Lecture 11

Operator Splitting with Discretisations

We continue the study of operator splitting procedures and present two more topics concerning these.

In Lecture 10 we investigated the convergence of the splitting procedures in the case when the sub-problems are solved *exactly*. In concrete problems, however, the exact solutions are usually not known. Therefore the use of certain approximation schemes is needed to solve the sub-problems. When a partial differential equation is to be solved by applying operator splitting together with approximation schemes, one usually follows the next steps:

1. The spatial differential operator is split into sub-operators of simpler form.
2. Each sub-operator is approximated by an appropriate space discretisation method (called semi-discretisation). Then we obtain systems of ordinary differential equations corresponding to the sub-operators.
3. Each solution of the semi-discretised system is obtained by using a time discretisation method.

In this lecture we investigate the cases where the solutions of the sub-problems are approximated by using only a time discretisation method or only a space discretisation method. More general cases, where all the steps 1, 2, and 3 are considered, will be investigated in the project phase of this seminar.

As in Lecture 10, we consider the abstract Cauchy problem on the Banach space X of the form

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}u(t) = (A+B)u(t), & t > 0 \\ u(0) = u_0 \end{cases} \tag{11.1}$$

with densely defined, closed linear operators A and B. Throughout the lecture we suppose that the closure C of $A+B$ with domain $D(A) \cap D(B)$ is a generator, and also that the initial value u_0 is taken from $D(A) \cap D(B)$.

Although we started our study of splitting procedures by giving the operators A and B explicitly in the abstract Cauchy problem (11.1), in real-life applications the sum operator C is given. The natural question arises how to split the operator C into the sub-operators A and B (cf. step 1 above). In practice, there are several ways to do this:

a) First we discretise the operator C in space, and split the matrix appearing in the semi-discretised problem (according to some of its structural properties).

b) The operator C describes the combined effect of several different phenomena and is already written as a sum of the corresponding sub-operators.

c) We split the operator according to the space directions (**dimension splitting**).[1]

[1] Sometimes also called as coordinate splitting.

Procedure a) leads to the matrix case which was already presented in Section 10.1, and b) was investigated in Section 10.2. First, we develop some abstract results that are applicable for the dimension splitting. In contrast to the results presented in Section 10.2 (when operator B was bounded), in this case we will have two unbounded operators.

11.1 Resolvent splittings

We turn our attention to the resolvent splittings which have already been introduced in Definition 10.1:

$$\begin{aligned} &\text{Lie splitting:} && F_{\text{Lie}}(h) = \big(I - hB\big)^{-1}\big(I - hA\big)^{-1}, \\ &\text{Peaceman–Rachford splitting:} && F_{\text{PR}}(h) = \big(I - \tfrac{h}{2}B\big)^{-1}\big(I + \tfrac{h}{2}A\big)\big(I - \tfrac{h}{2}A\big)^{-1}\big(I + \tfrac{h}{2}B\big). \end{aligned}$$

We note that for closed and linear operators A, B , there appear the corresponding resolvents in $F_{\text{Lie}}(h)$ and $F_{\text{PR}}(h)$. We only have to assume that $\frac{1}{h}$ and $\frac{2}{h}$ belong to the resolvent sets of both A and B for the Lie and the Peaceman–Rachford splittings, respectively. We are interested in, however, the convergence of the splitting procedures. Since h has to be small then, we may suppose without loss of generality that $\frac{1}{h}$ (and therefore $\frac{2}{h}$) is large enough.

Observe that the terms in Lie and Peaceman–Rachford splittings correspond to the explicit and implicit Euler methods with time steps h or $\frac{h}{2}$. Thus, the resolvent splittings can be considered as the application of operator splitting together with special time discretisation methods.

We prove next the first-order convergence of the Lie splitting following the idea presented by Hansen and Ostermann.[2]

Theorem 11.1. *Let A and B be linear operators, and suppose that there is a $\lambda_0 > 0$ such that $\lambda \in \rho(A) \cap \rho(B)$ for all $\lambda \geq \lambda_0$ (i.e., the Lie splitting is well-defined), and that there exist constants $M \geq 1$ and $\omega \in \mathbb{R}$ such that*

$$\big\|\big(F_{\text{Lie}}(h)\big)^k\big\| \leq M\mathrm{e}^{k\omega h} \tag{11.2}$$

holds for all $k \in \mathbb{N}$ and $h \in [0, \frac{1}{\lambda_0}]$ (i.e., the Lie splitting is stable). Suppose further that the closure C of $A + B$ (with $A + B$ having the natural domain $D(A) \cap D(B)$) generates a strongly continuous semigroup on the Banach space X of type (M, ω), and that $D(C^2) \subseteq D(AB) \cap D(A)$ and $AB(\lambda_0 - C)^{-2} \in \mathscr{L}(X)$ hold. Then the Lie splitting is first-order convergent on $D(C^2)$. That is to say, for all $t_0 \geq 0$ we have

$$\|u_{\text{Lie},h}(nh) - u(nh)\| \leq hK\big(\|u_0\| + \|Cu_0\| + \|C^2 u_0\|\big)$$

for all $u_0 \in D(C^2)$ and $nh \in [0, t_0]$, $n \in \mathbb{N}$, $h \in [0, h_0]$, where the constant K may depend on t_0, but not on n and h.

Proof. For better readability we denote the semigroup operators by e^{tC}. By using the definition of the split solution, $u_{\text{Lie},h}(nh) = F_{\text{Lie}}(h)^n u_0$, and the telescopic identity, the error term can be rewritten as

$$u_{\text{Lie},h}(nh) - u(nh) = F_{\text{Lie}}(h)^n u_0 - \mathrm{e}^{nhC}u_0 = \sum_{j=0}^{n-1} F_{\text{Lie}}(h)^{n-j-1}\big(F_{\text{Lie}}(h) - \mathrm{e}^{hC}\big)\mathrm{e}^{jhC}u_0. \tag{11.3}$$

[2] E. Hansen and A. Ostermann, "Dimension splitting for evolution equations," Numer. Math. **108** (2008), 557–570.

Next we estimate the local error, i.e., the term $F_{\text{Lie}}(h) - \mathrm{e}^{hC}$ in the expression above. To do that we need to introduce some auxiliary functions, which will also come handy in later lectures.

For all $h > 0$ and $j \in \mathbb{N}$ we define the bounded linear operators $\varphi_j(hC)$ by

$$\varphi_j(hC)f := \frac{1}{h^j}\int_0^h \frac{\tau^{j-1}}{(j-1)!}\mathrm{e}^{(h-\tau)C}f\,\mathrm{d}\tau, \tag{11.4}$$

for all $f \in X$ and $\varphi_0(hC) = \mathrm{e}^{hC}$. These operators are uniformly bounded for $h \in (0, t_0]$. Indeed, this is trivial for $\varphi_0(jC)$, whereas for $j \geq 1$ we have

$$\begin{aligned}\|\varphi_j(hC)f\| &\leq \frac{1}{h^j}\int_0^h \frac{\tau^{j-1}}{(j-1)!}\big\|\mathrm{e}^{(h-\tau)C}f\big\|\mathrm{d}\tau \leq \frac{1}{h^j(j-1)!}\|f\|\int_0^h r^{j-1}M\mathrm{e}^{\omega(h-r)}\mathrm{d}r\\ &\leq \frac{1}{h^j(j-1)!}M_C\mathrm{e}^{\max\{0,\omega\}h}\|f\|\int_0^h r^{j-1}\mathrm{d}r \leq \frac{1}{j!}M\mathrm{e}^{\max\{0,\omega\}t_0}\|f\| = \text{const.}\cdot\|f\|.\end{aligned}$$

Moreover, the operators $\varphi_j(hC)$ satisfy the recurrence relation

$$\varphi_j(hC)f = \tfrac{1}{j!}f + hC\varphi_{j+1}(hC)f \tag{11.5}$$

for all $j = 0, 1, 2, ...$ and $f \in X$, see Exercise 1. They also leave $D(C)$ and $D(C^2)$ invariant. In the rest of the proof, we shall use only the following two consequences of (11.5):

$$\big(I - \varphi_0(hC)\big)f = -hC\varphi_1(hC)f \quad \text{and} \quad \big(\varphi_1(hC) - \varphi_0(hC)\big)f = hC\big(\varphi_2(hC) - \varphi_1(hC)\big)f.$$

For $f \in D(C^2)$ we shall derive a form for the local error $F_{\text{Lie}}(h)f - \mathrm{e}^{hC}f$ that allows for the appropriate estimates. So we take $f \in D(C^2) \subset D(A) \cap D(B)$. For the sake of brevity we introduce the following abbreviations for the resolvents:

$$R_A = R(\tfrac{1}{h}, A) \qquad \text{and} \qquad R_B = R(\tfrac{1}{h}, B),$$

then we have

$$F_{\text{Lie}}(h) = \tfrac{1}{h^2}R_BR_A.$$

By using the identity $I = \lambda R(\lambda, A) - AR(\lambda, A)$ for all $\lambda \in \rho(A)$, i.e. $I = \frac{1}{h}R_A - AR_A$ in our case, we express now the local error, i.e., the middle term in the telescopic sum (11.3) in the following form:

$$\begin{aligned}\big(F_{\text{Lie}}(h) - \mathrm{e}^{hC}\big)f &= \big(F_{\text{Lie}}(h) - \varphi_0(hC)\big)f\\ &= F_{\text{Lie}}(h)f - \big(\tfrac{1}{h}R_B - BR_B\big)\big(\tfrac{1}{h}R_A - AR_A\big)\varphi_0(hC)f\\ &= F_{\text{Lie}}(h)f - \big(\tfrac{1}{h^2}R_BR_A - \tfrac{1}{h}BR_BR_A - \tfrac{1}{h}R_BAR_A + BR_BAR_A\big)\varphi_0(hC)f\\ &= F_{\text{Lie}}(h)\big(I - \varphi_0(hC)\big)f + \big(\tfrac{1}{h}BR_BR_A + \tfrac{1}{h}R_BAR_A - BR_BAR_A\big)\varphi_0(hC)f.\end{aligned}$$

Since $f \in D(C^2) \subseteq D(A)$, we can write

$$\begin{aligned}&\big(F_{\text{Lie}}(h) - \mathrm{e}^{hC}\big)f\\ &\quad= F_{\text{Lie}}(h)\big(I - \varphi_0(hC)\big)f + \big(hBF_{\text{Lie}}(h) + hF_{\text{Lie}}(h)A - h^2BF_{\text{Lie}}(h)A\big)\varphi_0(hC)f.\end{aligned} \tag{11.6}$$

Observe that for every $f \in D(C^2) \subseteq D(A)$ we have the following relation:

$$\begin{aligned}\big(hBF_{\mathrm{Lie}}(h) - h^2BF_{\mathrm{Lie}}(h)A\big)f &= hBF_{\mathrm{Lie}}(h)(I-hA)f\\ &= hB(I-hB)^{-1}(I-hA)^{-1}(I-hA)f = hB(I-hB)^{-1}f\\ &= BR_Bf.\end{aligned}$$

For $f \in D(C^2) \subseteq D(B)$ we can rewrite this as:

$$\begin{aligned}\big(hBF_{\mathrm{Lie}}(h) - h^2BF_{\mathrm{Lie}}(h)A\big)f &= BR_Bf\\ &= R_BBf = R_B(\tfrac{1}{h}R_A - AR_A)Bf = \tfrac{1}{h}R_BR_ABf - R_BAR_ABf\\ &= \tfrac{1}{h}R_BR_ABf - R_BR_AABf = hF_{\mathrm{Lie}}(h)Bf - h^2F_{\mathrm{Lie}}(h)ABf.\end{aligned}$$

By inserting this last expression into (11.6), we obtain

$$\big(F_{\mathrm{Lie}}(h) - \mathrm{e}^{hC}\big)f = F_{\mathrm{Lie}}(h)\big(I - \varphi_0(hC)\big)f + hF_{\mathrm{Lie}}(h)(A+B)\varphi_0(hC)f - h^2F_{\mathrm{Lie}}(h)AB\varphi_0(hC)f.$$

By using the identity $\big(I - \varphi_0(hC)\big)f = -hC\varphi_1(hC)f$ we obtain

$$\begin{aligned}\big(F_{\mathrm{Lie}}(h) - \mathrm{e}^{hC}\big)f &= -F_{\mathrm{Lie}}(h)hC\varphi_1(hC)f + hF_{\mathrm{Lie}}(h)C\varphi_0(hC)f - h^2F_{\mathrm{Lie}}(h)AB\varphi_0(hC)f\\ &= F_{\mathrm{Lie}}(h)hC\big(\varphi_0(hC) - \varphi_1(hC)\big)f - h^2F_{\mathrm{Lie}}(h)AB\varphi_0(hC)f.\end{aligned}$$

On the other hand, we have $\big(\varphi_0(hC) - \varphi_1(hC)\big)f = hC\big(\varphi_1(hC) - \varphi_2(hC)\big)f$, and therefore

$$\big(F_{\mathrm{Lie}}(h) - \mathrm{e}^{hC}\big)f = F_{\mathrm{Lie}}(h)h^2C^2\big(\varphi_1(hC) - \varphi_2(hC)\big)f - h^2F_{\mathrm{Lie}}(h)AB\varphi_0(hC)f$$

for all $f \in D(C^2)$.

We now return to the error term. By using the equality above for $f = \mathrm{e}^{jhC}u_0$, and that

$$\varphi_0(hC)f = \mathrm{e}^{hC}f = (\lambda_0 - C)^{-2}\mathrm{e}^{hC}(\lambda_0 - C)^2 f$$

holds for all $f \in D(C^2)$, we obtain the following expression for the error term in (11.3):

$$\begin{aligned}&F_{\mathrm{Lie}}(h)^n u_0 - \mathrm{e}^{nhC}u_0\\ &\quad= h^2\sum_{j=0}^{n-1} F_{\mathrm{Lie}}(h)^{n-j}\Big(\big(\varphi_1(hC) - \varphi_2(hC)\big)C^2 - AB(\lambda_0 - C)^{-2}\mathrm{e}^{hC}(\lambda_0 - C)^2\Big)\,\mathrm{e}^{jhC}u_0\\ &\quad= h^2\sum_{j=0}^{n-1} F_{\mathrm{Lie}}(h)^{n-j}\Big(\big(\varphi_1(hC) - \varphi_2(hC)\big)\mathrm{e}^{jhC}C^2 - AB(\lambda_0 - C)^{-2}\mathrm{e}^{(j+1)hC}(\lambda_0 - C)^2\Big)\,u_0\end{aligned}$$

for all $u_0 \in D(C^2)$. Since the terms $\varphi_1(hC)$, $\varphi_2(hC)$, and e^{hC} are uniformly bounded for $h \in (0, t_0]$ and since by assumption the operator $AB(\lambda_0 - C)^{-2}$ is bounded, we obtain the desired estimate:

$$\begin{aligned}\big\|F_{\mathrm{Lie}}(h)^n u_0 - \mathrm{e}^{nhC}u_0\big\| &\le h^2\sum_{j=0}^{n-1}\|F_{\mathrm{Lie}}(h)^{n-j}\|\,\Big((\|\varphi_1(hC)\| + \|\varphi_2(hC)\|)\cdot\big\|\mathrm{e}^{hC}\big\|^j\cdot\|C^2u_0\|\\ &\qquad+\|AB(\lambda_0 - C)^{-2}\|\cdot\big\|\mathrm{e}^{hC}\big\|^{j+1}\cdot\|(\lambda_0^2 - 2\lambda_0C + C^2)u_0\|\Big)\\ &\le h^2nM\mathrm{e}^{\max\{0,\omega\}t}\Big(\mathrm{const.}\cdot M\mathrm{e}^{\max\{0,\omega\}t}\cdot\|C^2u_0\|\\ &\qquad+\mathrm{const.}\cdot M\mathrm{e}^{\max\{0,\omega\}t}\big(\|u_0\| + \|Cu_0\| + \|C^2u_0\|\big)\Big)\\ &\le h^2n\widetilde{K}\left(\|u_0\| + \|Cu_0\| + \|C^2u_0\|\right) = hK\left(\|u_0\| + \|Cu_0\| + \|C^2u_0\|\right)\end{aligned}$$

with a positive constant K depending on t_0. This completes the proof. □

Remark 11.2. The stability condition (11.2) in Theorem 11.1 is satisfied if

$$\big\|(I-hA)^{-1}\big\|\le 1\quad\text{and}\quad\big\|(I-hB)^{-1}\big\|\le 1$$

hold for all $h>0$. In this case, by the Lumer–Phillips theorem, Theorem 6.3 both A and B generate contraction semigroups on X. As a consequence of the convergence result above the operator C generates a contraction semigroup, too.

Following the idea of the proof above, the second-order convergence of the Peaceman–Rachford splitting can also be shown (see the already mentioned paper of Hansen and Ostermann).

Theorem 11.3. *Let A and B be linear operators, and suppose that there is a $\lambda_0>0$ such that $\lambda\in\rho(A)\cap\rho(B)$ for all $\lambda\ge\lambda_0$ (i.e., the Peaceman–Rachford splitting is well-defined on $D(B)$), and that there exist constants $M\ge 1$ and $\omega\in\mathbb{R}$ such that*

$$\big\|\big(F_{\mathrm{PR}}(h)\big)^k(I-\tfrac{h}{2}B)^{-1}\big\|\le M\mathrm{e}^{k\omega h}\tag{11.7}$$

holds for all $k\in\mathbb{N}$ and $h\in[0,\frac{1}{\lambda_0}]$ (i.e., the Peaceman–Rachford splitting is stable). Suppose further that the closure C of $A+B$ (with $A+B$ having the natural domain $D(A)\cap D(B)$) generates a strongly continuous semigroup on the Banach space X, and that $D(C^2)\subseteq D(AB)\cap D(A)$ and $AB(\lambda_0-C)^{-2}\in\mathscr{L}(X)$ hold. Then the Peaceman–Rachford splitting is second-order convergent on $D(C^3)$. That is, for all $t_0\ge 0$ we have

$$\|u_{\mathrm{PR},h}(t)-u(t)\|\le h^2K\sum_{j=0}^{3}\|C^ju_0\|$$

for all $u_0\in D(C^3)$ and $nh\in[0,t_0]$, $n\in\mathbb{N}$, $h\in[0,\frac{1}{\lambda_0}]$, where the constant K depends on t_0, but not on n and t.

Remark 11.4. The stability condition in Theorem 11.3 is satisfied for example if A and B generate contraction semigroups, and $X=H$ is a Hilbert space. Indeed, in this case we have

$$\big\|(I-hA)^{-1}\big\|\le 1\quad\text{and}\quad\big\|(I-hB)^{-1}\big\|\le 1\quad\text{for all } h>0,$$

and one can prove that also

$$\big\|(I+\tfrac{h}{2}A)(I-\tfrac{h}{2}A)^{-1}\big\|\le 1\quad\text{and}\quad\big\|(I+\tfrac{h}{2}B)(I-\tfrac{h}{2}B)^{-1}\big\|\le 1$$

hold for all $h>0$, see Exercise 4.

As an illustration we give an example of operators A,B,C that satisfy the conditions of Theorem 11.1, hence, for which the Lie splitting is first-order convergent. More details and examples are left to the project phase.

Example 11.5 (**Dimension splitting**)**.** Consider the heat equation in two dimensions on $\Omega=(0,1)\times(0,1)$ with homogeneous Dirichlet boundary condition:

$$\begin{cases}\partial_t w(t,x,y)=\partial_x\big(a(x,y)\partial_x w(t,x,y)\big)+\partial_y\big(b(x,y)\partial_y w(t,x,y)\big), & (x,y)\in\Omega,\ t>0\\ w(0,x,y)=w_0(x,y), & (x,y)\in\Omega\\ w(t,x,y)=0, & (x,y)\in\partial\Omega,\ t>0\end{cases}\tag{11.8}$$

with some given initial function w_0 and functions $a, b \in \mathrm{C}^2(\overline{\Omega})$ being positive on $\overline{\Omega}$. Problem (11.8) can be formulated as an abstract Cauchy problem on the Banach space $X = \mathrm{L}^2(\Omega)$:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} u(t) = Cu(t), \quad t > 0 \\ u(0) = u_0 \end{cases}$$

where the operator equals $Cf = \partial_x(a\partial_x f) + \partial_y(b\partial_y f)$ for all $f \in D(C) = \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$ and generates an analytic semigroup on X. First of all, we note that by the Sobolev embedding theorem[3] $\mathrm{H}^2(\Omega)$ is continuously embedded in $\mathrm{C}(\overline{\Omega})$ and $\mathrm{H}^4(\Omega)$ is continuously embedded in $\mathrm{C}^2(\overline{\Omega})$. These imply that the boundary conditions can be verified pointwise.

Now we split the operator according to the space directions, i.e., $C = A + B$ with

$$Af = \partial_x(a\partial_x f) \qquad \text{and} \qquad Bf = \partial_y(b\partial_y f)$$

for all $f \in D(C)$. In this special case the domains can be chosen as:[4]

$$D(A) = \big\{ f \in X : \partial_{xx} f, \partial_x f \in X, \text{ and } f(0,y) = f(1,y) = 0 \text{ for almost every } y \in (0,1) \big\}$$
$$\text{and } \; D(B) = \big\{ f \in X : \partial_{yy} f, \partial_y f \in X, \text{ and } f(x,0) = f(x,1) = 0 \text{ for almost every } x \in (0,1) \big\}.$$

Then operators A, B with the corresponding domains generate analytic semigroups as well.

We have to verify now the assumptions of Theorem 11.1. The operators A, B and $A+B$ generate analytic contraction semigroups on X. To prove the domain condition, we have to show that for all $f \in D(C^2)$ we have $Bf \in D(A)$. For $f \in D(C^2) \subseteq \mathrm{H}^4(\Omega)$ we have $Bf \in \mathrm{H}^2(\Omega)$ and $f = Cf = 0$ on the boundary $\partial\Omega$ of Ω. Then $\partial_x(Bf), \partial_{xx}(Bf) \in X$. Since $f = 0$ on the boundary, its first and second weak tangential derivatives equal zero on $\partial\Omega$. Then from the continuity of Bf it follows that on the horizontal lines of $\partial\Omega$ we have $Bf = \partial_y(b\partial_y f) = Cf - \partial_x(a\partial_x f) = (\partial_x a)(\partial_x f) + a\partial_{xx} f = 0 - 0 = 0$, and on the vertical lines $Bf = \partial_y(b\partial_y f) = (\partial_y b)(\partial_y f) + b\partial_{yy} f = 0$. This yields $X \ni Bf = 0$ on $\partial\Omega$, therefore, $Bf \in D(A)$.

The space $Y := D(C^2)$ equipped with the H^4-norm becomes a Banach space, see Exercise 2. The boundedness of operator $AB(I-C)^{-2}$ follows now from the estimate

$$\big\| AB(I-C)^{-2} \big\| \le \|AB\|_{\mathscr{L}(Y,X)} \cdot \big\| (I-C)^{-2} \big\|_{\mathscr{L}(X,Y)}. \tag{11.9}$$

Indeed, for $f \in Y = D(C^2)$ we have

$$\|ABf\| = \big\| \partial_x \big(a\partial_x \big(\partial_y (b\partial_y f) \big) \big) f \big\| \le \text{const.} \cdot \|f\|_Y.$$

On the other hand for $f \in Y = D(C^2)$ we have

$$\begin{aligned} \|(I-C)^2 f\| &\le \|(I - 2C + C^2) f\| \le \|f\| + 2\|Cf\| + \|C^2 f\| \le \text{const.} \cdot (\|f\| + \|Cf\| + \|C^2 f\|) \\ &\le \text{const.} \cdot \|f\|_Y. \end{aligned}$$

Thus, we have $(I-C)^2 \in \mathscr{L}(Y,X)$ and $(I-C)^2$ is closed. By Proposition 2.10 $(I-C)^{-2}$ is closed, too, and by the closed graph theorem it is bounded, i.e., $(I-C)^{-2} \in \mathscr{L}(X,Y)$. Then estimate (11.9) yields the desired boundedness.

[3] See, e.g., Theorem 4.12 in R. A. Adams, J. J. F. Fournier, Sobolev Spaces, Elsevier, 2003.

[4] A. Ostermann, K. Schratz, "Stability of exponential operator splitting methods for non-contractive semigroups," preprint.

11.2 Operator splitting with space discretisation

In this section we consider operator splittings applied together with a space discretisation method (steps 1 and 2). That is, we assume that the semigroups generated by the sub-operators A, B are approximated by some approximate semigroups.

Consider the abstract Cauchy problem (11.1) on the Banach space X for the sum of the generators A and B. As we did in Assumption 3.2, we define approximate spaces and projection-like operators between the approximate spaces and the original space X.

Assumption 11.6. Let X_m, X be Banach spaces and assume that there are bounded linear operators $P_m : X \to X_m$, $J_m : X_m \to X$ with the following properties:

a) There is a constant $K > 0$ with $\|P_m\|$, $\|J_m\| \leq K$ for all $m \in \mathbb{N}$,

b) $P_m J_m = I_m$, the identity operator on X_m, and

c) $J_m P_m f \to f$ as $m \to \infty$ for all $f \in X$.

As already illustrated in Examples 3.3 and 3.4, the operators P_m together with the spaces X_m usually refer to a kind of space discretisation (cf. Appendix A), the spaces X_m are usually finite dimensional spaces, and the operators J_m refer to the interpolation method how we associate specific elements of the function space to the elements of the approximating spaces.

First we split the operator $C = A + B$ appearing in the original problem (11.1) into the sub-operators A and B. In order to obtain the semi-discretised systems, the sub-operators A and B have to be approximated by operators A_m and B_m for $m \in \mathbb{N}$ fixed. Suppose that the operators A_m and B_m generate the strongly continuous semigroups T_m and S_m on the space X_m, respectively. For the analysis of the convergence, we need to recall the following from Lecture 3.

Assumption 11.7. Suppose that for $m \in \mathbb{N}$ the semigroups T_m and S_m and their generators A_m, B_m satisfy the following conditions:

a) there exist constants $M \geq 1$ and $\omega \in \mathbb{R}$ such that T_m and S_m are all of type (M, ω), and for all $h > 0$, $k, m \in \mathbb{N}$ we have
$$\left\| \left(S_m(h) T_m(h)\right)^k \right\| \leq M \mathrm{e}^{k\omega h}. \tag{11.10}$$

b) We have

$$\lim_{m\to\infty} J_m A_m P_m f = Af \qquad \text{for all } f \in D(A)$$

and
$$\lim_{m\to\infty} J_m B_m P_m f = Bf \qquad \text{for all } f \in D(B).$$

The semigroups T_m, S_m, $m \in \mathbb{N}$, are called **approximate semigroups**, and their generators A_m, B_m, $m \in \mathbb{N}$, are called **approximate generators** if they possess the above properties.

Remark 11.8. From the assumption above and from the first Trotter–Kato approximation theorem, Theorem 3.14, it follows that

$$\lim_{m\to\infty} J_m T_m(h) P_m f = T(h) f$$

and
$$\lim_{m\to\infty} J_m S_m(h) P_m f = S(h) f$$

for all $f \in X$ and locally uniformly in h, where T and S are the semigroups generated by A and B, respectively.

From now on we consider the exponential splittings, that is, the sequential and Marchuk–Strang splittings. We remark that, analogously to the case of exact solutions (i.e., splitting without approximation), the stability condition (11.10) implies the stability of the reversed order and the Marchuk–Strang splittings (cf. Exercise 10.5). This means that the sequential and the Marchuk–Strang splittings fulfill their stability condition (with space discretisation) if the stability condition (11.10) holds. Therefore, it suffices to control only this condition in both cases.

Definition 11.9. For the case of spatial approximation, we define the split solutions of (11.1) as

$$u_{\mathrm{spl},n,m}(t) = J_m\big(F_{\mathrm{spl},m}(h)\big)^n P_m u_0, \quad (t = nh)$$

for $m, n \in \mathbb{N}$ fixed and for $u_0 \in X$. The operators $F_{\mathrm{seq},m}$, describing the splitting procedures together with the space discretisation method, have the form (cf. Definition 10.1):

Sequential splitting: $F_{\mathrm{seq},m}(h) = S_m(h)T_m(h)$,

Marchuk–Strang splitting: $F_{\mathrm{MS},m}(h) = T_m(h/2)S_m(h)T_m(h/2)$.

Definition 11.10. The numerical method for solving problem (11.1) described above is **convergent at a fixed time level** $t > 0$ if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have

$$\|u_{\mathrm{spl},n,m}(t) - u(t)\| \leq \varepsilon.$$

This is the usual well-known notion of the convergence of a sequence with two indices and we will use the notation

$$\lim_{n,m\to\infty} u_{\mathrm{spl},n,m}(t) = u(t)$$

to express this.

In order to prove the convergence of operator splitting in this case, we state a modified version of Chernoff's theorem, Theorem 5.12, which is applicable for approximate semigroups as well.

Theorem 11.11 (Modified Chernoff Theorem)**.** *Consider a sequence of strongly continuous functions* $F_m : [0, \infty) \to \mathscr{L}(X_m)$, $m \in \mathbb{N}$, *satisfying*

$$F_m(0) = I_m \tag{11.11}$$

for all $m \in \mathbb{N}$, *and suppose that there exists constants* $M \geq 1$, $\omega \in \mathbb{R}$ *such that*

$$\big\|\big(F_m(t)\big)^k\big\| \leq M\mathrm{e}^{k\omega t} \tag{11.12}$$

holds for all $t \geq 0$ *and* $m, k \in \mathbb{N}$. *Suppose further that the limit*

$$\lim_{m\to\infty} \frac{J_m F_m(h) P_m f - J_m P_m f}{h}$$

exists uniformly in $h \in (0, t_0]$, *and that*

$$Gf := \lim_{h\searrow 0} \lim_{m\to\infty} \frac{J_m F_m(h) P_m f - J_m P_m f}{h} \tag{11.13}$$

exists for all $f \in Y \subset X$, *where* Y *and* $(\lambda_0 - G)Y$ *are dense subspaces in* X *for* $\lambda_0 > 0$. *Then the closure* $C = \overline{G}$ *of* G *generates a bounded strongly continuous semigroup* U, *which is given by*

$$U(t)f = \lim_{n,m\to\infty} J_m\big(F_m(h)\big)^n P_m f \tag{11.14}$$

for all $f \in X$ *uniformly for* t *in compact intervals* $(t = nh)$*.*

Proof. For $h > 0$ we define

$$G_{h,m} := \frac{F_m(h) - I_m}{h} \in \mathscr{L}(X_m)$$

for all fixed $h \in (0, t_0]$ and $m \in \mathbb{N}$. Observe that for all $f \in Y$ we have

$$\lim_{h \searrow 0} \lim_{m \to \infty} J_m G_{h,m} P_m f = Gf.$$

Then every semigroup $(\mathrm{e}^{tG_{h,m}})_{t \geq 0}$ satisfies

$$\left\| \mathrm{e}^{tG_{h,m}} \right\| \leq \mathrm{e}^{-\frac{t}{h}} \left\| \mathrm{e}^{\frac{t}{h} F_m(h)} \right\| \leq \mathrm{e}^{-\frac{t}{h}} \sum_{k=0}^{\infty} \frac{t^k}{h^k k!} \left\| \left(F_m(h)\right)^k \right\| \leq M \mathrm{e}^{\omega' t} \tag{11.15}$$

for some $\omega' \in \mathbb{R}$ and for every fixed h and m. This shows that the assumptions of the first Trotter–Kato approximation theorem, Theorem 3.14, are fulfilled. Hence we can take the limit in $m \to \infty$ (which is uniform in $h \in (0, t_0]$), and then take limit as $h \to 0$ obtaining that the closure $\overline{G}$ of G generates a strongly continuous semigroup U given by

$$\lim_{h \searrow 0} \lim_{m \to \infty} \left\| U(t) f - J_m \mathrm{e}^{tG_{h,m}} P_m f \right\| = 0 \tag{11.16}$$

for all $f \in X$ and uniformly for t in compact intervals. On the other hand, we have for $t = nh$ by Lemma 5.7 that

$$\begin{aligned} \left\| J_m \mathrm{e}^{tG_{h,m}} P_m f - J_m \left(F_m(h)\right)^n P_m f \right\| &= \left\| J_m \mathrm{e}^{n(F_m(h) - I_m)} P_m f - J_m \left(F_m(h)\right)^n P_m f \right\| \\ &\leq \sqrt{n} M \| J_m F_m(h) P_m f - J_m P_m f \| = \frac{\sqrt{n}}{h} M \left\| \frac{J_m F_m(h) P_m f - J_m P_m f}{h} \right\|. \end{aligned} \tag{11.17}$$

Using that $h = \frac{t}{n}$ for some $t \in [0, t_0]$ and taking the limit, we obtain

$$\lim_{h \searrow 0} \lim_{m \to \infty} \left\| J_m \mathrm{e}^{tG_{h,m}} P_m f - J_m \left(F_m(h)\right)^n P_m f \right\| = \lim_{h \searrow 0} \lim_{m \to \infty} \frac{tM}{\sqrt{n}} \left\| \frac{J_m F_m(h) P_m f - J_m P_m f}{h} \right\| = 0 \tag{11.18}$$

for all $f \in Y$, and uniformly for t in compact intervals. The combination of (11.16) and (11.18) yields

$$\begin{aligned} &\lim_{h \searrow 0} \lim_{m \to \infty} \left\| U(t) f - J_m \left(F_m(h)\right)^n P_m f \right\| \\ &\quad \leq \lim_{h \searrow 0} \lim_{m \to \infty} \left\| U(t) f - J_m \mathrm{e}^{tG_{h,m}} P_m f \right\| + \lim_{h \searrow 0} \lim_{m \to \infty} \left\| J_m \mathrm{e}^{tG_{h,m}} P_m f - J_m \left(F_m(h)\right)^n P_m f \right\| = 0 \end{aligned}$$

$(t = nh)$ for all $f \in Y$, and uniformly for t in compact intervals. By Theorem 2.30, the statement follows for all $f \in X$. □

In the rest of this lecture, we consider the convergence of the sequential splitting applied together with a space discretisation method.

Lemma 11.12. *Let J_m, P_m, T_m be operators introduced in Assumptions 11.6 and 11.7. Then*

$$\lim_{m \to \infty} \frac{J_m T_m(h) P_m f - J_m P_m f}{h} = \tfrac{1}{h}\left(T(h) f - f\right)$$

holds for all $f \in D(A)$ uniformly in $h \in (0, t_0]$, and

$$\lim_{h \searrow 0} \tfrac{1}{h}\left(T(h) f - f\right) = Af.$$

Proof. Let us investigate the following difference for all $f \in D(A)$:

$$\left\| \frac{J_m T_m(h) P_m f - J_m P_m f}{h} - \frac{T(h)f - f}{h} \right\| = \frac{1}{h} \left\| \int_0^h J_m A_m T_m(s) P_m f \mathrm{d}s - \int_0^h AT(s) f \mathrm{d}s \right\|$$
$$\le \sup_{s\in[0,t_0]} \|J_m A_m T_m(s) P_m f - AT(s) f\| = \sup_{s\in[0,t_0]} \|J_m T_m(s) A_m P_m f - T(s) Af\|$$
$$\le \sup_{s\in[0,t_0]} \big\|J_m T_m(s) P_m (J_m A_m P_m f - Af) + \big(J_m T_m(s) P_m - T(s)\big) Af\big\|$$
$$\le \sup_{s\in[0,t_0]} \|J_m\| \cdot \|T_m(s)\| \cdot \|P_m\| \cdot \|J_m A_m P_m f - Af\| + \sup_{s\in[0,t_0]} \big\|\big(J_m T_m(s) P_m - T(s)\big) Af\big\|.$$

By Assumption 11.7, the term $\|J_m A_m P_m f - Af\|$ tends to 0 as m tends to infinity. Since $g := Af$ is a fixed element in the Banach space X, $\|J_m T_m(s) P_m g - T(s) g\|$ tends to 0 uniformly in h as $m \to \infty$ because of Remark 11.8. Operators J_m and P_m were assumed to be bounded. The semigroups T_m are of type (M, ω), independently of m. Therefore,

$$\sup_{s\in[0,t_0]} \|T_m(s)\| \le \sup_{s\in[0,t_0]} M\mathrm{e}^{\omega s} \le M\mathrm{e}^{\max\{0,\omega\} t_0} = \text{const.} < \infty.$$

Hence, the difference above tends to 0 uniformly in h. The second limit as $h \searrow 0$ can be obtained by using the definition of the generator:

$$\lim_{h\searrow 0} \frac{T(h)f - f}{h} = Af \qquad \text{for all } f \in D(A).$$

Thus, the statement is proved. □

The same result is true for the semigroup S generated by the operator B, that is,

$$\lim_{h\searrow 0} \lim_{m\to\infty} \frac{J_m S_m(h) P_m f - J_m P_m f}{h} = Bf \tag{11.19}$$

holds for all $f \in D(B)$, where the limit as $m \to \infty$ is locally uniform in h.

Theorem 11.13. *The sequential splitting is convergent at time level $t > 0$ if the stability condition* (11.10) *holds for the approximate semigroups, and the approximate generators satisfy Assumption 11.7.*

Proof. According to the modified Chernoff theorem, Theorem 11.11, the sequential splitting is convergent if the stability (11.12) and the consistency (11.13) hold for the operator

$$F_m(h) = S_m(h) T_m(h). \tag{11.20}$$

The stability condition (11.12) is fulfilled, since we assumed that (11.10) holds. In order to prove the consistency criterion (11.13), we investigate the following limit:

$$\lim_{h\searrow 0} \lim_{m\to\infty} \frac{J_m S_m(h) T_m(h) P_m f - J_m P_m f}{h}$$
$$= \lim_{h\searrow 0} \lim_{m\to\infty} J_m S_m(h) P_m \frac{J_m T_m(h) P_m f - J_m P_m f}{h}$$
$$+ \lim_{h\searrow 0} \lim_{m\to\infty} \frac{J_m S_m(h) P_m f - J_m P_m f}{h}.$$

Remark 11.8 implies

$$\lim_{m\to\infty} J_m S_m(h) P_m f = S(h)f \quad \text{for all } f \in X \text{ and uniformly for } h \in [0, t_0]$$

and
$$\lim_{h\searrow 0} S(h)f = f \quad \text{for all } f \in X.$$

Notice further that the set $\left\{\frac{1}{h}(J_m T_m(h) P_m f - J_m P_m f) : h \in (0, t_0]\right\}$ is relatively compact for all $f \in D(A)$, and that on compact sets the strong and the uniform convergence is equivalent due to Theorem 2.30. Then Lemma 11.12 and (11.19) imply that

$$\lim_{h\searrow 0} \lim_{m\to\infty} \frac{J_m F_m(h) P_m f - J_m P_m f}{h} = (A+B)f$$

holds for all $f \in D(A) \cap D(B)$ (see also the proof of Corollary 4.10). This completes the proof. □

We state now the convergence of the Marchuk–Strang splitting, and leave the proof as Exercise 5.

Theorem 11.14. *The Marchuk–Strang splitting is convergent at time level $t > 0$ if the stability condition* (11.10) *holds for the approximate semigroups, and the approximate generators satisfy Assumption 11.7.*

Exercises

1. Prove the recurrence relation (11.5).

2. Let C be the operator from Example 11.5. Prove that the H^4-norm makes $D(C^2)$ a Banach space.

3. Consider the operators A, B and C from Example 11.5 with $a = b = 1$. Show that they generate analytic contraction semigroups.

4. Suppose A generates a contraction semigroup on the Hilbert space H. Prove that the **Cayley transform**

$$G = (I + A)(I - A)^{-1}$$

of A is a contraction.

5. Prove Theorem 11.14.