Lecture 3

Approximation of Semigroups – Part 1

The main topic of this lecture will be to establish approximation theorems for operator semigroups. Consider the abstract initial value problem (Cauchy problem)

$$\begin{cases} \dot{u}(t) = Au(t), & t \geq 0 \\ u(0) = u_0 \in X, \end{cases} \tag{ACP}$$

where we suppose that A generates the strongly continuous semigroup T on X. In many applications we are able to construct a sequence of approximating operators A_n which generate strongly continuous semigroups T_n and converge to A in some sense. The question is if this implies the convergence of the semigroups T_n to T?

These approximations usually involve some numerical methods: Either some approximation of the operator A (for example finite differences, see Example 3.7 below), or an approximation of the solution u of a stationary problem $Au = f$ (for example finite element or spectral method, see Example 3.6).

After working through the examples and the exercises, we will see that operator norm convergence would be simply too much to expect, and to weaken this type of convergence, the pointwise one is our next bet. Therefore, in this lecture we shall investigate the *strong convergence* of semigroups, i.e., the property

$$T_n(t)f \to T(t)f \quad \text{as } n \to \infty \quad \text{for all } f \in X \tag{3.1}$$

uniformly for t in compact intervals of $[0, \infty)$.

Remark 3.1. If the convergence stated above holds for the semigroups T_n, T, then the uniform boundedness principle, Theorem 2.28, immediately implies that $\|T_n(t)\|$ has to remain bounded as $n \to \infty$ for all $t \geq 0$. More is true: There exist constants $M \geq 1$, $\omega \in \mathbb{R}$ such that

$$\|T_n(t)\| \leq M\mathrm{e}^{\omega t} \qquad \text{holds for all } n \in \mathbb{N},\ t \geq 0. \tag{3.2}$$

We leave the proof as Exercise 1. This exponential inequality, called *stability condition*, provides a necessary condition to have convergence of the semigroups as in (3.1).

3.1 Generator approximations

Usually, after discretising a differential operator, we end up with a matrix, hence not an operator on the original space, but rather on $\mathbb{C}^n$. So the next important point is that we not only have approximating operators, but also approximating spaces. This motivates our general setup.

Assumption 3.2. Let X_n, X be Banach spaces and assume that there are bounded linear operators $P_n : X \to X_n$, $J_n : X_n \to X$ with the following properties:

- There is a constant $K > 0$ with $\|P_n\|, \|J_n\| \leq K$ for all $n \in \mathbb{N}$,

- $P_nJ_n = I_n$, the identity operator on X_n, and
- $J_nP_nf \to f$ as $n \to \infty$ for all $f \in X$.

An important remark on our notation. The symbol $\|\cdot\|$ refers here to the operator norm in $\mathscr{L}(X, X_n)$ and $\mathscr{L}(X_n, X)$, respectively. We use the convention that if it is clear from the context, we often do not distinguish in the notation between the norms on different spaces.

Example 3.3 (spectral method). Consider the spaces $X = \ell^2$ and $X_n = \mathbb{C}^n$ with the Euclidian norm and define for $f = (f_k) \in \ell^2$ the operator

$$P_n : \ell^2 \to \mathbb{C}^n, \quad P_nf := (f_1, \dots, f_n),$$

and for $y = (y_1, \dots, y_n) \in \mathbb{C}^n$ the operator

$$J_n : \mathbb{C}^n \to \ell^2, \quad J_n(y_1, y_2, \dots, y_n) := (y_1, \dots, y_n, 0, \dots).$$

Clearly, J_nP_n equals the projection onto the first n coordinates. For this example all the above mentioned properties are satisfied, in particular, the last one because

$$\|J_nP_nf - f\|^2 = \sum_{k=n+1}^{\infty} |f_k|^2 \to 0 \text{ as } n \to \infty.$$

This is essentially the same example as if we take $X = L^2(0,1)$, $X_n = \mathbb{C}^n$, $P_nf :=$ the first n Fourier coefficients of f, and $J_n(y_1, \dots, y_n) :=$ the finite trigonometric sum built from the coefficients $y_1, \dots, y_n$ (*spectral method*, see Appendix A.2).

Example 3.4 (finite difference). In this example we try to capture the standard discretisation of continuous functions through grid points in an abstract way (*finite difference method*, see Appendix A.1). Let

$$X := \{f \in \mathrm{C}([0,1]) : f(1) = 0\} = \mathrm{C}_{(0)}([0,1]) \quad \text{and} \quad X_n = \mathbb{C}^n,$$

both with the respective maximum norm, see also Exercise 7 in Lecture 1. We define

$$(P_nf)_k := f(\tfrac{k}{n}), \quad k = 0, \dots, n-1,$$

and

$$J_n(y_0, \dots, y_{n-1}) := \sum_{k=0}^{n-1} y_k B_{n,k}(x),$$

where for $k \in \{0, \dots n-1\}$ and $x \in [0,1]$ we have

$$B_{n,k}(x) = \begin{cases} n\big(x - \frac{k}{n}\big) & \text{if } x \in \big[\frac{k-1}{n}, \frac{k}{n}\big), \\ n\big(\frac{k+1}{n} - x\big) & \text{if } x \in \big[\frac{k}{n}, \frac{k+1}{n}\big), \\ 0 & \text{otherwise,} \end{cases}$$

the hat functions. Then $P_nJ_n = I_{\mathbb{C}^n}$ and for $n \to \infty$ the convergence $J_nP_nf \to f$ hold true, see Exercise 2.

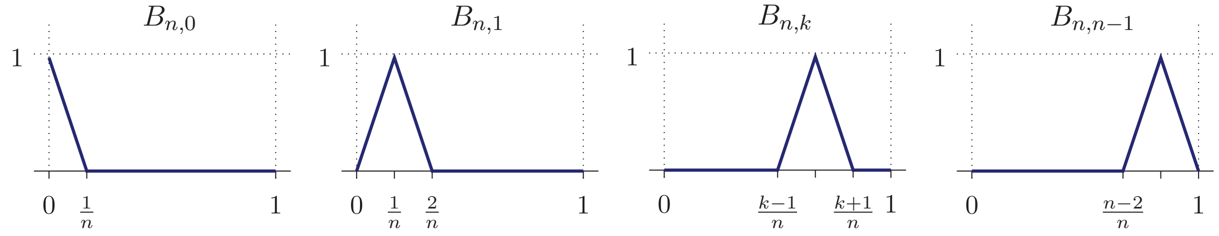

Figure 3.1: The hat functions

After setting the stage for the problems, let us turn our attention back to approximation problems and introduce a general assumption on the convergence of the generators. When examining the examples, we learn that the following setup is quite natural.

Assumption 3.5. Suppose that the operators A_n, A generate strongly continuous semigroups on X_n and X, respectively, and that there are constants $M \geq 0$, $\omega \in \mathbb{R}$ such that the stability condition (3.2) holds. Further suppose that there is a dense subset $Y \subset D(A)$ such that for all $g \in Y$ there is a sequence $y_n \in D(A_n)$ satisfying

$$\|y_n - P_n g\|_{X_n} \to 0 \quad \text{and} \quad \|A_n y_n - P_n A g\|_{X_n} \to 0 \quad \text{as } n \to \infty. \tag{3.3}$$

Clearly, by Assumption 3.2, the convergence stated above is equivalent to

$$\|J_n A_n y_n - Ag\| \to 0 \quad \text{as } n \to \infty$$

for all $g \in Y$. We will freely make use of this equivalence later on, depending on which formulation is more convenient in the given situation. In applications we typically (but not necessarily) have $P_n Y \subset D(A_n)$ and $y_n = P_n g$.

Example 3.6 (spectral method). Taking the setup of Example 3.3 and motivated by the heat equation introduced in Section 1.1, we define the operator on $f = (f_k) \in X$ as

$$Af := (-k^2 f_k)_{k \in \mathbb{N}} \quad \text{with} \quad D(A) = \big\{(f_k) \in X \,:\, (k^2 f_k) \in X\big\}.$$

Further, for $f \in D(A)$, we define

$$A_n(P_n f) := P_n A f, \quad \text{i.e.,} \quad A_n(y_1, y_2, \dots, y_n) := (-y_1, -4y_2, \dots, -n^2 y_n).$$

See also Appendix A for some details about the spectral method.

Analogously, motivated by the Schrödinger equation, we define the operator B by

$$Bf := (\mathrm{i}k^2 f_k) \quad \text{for } f = (f_k) \in D(B) := D(A).$$

The approximating operators are then $B_n(y_1, \dots, y_n) := (\mathrm{i}y_1, \mathrm{i}4y_2, \dots, \mathrm{i}n^2 y_n)$.

Example 3.7 (finite difference). Continuing Example 3.4, define the generator

$$Af := f' \quad \text{with} \quad D(A) := \big\{f \in \mathrm{C}^1([0,1]) : f(1) = 0\big\}.$$

For $y = (y_0, \dots y_{n-1}) \in X_n$, we define

$$(A_n y)_k := n(y_{k+1} - y_k) \quad \text{for } k := 0, \dots n-2 \quad \text{and} \quad (A_n y)_{n-1} := -n y_{n-1},$$

being the standard first-order finite difference scheme. By using that $y = P_n f$, we can write it in a slightly different form:

$$(A_n P_n f)_k := n\big(f(\tfrac{k+1}{n}) - f(\tfrac{k}{n})\big) \quad \text{for } k = 0, \ldots, n-1.$$

Then, by the mean value theorem, we obtain

$$\begin{aligned}
\|J_n A_n P_n f - Af\|_\infty &= \left\| \sum_{k=0}^{n-1} \frac{f(\frac{k+1}{n}) - f(\frac{k}{n})}{\frac{1}{n}} B_{n,k} - f' \right\|_\infty = \left\| \sum_{k=0}^{n-1} f'(\xi_k) B_{n,k} - f' \right\|_\infty \\
&\le \left\| f' - \sum_{k=0}^{n-1} f'(\tfrac{k}{n}) B_{n,k} \right\|_\infty + \max_{k=0,\ldots,n-1} |f'(\tfrac{k}{n}) - f'(\xi_k)| \\
&\le \left\| f' - \sum_{k=0}^{n-1} f'(\tfrac{k}{n}) B_{n,k} \right\|_\infty + \omega(f', \tfrac{1}{n}) \to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

Here, $\omega(f', s)$ is the modulus of (uniform) continuity of the function f' defined as usual by

$$\omega(f', s) := \sup\big\{ |f'(x) - f'(y)| \sup |x - y| \le s \big\}.$$

An important observation concerning this example is the following. If we assume a bit more regularity and take $f \in \mathrm{C}^2([0,1])$, then we obtain

$$(A_n P_n f - P_n A f)_k = \frac{f(\frac{k+1}{n}) - f(\frac{k}{n})}{\frac{1}{n}} + f'(\tfrac{k}{n}) = f''(\xi_k) \frac{1}{2n},$$

by Taylor's formula, and hence,

$$\|A_n P_n f - P_n A f\| \le \frac{\|f''\|_\infty}{2n}.$$

This means that, in this example, we not only have convergence, but even *first-order convergence* for twice differentiable functions.

The following Proposition 3.8 will guarantee that for all $f \in \mathrm{C}^2([0,1])$ which remain in C^2 under the semigroup, this convergence carries over to the convergence of the semigroups. More precisely, we have that

for all $$f \in D(A^2) = \big\{ f \in \mathrm{C}^2([0,1]) : f(0) = f'(0) = f''(0) = 0 \big\}$$

and for all $t > 0$ there is $C > 0$ such that

$$\|T_n(t) P_n f - P_n T(t) f\| \le C \frac{\|f\|_{\mathrm{C}^2}}{n}.$$

Motivated by the previous example, we can formulate our first approximation result on semigroups.

Proposition 3.8. *Suppose that Assumptions 3.2 and 3.5 hold, that $P_n Y \subset D(A_n)$, and that Y is a Banach space invariant under the semigroup T satisfying*

$$\|T(t)\|_Y \le M \mathrm{e}^{\omega t}.$$

If there are constants $C > 0$ and $p \in \mathbb{N}$ with the property that for all $f \in Y$

$$\|A_n P_n f - P_n A f\|_{X_n} \leq C \frac{\|f\|_Y}{n^p},$$

then for all $t > 0$ there is $C' > 0$ such that

$$\|T_n(t) P_n f - P_n T(t) f\|_{X_n} \leq C' \frac{\|f\|_Y}{n^p}.$$

Moreover, this convergence is uniform in t on compact intervals.

In this case we say that we have **convergence of order** p. Also notice that, as discussed in Proposition 2.20, the subspace Y will be a core for the generator A (being dense and invariant under the semigroup), and hence $(\lambda - A)Y \subset X$ will be also a dense set for some/all $\lambda \in \rho(A)$. In many examples we will take $Y := D(A^l)$ for some $l \in \mathbb{N}$.

Proof. For simplicity, we first carry out the proof in the special case $X_n = X$ and $J_n = P_n = I$. It is clear that for $f \in Y$, we have $Af = A_n f + (A - A_n)f$. Application of the fundamental theorem of calculus to the continuously differentiable function $[0, t] \ni s \mapsto T_n(t-s)T(s)f$ implies that the variation of constants formula

$$T(t)f = T_n(t)f + \int_0^t T_n(t-s)(A - A_n)T(s)f \, \mathrm{d}s$$

holds. Therefore, we have

$$\begin{aligned} \|T(t)f - T_n(t)f\| &\leq \int_0^t M \mathrm{e}^{\omega(t-s)} \|(A - A_n)T(s)f\| \, \mathrm{d}s \\ &\leq \int_0^t M \mathrm{e}^{\omega(t-s)} \frac{C}{n^p} \|T(s)f\|_Y \, \mathrm{d}s \leq M^2 \mathrm{e}^{\omega t} \cdot t \cdot \frac{C}{n^p} \|f\|_Y. \end{aligned}$$

From this the assertion follows. The general case can be considered by applying the fundamental theorem of calculus to the modified function $[0, t] \ni s \mapsto T_n(t-s)P_n T(s)f$ to obtain the variation of constants formula

$$P_n T(t) f = T_n(t) P_n f + \int_0^t T_n(t-s) \big(P_n A - A_n P_n\big) T(s) f \, \mathrm{d}s.$$

From here, the argument is the same as above. □

3.2 Resolvent approximations

We will see that in many applications the situation is slightly more complicated than we had before.

Example 3.9 (spectral method). Going back to Example 3.6, we can see quickly that it is difficult and unnatural to obtain similar estimates on the convergence of the generators. But we can immediately infer the following.

If $g = (g_n) \in X$, then there is $f = (f_n) \in D(A)$ such that $g = Af$, $f = A^{-1}g$. Hence,

$$\begin{aligned}\|J_nP_nf - f\| &= \|J_nP_nA^{-1}g - A^{-1}g\| = \|(0,\dots,0,(n+1)^{-2}g_{n+1},\dots)\| \\ &\le \frac{1}{(n+1)^2}\|J_nP_ng - g\| \le \frac{1}{(n+1)^2}\|g\|.\end{aligned}$$

Since, by definition, $A_nP_n = P_nA$, this implies that

$$\|J_nA_n^{-1}P_n - A^{-1}\| \le \frac{1}{(n+1)^2},$$

which means that it is natural to expect the convergence of the resolvent operators.[1]

Hence, in order to infer the convergence of the semigroups, we have to prove the analogue of Proposition 3.8 but now with resolvent convergence.

Proposition 3.10. *Assume that Assumptions 3.2 and 3.5 hold. If there are $C > 0$ and $p \in \mathbb{N}$ such that*

$$\|A_n^{-1}P_n - P_nA^{-1}\| \le \frac{C}{n^p},$$

then for all $t > 0$ there is $C' > 0$ such that

$$\|T_n(t)P_ng - P_nT(t)g\|_{X_n} \le C'\frac{\|g\|_{A^2}}{n^p}$$

for all $g \in D(A^2)$. Furthermore, this convergence is uniform in t on compact intervals[2].

A surprising and important observation here is that although we assume operator norm convergence of the resolvents, we do not get back norm convergence of the approximating semigroups in general. This observation will be illustrated in an example before the proof.

Remark 3.11. 1. It is clear that instead of the convergence of A_n^{-1} to A^{-1} we can assume $J_n(\lambda - A_n)^{-1}P_n \to (\lambda - A)^{-1}$ for some $\lambda \in \rho(A) \cap \rho(A_n)$.

2. In concrete cases, as we shall also see in the examples below, these results are far from being sharp. To obtain better results, we also need more structural properties of the approximation.

Example 3.12 (spectral method). We refer to Example 3.6 again, and define the semigroups generated by A and B as

$$T(t)f := \mathrm{e}^{tA}f = (\mathrm{e}^{-t}f_1, \mathrm{e}^{-4t}f_2, \dots, \mathrm{e}^{-k^2t}f_k, \dots)$$

and

$$S(t)f := \mathrm{e}^{tB}f = (\mathrm{e}^{\mathrm{i}t}f_1, \mathrm{e}^{\mathrm{i}4t}f_2, \dots, \mathrm{e}^{\mathrm{i}k^2t}f_k, \dots).$$

The approximating semigroups are $T_n(t) = \mathrm{diag}(\mathrm{e}^{-t}, \mathrm{e}^{-4t}, \dots, \mathrm{e}^{-n^2t})$, that is,

$$J_nT_n(t)P_n = J_nP_nT(t).$$

Similarly, we have $S_n(t) = \mathrm{diag}(\mathrm{e}^{\mathrm{i}t}, \mathrm{e}^{\mathrm{i}4t}, \dots, \mathrm{e}^{\mathrm{i}n^2t})$. We conclude that

$$\|J_nP_nT(t)f - T(t)f\|^2 = \sum_{k=n+1}^{\infty} \mathrm{e}^{-2k^2t}|f_k|^2 \le \mathrm{e}^{-2(n+1)^2t}\sum_{k=n+1}^{\infty}|f_k|^2 \le \mathrm{e}^{-2(n+1)^2t}\|f\|^2,$$

[1] Recall that $A^{-1} = -R(0, A)$, the resolvent defined in Lecture 2.

[2] We use the notation $\|g\|_{A^2} := \|g\| + \|A^2f\|$ for the graph norm of A^2.

which shows that for $t > 0$ we get a convergence in operator norm being quicker than any polynomial.

For the other example, however, we observe that

$$\|J_nP_nS(t)f - S(t)f\|^2 = \sum_{k=n+1}^{\infty} |\mathrm{e}^{\mathrm{i}k^2t} f_k|^2 = \sum_{k=n+1}^{\infty} |f_k|^2.$$

Let us introduce again $f = (f_n) \in X$ for $g = (g_n) \in D(B)$ such that $f = Bg$, $g = B^{-1}f$. Then we can repeat the argument:

$$\begin{aligned}\|J_nP_nS(t)g - S(t)g\|^2 &= \sum_{k=n+1}^{\infty} |\mathrm{e}^{\mathrm{i}k^2t} g_k|^2 = \sum_{k=n+1}^{\infty} |\tfrac{1}{k^2} f_k|^2 \\ &\le \frac{1}{(n+1)^4} \sum_{k=n+1}^{\infty} |f_k|^2 \le \frac{1}{(n+1)^4}\|f\|^2 = \frac{1}{(n+1)^4}\|Bg\|^2.\end{aligned}$$

This shows that in this case we can only recover the convergence order $p = 2$ for $g \in D(B)$. Thus, we have to be careful even with this simple example.

Proof of Proposition 3.10. In order to simplify matters, we again concentrate first on the calculations in the case $X_n = X$, $J_n = P_n = I$. Let us start by fixing some $t_0 > 0$. Then for all $t \in [0, t_0]$ we obtain that

$$\begin{aligned}&\big(T_n(t) - T(t)\big)A^{-1}f \\ &\quad = \underbrace{T_n(t)(A^{-1} - A_n^{-1})f} + A_n^{-1}(T_n(t) - T(t))f + \underbrace{(A_n^{-1} - A^{-1})T(t)f}.\end{aligned} \tag{3.4}$$

It is clear from the stability assumption that the first and the last term of this sum converge to 0 in the operator norm and at the desired rate, i.e.,

$$\|T_n(t)(A^{-1} - A_n^{-1})f\| \le M\mathrm{e}^{\omega t_0}\|A^{-1} - A_n^{-1}\| \cdot \|f\|,$$

and

$$\|(A^{-1} - A_n^{-1})T(t)f\| \le M\mathrm{e}^{\omega t_0}\|A^{-1} - A_n^{-1}\| \cdot \|f\|.$$

Hence, we have to concentrate on the middle term. Instead of this term, we consider first a more symmetric one where the fundamental theorem of calculus comes to help. Indeed, let us first show that

$$A_n^{-1}\big(T(t) - T_n(t)\big)A^{-1}h = \int_0^t T_n(t-s)\left(A^{-1} - A_n^{-1}\right)T(s)h\,\mathrm{d}s \tag{3.5}$$

holds for all $h \in X$ and $t > 0$. To this end, observe that the function

$$[0, t] \ni s \mapsto T_n(t-s)A_n^{-1}T(s)A^{-1}h \in X$$

is differentiable by Theorem 2.31, and its derivative is

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}s}\left(T_n(t-s)A_n^{-1}T(s)A^{-1}h\right) &= T_n(t-s)\left(-A_nA_n^{-1}T(s) + A_n^{-1}T(s)A\right)A^{-1}h \\ &= T_n(t-s)\left(A^{-1} - A_n^{-1}\right)T(s)h.\end{aligned}$$

Hence, the fundamental theorem of calculus yields formula (3.5).

Now we can see that for $h \in X$ the inequality

$$\begin{aligned}\|A_n^{-1}(T(t) - T_n(t))A^{-1}h\| &\le \int_0^t M\mathrm{e}^{\omega(t-s)} \cdot \|A^{-1} - A_n^{-1}\| \cdot \|T(s)h\| \,\mathrm{d}s \\ &\le t_0 M^2 \mathrm{e}^{\omega t_0} \|A^{-1} - A_n^{-1}\| \cdot \|h\|\end{aligned}$$

holds. Summarising the estimates from above, we conclude that for $g \in D(A^2)$ we can introduce $f = Ag$ and $h = Af$ to obtain that

$$\|T_n(t)g - T(t)g\| \le \|A^{-1} - A_n^{-1}\| M \mathrm{e}^{\omega t_0}(t_0 M \|A^2 g\| + 2\|Ag\|),$$

which yields the desired estimate. □

3.3 The First Trotter–Kato Theorem

We turn our attention now to the general approximation theorems with the weakest possible assumptions. We start by investigating convergence of generators. As we have seen before, convergence of operators and convergence of the corresponding resolvents are connected.

Lemma 3.13. *Suppose that Assumption 3.2 is satisfied and that A_n, A are closed operators on X_n and X, respectively, such that there is some $\lambda \in \rho(A_n) \cap \rho(A)$ for all $n \in \mathbb{N}$ and there is a constant $M \ge 0$ with the property*

$$\|R(\lambda, A_n)\| \le M \quad \text{for all } n \in \mathbb{N}.$$

Then the following assertions are equivalent.

(i) There is a dense subset $Y \subset D(A)$ such that $(\lambda - A)Y$ is dense in X, and for all $f \in Y$ there is a sequence $f_n \in D(A_n)$ satisfying $\|f_n - P_n f\|_{X_n} \to 0$ for which

$$\|A_n f_n - P_n A f\|_{X_n} \to 0 \quad \text{as } n \to \infty. \tag{3.6}$$

(ii) $\|R(\lambda, A_n)P_n f - P_n R(\lambda, A) f\|_{X_n} \to 0$ as $n \to \infty$ for all $f \in X$.

Proof. (i) ⇒ (ii): It is sufficient to show the convergence for the dense subspace $(\lambda - A)Y$. So let us take $f \in Y$ and $g := (\lambda - A)f$. By assumption, we can choose a sequence $f_n \in D(A_n)$ so that $(f_n - P_n f) \to 0$ and $(A_n f_n - P_n A f) \to 0$, hence

$$g_n := (\lambda - A)g_n$$

satisfies $(g_n - P_n g) \to 0$. Therefore, we obtain

$$\begin{aligned}\|R(\lambda, A_n)P_n g - P_n R(\lambda, A)g\| &\le \|R(\lambda, A_n)P_n g - R(\lambda, A_n)g_n\| + \|R(\lambda, A_n)g_n - P_n R(\lambda, A)g\| \\ &\le \|R(\lambda, A_n)\| \cdot \|P_n g - g_n\| + \|f_n - f\| \to 0 \qquad \text{as } n \to \infty.\end{aligned}$$

(ii) ⇒ (i): We set $Y := D(A)$. For given $f \in D(A)$ let $g := (\lambda - A)f$ and $f_n := R(\lambda, A_n)P_n g$. We can see that

$$A_n f_n = A_n R(\lambda, A_n)P_n g = \lambda R(\lambda, A_n)P_n g - P_n g,$$

and

$$P_n A f = P_n A R(\lambda, A)g = \lambda P_n R(\lambda, A)g - P_n g.$$

Hence, from the assumption it follows

$$(A_n f_n - P_n A f) \to 0.$$

□

Now we can show that, assuming stability and using ideas presented in the previous section, this convergence of the generators is equivalent to the (strong) convergence of the approximating semigroups. Though unnecessary, for the sake of completeness we formulate as appears in textbooks.

Theorem 3.14 (First Trotter–Kato Approximation Theorem)**.** *Suppose that Assumption 3.2 is satisfied and that A_n, A generate strongly continuous semigroups in X_n and X, respectively, and that there are $M \geq 0$, $\omega \in \mathbb{R}$ such that the stability condition* (3.2) *holds. Then the following are equivalent.*

(i) There is a dense subspace $Y \subset D(A)$ such that there is $\lambda > 0$ with $(\lambda - A)Y$ being dense in X. Furthermore, for all $f \in Y$ there is a sequence $f_n \in D(A_n)$ satisfying

$$\|f_n - P_n f\| \to 0 \text{ and } \quad \|A_n f_n - P_n A f\| \to 0 \quad \text{as } n \to \infty. \tag{3.7}$$

(ii) $\|R(\lambda, A_n)P_n f - P_n R(\lambda, A)f\| \to 0$ as $n \to \infty$ for all $f \in X$ and some/all $\lambda > \omega$.

(iii) $\|T_n(t)P_n f - P_n T(t)f\| \to 0$ as $n \to \infty$ for all $f \in X$ uniformly for t in compact intervals.

Proof. In view of Lemma 3.13, only the equivalence of (ii) and (iii) has to be shown.

(iii) ⇒ (ii): From the integral representation of the resolvent we see that

$$\|R(\lambda, A_n)P_n f - P_n R(\lambda, A)f\| \leq \int_0^\infty \mathrm{e}^{-\lambda t}\|T(t)f - T_n(t)f\|\,\mathrm{d}t.$$

The desired convergence follows then from the Lebesgue dominated convergence theorem.

(ii) ⇒ (iii): We repeat here the arguments from the proof of Proposition 3.10 in a careful way. Then we can see how we can refine our arguments. From the uniform boundedness principle, Theorem 2.28 and from the stability condition it follows that it suffices to show strong convergence on a dense subset. Fixing $t_0 > 0$ we obtain for all $t \in [0, t_0]$ that

$$\begin{aligned}(T_n(t) - T(t))R(\lambda, A)f &= \\ = \underbrace{T_n(t)(R(\lambda, A) - R(\lambda, A_n))f} &+ R(\lambda, A_n)(T_n(t) - T(t))f + \underbrace{(R(\lambda, A_n) - R(\lambda, A))T(t)f}.\end{aligned}$$

It is clear from the stability assumption that the first and the last term of this sum converge to 0, that is,

$$\|T_n(t)(R(\lambda, A) - R(\lambda, A_n))f\| \leq M\mathrm{e}^{\omega t_0}\|R(\lambda, A) - R(\lambda, A_n)f\| \to 0,$$

and

$$\|(R(\lambda, A) - R(\lambda, A_n))T(t)f\| \leq \|(R(\lambda, A) - R(\lambda, A_n))T(t)f\| \to 0$$

as $n \to \infty$. Note that, since the set $\{T(t)f : t \in [0, t_0]\}$ is compact, the second term converges uniformly (i.e., independently of t) to zero, see Proposition 2.30.

Hence, we have to concentrate on the middle term. Here, repeating previous arguments from the proof of Proposition 3.10, we obtain that for all $h \in X$

$$\|R(\lambda, A_n)(T(t) - T_n(t))R(\lambda, A)h\| \leq \int_0^t M\mathrm{e}^{\omega(t-s)} \cdot \|(R(\lambda, A) - R(\lambda, A_n))T(s)h\|\,\mathrm{d}s.$$

Observe again that the set $\{T(t)h : t \in [0, t_0]\}$ is compact and hence the integrand converges uniformly in $s \in [0, t_0]$. Thus, we obtain that the middle term also converges to zero in the desired way. □

3.4 Exercises

1. Prove the exponential estimate, the stability condition, from Remark 3.1.

2. Consider the operators J_n and P_n in Example 3.4 and show that for each $f \in X$, $J_n P_n f \to f$, i.e., for each $f \in \mathrm{C}_{(0)}([0,1])$ we have

$$\sum_{k=0}^{n-1} f(\tfrac{k}{n}) B_{n,k}(x) \to f(x)$$

as $n \to \infty$, uniformly in $x \in [0,1]$.

3. Let $X := L^1(0,1)$, $X_n = \mathbb{C}^n$, and define the operators

$$J_n(y_1, \ldots, y_n) := \sum_{k=1}^{n} y_k \cdot \chi_{[(k-1)/n, k/n]},$$

$$(P_n f)_k := n \cdot \int_{\frac{k-1}{n}}^{\frac{k}{n}} f(x) dx,$$

and the norm

$$\|(y_k)\|_n := \frac{1}{n} \sum_{k=1}^{n} |y_k|$$

for $(y_k) \in X_n$. Here χ stands for the characteristic function of a set. Prove that this scheme satisfies the conditions of Assumptions 3.2. Perform analogous calculations to Example 3.7.

4. Finish the proof of Proposition 3.10.

5. Solve the exercises in Appendix A.