrak{A}$ and $\mathfrak{A}_{|Y}$ is a uniformly continuous semigroup on Y whose generator is $A_{|Y}$.*

Proof The operator $A_{|Y}$ is closed since A is closed. Its domain is all of Y, and therefore, by the closed graph theorem, $A_{|Y} \in \mathcal{B}(Y)$.

Let $x \in Y$. Then $x \in \mathcal{D}(A)$ (since $Y \subset \mathcal{D}(A)$), and $Ax \in Y$ (since Y is invariant). Repeating the same argument with x replaced by Ax we find that $Ax \in \mathcal{D}(A)$, i.e, $x \in \mathcal{D}\left(A^2\right)$, and that $(A^2)_{|Y} = (A_{|Y})^2$. Continuing in the same way we find that $Y \subset \cap_{n=1}^{\infty}\mathcal{D}(A^n)$ and that $(A^n)_{|Y} = (A_{|Y})^n$ for all $n = 1, 2, 3, \ldots$.

That Y is invariant under $\mathfrak{A}$ follows from Theorem 3.14.4 (every λ with $|\lambda| > \|A_{|Y}\|$ belongs to the resolvent set of $A_{|Y}$). That $\mathfrak{A}_{|Y}$ is uniformly continuous follows from the fact that its generator $A_{|Y}$ is bounded. □

We now turn our attention to subspaces which are *reducing* and not just invariant.

Lemma 3.14.7 *Let $X = Y \dotplus Z$, let π be the projection of X onto Y along Z, and let $A \colon X \supset \mathcal{D}(A) \to X$ be a linear operator. Then Y and Z is a pair of reducing subspaces of A if and only if π maps $\mathcal{D}(A)$ into $\mathcal{D}(A)$ and $\pi Ax = A\pi x$ for all $x \in \mathcal{D}(A)$.*

Proof Assume that π maps $\mathcal{D}(A)$ into itself and that $\pi Ax = A\pi x$ for all $x \in \mathcal{D}(A)$. Then every $x \in \mathcal{D}(A)$ can be split into $x = \pi x + (1-\pi)x$ where $\pi x \in \mathcal{D}(A) \cap Y$ and $(1-\pi)x \in \mathcal{D}(A) \cap Z$. Moreover, for all $y \in \mathcal{D}(A) \cap Y$, $\pi Ay = A\pi y = Ay$, hence $Ay \in Y$, and for all $z \in \mathcal{D}(A) \cap Z$, $\pi Az = A\pi z = 0$, hence $Az \in Z$. Thus both Y and Z are invariant subspaces of A, and Y and Z is a pair of reducing subspaces of A.

Conversely, suppose that Y and Z is a pair of reducing subspaces of A. Then π maps $\mathcal{D}(A)$ into $\mathcal{D}(A)$ (since every $x \in X$ has a unique representation $x = y + z$ with $y \in \mathcal{D}(A) \cap Y$ and $z \in \mathcal{D}(A) \cap Z$, and $\pi x = y$). Moreover, for all $x \in \mathcal{D}(A)$, $\pi A(1-\pi)x = 0$ and $(1-\pi)A\pi = 0$ (since $1-\pi$ is the complementary projection of X onto Z along Y), and

$$\pi A = \pi A\pi = A\pi.$$

□

Theorem 3.14.8 *Let $X = Y \dotplus Z$, let π be the projection of X onto Y along Z, and let $A \colon X \supset \mathcal{D}(A) \to X$ be a linear operator with a nonempty resolvent set. Then the following conditions are equivalent:*

(i) *Y and Z are reducing subspaces of A,*
(ii) *Y and Z are reducing subspaces of $(\lambda - A)^{-1}$ for some $\lambda \in \rho(A)$.*
(iii) *Y and Z are reducing subspaces of $(\lambda - A)^{-1}$ for all $\lambda \in \rho(A)$.*

If, in addition, A is the generator of a C_0 semigroup $\mathfrak{A}$, then (i)–(iii) are equivalent to

(iv) *Y and Z are reducing subspaces of $\mathfrak{A}$.*

Proof Let π denote the projection of X onto Y along Z.

(i) ⇒ (iii): Assume (i). By Lemma 3.14.7, $(\lambda - A)\pi x = \pi(\lambda - A)x$ for all $\lambda \in \mathbb{C}$ and all $x \in \mathcal{D}(A)$. In particular, if we take $\lambda \in \rho(A)$, then we can apply $(\lambda - \widetilde{A})^{-1}$ to both sides of this identity and replace x by $(\lambda - A)^{-1}x$ to get for all $x \in X$,

$$\pi(\lambda - A)^{-1}x = (\lambda - A)^{-1}\pi x.$$

By Lemma 3.14.7, this implies (iii).

(iii) ⇒ (ii): This is trivial.

(ii) ⇒ (i): Assume (ii). Then, for all $x \in \mathcal{D}(A)$,

$$\pi x = (\lambda - A)^{-1}\pi(\lambda - A)x.$$

This implies that π maps $\mathcal{D}(A)$ into itself. Applying $(\lambda - A)$ to both sides of this identity we get $(\lambda - A)\pi x = \pi(\lambda - A)x$, or equivalently, $\pi Ax = A\pi x$. By Lemma 3.14.7, this implies (i).

(iv) ⇒ (ii): This follows from the fact that (i) implies (ii) in Theorem 3.14.4.

(iii) ⇒ (iv): If (iii) holds, then by Theorem 3.7.5, for all $t \geq 0$ and all $x \in X$,

$$\mathfrak{A}^t\pi x = \lim_{n\to\infty}\Big(1 - \frac{t}{n}A\Big)^{-n}\pi x = \pi \lim_{n\to\infty}\Big(1 - \frac{t}{n}A\Big)^{-n} x = \pi\mathfrak{A}^t x.$$

By Lemma 3.14.7, this implies (iv). □

Corollary 3.14.9 *If the equivalent conditions (i)–(iii) in Theorem 3.14.8 hold, then $\rho(A) = \rho(A_{|Y}) \cap \rho(A_{|Z})$, and for all $\lambda \in \rho(A)$,*

$$\begin{aligned}(\lambda - A_{|Y})^{-1} &= (\lambda - A)^{-1}_{|Y}, \quad (\lambda - A_{|Z})^{-1} = (\lambda - A)^{-1}_{|Z},\\ (\lambda - A)^{-1} &= (\lambda - A_{|Y})^{-1}\pi + (\lambda - A_{|Z})^{-1}(1-\pi),\end{aligned} \tag{3.14.1}$$

where π is the projection of X onto Y along Z.

Proof Let $\lambda \in \mathbb{C}$. Then $\lambda \in \rho(A)$ if and only if the equation $(\lambda - A)x = w$ has a unique solution x for every $w \in X$, and x depends continuously on w. By projecting this equation onto Y and Z (i.e., we multiply the equation by π and $1 - \pi$, where π is the projection onto Y along Z, and recall that π commutes with A) we get the two independent equations

$$(\lambda - A_{|Y})y = \pi w, \qquad (\lambda - A_{|Y})z = (1-\pi)w,$$

where $y = \pi x \in Y$ and $z = x - y \in Z$. The original equation is solvable if and only if both of these equations are solvable, and this implies that $\rho(A) = \rho(A_+) \cap \rho(A_-)$. Furthermore,

$$\begin{aligned} x &= (\lambda - A)^{-1}w = y + z \\ &= (\lambda - A_{|Y})^{-1}\pi w + (\lambda - A_{|Y})^{-1}(1-\pi)w,\end{aligned}$$

which gives us (3.14.1). □

One common way to construct invariant subspaces of a semigroup is to use a *spectral projection*. This is possible whenever the spectrum of the generator is not connected.

Theorem 3.14.10 *Let $A \colon X \supset \mathcal{D}(A) \to X$ be a densely defined linear operator with a nonempty resolvent set $\rho(A)$. Let Γ be a positively oriented piecewise continuously differentiable Jordan curve contained in $\rho(A)$ which separates*

$\sigma(A)$ into two nontrivial parts $\sigma(A) = \sigma_+(A) \cup \sigma_-(A)$, where $\sigma_+(A)$ lies inside Γ and $\sigma_-(A)$ lies outside Γ (in particular, $\infty \notin \sigma_+(A)$). Then the operator $\pi \in \mathcal{B}(X)$ defined by

$$\pi = \frac{1}{2\pi j} \oint_\Gamma (\lambda - A)^{-1}\, d\lambda. \tag{3.14.2}$$

is a projection which maps X into $\mathcal{D}(A)$. Denote $X_+ = \mathcal{R}(\pi)$ and $X_- = \mathcal{N}(\pi)$, $A_+ = A_{|X_+}$ and $A_- = A_{|X_-}$. Then the following claims are true.

(i) *$X = X_+ \dotplus X_-$, and X_+ and X_- are reducing subspaces of A and of $(\lambda - A)^{-1}$ for all $\lambda \in \rho(A)$.*
(ii) *$X_+ \subset \cap_{n=1}^\infty \mathcal{D}(A^n)$, $A_+ \in \mathcal{B}(X_+)$, and for all $n = 1, 2, 3, \ldots$, $(A_+)^n = A^n_{|X_+}$ and $(A_-)^n = A^n_{|\mathcal{D}(A^n)\cap X_-}$.*
(iii) *$\sigma(A_+) = \sigma_+(A)$ and $\sigma(A_-) = \sigma_-(A)$.*
(iv) *If A is the generator of a C_0 semigroup $\mathfrak{A}$, then A_+ is the generator of the norm-continuous semigroup $\mathfrak{A}_+ = \mathfrak{A}_{|X_+}$ and A_- is the generator of the C_0 semigroup $\mathfrak{A}_- = \mathfrak{A}_{|X_-}$.*

The projection π constructed above is often referred to as the *Riesz projection* corresponding to the part $\sigma_+(A)$ of $\sigma(A)$.

Proof As in Section 3.6, let us denote $\mathcal{D}(A)$ by X_1. Then the function $\lambda \mapsto (\lambda - A)^{-1}$ is bounded and (uniformly) continuous on Γ with values in $\mathcal{B}(X; X_1)$, so $\pi \in \mathcal{B}(X; X_1)$.

Let $\mu \in \rho(A)$. Then $(\mu - A)^{-1}$ is analytic on Γ, so

$$\begin{aligned}(\mu - A)^{-1}\pi &= (\mu - A)^{-1}\frac{1}{2\pi j}\oint_\Gamma (\lambda - A)^{-1}\, d\lambda\\ &= \frac{1}{2\pi j}\oint_\Gamma (\mu - A)^{-1}(\lambda - A)^{-1}\, d\lambda = \pi(\mu - A)^{-1},\end{aligned}$$

since $(\mu - A)^{-1}$ and $(\lambda - A)^{-1}$ commute. Thus, π commutes with $(\mu - A)^{-1}$. Furthermore, by the resolvent identity (3.2.1),

$$\begin{aligned}(\mu - A)^{-1}\pi &= \frac{1}{2\pi j}\oint_\Gamma \frac{(\mu - A)^{-1} - (\lambda - A)^{-1}}{\lambda - \mu}\, d\lambda\\ &= \frac{(\mu - A)^{-1}}{2\pi j}\oint_\Gamma (\lambda - \mu)^{-1}\, d\lambda - \frac{1}{2\pi j}\oint_\Gamma \frac{(\lambda - A)^{-1}}{\lambda - \mu}\, d\lambda\\ &= \begin{cases}(\mu - A)^{-1} - \frac{1}{2\pi j}\oint_\Gamma \frac{(\lambda - A)^{-1}}{\lambda - \mu}\, d\lambda, & \text{if } \mu \text{ lies inside } \Gamma,\\ -\frac{1}{2\pi j}\oint_\Gamma \frac{(\lambda - A)^{-1}}{\lambda - \mu}\, d\lambda, & \text{if } \mu \text{ lies outside } \Gamma.\end{cases}\end{aligned} \tag{3.14.3}$$

We next show that π is a projection. If we perturb the path Γ in such a way that the new path Γ_1 lies outside Γ, and so that both Γ_1 and Γ together with the area in between lie in $\rho(A)$, then, by the analyticity of the resolvent $(\lambda - A)^{-1}$ in this area,

$$\pi = \frac{1}{2\pi j} \oint_{\Gamma_1} (\mu - A)^{-1}\, d\mu.$$

Therefore, by (3.14.2), (3.14.3) (note that the index of λ with respect to Γ_1 is one)

$$\begin{aligned}
\pi^2 &= \frac{1}{2\pi j} \oint_{\Gamma_1} (\mu - A)^{-1}\pi\, d\mu \\
&= -\frac{1}{2\pi j} \oint_{\Gamma_1} \frac{1}{2\pi j} \oint_{\Gamma} \frac{(\lambda - A)^{-1}}{\lambda - \mu}\, d\lambda\, d\mu \\
&= -\frac{1}{2\pi j} \oint_{\Gamma} (\lambda - A)^{-1} \frac{1}{2\pi j} \oint_{\Gamma_1} \frac{d\mu}{\lambda - \mu}\, d\lambda \\
&= \frac{1}{2\pi j} \oint_{\Gamma} (\lambda - A)^{-1}\, d\lambda = \pi.
\end{aligned}$$

Thus $\pi^2 = \pi$, and we have proved that π is a projection. By Lemma 3.14.2, $X = X_+ \dotplus X_-$. (In particular, X_+ and X_- are closed.)

We now proceed to verify properties (i)–(iv).

(i) We know from the argument above that π commutes with $(\lambda - A)^{-1}$ for all $\lambda \in \rho(A)$, and we get (i) from Lemma 3.14.7 (applied to $(\lambda - A)^{-1}$) and Theorem 3.14.8.

(ii) This follows from Theorem 3.14.6.

(iii) By Corollary 3.14.9, $\rho(A) = \rho(A_+) \cap \rho(A_-)$. Thus, to prove (iii) it suffices to show that A_+ has no spectrum outside Γ and that A_- has no spectrum inside Γ.

Let $\mu \in \rho(A)$ lie outside Γ. Then $\mu \in \rho(A_+)$, and by (3.14.3) and Corollary 3.14.9, for all $x \in X_+$,

$$(\mu - A_+)^{-1}x = -\frac{1}{2\pi j} \oint_{\Gamma} \frac{(\lambda - A)^{-1}}{\lambda - \mu}\, x\, d\lambda.$$

Thus, in particular,

$$\left\|(\mu - A_+)^{-1}\right\| \le \frac{1}{2\pi} \oint_{\Gamma} |\lambda - \mu|^{-1} \left\|(\lambda - A)^{-1}\right\| |\, d\lambda|.$$

We know that neither A nor A_+ has any spectrum in a neighborhood of Γ (since $\Gamma \in \rho(A)$), and away from Γ the right-hand side of the above inequality is bounded, uniformly in μ. This implies that A_+ cannot have any spectrum outside Γ, because by Lemma 3.2.8(iii), $\left\|(\mu - A_+)^{-1}\right\| \to \infty$ as μ approaches a point in $\sigma(A_+)$. In the same way it can be shown that A_- cannot have any spectrum inside Γ: note that by (3.14.3),

$$(\mu - A_-)^{-1}x = \frac{1}{2\pi j}\oint_\Gamma \frac{(\lambda - A)^{-1}}{\lambda - \mu}\, x\, d\lambda$$

for all $x \in X_-$ and all $\mu \in \rho(A)$ which lie inside Γ.

(iv) See Theorems 3.14.6 and 3.14.8. □

In the case of a *normal* semigroup it is possible to use a different type of spectral projection, which does not require the spectrum of the generator to be disconnected.

Theorem 3.14.11 *Let A be a closed and densely defined normal operator on a Hilbert space X (i.e., $A^*A = AA^*$), and let E be the corresponding spectral resolution of A, so that*

$$\langle Ax, y\rangle_X = \int_{\sigma(A)} \lambda\langle E(d\lambda)x, y\rangle, \quad x \in \mathcal{D}(A), \quad y \in X.$$

Let F be a bounded Borel set in $\mathbb{C}$, and let $\pi = E(F)$, i.e.,

$$\langle \pi x, y\rangle_X = \int_{\sigma(A)\cap F} \langle E(d\lambda)x, y\rangle, \quad x \in X, \quad y \in X.$$

Then π is an orthogonal projection which maps X into $\mathcal{D}(A)$. Denote $X_+ = \mathcal{R}(\pi)$ and $X_- = \mathcal{N}(\pi)$, $A_+ = A_{|X_+}$ and $A_- = A_{|X_-}$. Then the claims (i), (ii), and (iv) in Theorem 3.14.10 hold, and the claim (iii) is replaced by

(iii′) $\sigma(A_+) = \overline{\sigma(A) \cap F}$, *and* $\sigma(A_-) = \overline{\sigma(A) \setminus F}$.

We leave the proof of this theorem to the reader. (That π is a self-adjoint projection follows from the definition of a spectral resolution, and the rest follows either directly from the properties of a spectral resolution or from an argument similar to the one used in the proof of Theorem 3.14.10. Note, in particular, that $\mathcal{R}(\pi) \subset \mathcal{D}(A)$ since F is bounded.)

So far we have primarily looked at *closed* invariant subspaces. There is also another class of subspaces that play an important role in the theory, namely invariant Banach spaces which are *continuously embedded* in the state space. An example of this are the spaces X_α with $\alpha > 0$ introduced in Sections 3.6 and 3.9. These spaces are typically invariant under the semigroup $\mathfrak{A}$, but not under its generator A.

Definition 3.14.12 Let $A\colon X \supset \mathcal{D}(A) \to A$ be a linear operator, and let Y be a subspace of X (not necessarily dense). By the *part of A in Y* we mean the operator $\widetilde{A}$ which is the restriction of A to

$$\mathcal{D}(\widetilde{A}) = \{x \in \mathcal{D}(A) \cap Y \mid Ax \in Y\}.$$

Definition 3.14.13 Let $\mathfrak{A}$ be a C_0 semigroup on the Banach space X with generator A, and let let Y be another Banach space embedded in X (not necessarily densely). We call Y *A-admissible* if Y is invariant under $\mathfrak{A}$ and the restriction of $\mathfrak{A}$ to Y is strongly continuous in the norm of Y (i.e, $\mathfrak{A}_{|Y}$ is a C_0 semigroup on Y).

Theorem 3.14.14 *Let $\mathfrak{A}$ be a C_0 semigroup on the Banach space X with generator A and growth bound $\omega_{\mathfrak{A}}$, and let Y be another Banach space embedded in X (not necessarily densely). Then Y is A-admissible if and only if the following two conditions hold:*

(i) *Y is an invariant subspace of $(\lambda - A)^{-1}$ for all real $\lambda > \omega_{\mathfrak{A}}$,*
(ii) *The part of A in Y is the generator of a C_0 semigroup on Y.*

When these conditions hold, then the generator of $\mathfrak{A}_{|Y}$ is the part of A in Y.

Proof Assume that Y is A-admissible. Then Theorem 3.2.9(i) (applied to $\mathfrak{A}_{|Y}$) implies (i) (and it is even true that Y is an invariant subspace of $(\lambda - A)^{-1}$ for all real $\lambda \in \mathbb{C}^+_{\omega_{\mathfrak{A}}}$). Denote the generator of $\mathfrak{A}_{|Y}$ by A_1, and denote the part of A in Y by $\widetilde{A}$. Since the norm in Y is stronger than the norm in X, it follows easily that $\mathcal{D}(A_1) \subset \mathcal{D}(A) \cap Y$, and that for $x \in \mathcal{D}(A_1)$, $Ax = A_1x \in Y$. Thus, $\widetilde{A}$ is an extension of A_1. On the other hand, if $x \in \mathcal{D}(\widetilde{A})$, then $Ax \in Y$, and each term in the identity

$$\mathfrak{A}^t x - x = \int_0^t \mathfrak{A}^s Ax\, ds, \qquad t \geq 0,$$

belongs to Y. Dividing by t and letting $t \downarrow 0$ we find that $x \in \mathcal{D}(A_1)$. Thus, $A_1 = \widetilde{A}$, and $\widetilde{A}$ is the generator of the C_0 semigroup $\mathfrak{A}_{|Y}$ on Y.

Conversely, suppose that (i) and (ii) hold. Denote the C_0 semigroup generated by $\widetilde{A}$ by $\widetilde{\mathfrak{A}}$. For all $x \in \mathcal{D}(\widetilde{A})$ and all $\lambda > \omega_{\mathfrak{A}}$ we have (since $\lambda \in \rho(A)$)

$$(\lambda - A)^{-1}(\lambda - \widetilde{A})x = (\lambda - A)^{-1}(\lambda - A)x = x,$$

and for all $y \in Y$ we have (because of (ii))

$$(\lambda - \widetilde{A})(\lambda - A)^{-1}x = (\lambda - A)(\lambda - A)^{-1}x = x.$$

Thus, $(\lambda - \widetilde{A})$ maps $\mathcal{D}(\widetilde{A})$ one-to-one onto Y, and $(\lambda - \widetilde{A})^{-1}$ is the restriction of $(\lambda - A)^{-1}$ to Y. Fix $t > 0$, and choose n so large that $n/t > \omega_{\mathfrak{A}}$. Then, for all $\lambda > \omega_{\mathfrak{A}}$ and all $y \in Y$,

$$\left(1 - \frac{t}{n}A\right)^{-n} y = \left(1 - \frac{t}{n}\widetilde{A}\right)^{-n} y.$$

Let $n \to \infty$. By Theorem 3.7.5, the left-hand side tends to $\mathfrak{A}^t y$ in X and the right-hand side tends to $\widetilde{\mathfrak{A}}^t y$ in Y, hence in X. Thus, $\widetilde{\mathfrak{A}} = \mathfrak{A}_{|Y}$. This implies both that Y is invariant under $\mathfrak{A}$, and that $\mathfrak{A}_{|Y}$ is strongly continuous. □

Let us end this section by proving the following extension of Theorem 3.14.8 (to get that theorem we take $\widetilde{X} = X$, $E = \pi$, $\widetilde{A} = A$, and $\widetilde{\mathfrak{A}} = \mathfrak{A}$).

Theorem 3.14.15 *Let A be the generator of a C_0 semigroup $\mathfrak{A}$ on X, let $\widetilde{A}$ be the generator of a C_0 semigroup $\widetilde{\mathfrak{A}}$ on $\widetilde{X}$, and let $E \in \mathcal{B}(X; \widetilde{X})$. Then the following conditions are equivalent.*

(i) $E\mathfrak{A}^t = \widetilde{\mathfrak{A}}^t E$ *for all* $t \geq 0$.
(ii) $E(\lambda - A)^{-1} = (\lambda - \widetilde{A})^{-1} E$ *for some* $\lambda \in \rho(A) \cap \rho(\widetilde{A})$.
(iii) $E(\lambda - A)^{-1} = (\lambda - \widetilde{A})^{-1} E$ *for all* $\lambda \in \rho(A) \cap \rho(\widetilde{A})$.
(iv) *E maps $\mathcal{D}(A)$ into $\mathcal{D}(\widetilde{A})$, and $EAx = \widetilde{A}Ex$ for all $x \in \mathcal{D}(A)$.*

If, in addition, $\mathfrak{A}$ and $\widetilde{\mathfrak{A}}^t$ are groups, then these conditions are further equivalent to

(v) $E\mathfrak{A}^t = \widetilde{\mathfrak{A}}^t E$ *for all* $t \in R$.

Proof (i) ⇒ (ii): If (i) holds, then by Theorem 3.2.9(i), for all $\lambda > \max\{\omega_{\mathfrak{A}}, \omega_{\widetilde{\mathfrak{A}}}\}$ and all $x \in X$,

$$E(\lambda - A)^{-1}x = \int_0^\infty \mathrm{e}^{-\lambda s} E\mathfrak{A}^s x = \int_0^\infty \mathrm{e}^{-\lambda s} \widetilde{\mathfrak{A}}^s Ex = (\lambda - \widetilde{A})^{-1} Ex.$$

(ii) ⇒ (iv): By (ii), for all $x \in \mathcal{D}(A)$,

$$Ex = (\lambda - \widetilde{A})^{-1} E(\lambda - A)x.$$

This implies that E maps $\mathcal{D}(A)$ into $\mathcal{D}(\widetilde{A})$. Applying $(\lambda - \widetilde{A})$ to both sides of this identity we get $(\lambda - \widetilde{A})Ex = E(\lambda - A)x$, or equivalently, $EAx = \widetilde{A}Ex$.

(iv) ⇒ (iii): If (iv) holds, then $(\lambda - \widetilde{A})Ex = E(\lambda - A)x$ for all $\lambda \in \mathbb{C}$ and all $x \in \mathcal{D}(A)$. In particular, if we take $\lambda \in \rho(A) \cap \rho(\widetilde{A})$, then we can apply $(\lambda - \widetilde{A})^{-1}$ to both sides of this identity and replace x by $(\lambda - A)^{-1}x$ to get for all $x \in X$,

$$E(\lambda - A)^{-1}x = (\lambda - \widetilde{A})^{-1} Ex.$$

(iii) ⇒ (i): If (iii) holds, then by Theorem 3.7.5, for all $t \geq 0$ and all $x \in X$,

$$\mathfrak{A}^t Ex = \lim_{n\to\infty} \Big(1 - \frac{t}{n} A\Big)^{-n} Ex = E \lim_{n\to\infty} \Big(1 - \frac{t}{n} \widetilde{A}\Big)^{-n} x = E\widetilde{\mathfrak{A}}^t x.$$

(v) We get $E\mathfrak{A}^{-t} = \widetilde{\mathfrak{A}}^{-t} E$ for all $t \geq 0$ by multiplying the identity in (i) by $\widetilde{\mathfrak{A}}^{-t}$ to the left and by $\mathfrak{A}^{-t}$ to the right.

□

3.15 Comments

By now, most of the results in this chapter are classic. We refer the reader to Davies [1980], Dunford and Schwartz [1958, 1963, 1971], Goldstein [1985], Lunardi [1995], Nagel [1986], Hille and Phillips [1957], Pazy [1983], and Yosida [1974] for the history and theory of C_0 semigroups beyond what we have presented here.

Sections 3.2–3.3 The generators of the shift semigroups and their resolvents have been studied in, e.g., Hille and Phillips [1957, Sections 19.2–19.4]. Diagonal semigroups appear frequently in the theory of *parabolic* and *hyperbolic* partial differential equations. Sometimes the basis of eigenvectors of the generator is not orthonormal, but instead a *Riesz basis* in the sense of Curtain and Zwart (1995, Definition 2.3.1). However, a Riesz basis can be transformed into an orthonormal basis by means of a similarity transformation of the type described in Example 2.3.7. This makes it easy to extend the theory for diagonal semigroups presented in Examples 3.3.3 and 3.3.5 to semigroups whose generator has a set of eigenvectors which are a Riesz basis for the state space. These types of semigroups are studied in some detail by Curtain and Zwart [1995].

Sections 3.4, 3.7, and 3.8 See, for example, Pazy [1983], for the history of the Hille–Yosida and Lumer–Phillips theorems, the different approximation theorems, and the Cauchy problem.

Section 3.5 Our presentation of the dual semigroup follows roughly Hille and Phillips (1957, Chapter 14), except for the fact that we use the conjugate-linear dual instead of the linear dual.

Section 3.6 Rigged spaces of the type discussed in Section 3.6 are part of the traditional semigroup formulation of partial differential equations of *parabolic* and *hyperbolic* type. See, for example, Lions [1971], Lunardi [1995], and Lasiecka and Triggiani [2000a, b]. In the reflexive case the space X_{-1} is usually defined to be the dual of the domain of the adjoint operator (see Remark 3.6.1). The spaces X_1 and X_{-1} have been an important part of the theory of well-posed linear systems since Helton [1976]. Spaces X_α of fractional order are introduced in Section 3.9.

Section 3.9 This section is based in part on Dunford and Schwartz (1958, Section VII.9), Rudin (1973, Chapter 10), and Pazy (1983, Section 2.6).

Section 3.10 The results of this section are fairly standard. Analytic semigroups most frequently arise from the solution of *parabolic* or heavily damped *hyperbolic* partial differential equations. Theorem 3.10.11 has been modeled after Lunardi (1995, Proposition 2.4.1) and Mikkola (2002, Lemma 9.4.3). For further results about analytic semigroups we refer the reader to standard text books, such as Goldstein (1985, Sections 1.5, 2.4 and 2.5), Hille and Phillips (1957, Chapter 17), Lunardi [1995], Lasiecka and Triggiani [2000a, b], and Pazy (1983, Sections 2.5, 7.2 and 7.3).

Section 3.11 The proof of Theorem 3.11.4 is based on Davies (1980, Theorem 2.19), which goes back to Hille and Phillips (1957, Theorem 16.4.1, p. 460). This theorem is also found in Nagel (1986, p. 87). The reduction of the spectrum determined growth property to the corresponding spectral inclusion that we use in the proof of Theorem 3.11.4 is classic; see, e.g., Slemrod (1976, pp. 783–784), Triggiani (1975, p. 387), Zabczyk (1975), or Nagel (1986, p. 83). Corollary 3.11.5(i) is due to Triggiani [1975, pp. 387–388] and Corollary 3.11.5(ii) is due to Zabczyk [1975]. Theorem 3.11.6 was proved independently by Herbst [1983], Huang [1985], and Prüss [1984]. Lemma 3.11.7 has been modeled after Weiss (1988b, Theorem 3.4). The implication (v) $\Rightarrow$ (i) in Theorem 3.11.8 is often called *Datko's theorem* after Datko (1970, p. 615) who proves this implication in the Hilbert space case with $p = 2$. The general version of the same implication was proved by Pazy (1972) (see Pazy (1983, Theorem 4.1 and p. 259)). Our proof follows the one given by Pritchard and Zabczyk [1981, Theorem 3.4]. In the reflexive case the implication (vii) $\Rightarrow$ (i) can be reduced to the implication (v) $\Rightarrow$ (i) through duality, but we prefer to give a direct proof, which is valid even in the non reflexive case (the non reflexive part of the implication (vii) $\Rightarrow$ (i) may be new). Our proof of the implication (ix) $\Rightarrow$ (i) in Theorem 3.11.8 has been modeled after Weiss (1988b, Theorem 4.2) (we have not been able to find this explicit implication in the existing literature). Examples of semigroups which do *not have the spectrum determined growth property* are given in Curtain and Zwart (1995, Example 5.1.4 and Exercise 5.6), Greiner *et al.* (1981, Example 4.2), Davies (1980, Theorem 2.17), Hille and Phillips (1957, p. 665), and Zabczyk [1975]. For additional results about the spectral determined growth property, and more generally, the spectral mapping property, we refer the reader to van Neerven [1996], which is devoted to this very question.

Section 3.12 The frequency domain plays a very important role in many texts, including this one. Most mathematical texts use the Fourier transform instead of the Laplace transform, and they replace the right half-plane by the upper half-plane. However, we prefer to stick to the engineering tradition at this point. It is possible to develop a symbolic calculus for generators of semigroups based on the representation of $\hat{u}(A)$ in Theorem 3.12.6 (this is done in Dunford and Schwartz (1958, Section VIII.2) in the case where A generates a group rather than a semigroup), and our definition of $(\gamma - A)^{-\alpha}$ given in Section 3.9 is based on a special case of that calculus.

Section 3.13 Most books on operator theory *define* the shift semigroups using the characterization given in Propositions 3.13.1 and 3.13.2 (and they completely ignore the time-domain versions of these semigroups). This is especially true in the Hilbert space L^2-well-posed case. In our context the time-domain versions are more natural to work with.

Section 3.14 The results in this section are classic (but not always that easy to find in the literature). Our presentation follows loosely Curtain and Zwart [1995] and Pazy [1983].