MATHEMATICS OF COMPUTATION
Volume 67, Number 221, January 1998, Pages 21–44
S 0025-5718(98)00915-6

# THE TROTTER-KATO THEOREM AND APPROXIMATION OF PDEs

KAZUFUMI ITO AND FRANZ KAPPEL

ABSTRACT. We present formulations of the Trotter-Kato theorem for approximation of linear $C_0$-semigroups which provide very useful framework when convergence of numerical approximations to solutions of PDEs are studied. Applicability of our results is demonstrated using a first order hyperbolic equation, a wave equation and Stokes' equation as illustrative examples.

## 1. INTRODUCTION

In this paper versions of the Trotter-Kato theorem [8], [15] for approximating a linear $C_0$-semigroup $T(t)$ on a Banach space $X$ are derived, which are useful for studying convergence of numerical approximations of solutions to partial differential equations. Our study is motivated by the version of the Trotter-Kato theorem discussed in [11, Section 3.6]. The goal is to provide a general approach, which is flexible enough to cover a variety of approximation schemes for infinite dimensional systems. Of course it is not possible to get precise error estimates at this level of generality. In order to get those one usually has to exploit the special structure of a system, what we shall demonstrate in a few situations.

In Section 2 we present a version of the Trotter-Kato theorem which is standard except for the fact that the state space on which the semigroup is defined is a closed proper subspace of an ambient Banach or Hilbert space. The approximating spaces are isomorphic to subspaces of this ambient space but not necessarily of the state space. Furthermore, we present in this section error estimates for smooth initial data in the general case and also for analytic semigroups. In Section 3 we discuss possibilities to verify the basic assumptions of the Trotter-Kato theorem, i.e., how to establish the stability and the consistency property. Applicability of the results is demonstrated in Section 4 for a first order wave equation, a second order wave equation in one space dimension and Stokes' equation as illustrative examples.

## 2. THE TROTTER-KATO THEOREM

### 2.1. Statement and proof of the theorem.

Let $Z$ and $X_n$ be Banach spaces with norms $\|\cdot\|$, $\|\cdot\|_n$, $n = 1, 2, \dots$, respectively, and $X$ be a closed linear subspace

---

Received by the editor August 18, 1995 and, in revised form, August 1, 1996.

1991 *Mathematics Subject Classification.* Primary 47D05, 47H05, 65J10, 35K22, 35L99.

*Key words and phrases.* Semigroups of transformations, Trotter-Kato-Theorems, numerical approximation of linear evolutionary equations.

Research of the first author was supported in part by the NSF under Grant UINT-8521208 and DMS-8818530 and by the Air Force Office of Scientific Research under contract AFOSR-90-0091.

Research by the second author was supported in part by FWF(Austria) under Grants P6005, P8146-PHY and under F003.

of $Z$. On $X$ a $\mathrm{C}_0$-semigroup $T(\cdot)$ with infinitesimal generator $A$ is given. The goal is to construct approximating generators $A_n$ on the spaces $X_n$ such that the $\mathrm{C}_0$-semigroups $T_n(\cdot)$ generated by $A_n$ approximate $T(\cdot)$ in a sense which will be made precise below. We will make the following assumptions:

For every $n = 1, 2, \dots$ there exist bounded linear operators $P_n : Z \to X_n$ and $E_n : X_n \to Z$ satisfying
(A1) $\|P_n\| \le M_1$, $\|E_n\| \le M_2$, where $M_1$, $M_2$ are independent of $n$,
(A2) $\|E_n P_n x - x\| \to 0$ as $n \to \infty$ for all $x \in X$,
(A3) $P_n E_n = I_n$, where $I_n$ is the identity operator on $X_n$.

Assumption (A2) is a consequence of each of the two equivalent statements in the Trotter-Kato theorem. Therefore when choosing the spaces $X_n$ and the operators $P_n$, $E_n$ one has to make sure that (A2) is also satisfied. However, (A2) need not be assumed explicitly in the theorem. In many situations one has $X = Z$, but Section 4.3, where we consider Stokes' equation, presents an example where it is advantageous to define the operators $P_n$, $E_n$ first for an ambient space $Z$ which contains the actual state space for the equation as a proper closed subspace.

The general setting can be phrased in an equivalent way for subspaces of $Z$. In order to see this define the subspaces $Z_n$ of $Z$ and the mappings $\pi_n : Z \to Z_n$ by

$$Z_n = \operatorname{range} E_n \quad \text{and} \quad \pi_n = E_n P_n, \quad n = 1, 2, \dots .$$

The subspaces $Z_n$ are endowed with the $Z$-norm. It is easy to see that the $Z_n$ are closed subspaces of $Z$ and that $\pi_n$ are projections $Z \to Z_n$, i.e., $\pi_n^2 = \pi_n$ and $\operatorname{range} \pi_n = Z_n$. Furthermore, $\tilde{T}_n(t) = E_n T_n(t) P_n \mid_{Z_n}$, $t \ge 0$, defines a $\mathrm{C}_0$-semigroup on $Z_n$ with infinitesimal generator $\tilde{A}_n$ given by $\operatorname{dom} \tilde{A}_n = E_n \operatorname{dom} A_n$ and $\tilde{A}_n = E_n A_n P_n \mid_{Z_n}$. Assumption (A1) implies that there exists a constant $\tilde{M} > 0$ such that

$$\text{(B1)} \qquad \|\pi_n\| \le \tilde{M}, \quad n = 1, 2, \dots,$$

is true, whereas from assumption (A2) we get

$$\text{(B2)} \qquad \lim_{n \to \infty} \pi_n z = z \quad \text{for all } z \in X.$$

Note that by the uniform boundedness principle assumption (B1) is automatically satisfied if (B2) holds for *all* $z \in Z$. In general we do not have $Z_n \subset X$. See Section 4.3 for an example. If one has numerical approximation in mind, then the spaces $Z_n$ are finite dimensional, of course.

Conversely, let $Z_n$, $n = 1, 2, \dots$ , be a sequence of subspaces of $Z$ with projections $\pi_n : Z \to Z_n$ and canonical injections $\iota_n : Z_n \to Z$. We assume that (B1) and (B2) are satisfied. Then obviously assumption (B1) implies (A1) and (B2) implies (A2) for $X_n = Z_n$, $P_n = \pi_n$ and $E_n = \iota_n$. (A3) is trivially satisfied.

The most frequent situation where the setting introduced at the beginning of this section occurs is when we start with a sequence of finite dimensional subspaces of $Z_n$, $\dim Z_n = k_n$. For each subspace $Z_n$ we choose a basis $z_1^n, \dots, z_{k_n}^n$ and define the mapping $p_n : Z_n \to X_n := \mathbb{R}^{k_n}$ by $p_n z = (\alpha_1, \dots, \alpha_{k_n})^{\mathrm{T}}$ for $z = \sum_{j=1}^{k_n} \alpha_j z_j \in Z_n$. The norm on $X_n$ is defined by $\|x\|_{X_n} = \|p_n^{-1} x\|_Z$. If we define the mappings

$P_n : Z \to X_n$, $E_n : X_n \to Z$ by

$$\begin{aligned} P_n z &= p_n \pi_n z, \quad z \in Z, \\ E_n x &= \iota_n p_n^{-1} x, \quad x \in X_n, \end{aligned}$$

then assumptions (A1) – (A3) are satisfied.

Before we state the Trotter-Kato theorem we introduce the following notation: $A \in G(M, \omega, X)$, $M \geq 1$, $\omega \in \mathbb{R}$, means that $A$ is the infinitesimal generator of a $C_0$-semigroup $T(t)$, $t \geq 0$, satisfying $\|T(t)\| \leq Me^{\omega t}$, $t \geq 0$. Of course, if $A$ is the infinitesimal generator of a $C_0$-semigroup, then $A \in G(M, \omega, X)$ for some $M \geq 1$ and $\omega \in \mathbb{R}$.

**Theorem 2.1** (Trotter-Kato). *Assume that* (A1) and (A3) *are satisfied. Let $A$ resp. $A_n$ be in $G(M, \omega, X)$ resp. in $G(M, \omega, X_n)$ and let $T(t)$ and $T_n(t)$ be the semigroups generated by $A$ and $A_n$ on $X$ and $X_n$, respectively. Then the following statements are equivalent:*

(a) *There exists a $\lambda_0 \in \rho(A) \cap \bigcap_{n=1}^{\infty} \rho(A_n)$ such that, for all $x \in X$,*

$$\|E_n(\lambda_0 I_n - A_n)^{-1} P_n x - (\lambda_0 I - A)^{-1} x\| \to 0 \quad \textit{as } n \to \infty.$$

(b) *For every $x \in X$ and $t \geq 0$,*

$$\|E_n T_n(t) P_n x - T(t)x\| \to 0 \quad \textit{as } n \to \infty$$

*uniformly on bounded $t$-intervals.*

*If* (a) *or* (b) *is true, then* (a) *holds for all $\lambda$ with* $\operatorname{Re}\lambda > \omega$.

*Proof.* If we set $Z_n = \operatorname{range} E_n$ and $\pi_n = E_n P_n$, $n = 1, 2, \ldots$, then the theorem is proved if we establish equivalence of the following two statements:

($\tilde{\text{a}}$) There exists a $\lambda_0 \in \rho(A) \cap \bigcap_{n=1}^{\infty} \rho(\tilde{A}_n)$ such that, for all $x \in X$,

$$\|(\lambda_0 \tilde{I}_n - \tilde{A}_n)^{-1} \pi_n x - (\lambda_0 I - A)^{-1} x\| \to 0 \quad \text{as} \quad n \to \infty.$$

($\tilde{\text{b}}$) For every $x \in X$ and $t \geq 0$,

$$\|\tilde{T}_n(t) \pi_n x - T(t)x\| \to 0 \quad \text{as} \quad n \to \infty$$

uniformly on bounded $t$-intervals.

For the rest of the proof we shall write $T_n(t)$ and $A_n$ instead of $\tilde{T}_n(t)$ and $\tilde{A}_n$, respectively. It is no loss of generality if we assume that ($\tilde{\text{a}}$) holds for $\lambda_0 = 0$.

a) We first show that ($\tilde{\text{a}}$) implies ($\tilde{\text{b}}$). For $x \in X$ we define

$$e_n(t) = \big(T_n(t)\pi_n - \pi_n T(t)\big)x, \quad n = 1, 2, \ldots, \ t \geq 0.$$

For $x \in \operatorname{dom} A$, the function $u_n(t)$ defined by

$$u_n(t) = A_n^{-1} e_n(t), \quad t \geq 0, \ n = 1, 2, \ldots,$$

is in $C^1(0, \infty; Z_n)$ and satisfies

$$\begin{aligned} \dot{u}_n &= A_n u_n + \pi_n \Delta_n A T(t) x, \\ u_n(0) &= 0, \end{aligned} \tag{2.1}$$

where we have set

$$\Delta_n = A^{-1} - A_n^{-1} \pi_n.$$

Indeed, $A_n^{-1} T_n(t) \pi_n x = T_n(t) A_n^{-1} \pi_n x$ is continuously differentiable on $[0, \infty)$, because $A_n^{-1} \pi_n x$ is in $\operatorname{dom} A_n$, whereas $A_n^{-1} \pi_n T(t) x$ is continuously differentiable, because $x \in \operatorname{dom} A$ and $A_n^{-1}\pi_n$ is a bounded operator $X \to Z_n$. An easy calculation proves (2.1).

From (2.1) we obtain by the variation of parameter formula that, for $t \geq 0$, $x \in \operatorname{dom} A$,

$$u_n(t) = \int_0^t T_n(t-\tau)\pi_n\Delta_n AT(\tau)x\,d\tau.$$

For $x \in \operatorname{dom} A^2$ integration by parts implies

$$\begin{aligned} u_n(t) &= -A_n^{-1}\pi_n\Delta_n AT(t)x + A_n^{-1}T_n(t)\pi_n\Delta_n Ax \\ &\quad + A_n^{-1}\int_0^t T_n(t-\tau)\pi_n\Delta_n A^2T(\tau)x\,d\tau, \quad t \geq 0. \end{aligned} \tag{2.2}$$

Here we have used $A_nT_n(t-\tau)y = -\frac{d}{d\tau}T_n(t-\tau)y$, $y \in Z_n$, and $\frac{d}{d\tau}T(\tau)x = AT(\tau)x$, $x \in \operatorname{dom} A$. From this representation of $u_n(t)$ we obtain the error representation:

$$\begin{aligned} e_n(t) &= -\pi_n\Delta_n AT(t)x + T_n(t)\pi_n\Delta_n Ax \\ &\quad + \int_0^t T_n(t-\tau)\pi_n\Delta_n A^2T(\tau)x\,d\tau, \quad t \geq 0,\ x \in \operatorname{dom} A^2. \end{aligned} \tag{2.3}$$

In order to prove $\lim_{n\to\infty} e_n(t) = 0$ uniformly for $t$ in bounded intervals, we consider the terms on the right-hand side of (2.3) separately. For any $\overline{T} > 0$ the set $\{T(t)Ax \mid 0 \leq t \leq \overline{T}\}$ is compact. Therefore we have

$$\pi_n\Delta_n AT(t)x \to 0 \quad \text{as} \quad n \to \infty$$

uniformly on $[0,\overline{T}]$. For the second term on the right-hand side of (2.3) this is obvious, because $\|T_n(t)\| \leq Me^{\omega t}$, $t \geq 0$, $n = 1, 2, \ldots$ .

Since, for $x \in \operatorname{dom} A^2$, the set $\{A^2T(\tau)x \mid 0 \leq \tau \leq \overline{T}\}$ is compact, we see that $\|\Delta_n A^2T(\tau)x\| \to 0$ as $n \to \infty$ uniformly on $[0,\overline{T}]$. Therefore also the integral on the right-hand side of (2.3) converges to zero uniformly on $[0,\overline{T}]$. Thus we have proved that $\lim_{n\to\infty} e_n(t) = 0$ uniformly on $0 \leq t \leq \overline{T}$ for any $x \in \operatorname{dom} A^2$. By a standard density argument we see that this is true for all $x \in X$ (note that, by definition of $e_n(t)$, there exists a constant $c_0 > 0$ such that $\sup_{0\leq t\leq\overline{T}} \|e_n(t)\| \leq c_0e^{\omega\overline{T}}\|x\|$, $x \in X$, $n = 1, 2, \ldots$).

It remains to prove that

$$\lim_{n\to\infty} \|\pi_nT(t)x - T(t)x\| = 0 \quad \text{uniformly on } [0,\overline{T}].$$

By compactness of $\{T(t)x \mid 0 \leq t \leq \overline{T}\}$ we only have to prove $\lim_{n\to\infty} \pi_nx = x$ for all $x \in X$. For $x \in \operatorname{dom} A$ we get (observing that $\ker(I - \pi_n) = Z_n$)

$$\pi_nx - x = (\pi_n - I)\Delta_n Ax. \tag{2.4}$$

This implies $\lim_{n\to\infty} \pi_nx = x$ for $x \in \operatorname{dom} A$. The result for $x \in X$ follows by a density argument.

b) Assume now that ($\tilde{\text{b}}$) holds and that $\operatorname{Re}\lambda > \omega$. Then

$$\|(\lambda I_n - A_n)^{-1}\pi_nx - (\lambda I - A)^{-1}x\| \leq \int_0^\infty e^{-\operatorname{Re}\lambda t}\|T_n(t)\pi_nx - T(t)x\|\,dt.$$

The right-hand side of this inequality tends to zero as $n \to \infty$ by ($\tilde{\text{b}}$), the choice of $\lambda$ and Lebesgue's dominated convergence theorem. □

*Remarks.* 1. The proof of Theorem 2.1 as given above is a slight modification of Kato's proof putting more emphasis on the representation of the error $e_n(t)$ which will be useful in the next subsection.

2. The assumption $A_n \in G(M, \omega, X_n)$, $n = 1, 2, \ldots$, or equivalently $\|T_n(t)\|_n \leq Me^{\omega t}$, $n = 1, 2, \ldots$, usually is called the *stability property* of the approximations, whereas statement (a) is called the *consistency property* of the approximations. With this terminology the Trotter-Kato theorem essentially states that, under the assumption of stability, consistency is equivalent to convergence (as characterized in statement (b)).

Using the uniform boundedness principle and the standard proof for the fact that any $C_0$-semigroup $T(\cdot)$ satisfies an estimate of the form $\|T(t)\| \leq Me^{\omega t}$, $t \geq 0$, it is easy to see that on the other hand convergence implies stability (and consequently also consistency). Compare Theorem 4.4 in [10].

3. Consider the setting used in the proof of Theorem 2.1. With the operators $A$ and $\tilde{A}_n$, $n = 1, 2, \ldots$, we can associate the steady state problems

$$Au = \lambda_0 u - y, \quad y \in X, \tag{2.5}$$

on $X$ and

$$\tilde{A}_n u_n = \lambda_0 u_n - \pi_n y \tag{2.6}$$

on $Z_n$. The consistency hypothesis (ã) just means that these steady state problems, for all $y \in X$, have unique solutions $u$ resp. $u_n$ which depend continuously on $y$ and

$$\lim_{n \to \infty} u_n = u. \tag{2.7}$$

Indeed, the assumptions on the solvability of the steady state problems are equivalent to $\lambda_0 \in \rho(A) \cap \bigcap_{n=1}^{\infty} \rho(\tilde{A}_n)$ and (2.7) is just the strong convergence of the resolvent operators, because $u = (\lambda_0 I - A)^{-1} y$ and $u_n = (\lambda_0 \tilde{I}_n - \tilde{A}_n)^{-1} \pi_n y$.

In view of these considerations the Trotter-Kato theorem states that, under the assumption of stability, convergence of the solutions of the steady state problems associated with the semigroup generators implies convergence of the semigroups. This point of view was stressed in [10], where it was also shown that convergence rates are also preserved. We shall address this question in the next subsection.

4. The error function $e_n(t)$ is continuously differentiable on $[0, \infty)$, if $x \in \operatorname{dom} A$ and $\pi_n x \in \operatorname{dom} A_n$, which is certainly the case if the $A_n$'s are bounded. The most common situation where the $A_n$'s are bounded occurs when the spaces $X_n$ are finite dimensional. Then $e_n(t)$ is the solution of

$$\begin{aligned} \dot{e}_n &= A_n e_n + A_n \pi_n \Delta_n A T(t) x, \quad t \geq 0, \\ e_n(0) &= 0. \end{aligned} \tag{2.8}$$

This implies

$$e_n(t) = \int_0^t A_n T_n(t - \tau) \pi_n \Delta_n A T(\tau) x \, d\tau, \quad t \geq 0. \tag{2.9}$$

From this representation we can get (2.3) by integration by parts directly provided $x \in \operatorname{dom} A^2$. Thus the introduction of $u_n(t)$ is not necessary in cases where $e_n(t)$ is differentiable.

5. A somewhat different proof of Theorem 2.1 can be given using the approach followed in [10]. Let the setting be that used in the proof of Theorem 2.1 and define the "elliptic" projections $q_n : \operatorname{dom} A \to Z_n$ by

$$q_n = \tilde{A}_n^{-1}\pi_n A, \quad n = 1, 2, \dots .$$

For $x \in \operatorname{dom} A^2$ we introduce the error

$$f_n(t) = \tilde{T}_n(t)q_n x - q_n T(t)x, \quad t \geq 0,$$

which is continuously differentiable. This follows from $q_n x \in \operatorname{dom} \tilde{A}_n$ and $q_n T(t)x = \tilde{A}_n^{-1}\pi_n T(t)Ax$. We have $f_n(0) = 0$ and

$$\dot{f}_n(t) = \tilde{A}_n f_n(t) + \pi_n(I - q_n)T(t)Ax, \quad t \geq 0.$$

This gives

$$f_n(t) = \int_0^t \tilde{T}_n(t-s)\pi_n(I - q_n)T(s)Ax\, ds$$

and

$$T(t)x - \tilde{T}_n(t)q_n x = (I - q_n)T(t)x - \int_0^t \tilde{T}_n(t-s)\pi_n(I - q_n)T(s)Ax\, ds, \quad t \geq 0.$$

Observing that, for $y \in \operatorname{dom} A$, we have $(I - q_n)y = (A^{-1} - \tilde{A}_n^{-1}\pi_n)Ay$ we see that the same arguments used in the proof of Theorem 2.1 give

$$\lim_{n\to\infty} \tilde{T}_n(t)q_n x = T(t)x, \quad x \in \operatorname{dom} A,$$

uniformly on bounded $t$-intervals. In order to get $\tilde{T}_n(t)\pi_n x \to T(t)x$ uniformly on bounded $t$-intervals for any $x \in X$ one has to choose a sequence $(x_k) \subset \operatorname{dom} A$ with $x_k \to x$ and to apply the standard arguments to the estimate

$$\begin{aligned} \|T(t)x - \tilde{T}_n(t)\pi_n x\| &\leq \|T(t)(x - x_k)\| + \|(T(t) - \tilde{T}_n(t)q_n)x_k\| \\ &\quad + \|\tilde{T}_n(t)(q_n x_k - \pi_n x)\| \\ &\leq Me^{\omega t}\|x - x_k\| + \|(T(t) - \tilde{T}_n(t)q_n)x_k\| \\ &\quad + Me^{\omega t}\big(\|q_n x_k - x_k\| + \|x_k - \pi_n x_k\| + \|\pi_n\|\,\|x_k - x\|\big). \end{aligned}$$

2.2. **Error estimates for smooth initial data.** The proof of the Trotter-Kato theorem as given in the previous subsection offers also the possibility to obtain error estimates for the approximations. However, because of the generality of Theorem 2.1 we cannot expect to get error estimates which are sharp in specific situations. In order to get sharp estimates one has to exploit the special structure of the problem at hand. See for instance [2] for parabolic equations and [9], [5] for delay equations of retarded type. In the following let $\|\cdot\|_{\operatorname{dom} A^\alpha}$ denote the graph norm on $\operatorname{dom} A^\alpha$, $\alpha > 0$.

**Proposition 2.2.** *Let the assumptions of Theorem* 2.1 *be satisfied and, for any* $\lambda_0 \in \rho(A) \cap \bigcap_{n=1}^\infty \rho(A_n)$, *set* $\Delta_n(\lambda_0) = E_n(\lambda_0 I_n - A_n)^{-1}P_n - (\lambda_0 I - A)^{-1}$ *Then the following is true:*

a) *For any* $\overline{T} > 0$ *and any* $\alpha > 0$, *there exists a constant* $\gamma = \gamma(\overline{T}, \alpha) > 0$ *such that*

$$\|E_n T_n(t)P_n x - T(t)x\| \leq \gamma\|\Delta_n(\lambda_0)\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)}\|x\|_{\operatorname{dom} A^{\alpha+2}}, \quad 0 \leq t \leq \overline{T},$$

*for all* $x \in \operatorname{dom} A^{\alpha+2}$ *and* $n = 1, 2, \dots$ .

b) *If, in addition, the semigroup* $T(\cdot)$ *is analytic, then for any* $\overline{T} > 0$, $\varepsilon > 0$ *and* $\alpha > 0$, *there exists a constant* $\gamma = \gamma(\overline{T}, \varepsilon, \alpha) > 0$ *such that*

$$\|E_n T_n(t) P_n x - T(t)x\| \le \gamma \|\Delta_n(\lambda_0)\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)} \|x\|_{\operatorname{dom} A^{\alpha+1+\varepsilon}}, \quad 0 \le t \le \overline{T},$$

*for all* $x \in \operatorname{dom} A^{\alpha+1+\varepsilon}$ *and* $n = 1, 2, \dots$ .

*Proof.* As in the proof of Theorem 2.1 we can assume without restriction of generality that $0 \in \rho(A) \cap \bigcap_{n=1}^\infty \rho(A_n)$. Furthermore, for the proof we adopt the same setting as in the proof of Theorem 2.1 and write again $T_n(t)$ and $A_n$ instead of $\tilde{T}_n(t)$ and $\tilde{A}_n$, respectively. Correspondingly we also set $\Delta_n = A^{-1} - A_n^{-1}\pi_n$, $n = 1, 2, \dots$ . In the following 'const.' always denotes a positive constant which does not depend on $x$ or $t$ (in the given sets) and may have different values at different occurrences.

The proof for part a) is straightforward, estimating the terms on the right-hand sides of (2.3) and (2.4). We have to observe that the restriction of the semigroup $T(\cdot)$ to $(\operatorname{dom} A^\alpha, \|\cdot\|_{\operatorname{dom} A^\alpha})$ is also of type $G(M, \omega, \operatorname{dom} A^\alpha)$ and $\|A^\beta x\|_{\operatorname{dom} A^\alpha} \le \text{const.}\|x\|_{\operatorname{dom} A^{\alpha+\beta}}$.

For the proof of part b) we observe first that in case of an analytic semigroup the representation (2.3) of $e_n(t)$ is valid for $x \in \operatorname{dom} A^{1+\delta}$, $\delta > 0$. The integration by parts which leads to (2.2) can also be performed under the present conditions. We only have to observe that for an analytic semigroup we have $T(\tau)x \in \operatorname{dom} A^k$, $k = 1, 2, \dots$, $x \in X$ and $\tau > 0$. Furthermore, we have to use the estimate $\|A^2 T(\tau)x\| = \|A^{1-\delta} T(\tau) A^{1+\delta} x\| \le \text{const.}\tau^{-1+\delta}\|T(\tau)A^{1+\delta}x\|$, $\tau > 0$, $x \in \operatorname{dom} A^{1+\delta}$.

We only have to consider the integral term on the right-hand side of (2.3), because for the other two terms and the term on the right-hand side of (2.4) we see immediately that, for $\varepsilon > 0$,

$$\|Ax\|_{\operatorname{dom} A^\alpha} \le \text{const.}\|x\|_{\operatorname{dom} A^{\alpha+1}}, \quad x \in \operatorname{dom} A^{\alpha+1},$$

and

$$\|T(t)Ax\|_{\operatorname{dom} A^\alpha} \le M e^{\omega \overline{T}} \|Ax\|_{\operatorname{dom} A^\alpha} \le \text{const.}\|x\|_{\operatorname{dom} A^{\alpha+1}},$$

for $x \in \operatorname{dom} A^{\alpha+1}$ and $0 \le t \le \overline{T}$. For the integral term we get the estimate

$$\begin{aligned}
&\Big\| \int_0^t T_n(t-\tau)\pi_n \Delta_n A^2 T(\tau)x \, d\tau \Big\| \\
&\quad \le M e^{\omega \overline{T}} \|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)} \int_0^{\overline{T}} \|A^2 T(\tau)x\|_{\operatorname{dom} A^\alpha} d\tau \\
&\quad \le \text{const.}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)} \int_0^{\overline{T}} \|A^{1-\varepsilon} A^{1+\varepsilon} T(\tau)x\|_{\operatorname{dom} A^\alpha} \, d\tau \\
&\quad \le \text{const.}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)} \int_0^{\overline{T}} \frac{1}{\tau^{1-\varepsilon}} \|A^{1+\varepsilon} T(\tau)x\|_{\operatorname{dom} A^\alpha} d\tau \\
&\quad \le \text{const.}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)} \|x\|_{\operatorname{dom} A^{\alpha+1+\varepsilon}}
\end{aligned}$$

for $x \in \operatorname{dom} A^{\alpha+1+\varepsilon}$. With respect to properties of fractional powers of closed operators which have been used in this proof we refer to [11], for instance. □

In case of second order parabolic equations with a selfadjoint uniformly elliptic operator it was shown in [2] that we can take $\varepsilon = 0$ in part b) of Proposition 2.2. Using basically the same ideas as in [2] we can prove an analogous result for analytic semigroups on a Hilbert space with arbitrary selfadjoint infinitesimal generator.

**Proposition 2.3.** *Let $Z$ and $X_n$ be Hilbert spaces and assume that* (A1), (A3) *are satisfied. Furthermore, assume that $A$ generates an analytic semigroup on $X$ and that the $A_n$ are selfadjoint bounded operators on $X_n$ with the property that, for a $\lambda_0 \in \rho(A) \cap \bigcap_{n=1}^{\infty} \rho(A_n)$, the operators $A_n - \lambda_0 I_n$ are dissipative, i.e., $A_n \in G(1, \lambda_0, X_n)$. Then the following is true:*

a) *For any $\overline{T} > 0$ and any $\alpha > 0$ there exists a $\gamma = \gamma(\overline{T}, \alpha) > 0$ such that*

$$\|E_n T_n(t) P_n x - T(t)x\| \le \gamma \|\Delta_n(\lambda_0)\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)} \|x\|_{\operatorname{dom} A^{\alpha+1}}$$

*for all $t \in [0, \overline{T}]$ and all $x \in \operatorname{dom} A^{\alpha+1}$.*

b) *Assume that in addition statement* (a) *of Theorem* 2.1 *is true. Then for any $\delta > 1$ and $\alpha > 0$ there exists a $\gamma = \gamma(\delta, \alpha) > 0$ such that*

$$\|E_n T_n(t) P_n x - T(t)x\| \le \gamma \|\Delta_n(\lambda_0)\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)} \|x\|_{\operatorname{dom} A^{\alpha+1/2}}$$

*for all $t \in [1/\delta, \delta]$ and all $x \in \operatorname{dom} A^{\alpha+1/2}$.*

*Proof.* The general setting is as in the proofs of Theorem 2.1 and Proposition 2.2, respectivley. Instead of equation (2.8) we first define $v_n(t) = te_n(t)$, which, for $x \in \operatorname{dom} A^\alpha$, $\alpha > 0$, satisfies

$$\begin{aligned} \dot{v}_n &= A_n v_n + e_n(t) + tA_n \pi_n \Delta_n AT(t)x, \quad t > 0, \\ v_n(0) &= 0. \end{aligned}$$

By the variation of constants formula we get

$$v_n(t) = \int_0^t T_n(t-\tau) e_n(\tau)\, d\tau + \int_0^t A_n T_n(t-\tau) \pi_n \Delta_n \tau A T(\tau) x\, d\tau, \quad t \ge 0.$$

Integration by parts in the second integral gives

$$\begin{aligned} e_n(t) = {} & \frac{1}{t} \int_0^t T_n(t-\tau) e_n(\tau)\, d\tau - \pi_n \Delta_n AT(t)x \\ & + \frac{1}{t} \int_0^t T_n(t-\tau) \pi_n \Delta_n AT(\tau) x\, d\tau \\ & + \frac{1}{t} \int_0^t T_n(t-\tau) \pi_n \Delta_n \tau A^2 T(\tau) x\, d\tau, \quad t > 0,\ x \in \operatorname{dom} A^\alpha. \end{aligned} \tag{2.10}$$

Observe that $\|\tau A^2 T(\tau) x\| = \|\tau A^{2-\alpha} T(\tau) A^\alpha x\| \le \text{const.} \tau^{-1+\alpha} \|A^\alpha x\|$ for $x \in \operatorname{dom} A^\alpha$, which guarantees that the last integral in (2.10) exists. Analogously one sees that the other integrals also exist. From equation (2.8) we get, for $x \in \operatorname{dom} A^\alpha$,

$$e_n(t) = A_n^{-1} \dot{e}_n(t) - \pi_n \Delta_n AT(t)x, \quad t > 0.$$

Taking inner products with $e_n(t)$ on both sides, observing that by selfadjointness of $A_n$ we have

$$\operatorname{Re}\langle e_n(t), A_n^{-1} \dot{e}_n(t) \rangle = \frac{1}{2} \frac{d}{dt} \langle e_n(t), A_n^{-1} e_n(t) \rangle$$

and integrating from 0 to $t$, we get

$$\begin{aligned} \int_0^t \|e_n(\tau)\|^2 d\tau &= \frac{1}{2} \langle e_n(t), A_n^{-1} e_n(t) \rangle - \operatorname{Re} \int_0^t \langle e_n(\tau), \pi_n \Delta_n AT(\tau)x \rangle\, d\tau \\ &\le \frac{1}{2} \int_0^t \|e_n(\tau)\|^2 d\tau + \frac{1}{2} \int_0^t \|\pi_n \Delta_n AT(\tau) x\|^2 d\tau, \quad t \ge 0, \end{aligned}$$

where we have also used dissipativeness of $A_n$. Consequently we have

$$\int_0^t \|e_n(\tau)\|^2 d\tau \le \int_0^t \|\pi_n \Delta_n A T(\tau)x\|^2 d\tau, \quad t \ge 0,\ x \in \operatorname{dom} A^\alpha.$$

From this we get

(2.11)
$$\begin{aligned}\Big\|\frac{1}{t}\int_0^t T_n(t-\tau)e_n(\tau)\,d\tau\Big\| &\le \frac{1}{t}\Big(\int_0^t \|T_n(\tau)\|^2 d\tau\Big)^{1/2}\Big(\int_0^t \|e_n(\tau)\|^2 d\tau\Big)^{1/2}\\ &\le M\tilde{M}t^{-1/2}e^{|\omega|t}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)}\Big(\int_0^t \|AT(\tau)\|^2_{\operatorname{dom} A^\alpha} d\tau\Big)^{1/2}, \quad t > 0.\end{aligned}$$

In order to prove part a) we choose $x \in \operatorname{dom} A^{\alpha+1}$. Then $\|AT(\tau)x\|_{\operatorname{dom} A^\alpha} \le Me^{|\omega|\overline{T}}\|x\|_{\operatorname{dom} A^{\alpha+1}}$ which shows that

$$\Big\|\frac{1}{t}\int_0^t T_n(t-\tau)e_n(\tau)\,d\tau\Big\| \le \text{const.}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)}\|x\|_{\operatorname{dom} A^{\alpha+1}},$$

for $t \in [0, \overline{T}]$.

The second and third term on the right-hand side of (2.10) can easily be estimated by $\text{const.}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)}\|x\|_{\operatorname{dom} A^{\alpha+1}}$ for all $t \in [0, \overline{T}]$ and $x \in \operatorname{dom} A^{\alpha+1}$. For the fourth term we get

$$\begin{aligned}&\Big\|\frac{1}{t}\int_0^t T_n(t-\tau)\pi_n\Delta_n \tau A^2 T(\tau)x\,d\tau\Big\|\\ &\qquad\le \text{const.}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)}\frac{1}{t}\int_0^t \tau\|A^2T(\tau)x\|_{\operatorname{dom} A^\alpha} d\tau\\ &\qquad\le \text{const.}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)}\frac{1}{t}\int_0^t \|T(\tau)Ax\|_{\operatorname{dom} A^\alpha} d\tau\\ &\qquad\le \text{const.}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)}\|x\|_{\operatorname{dom} A^{\alpha+1}}\end{aligned}$$

for $0 < t \le \overline{T}$ and $x \in \operatorname{dom} A^{\alpha+1}$.

For the proof of b) we choose $x \in \operatorname{dom} A^{\alpha+1/2}$ and observe first that selfadjointness of the $A_n$ together with ($\tilde{\text{a}}$) implies that also $A$ has to be selfadjoint. Consequently we have

$$\begin{aligned}\|AT(\tau)x\|^2_{\operatorname{dom} A^\alpha} &= \|AT(\tau)x\|^2 + \|AT(\tau)A^\alpha x\|^2\\ &= \frac{1}{2}\frac{d}{d\tau}\langle T(\tau)A^{1/2}x, T(\tau)A^{1/2}x\rangle\\ &\qquad + \frac{1}{2}\frac{d}{d\tau}\langle T(\tau)A^{\alpha+1/2}x, T(\tau)A^{\alpha+1/2}x\rangle, \quad \tau \ge 0,\end{aligned}$$

and consequently

$$\int_0^t \|AT(\tau)x\|^2_{\operatorname{dom} A^\alpha} d\tau \le \frac{1}{2}\|T(t)A^{1/2}x\|^2_{\operatorname{dom} A^\alpha} \le \text{const.}\|x\|^2_{\operatorname{dom} A^{\alpha+1/2}}, \quad t \ge 0.$$

This and (2.11) prove that

$$\Big\|\frac{1}{t}\int_0^t T_n(t-\tau)e_n(\tau)\,d\tau\Big\| \le \text{const.}\|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)}\|x\|_{\operatorname{dom} A^{\alpha+1/2}}$$

for all $t \in [1/\delta, \delta]$ and all $x \in \operatorname{dom} A^{\alpha+1/2}$. For the other terms in (2.10) we get the analogous estimates if we observe

$$\begin{aligned}\int_0^t \|AT(\tau)x\|_{\operatorname{dom} A^\alpha} d\tau &\le \text{const.}\|A^{1/2}x\|_{\operatorname{dom} A^\alpha} \int_0^t \tau^{-1/2} d\tau,\\ \int_0^t \tau\|A^2T(\tau)x\|_{\operatorname{dom} A^\alpha} d\tau &= \int_0^t \tau\|A^{3/2}T(\tau)A^{1/2}x\|_{\operatorname{dom} A^\alpha} d\tau\\ &\le \text{const.}\|A^{1/2}x\|_{\operatorname{dom} A^\alpha} \int_0^t \tau^{-1/2} d\tau\end{aligned}$$

for $t \in [0, \delta]$. □

*Remarks.* 1. Note the difference in the two statements of Proposition 2.3. The first one requires $x \in \operatorname{dom} A^{\alpha+1}$ but gives an estimate on intervals $[0, \bar{T}]$, whereas the second one requires $x \in \operatorname{dom} A^{\alpha+1/2}$ only and gives an estimate on compact $t$-intervals which exclude $t = 0$. The assumption, that the consistency property (a) is satisfied, in part b) is only used in order to prove that $A$ is also selfadjoint.

2. Without restriction of generality we take the setting used in the proof of Proposition 2.2 and set (note that $\Delta_n = A^{-1} - A_n^{-1}$)

$$F(n) = \|\Delta_n\|_{\mathcal{L}(\operatorname{dom} A^\alpha, X)}.$$

This means that for any $x \in \operatorname{dom} A^\alpha$ we have

$$\|A^{-1}x - A_n^{-1}\pi_n x\| \le F(n)\|x\|_{\operatorname{dom} A^\alpha}. \tag{2.12}$$

If we observe that $u = A^{-1}x$ resp. $u_n = A_n^{-1}\pi_n x$ are the unique solutions of the steady state problems

$$Au = x \quad \text{resp.} \quad A_n u_n = \pi_n x$$

we can rewrite inequality (2.12) as

$$\|u - u_n\| = \|u - q_n u\| \le F(n)\|Au\|_{\operatorname{dom} A^\alpha} \le F(n)\|u\|_{\operatorname{dom} A^{\alpha+1}}. \tag{2.13}$$

Therefore Proposition 2.2, a) can be stated as follows: If the estimate (2.13) is true for an approximation scheme for the steady state problem $Au = x$, then we have the same rate estimate for the corresponding approximation scheme for the Cauchy problem $\dot{u} = Au$, $u(0) = x$, provided $x \in \operatorname{dom} A^{\alpha+2}$ (i.e., $x \in \operatorname{dom} A$ and $Ax \in \operatorname{dom} A^{\alpha+1}$). This shows that the results of this section are closely related to results in [10]. For instance, the assumption that the estimate (2.13) is satisfied is exactly the assumption in [10] that "Theorem T" is true for $\tilde{X} = \operatorname{dom} A^{\alpha+1}$ (see [10, p. 130]). Proposition 2.2, a) essentially is Theorem 4.2 in [10] with the difference that in [10] the estimate is for $T_n(t)q_n x - T(t)x$ instead of $T_n(t)\pi_n x - T(t)x$. Furthermore, the results of this section show that the smoothness assumption $x \in \operatorname{dom} A^{\alpha+2}$ can be relaxed considerably. In case of general analytic semigroups in Banach spaces we need $x \in \operatorname{dom} A^{\alpha+1+\epsilon}$ (Proposition 2.2, b)). If in addition we assume that the spaces are Hilbert spaces and the generators are selfadjoint, then $x \in \operatorname{dom} A^{\alpha+1}$ resp. $x \in \operatorname{dom} A^{\alpha+1/2}$ is sufficient (Proposition 2.3).

## 3. HOW TO ESTABLISH STABILITY AND CONSISTENCY

In order to apply Theorem 2.1 one faces the following major difficulties:

a) In general it is very difficult to verify the stability property, i.e., to prove that $A_n \in G(M, \omega, X_n)$, $n = 1, 2, \ldots$, for some $M \geq 1$, $\omega \in \mathbb{R}$, when $M > 1$ is necessary.
b) Direct verification of the consistency property (a) involves computation of the resolvents $(\lambda I_n - A_n)^{-1}$, which in general is almost impossible.

Of course, the Hille-Yosida generation theorem for $C_0$-semigroups tells us among other things that $A_n \in G(M, \omega, X_n)$ if $\lambda \in \rho(A_n)$ for $\operatorname{Re}\lambda > \omega$ and

$$\|(A_n - \lambda I_n)^{-k}\| \leq \frac{M}{(\operatorname{Re}\lambda - \omega)^k}, \quad \operatorname{Re}\lambda > \omega, \; k = 1, 2, \ldots .$$

But to establish these inequalities for the powers of the resolvent operators in particular for the approximating generators $A_n$ is in most cases (i.e., except $M = 1$) impossible. In general, the only way to verify the stability property is to use dissipativity estimates possibly after renorming the spaces $X_n$ with uniformly equivalent norms.

Concerning the consistency property one tries at any case to avoid computation of the resolvent operators $(\lambda I_n - A_n)^{-1}$ and direct verification of condition (a). Usually it is very easy to compute explicit representations of the approximating generators $A_n$. Therefore one would like to replace (a) by a condition involving convergence of the operators $A_n$ to $A$ in some sense. The following result is well known, the proofs perhaps are different (see for instance [11]):

**Proposition 3.1.** *Let the assumptions of Theorem* 2.1 *be satisfied. Then statement* (a) *of Theorem* 2.1 *is equivalent to* (A2) *and the following two statements:*

(C1) *There exists a subset $D \subset \operatorname{dom} A$ such that $\overline{D} = X$ and $\overline{(\lambda_0 I - A)D} = X$ for a $\lambda_0 > \omega$.*

(C2) *For all $u \in D$ there exists a sequence $(\overline{u}_n)_{n\in\mathbb{N}}$ with $\overline{u}_n \in \operatorname{dom} A_n$ such that*

$$\lim_{n\to\infty} E_n\overline{u}_n = u \quad \text{and} \quad \lim_{n\to\infty} E_nA_n\overline{u}_n = Au.$$

*Proof.* Without restriction of generality we can assume $\lambda_0 = 0$ for the proof. We first prove that (a) implies (A2) and (C1), (C2). To this end we first set $D = \operatorname{dom} A$ which implies $AD = X$, i.e., (C1) is satisfied. In the proof of Theorem 2.1 we have already shown that (a) implies (A2) (compare (2.4)).

We next fix $u \in \operatorname{dom} A$, choose $x \in X$ with $u = A^{-1}x$ and set $\overline{u}_n = A_n^{-1}P_nAu$. Then we have

$$E_n\overline{u}_n - u = E_nA_n^{-1}P_nx - A^{-1}x \to 0$$

as $n \to \infty$ by (a). Furthermore, we have (using (A2))

$$E_nA_n\overline{u}_n - Au = E_nA_nA_n^{-1}P_nx - AA^{-1}x = E_nP_nx - x \to 0$$

as $n \to \infty$. Thus we see that (C2) is also true.

In order to prove that (A2) and (C1), (C2) imply (a) we use the identity

$$E_nA_n^{-1}P_n - A^{-1} = E_n(A_n^{-1}P_nA - P_n)A^{-1} + (E_nP_n - I)A^{-1}. \tag{3.1}$$

For $x \in AD$ we choose $u \in D$ with $x = Au$ and set $u_n = A_n^{-1}P_nx = A_n^{-1}P_nAu$. Furthermore, for $u$, we choose $\overline{u}_n$ according to (C2). Then we get

$$\|\overline{u}_n - P_nu\|_n = \|P_n(E_n\overline{u}_n - u)\|_n \leq M_1\|E_n\overline{u}_n - u\| \to 0$$

as $n \to \infty$ and

$$\begin{aligned}\|\overline{u}_n - u_n\|_n &\le \|A_n^{-1}\|\,\|A_n\overline{u}_n - P_nAu\|_n\\ &\le \|A_n^{-1}\|\,\|P_n\|\,\|E_nA_n\overline{u}_n - Au\| \to 0\end{aligned}$$

as $n \to \infty$. Note that $\|A_n^{-1}\|$ is uniformly bounded, because $A_n \in G(M,\omega,X_n)$ for all $n$. The last two estimates prove that

$$\|P_nu - u_n\|_n \le \|P_nu - \overline{u}_n\|_n + \|\overline{u}_n - u_n\|_n \to 0$$

as $n \to \infty$. This estimate together with (3.1) and (A2) implies

$$\begin{aligned}\|E_nA_n^{-1}P_nx - A^{-1}x\| &\le \|E_n(u_n - P_nu)\| + \|E_nP_nu - u\|\\ &\le M_2\|u_n - P_nu\|_n + \|E_nP_nu - u\| \to 0\end{aligned}$$

as $n \to \infty$ for all $x \in AD$. A density argument finishes the proof for (a) (note that $\|E_nA_n^{-1}P_n\|$ is uniformly bounded). □

*Remark.* In fact, conditions (C1) and (C2) provide a formulation of the consistency property which is essentially the original one. See for instance [12, Chapter 3] for difference approximations.

The example in Subsection 4.1 below demonstrates the usefulness of conditions (C1) and (C2). However, in many applications, in particular if the abstract Cauchy problem is the abstraction of a PDE-problem, the generator $A$ is defined via a sesquilinear form $\sigma$, which is given on a densely and continuously embedded subspace $V$ of the state space $X$. Then the approximating generators $A_n$ usually are defined by sesquilinear forms $\sigma_n$ on the approximating state spaces $X_n$. These sesquilinear forms $\sigma_n$ are obtained from restrictions of $\sigma$ to appropriate subspaces $V_n$ of $V$ which are isomorphic to $X_n$. Of course, in such a case one would like to establish the stability and consistency property by using the approximating sesquilinear forms $\sigma_n$. Instead of formulating some general results in this direction we demonstrate the ideas by the examples in Subsections 4.2 and 4.3. The main reason for this approach to the problem is the fact that usually one has to exploit the special structure of the problem under consideration, which makes it very difficult to provide simple general conditions which cover a wide range of special cases.

Parabolic problems allow much stronger results, which will be presented in a different paper.

## 4. EXAMPLES

In this section we demonstrate applicability of the results developed in the previous sections. As already mentioned in the introduction the goal is to show that a variety of concrete situations is covered by the general framework presented in this paper.

### 4.1. A first order hyperbolic PDE.

In this example the role of the operators $P_n$ and $E_n$ appearing in conditions (A1)–(A3) and the usefulness of Proposition 3.1 are demonstrated. Consider the first order hyperbolic PDE

$$\begin{aligned}&\frac{\partial}{\partial t}u(t,x) + \frac{\partial}{\partial x}u(t,x) = 0, \quad x \in (0,1),\\ &u(t,0) = 0.\end{aligned} \tag{4.1}$$

The equation is studied in three different state spaces: $X = L^1(0,1)$, $X = L^2(0,1)$ and $X = C_0(0,1)$, respectively, where $C_0(0,1)$ is the space of continuous functions on [0,1] vanishing at $x = 0$. It is not difficult to show that the linear operator $A$ defined by

$$A\phi = -\phi', \quad \phi \in \operatorname{dom} A,$$

with $\operatorname{dom} A = \{\phi \in X \mid \phi \text{ is absolutely continuous on } [0,1] \text{ with } \phi' \in X \text{ and } \phi(0) = 0\}$ generates a $C_0$-semigroup on $X$. The numerical method analyzed here is the first order finite difference scheme

$$\begin{aligned} &\frac{d}{dt}u_k(t) = \frac{u_{k-1}(t) - u_k(t)}{\Delta x}, \quad k = 1, \dots, n, \\ &u_0(t) = 0, \end{aligned} \tag{4.2}$$

where $\operatorname{col}(u_1, \dots, u_n) \in X_n = \mathbb{R}^n$ and $u_k(t)$ represents an approximating value for $u(t,x)$ at the $k$-th nodal point $x_k = k\,\Delta x$ with $\Delta x = 1/n$. From equations (4.2) it is clear that the approximating generators $A_n$ on $\mathbb{R}^n$ are given by

$$(A_n u)_k = \frac{1}{\Delta x}(u_{k-1} - u_k), \quad k = 1, \dots, n,$$

where we set $u_0 = 0$.

*Case* 1. $X = L^1(0,1)$.

Let $P_n$, $E_n$ and $\|\cdot\|_n$ be defined as

$$\begin{aligned} E_n u &= \sum_{k=1}^{n} u_k \chi_{(x_{k-1}, x_k]}, \quad u \in X_n, \\ (P_n\phi)_k &= \frac{1}{\Delta x}\int_{x_{k-1}}^{x_k} \phi(x)\,dx, \quad 1 \le k \le n, \ \phi \in X, \\ \|u\|_n &= \Delta x \sum_{k=1}^{n} |u_k|, \quad u \in X_n. \end{aligned}$$

It is easy to show that the conditions (A1)–(A3) are satisfied. For an element $u \in X_n \setminus \{0\}$ the elements $v$ in the duality set $F_n(u) \subset X_n^*$ are given by

$$v = \Delta x \|u\|_n (\alpha_1, \dots, \alpha_n),$$

where $\alpha_k = \operatorname{sgn} u_k$ if $u_k \neq 0$ and $|\alpha_k| \le 1$ if $u_k = 0$. Then it is easy to see that

$$\langle A_n u, v\rangle \le 0 \quad \text{for all } v \in F_n(u),$$

which establishes the stability property.

In order to verify the consistency property we choose $D = \operatorname{dom} A = \{\phi \in C^1(0,1) \mid \phi(0) = 0\}$ which establishes condition (C1) in Proposition 3.1 with $\omega = 0$. For $u \in \operatorname{dom} A$ we define $\bar{u}_n \in X_n$ by

$$\bar{u}_n = \operatorname{col}\big(u(x_1), \dots, u(x_n)\big). \tag{4.3}$$

Then simple computations show that

$$\|E_n\bar{u}_n - u\|_{L^1} \le \Delta x\, \|u'\|_{L^1},$$

which proves $\lim_{n\to\infty} E_n\bar{u}_n = u$.

Furthermore we have

$$\|E_n A_n \bar{u}_n - Au\|_{L^1} \le \frac{1}{\Delta x}\sum_{k=1}^{n}\int_{x_{k-1}}^{x_k}\int_{x_{k-1}}^{x_k} |u'(\tau) - u'(\sigma)|\,d\sigma\,d\tau \le \omega(u'; \Delta x),$$

which proves $\lim_{n\to\infty} E_nA_n\bar{u}_n = Au$. Here $h \to \omega(u'; h)$ denotes the modulus of continuity for $u'$. Consistency now follows from Proposition 3.1.

*Case* 2. $X = L^2(0,1)$.

Let $P_n$, $E_n$ be as in Case 1 and $\|\cdot\|_n$ be given by

$$\|u\|_n^2 = \Delta x \sum_{k=1}^{n} |u_k|^2, \quad u \in X_n.$$

In this case the inner product on $X_n$ is defined by $\langle u, v\rangle_n = \langle E_n u, E_n v\rangle_{L^2}$. Then stability is obvious from

$$\langle A_n u, u\rangle_n = \sum_{k=1}^{n} \left(u_{k-1}u_k - |u_k|^2\right) \le 0, \quad u \in X_n.$$

In order to verify consistency let $D = \operatorname{dom} A$ and define $\bar{u}_n$ by (4.3). We have

$$\begin{aligned}
\|E_n\bar{u}_n - u\|_{L^2}^2 &\le \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} \Big(\int_\tau^{x_k} |u'(\sigma)| d\sigma\Big)^2 d\tau \\
&\le \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} (x_k - \tau) \int_\tau^{x_k} |u'(\sigma)|^2 d\sigma\, d\tau \\
&= \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} |u'(\sigma)|^2 \int_{x_{k-1}}^{\sigma} (x_k - \tau) d\tau\, d\sigma \\
&\le \frac{1}{2} \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} |u'(\sigma)|^2 d\sigma\, (\Delta x)^2 = \frac{1}{2}(\Delta x)^2 \|u'\|_{L^2}^2,
\end{aligned}$$

which tends to zero as $n \to \infty$. Concerning $A_n\bar{u}_n$ we have

$$\begin{aligned}
\|E_nA_n\bar{u}_n - Au\|_{L^2}^2 &\le \Big(\frac{1}{\Delta x}\Big)^2 \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} \Big(\int_{x_{k-1}}^{x_k} \big|u'(\tau) - u'(\sigma)\big|\, d\sigma\Big)^2 d\tau \\
&\le \Big(\frac{1}{\Delta x}\Big)^2 \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} \Delta x \int_{x_{k-1}}^{x_k} \big|u'(\tau) - u'(\sigma)\big|^2 d\sigma\, d\tau \\
&\le \omega(u'; \Delta x)^2 \to 0
\end{aligned}$$

as $n \to \infty$. This finishes the proof of (C2) in Proposition 3.1.

*Case* 3. $X = C_0(0,1)$.

Assume that $P_n$, $E_n$ and $\|\cdot\|_n$ are defined as

$$\begin{aligned}
&E_n u = \sum_{k=1}^{n} u_k B_k(x), \quad u \in X_n, \\
&(P_n\phi)_k = \phi(x_k), \quad 1 \le k \le n,\ \phi \in X, \\
&\|u\|_n = \max_{1\le k\le n} |u_k|, \quad u \in X_n,
\end{aligned}$$

where the first order $B$-spline $B_k(x)$, $k = 1, \dots, n$ , for $0 \le x \le 1$, is given by

$$B_k(x) = \begin{cases} n(x - x_{k-1}), & x \in [x_{k-1}, x_k], \\ n(x_{k+1} - x), & x \in [x_k, x_{k+1}], \\ 0 & \text{otherwise}. \end{cases} \tag{4.4}$$

Obviously, assumptions (A1)–(A3) are satisfied. For (A2) one has to observe that $E_nP_nu$ is the first order spline interpolating $u$ at the meshpoints.

For $u \in X_n$ the elements $v \in F_n(u)$ are given by

$$v_k = \begin{cases} \|u\|_n \operatorname{sgn} u_i & \text{for } k = i, \\ 0 & \text{for } k \neq i, \end{cases}$$

where $i$ is an index such that $|u_i| = \max_k |u_k|$. Then it is easy to see that $\langle A_nu, v\rangle \leq 0$ for all $v \in F_n(u)$, i.e., the stability property is satisfied.

For the consistency property, we again choose $D = \operatorname{dom} A$ and $\bar{u}_n \in X_n$ for $u \in \operatorname{dom} A$ as in the previous cases. Then we have

$$\lim_{n\to\infty} \|E_n\bar{u}_n - u\|_\infty = 0,$$

because $E_n\bar{u}_n$ is the first order spline interpolating the continuously differentiable function $u$ at the meshpoints. Moreover, we get for numbers $\xi_k \in (x_{k-1}, x_k)$ the estimate

$$\begin{aligned} \|E_nA_n\bar{u}_n - Au\|_\infty &= \|\sum_{k=1}^{n} \frac{u(x_{k-1}) - u(x_k)}{\Delta x} B_k + u'\|_\infty = \|u' - \sum_{k=1}^{n} u'(\xi_k)B_k\|_\infty \\ &\leq \|u' - \sum_{k=1}^{n} u'(x_k)B_k\|_\infty + \max_{k=1,\dots,n} |u'(x_k) - u'(\xi_k)| \\ &\leq \|u' - \sum_{k=1}^{n} u'(x_k)B_k\|_\infty + \omega(u'; \Delta x) \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

The first term on the right-hand side tends to zero, because $\sum_{k=1}^{n} u'(x_k)B_k$ is the first order spline interpolating the continuous function $u'$ at the meshpoints (note that $u'(0) = 0$). This finishes the proof for (C2).

4.2. **A second order wave equation in one space dimension.** This example demonstrates how to use sesquilinear forms in order to prove stability and consistency of approximations. We consider the wave equation

$$\frac{\partial^2}{\partial t^2}u(t,x) = \frac{\partial^2}{\partial x^2}u(t,x), \quad 0 \leq x \leq 1, \tag{4.5}$$

with boundary conditions

$$\begin{aligned} &u(t,0) = 0, \\ &k\frac{\partial}{\partial t}u(t,1) + \frac{\partial}{\partial x}u(t,1) = 0, \quad k > 0. \end{aligned}$$

Defining $z_1 = u$ and $z_2 = \frac{\partial}{\partial t}u$ one can write (4.5) as the system of first order equations

$$\frac{\partial}{\partial t}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{\partial^2}{\partial x^2} & 0 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{4.6}$$

with

$$z_1(t,0) = 0 \quad \text{and} \quad kz_2(t,1) + \frac{\partial}{\partial x}z_1(t,1) = 0.$$

a) *Well-posedness of the problem.* Let the Hilbert space $\tilde{V} = \{\phi \in H^1(0,1) \mid \phi(0) = 0\}$ be equipped with the norm $\|\phi\|_{\tilde{V}} = \left(\int_0^1 |\phi'|^2 dx\right)^{1/2}$ and set $Z = X = \tilde{V} \times L^2$ ($L^2 = L^2(0,1)$). The linear operator $A$ on $X$ is defined by

$$\begin{aligned}
&\operatorname{dom} A = \{(\phi_1, \phi_2) \in X \mid \phi_1 \in H^2(0,1),\ \phi_2 \in \tilde{V} \text{ and } k\phi_2(1) + \phi_1'(1) = 0\},\\
&A(\phi_1, \phi_2) = (\phi_2, \phi_1'') \quad \text{for } (\phi_1, \phi_2) \in \operatorname{dom} A.
\end{aligned}$$

It is easy to verify that $A$ is m-dissipative and thus generates a $\mathrm{C}_0$-semigroup on $X$ (note that $A$ is densely defined, because $X$ is a Hilbert space).

For $\phi = (\phi_1, \phi_2) \in \operatorname{dom} A$ and $\psi = (\psi_1, \psi_2) \in X$ we set

$$\begin{aligned}
\sigma(\phi, \psi) = \langle A\phi, \psi \rangle &= \langle \phi_2, \psi_1 \rangle_{\tilde{V}} + \langle \phi_1'', \psi_2 \rangle_{L^2}\\
&= \int_0^1 \left(\phi_2'\psi_1' + \phi_1''\psi_2\right) dx.
\end{aligned}$$

If also $\psi_2 \in \tilde{V}$, then we can integrate by parts and obtain (using also the boundary condition at $x = 1$)

$$\sigma(\phi, \psi) = \int_0^1 \left(\phi_2'\psi_1' - \phi_1'\psi_2'\right) dx - k\phi_2(1)\psi_2(1). \tag{4.7}$$

This equation makes sense for all $\phi, \psi \in V = \tilde{V} \times \tilde{V}$. Trivially $V$ is densely embedded in $X$. We define the sesquilinear form $\sigma : V \times V \to \mathbb{R}$ by (4.7). It is not difficult to see that $\phi \in V$ is in $\operatorname{dom} A$ if and only if $|\sigma(\phi, \psi)| \le K(\phi)\|\psi\|_X$ for all $\psi \in V$.

b) *The approximating spaces.* We consider a mixed finite element method and try to approximate solutions of (4.6) by

$$\begin{aligned}
z_n^{(1)}(t,x) &= \sum_{i=1}^n \alpha_i(t) B_i(x),\\
z_n^{(2)}(t,x) &= \sum_{i=1}^n \beta_i(t) S_i(x),
\end{aligned} \tag{4.8}$$

where $x_i = i/n$, $i = 0, \dots, n$, $B_i(x)$ are the first order $B$-splines defined by (4.4) and $S_i(\cdot) = \frac{1}{2}\chi_{(x_{i-1}, x_{i+1}) \cap (0,1)}$ for $i = 1, \dots, n$. We define $X_n = V_n \times H_n$, where

$$\begin{aligned}
V_n &= \{\phi \in \tilde{V} \mid \phi = \sum_{i=1}^n \alpha_i B_i,\ \alpha_i \in \mathbb{R}\},\\
H_n &= \{\psi \in L^2 \mid \psi = \sum_{i=1}^n \beta_i S_i,\ \beta_i \in \mathbb{R}\},
\end{aligned}$$

are equipped with the inner product induced from $\tilde{V}$ resp. $L^2$. As projections $X \to X_n$ we choose the orthogonal projections $P_n = (P_n^{(1)}, P_n^{(2)})$ and set $E_n = P_n^*$, i.e., $E_n$ is the canonical injection $X_n \to X$. Obviously, assumptions (A1)–(A3) are satisfied. Since $P_n^{(1)}$ is the orthogonal projection $\tilde{V} \to V_n$ with respect to the $\tilde{V}$-inner product, it is easy to see that, for $f \in \tilde{V}$, $P_n^{(1)} f$ is the first order spline which interpolates $f$ at the meshpoints $x_i = i/n$, $i = 0, \dots, n$ . Note that $f(0) = \left(P_n^{(1)} f\right)(0) = 0$.

c) *The approximating operators.* Since $X_n$ is not a subspace of $V$, we cannot define $\sigma_n$ to be the restriction of $\sigma$ to $X_n$. However, $D_n = V_n \times V_n$ is a subspace of $V$, so that we can define the sesquilinear forms $\tilde{\sigma}_n : D_n \times D_n \to \mathbb{R}$ by $\tilde{\sigma}_n = \sigma \mid_{D_n \times D_n}$. Moreover, the spaces $D_n$ are isomorphic to $X_n$, an isomorphism $i_n : D_n \to X_n$ given by

$$i_n(u,v) = \big(u, P_n^{(2)} v\big) \quad \text{for } (u,v) \in D_n.$$

A simple computation shows that

$$P_n^{(2)} \phi = \sum_{i=1}^{n} \alpha_i S_i \quad \text{for } \phi = \sum_{i=1}^{n} \alpha_i B_i \in V_n.$$

We define the sesquilinear forms $\sigma_n : X_n \times X_n$ by

$$\sigma_n(x,y) = \tilde{\sigma}_n(i_n^{-1}x, i_n^{-1}y) = \sigma(i_n^{-1}x, i_n^{-1}y), \quad x, y \in X_n,$$

and the approximating operators $A_n$ by

$$\langle A_n x, v\rangle = \sigma_n(x,v), \quad x, v \in X_n.$$

From this it is easy to compute the matrix representations for the operators $A_n$ with respect to the bases $B_1, \dots, B_n$ of $V_n$ and $S_1, \dots, S_n$ of $H_n$. Let

$$x = \Big(\sum_{i=1}^{n} \alpha_i B_i, \sum_{i=1}^{n} \beta_i S_i\Big) \quad \text{and} \quad A_n x = \big(\sum_{i=1}^{n} \gamma_i B_i, \sum_{i=1}^{n} \delta_i S_i\big).$$

We set $\alpha = \operatorname{col}(\alpha_1, \dots, \alpha_n)$, $\beta = \operatorname{col}(\beta_1, \dots, \beta_n)$, $\gamma = \operatorname{col}(\gamma_1, \dots, \gamma_n)$ and $\delta = \operatorname{col}(\delta_1, \dots, \delta_n)$. Then simple computations show that

$$Q_n \delta = -H_n \alpha - F_n \beta \quad \text{and} \quad \gamma = \beta, \tag{4.9}$$

where

$$Q_n = \frac{1}{4n}\begin{pmatrix} 2 & 1 & 0 & \cdots & 0 \\ 1 & 2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & 1 \\ 0 & \cdots & 0 & 1 & 1 \end{pmatrix}, \quad H_n = n\begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix},$$

$$F_n = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ \vdots & & 0 & 0 \\ 0 & \cdots & 0 & k \end{pmatrix}.$$

We note that the matrix

$$\begin{pmatrix} 0 & I_n \\ -Q_n^{-1} H_n & -Q_n^{-1} F_n \end{pmatrix}$$

is nonsingular, which follows from $\det Q_n \neq 0$, $\det H_n \neq 0$. This in particular implies $0 \in \rho(A_n)$, $n = 1, 2, \dots$ .

d) *The stability property.* For $x \in X_n$ we have

$$\langle A_n x, x\rangle = \sigma_n(x,x) = \sigma(i_n^{-1}x, i_n^{-1}x) \le 0$$

by (4.7), i.e., $A_n \in G(1,0,X_n)$ for all $n$.

e) *The consistency property.* We have already shown that $0 \in \rho(A) \cap \bigcap_{n=1}^{\infty} \rho(A_n)$. For $(f,g) \in X$ let $(u,v) = A^{-1}(f,g)$ and $(u_n, v_n) = A_n^{-1}P_n(f,g)$. From $A(u,v) = (v, u'') = (f,g)$ we conclude $v = f$ and $\langle u'', \psi\rangle = \langle g, \psi\rangle$ for all $\psi \in \tilde{V}$. Integration by parts and $(u,v) \in \operatorname{dom} A$ imply

$$-\langle u', \psi'\rangle = \langle g, \psi\rangle + kf(1)\psi(1) \quad \text{for all } \psi \in \tilde{V}. \tag{4.10}$$

We next derive an equation analogous to (4.10) for the approximations. By definition of the $A_n$ we have, for arbitrary $(\phi_n, \psi_n) \in X_n$,

$$\langle P_n(f,g), (\phi_n, \psi_n)\rangle = \langle A_n(u_n, v_n), (\phi_n, \psi_n)\rangle = \sigma(i_n^{-1}(u_n, v_n), i_n^{-1}(\phi_n, \psi_n)). \tag{4.11}$$

We define $\tilde{v}_n, \tilde{\psi}_n \in V_n$ by $i_n^{-1}(u_n, v_n) = (u_n, \tilde{v}_n)$ and $i_n^{-1}(\phi_n, \psi_n) = (\phi_n, \tilde{\psi}_n)$, i.e., $v_n = P_n^{(2)}\tilde{v}_n$ and $\psi_n = P_n^{(2)}\tilde{\psi}_n$. From (4.11) with $\psi_n = 0$ we obtain

$$\langle P_n^{(1)} f, \phi_n\rangle_{\tilde{V}} = \langle \tilde{v}_n, \phi_n\rangle_{\tilde{V}} \quad \text{for all } \phi_n \in V_n.$$

This proves $P_n^{(1)} f = \tilde{v}_n$ and consequently

$$v_n = P_n^{(2)} P_n^{(1)} f \ \Big(= \sum_{i=1}^{n} f(x_i) S_i\Big).$$

Again using (4.11) we get (also using $\tilde{v}_n(1) = (P_n^{(1)} f)(1) = f(1)$)

$$-\langle u_n', \tilde{\psi}_n'\rangle = \langle P_n^{(2)} g, P_n^{(2)}\tilde{\psi}_n\rangle + kf(1)\tilde{\psi}_n(1) \quad \text{for all } \tilde{\psi}_n \in V_n. \tag{4.12}$$

We choose $\bar{u}_n = P_n^{(1)} u \in V_n$. Then we get from (4.10) with $\psi = \tilde{\psi}_n$ and (4.12)

$$\begin{aligned} \langle \bar{u}_n' - u_n', \tilde{\psi}_n'\rangle &= \langle P_n^{(2)} g, P_n^{(2)}\tilde{\psi}_n\rangle - \langle g, \tilde{\psi}_n\rangle \\ &= \langle P_n^{(2)} g - g, P_n^{(2)}\tilde{\psi}_n\rangle + \langle g, P_n^{(2)}\tilde{\psi}_n - \tilde{\psi}_n\rangle \\ &= \langle g, P_n^{(2)}\tilde{\psi}_n - \tilde{\psi}_n\rangle, \quad \text{for all } \tilde{\psi}_n \in V_n. \end{aligned} \tag{4.13}$$

Here we have also used $\langle u', \tilde{\psi}_n'\rangle = \langle u, \tilde{\psi}_n\rangle_{\tilde{V}} = \langle P_n^{(1)} u, \tilde{\psi}_n\rangle_{\tilde{V}_n} = \langle \bar{u}_n', \tilde{\psi}_n'\rangle$. Equation (4.13) implies

$$\big|\langle \bar{u}_n - u_n, \tilde{\psi}_n\rangle_{\tilde{V}}\big| \le \|g\|_{L^2} \sup_{\substack{\tilde{\chi}_n \in V_n \\ \|\tilde{\chi}_n\|_{\tilde{V}} \le 1}} \|P_n^{(2)}\tilde{\chi}_n - \tilde{\chi}_n\|_{L^2}$$

for all $\tilde{\psi}_n \in V_n$ with $\|\tilde{\psi}_n\|_{\tilde{V}} \le 1$. Taking $\tilde{\psi}_n = \|\bar{u}_n - u_n\|^{-1}(\bar{u}_n - u_n)$ we get

$$\|\bar{u}_n - u_n\|_{\tilde{V}} \le \|g\|_{L^2} \sup_{\substack{\tilde{\chi}_n \in V_n \\ \|\tilde{\chi}_n\|_{\tilde{V}} \le 1}} \|P_n^{(2)}\tilde{\chi}_n - \tilde{\chi}_n\|_{L^2}. \tag{4.14}$$

By compactness of $\{\chi \in \tilde{V} \mid \|\chi\|_{\tilde{V}} \le 1\}$ in $L^2$ we see that the right-hand side of (4.14) tends to zero as $n \to \infty$. Thus we have

$$\|\bar{u}_n - u_n\|_{\tilde{V}} \to 0 \quad \text{as } n \to \infty. \tag{4.15}$$

Since $u \in H^2$ and $\bar{u}_n = P_n^{(1)} u$ is the first order spline which interpolates $u$ at the meshpoints, we also have

$$\|\bar{u}_n - u\|_{\tilde{V}} \to 0 \quad \text{as } n \to \infty. \tag{4.16}$$

Finally we get

$$\begin{aligned}\|E_nA_n^{-1}P_n(f,g) - A^{-1}(f,g)\|^2 &= \|(u_n,v_n)-(u,v)\|^2 = \|(u_n, P_n^{(2)}P_n^{(1)}f) - (u,f)\|^2\\ &= \|u_n - u\|_{\tilde V}^2 + \|P_n^{(2)}P_n^{(1)}f - f\|_{L^2}^2\\ &\le 2\|u-\bar u_n\|_{\tilde V}^2 + 2\|\bar u_n - u_n\|_{\tilde V}^2 + 2\|P_n^{(1)}f - f\|_{L^2}^2 + 2\|f - P_n^{(2)}f\|_{L^2}^2.\end{aligned}$$

From (4.15), (4.16), the fact that $P_n^{(1)}f$ is the interpolating first order spline for $f$ and that $P_n^{(2)}f$ is the orthogonal projection of $f$ onto $H_n$ we conclude that

$$\|E_nA_n^{-1}P_n(f,g) - A^{-1}(f,g)\| \to 0 \quad \text{as } n \to \infty$$

for all $(f,g) \in X$. Therefore we see from Theorem 2.1 that, for all initial conditions $\big(u(0,\cdot), \frac{\partial u}{\partial t}(0,\cdot)\big) \in X$,

$$\|E_n\big(z_n^{(1)}(t,\cdot), z_n^{(2)}(t,\cdot)\big) - \big(u(t,\cdot), \frac{\partial u}{\partial t}(t,\cdot)\big)\|_X \to 0 \quad \text{as } n \to \infty$$

uniformly on bounded $t$-intervals. The approximations $z_n^{(1)}(t,x)$, $z_n^{(2)}(t,x)$ are given by (4.8), where $\alpha(t) = \operatorname{col}(\alpha_1(t),\dots,\alpha_n(t))$ and $\beta(t) = \operatorname{col}(\beta_1(t),\dots,\beta_n(t))$ are solutions of

$$\begin{aligned}\frac{d\alpha}{dt} &= \beta,\\ Q_n\frac{d\beta}{dt} &= -H_n\alpha - F_n\beta\end{aligned}$$

with initial data $\alpha(0)$, $\beta(0)$ determined by

$$\Big(\sum_{i=1}^{n}\alpha_i(0)B_i, \sum_{i=1}^{n}\beta_i(0)S_i\Big) = P_n\big(u(0,\cdot), \frac{\partial u}{\partial t}(0,\cdot)\big).$$

4.3. **Stokes equation.** This example demonstrates that it is useful to consider situations where $X$ is a closed proper linear subspace of $Z$, because it can be very natural to choose the spaces $X_n$ as subspaces of $Z$ but not of $X$. Consider the homogeneous Stokes equation (e.g., see [14] resp. [4] for the stationary case)

$$\begin{aligned}&u_t + \operatorname{grad} p = \triangle u,\\ &\operatorname{div} u = 0, \quad x \in \Omega,\ t \ge 0, \qquad (4.17)\\ &u\,|_\Gamma = 0, \quad t \ge 0,\end{aligned}$$

where $\Omega$ is a connected bounded open set in $\mathbb{R}^N$, $N = 2,3$, with Lipschitz continuous boundary $\Gamma$. Of course, $\triangle$ denotes the Laplacian in $\mathbb{R}^N$.

a) *Well-posedness of the problem.* We shall consider solutions of (4.17) in a weak sense. We introduce the following spaces (see [4], [14]):

$$\begin{aligned}\mathcal{V} &= \{v \in \mathcal{D}(\Omega)^N \mid \operatorname{div} v = 0\},\\ V &= \text{closure of } \mathcal{V} \text{ in } W = H_0^1(\Omega)^N,\\ X &= \text{closure of } \mathcal{V} \text{ in } Z = L^2(\Omega)^N.\end{aligned}$$

Equivalently the spaces $V$ and $X$ can be defined as $V = \{v \in W \mid \operatorname{div} v = 0\}$ and $X = \{v \in Z \mid \operatorname{div} v = 0\}$, where the derivatives are understood in the distributional sense (see [4, Corollary I.2.5 and Theorem I.2.8]). On $W$ we introduce the inner product

$$\sigma(v,w)=\sum_{i=1}^{N}\int_{\Omega}\operatorname{grad}v_i\cdot\operatorname{grad}w_i\,dx,\qquad v,w\in W,$$

which is equivalent to the standard inner product. Of course, $V$ and $X$ are equipped with the inner products coming from $W$ resp. $Z$. Furthermore, $V$ is dense in $X$ with continuous injection. The inner product $\sigma(\cdot,\cdot)$ and therefore also its restriction to $V\times V$ satisfies the estimates

$$\begin{aligned}|\sigma(u,v)|&\le\|u\|_V\|v\|_V,\quad u,v\in V,\\ \sigma(u,u)&=\|u\|_V^2,\quad u\in V,\end{aligned}\tag{4.18}$$

which show that $\sigma$ is bounded and coercive. Therefore the operator $A$ defined by

$$\begin{aligned}\operatorname{dom}A=\{u\in V\mid{}&\text{there exists a }k=k(u)\text{ such that}\\ &|\sigma(u,\psi)|\le k\|\psi\|_X\text{ for all }\psi\in V\},\\ \langle Au,\psi\rangle={}&-\sigma(u,\psi),\quad u\in\operatorname{dom}A,\ \psi\in V,\end{aligned}$$

is the infinitesimal generator of an analytic semigroup $T(\cdot)$ on $X$ and, moreover, $0\in\rho(A)$. The operator $A$ is explicitly given by

$$\begin{aligned}\operatorname{dom}A&=V\cap H^2(\Omega)^N,\\ Au&=\pi\triangle u,\quad u\in\operatorname{dom}A,\end{aligned}$$

where $\pi$ is the orthogonal projection $Z\to X$ (see also [16, Section III.1]). In order to define the approximating generators we shall use the following variational formulation of (4.17) (see [4] for the stationary problem):

$$\begin{aligned}&\frac{d}{dt}\langle u(t,\cdot),\psi\rangle_Z=-\sigma(u(t,\cdot),\psi)+b(\psi,p(t,\cdot)),\quad t\ge0,\ \psi\in W,\\ &b(u(t,\cdot),\mu)=0,\quad t\ge0,\ L_0^2(\Omega),\end{aligned}\tag{4.19}$$

where $L_0^2(\Omega)=\{\chi\in L^2(\Omega)\mid\int_\Omega\chi\,dx=0\}$ and

$$b(v,\mu)=\int_\Omega\mu\operatorname{div}v\,dx,\qquad(v,\mu)\in W\times L_0^2(\Omega).$$

Note that "grad" is an isomorphism from $L_0^2(\Omega)$ onto $\{y\in H^{-1}(\Omega)^N\mid\langle y,v\rangle=0\text{ for all }v\in V\}$ (see [4, Corollary I.2.4]).

b) *Setting of the approximation framework.* For linearly independent elements $\phi_i^n\in H_0^1(\Omega)^N$, $i=1,\dots,k_n$, and $\mu_i^n\in L_0^2(\Omega)$, $i=1,\dots,m_n$, we define the spaces

$$W_n=\operatorname{span}\left(\phi_1^n,\dots,\phi_{k_n}^n\right),\qquad H_n=\operatorname{span}\left(\mu_1^n,\dots,\mu_{m_n}^n\right)$$

and the subspaces $V_n$ of $W_n$ by

$$V_n=\{\phi\in W_n\mid b(\phi,\mu)=0\text{ for all }\mu\in H_n\}$$

equipped with the $V$-norm. Furthermore, we define $X_n$ to be $V_n$ equipped with the $L^2(\Omega)^N$-norm. Note that neither $V_n$ is contained in $V$ nor is $X_n$ in $X$, because $H_n$ is a proper subspace of $L_0^2(\Omega)$.

Let $P_n$ be the orthogonal projection $Z \to X_n$ and $E_n$ be the canonical injection $X_n \to Z$. Then obviously (A1) and (A3) are satisfied. The sesquilinear forms $\sigma_n$ and the operators $A_n$ are defined by $\sigma_n = \sigma\,|_{V_n \times V_n}$ and

$$\langle A_n x, y\rangle = -\sigma_n(x,y), \quad x, y \in X_n.$$

Since (4.18) is also true for $\sigma_n$, we conclude that $0 \in \rho(A_n)$, $n = 1, 2, \ldots$, and

$$\langle A_n x, x\rangle = -\sigma(x,x) = -\|x\|_V^2 \le 0, \quad x \in X_n,$$

which establishes the stability property.

c) *The consistency property.* We impose the following conditions on the spaces $W_n$ and $H_n$:

(i) For all $u \in V$ there exist elements $w_n \in W_n$, $n = 1, 2, \ldots$, with

$$\|u - w_n\|_W \to 0 \quad \text{as } n \to \infty. \tag{4.20}$$

(ii) The uniform inf-sup condition (see [4]) is satisfied, i.e., there exists a constant $\beta > 0$ such that for all $n$

$$\sup_{w_n \in W_n \setminus \{0\}} \frac{b(w_n, \mu_n)}{\|w_n\|_W} \ge \beta \|\mu_n\|_{L^2(\Omega)} \quad \text{for all } \mu_n \in H_n. \tag{4.21}$$

We identify $W_n$ and $H_n$ with their duals and define the operator $\Psi_n : H_n \to W_n^* = W_n$ by

$$\langle \Psi_n \mu_n, w_n\rangle_W = b(w_n, \mu_n), \quad w_n \in W_n,\ \mu_n \in H_n.$$

From (4.21) we get

$$\begin{aligned} \|\Psi_n \mu_n\|_W &= \sup_{w_n \in W_n \setminus \{0\}} \frac{\langle \Psi_n \mu_n, w_n\rangle_W}{\|w_n\|_W} = \sup_{w_n \in W_n \setminus \{0\}} \frac{b(w_n, \mu_n)}{\|w_n\|_W} \\ &\ge \beta \|\mu_n\|_{L^2(\Omega)}, \quad \mu_n \in H_n. \end{aligned} \tag{4.22}$$

This proves that $\Psi_n$ is injective. The dual operator $\Psi_n^* : W_n \to H_n$ is given by $\Psi_n^* v_n = b(v_n, \cdot)$, $v_n \in W_n$. It is easy to see that $\ker \Psi_n^* = V_n$. Thus we have $\operatorname{range} \Psi_n^* = (\ker \Psi_n)^\perp = H_n$ and $\operatorname{range} \Psi_n = (\ker \Psi_n^*)^\perp = V_n^\perp$. Thus $\Xi_n^* = \Psi_n^*\,|_{V_n^\perp}$ is a bijective mapping $V_n^\perp \to H_n$ (see also [4, Lemma I.4.1]). Moreover, $\Xi_n^*$ is the adjoint of $\Xi_n : H_n \to V_n^\perp$ defined by $\Xi_n \mu_n = \Psi_n \mu_n$, $\mu_n \in H_n$. For the norms of $\Xi_n$ and $\Xi_n^*$ we get

$$\|\Xi_n\|_{\mathcal{L}(H_n, V_n^\perp)} = \|\Xi_n^*\|_{\mathcal{L}(V_n^\perp, H_n)} = \sup_{\substack{\mu_n \in H_n \setminus \{0\} \\ w_n \in V_n^\perp \setminus \{0\}}} \frac{b(w_n, \mu_n)}{\|\mu_n\|_{L^2(\Omega)} \|w_n\|_W}.$$

Therefore it follows from (4.22) that

$$\|(\Xi_n^*)^{-1}\|_{\mathcal{L}(H_n, V_n^\perp)} = \|\Xi_n^{-1}\|_{\mathcal{L}(V_n^\perp, H_n)} \le \frac{1}{\beta}. \tag{4.23}$$

Given $w_n \in W_n$ and $u \in V$ we define $f \in H_n$ by $\langle f, \mu_n\rangle_{L^2(\Omega)} = b(u - w_n, \mu_n)$ for all $\mu_n \in H_n$. We set $z_n = (\Xi_n^*)^{-1} f \in V_n^\perp$. This implies

$$\langle \Xi_n^* z_n, \mu_n\rangle_{L^2(\Omega)} = b(z_n, \mu_n) = \langle f, \mu_n\rangle_{L^2(\Omega)} = b(u - w_n, \mu_n), \quad \mu_n \in H_n.$$

Moreover, (4.23) implies

$$\begin{aligned}\|z_n\|_W &\le \frac{1}{\beta}\|f\|_{L^2(\Omega)} = \frac{1}{\beta}\sup_{\mu_n\in H_n\setminus\{0\}}\frac{\langle f,\mu_n\rangle_{L^2(\Omega)}}{\|\mu_n\|_{L^2(\Omega)}}\\ &= \frac{1}{\beta}\sup_{\mu_n\in H_n\setminus\{0\}}\frac{b(u-w_n,\mu_n)}{\|\mu_n\|_{L^2(\Omega)}} \le \frac{1}{\beta}\|b\|\,\|u-w_n\|_W,\end{aligned}$$

where $\|b\| = \sup_{v\in W\setminus\{0\},\,\mu\in L_0^2(\Omega)\setminus\{0\}} \dfrac{|b(v,\mu)|}{\|v\|_W\|\mu\|_{L_0^2(\Omega)}}$. The element $v_n = w_n + z_n$ satisfies (note that $u\in V$)

$$\begin{aligned}b(v_n,\mu_n) &= b(z_n,\mu_n) + b(w_n,\mu_n) = b(u-w_n,\mu_n) + b(w_n,\mu_n)\\ &= b(u,\mu_n) = 0 \quad \text{for all } \mu_n\in H_n,\end{aligned}$$

i.e., $v_n\in V_n$. Therefore we have

$$\|u-v_n\|_W \le \|u-w_n\|_W + \|z_n\|_W \le \Big(1+\frac{\|b\|}{\beta}\Big)\|u-w_n\|_W.$$

This and assumption (4.20) imply that for all $u\in V$ there exist $\bar v_n\in V_n$ such that

$$\lim_{n\to\infty}\|u-\bar v_n\|_W = 0. \tag{4.24}$$

Let $\phi\in X$ be given. Then by density of $V$ in $X$ there exists a sequence $(v_k)$ in $V$ with $\|\phi - v_k\|_Z\to 0$. By (4.24) there exists for each $v_k$ an element $\bar v_k\in V_k$ such that $\|v_k-\bar v_k\|_Z\le 1/k$. But then $\|\phi-\bar v_k\|_Z\to 0$ and consequently

$$\lim_{n\to\infty}\|\phi - P_n\phi\|_Z = 0 \tag{4.25}$$

(note that $X_n = V_n$ as sets). This also proves (A2). In order to establish the consistency property we first observe that $0\in\rho(A)\cap\bigcap_{n=1}^{\infty}\rho(A_n)$. For $\phi\in X$ we choose $u\in\operatorname{dom}A$ such that $\phi = Au$ and set $u_n = A_n^{-1}P_n\phi\in X_n$. For $u$ we choose $\bar v_n\in V_n$, $n=1,2,\dots$, such that (4.24) is true. By definition of $A$ and $A_n$ we have

$$\begin{aligned}\sigma(u,u_n-\bar v_n) &= \langle\phi,u_n-\bar v_n\rangle_Z,\\ \sigma(u_n,u_n-\bar v_n) &= \langle P_n\phi,u_n-\bar v_n\rangle_Z\end{aligned}$$

and consequently

$$\sigma(u_n-\bar v_n,u_n-\bar v_n) = \sigma(u-\bar v_n,u_n-\bar v_n) + \langle P_n\phi-\phi,u_n-\bar v_n\rangle_Z.$$

Observing (4.18) we get

$$\begin{aligned}\|u_n-\bar v_n\|_W &\le \|u-\bar v_n\|_W\|u_n-\bar v_n\|_W + \|P_n\phi-\phi\|_Z\|u_n-\bar v_n\|_Z\\ &\le \|u_n-\bar v_n\|_W\Big(\|u-\bar v_n\|_W + K\|P_n\phi-\phi\|_Z\Big),\end{aligned}$$

where $K$ is the embedding constant for the embedding $W\to Z$. This together with (4.24) and (4.25) implies

$$\lim_{n\to\infty}\|u_n-\bar v_n\|_W = 0. \tag{4.26}$$

Using the definitions of $u$ and $u_n$ we see that

$$\begin{aligned}\|A^{-1}\phi - E_nA_n^{-1}P_n\phi\|_W &= \|u-u_n\|_W\\ &\le \|u_n-\bar v_n\|_W + \|u-\bar v_n\|_W \to 0 \quad \text{as } n\to\infty\end{aligned}$$

by (4.24) and (4.26). This in particular implies that the consistency property (a) is true. By Theorem 2.1 we get

$$\lim_{n\to\infty} \|T_n(t)P_n\phi - T(t)\phi\|_Z = 0$$

uniformly on bounded $t$-intervals for each $\phi \in X$.

In order to see how one computes $T_n(t)P_n\phi$ set $\Phi^n = \big(\phi_1^n, \dots, \phi_{k_n}^n\big)$, $M^n = \big(\mu_1^n, \dots, \mu_{m_n}^n\big)$ and assume that $u_n(t) = \Phi^n\alpha_n(t)$ together with $p_n(t) = M^n\beta_n(t)$, $\alpha_n(t) \in \mathbb{R}^{k_n}$, $\beta_n(t) \in \mathbb{R}^{m_n}$, solve

$$\begin{aligned} &\frac{d}{dt}\langle u_n(t), \psi_n\rangle_Z = -\sigma(u_n(t), \psi_n) + b(\psi_n, p_n(t)), \quad t \ge 0,\\ &b(u_n(t), \mu_n) = 0, \quad t \ge 0, \\ &u_n(0) = P_n\phi, \end{aligned} \tag{4.27}$$

for all $\psi_n \in W_n$ and $\mu_n \in H_n$. The second equation in (4.27) implies that $u_n(t) \in X_n$, $t \ge 0$. If we take $\psi_n \in X_n$, then $b(\psi_n, p_n(t)) \equiv 0$ and (4.27) implies

$$\begin{aligned} \frac{d}{dt}\langle u_n(t), \psi_n\rangle_Z &= -\sigma(u_n(t), \psi_n) = \langle A_n u_n(t), \psi_n\rangle_Z, \quad t \ge 0,\\ u_n(0) &= P_n\phi, \end{aligned}$$

for all $\psi_n \in X_n$ or, equivalently,

$$\begin{aligned} \dot{u}_n(t) &= A_n u_n(t), \quad t \ge 0,\\ u_n(0) &= P_n\phi. \end{aligned}$$

This proves $u_n(t) = T_n(t)P_n\phi$, $t \ge 0$. Equations (4.27) imply that $\alpha_n(t)$ and $\beta_n(t)$ satisfy

$$\begin{aligned} &Q_n\dot{\alpha}_n(t) = -S_n\alpha_n(t) + B_n\beta_n(t), \quad t \ge 0,\\ &\alpha_n(t)^{\mathrm{T}}B_n = 0, \quad t \ge 0, \end{aligned} \tag{4.28}$$

where

$$Q_n = \Big(\langle \phi_i^n, \phi_k^n\rangle_Z\Big)_{i,k=1,\dots,k_n}, \quad S_n = \Big(\sigma(\phi_i^n, \phi_k^n)\Big)_{i,k=1,\dots,k_n},$$
$$B_n = \Big(b(\phi_i^n, \mu_k^n)\Big)_{\substack{i=1,\dots,k_n\\ k=1,\dots,m_n}}.$$

From (4.18) it is not difficult to conclude that $\operatorname{rank} B_n = m_n \le k_n$. The second equation together with the first equation in (4.28) implies

$$0 = B_n^{\mathrm{T}}\dot{\alpha}_n(t) = -B_n^{\mathrm{T}}Q_n^{-1}S_n\alpha_n(t) + B_n^{\mathrm{T}}Q_n^{-1}B_n\beta_n(t), \quad t \ge 0.$$

Because of $\operatorname{rank} B_n = m_n$, the matrix $R_n := B_n^{\mathrm{T}}Q_n^{-1}B_n$ is positive definite and therefore $R_n^{-1}$ exists. This implies

$$\beta_n(t) = R_n^{-1}B_n^{\mathrm{T}}Q_n^{-1}S_n\alpha_n(t), \quad t \ge 0. \tag{4.29}$$

Then the first equation in (4.28) gives

$$\dot{\alpha}_n(t) = \Big(I_n - Q_n^{-1}B_nR_n^{-1}B_n^{\mathrm{T}}\Big)Q_n^{-1}S_n\alpha_n(t), \quad t \ge 0. \tag{4.30}$$

An easy computation shows that $(\alpha_n^0)^{\mathrm{T}}B_n = 0$ implies $\alpha_n(t)^{\mathrm{T}}B_n \equiv 0$, where $\alpha_n(t)$ is the solution of (4.30) with initial value $\alpha_n^0$.

*Remark.* As already mentioned above we can prove stronger results in case the semigroup is analytic. Using the parabolic character of this problem one can show that

$$\lim_{n\to\infty} \|T_n(t)P_n\phi - T(t)\phi\|_W = 0$$

uniformly for $t$ in intervals $[1/\delta, \delta]$ for arbitrary $\delta > 1$.

## ACKNOWLEDGEMENT

The authors would like to thank an anonymous referee for his careful reading of the paper and his suggestions which improved the paper considerably.

## REFERENCES

[1] H. T. Banks and K. Ito, *A unified framework for approximation in inverse problems for distributed parameter systems*, Control-Theory and Advanced Technology **4** (1988), 73–90. MR **89d:**93034

[2] J. H. Bramble, A. H. Schatz, V. Thomée and L. B. Wahlbin, *Some convergence estimates for semidiscrete Galerkin type approximations for parabolic equations*, SIAM J. Numer. Anal. **14** (1977), 218–241. MR **56:**7231

[3] R. H. Fabiano and K. Ito, *Semigroup theory and numerical approximation for equations in linear viscoelasticity*, SIAM J. Math. Anal. **21** (1990), 374–393. MR **91b:**45021

[4] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer-Verlag, New York, 1986. MR **88b:**65129

[5] K. Ito and F. Kappel, *A uniformly differentiable approximation scheme for delay systems using splines*, Appl. Math. Optim. **23** (1991), 217–262. MR **92c:**65078

[6] K. Ito, F. Kappel and D. Salamon, *A variational approach to approximation of delay systems*, Integral and Differential Equations **4** (1991), 51–72. MR **91m:**47056

→ K. Ito and J. Turi, *Numerical methods for a class of singular integro-differential equations based on semigroup approximation*, SIAM J. Numerical Anal. **28** (1991), 1698–1722. MR **93b:**34022

[8] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1976. MR **53:**11389

[9] I. Lasiecka and A. Manitius, *Differentiability and convergence rates of approximating semigroups for retarded functional differential equations*, SIAM J. Numer. Anal. **25** (1988), 883–926. MR **89i:**65062

[10] S. V. Parter, *On the roles of "Stability" and "Convergence" in semidiscrete projection methods for initial-value problems*, Math. Comp. **34** (1980), 127–154. MR **81a:**65110

[11] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983. MR **85g:**47061

[12] R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-value Problems*, 2nd ed., Interscience Publ., New York 1967. MR **36:**3515

[13] H. Tanabe, *Equations of Evolution*, Pitman, London, 1979. MR **82g:**47032

[14] R. Temam, *Theory and Numerical Analysis of the Navier-Stokes Equations*, North-Holland, Amsterdam, 1977. MR **58:**29439

[15] H. F. Trotter, *Approximation of semi-groups of operators*, Pacific J. Math. **8** (1958), 887–919. MR **21:**2190

[16] W. von Wahl, *The Equations of Navier-Stokes and Abstract Parabolic Equations*, Vieweg, Braunschweig, 1985. MR **88a:**35195

DEPARTMENT OF MATHEMATICS, NORTH CAROLINA STATE UNIVERSITY, RALEIGH, NORTH CAROLINA 27695
*E-mail address*: `kito@eos.ncsu.edu`

INSTITUT FÜR MATHEMATIK, UNIVERSITÄT GRAZ, HEINRICHSTRASSE 36, A8010 GRAZ, AUSTRIA
*E-mail address*: `franz.kappel@kfunigraz.ac.at`