Lecture 4

The Lax Equivalence Theorem

We continue here the study of approximation theorems for semigroups by changing the field of space discretisations to time discretisations. Consider the initial value problem (abstract Cauchy problem)

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

where we suppose that A generates the strongly continuous semigroup T on the Banach space X. We saw in the previous Lecture 3 how to put spatial discretisations in an abstract setting. Generally, however, it is not usual to approximate the solution of an abstract Cauchy problem by exponential functions, but by some time discretisations, for example by using a finite difference scheme. They are in the focus of our interest this week.

Definition 4.1. Let T be semigroup with generator A, and consider the abstract Cauchy problem (ACP) on X. Consider further a strongly continuous function $F : [0, \infty) \to \mathscr{L}(X)$ with $F(0) = I$.

a) Suppose that there is $D \subseteq D(A)$ a dense subspace in X such that

$$\lim_{h \searrow 0} \frac{F(h)T(t)f - T(t+h)f}{h} = 0$$

holds for all $f \in D$ locally uniformly in t. Then we call F a **consistent finite difference scheme** (or finite difference method). To be more precise, we say that F is consistent with (ACP) on the subspace D.

b) A consistent finite difference scheme is called **stable**, if for all $t_0 > 0$ there is a constant $M \geq 1$ such that

$$\|F(h)^n\| \leq M$$

holds for all $h \geq 0$ and $n \in \mathbb{N}$ with $hn \leq t_0$.

c) A consistent finite difference scheme is called **convergent**, if for all $t > 0$, $h_k \to 0$, $n_k \to \infty$ with $h_k n_k \in [0, t]$ and $h_k n_k \to t$ we have

$$T(t)f = \lim_{k \to \infty} F(h_k)^{n_k} f$$

for all $f \in X$.

We mention the following examples.

Example 4.2. The semigroup T, i.e., the function $F(h) = T(h)$ is the best possible approximation. Certainly, this example has all the properties from the definition above, but is irrelevant from the numerical point of view.

The next example is an extremely important one, motivated by Euler's formula for the exponential function.

Example 4.3 (**Implicit Euler scheme**)**.** By Proposition 2.26.a) we know that there is $\omega \in \mathbb{R}$ such that every $\lambda \geq \omega$ belongs to the resolvent set $\rho(A)$ of A. For $h \in (0, \frac{1}{\omega}]$ we can define

$$F(h) = \tfrac{1}{h} R(\tfrac{1}{h}, A) = (I - hA)^{-1}, \quad \text{and} \quad F(0) = I.$$

(For $h > \frac{1}{\omega}$ we may set $F(h) = F(\frac{1}{\omega})$, but this is not important, because we shall be interested in small h values.) This numerical scheme is called **implicit Euler scheme**. We shall investigate its properties later.

Example 4.4 (**Crank–Nicolson scheme**)**.** We define the **Crank–Nicolson scheme** by

$$F(h) = (I + \tfrac{h}{2}A)(I - \tfrac{h}{2}A)^{-1} \quad \text{for } h \in (0, \tfrac{1}{\omega}] \quad \text{and} \quad F(0) = I.$$

Note that in our terminology the abstract Cauchy problem (ACP), and, in particular, information about the operator A is already incorporated in the finite difference scheme F. Hence, if we speak for example of the implicit Euler scheme, we mean the implicit Euler scheme for that particular problem, and not the implicit Euler scheme in general.

Consistency means in a way that the finite difference scheme is locally (i.e., for h small) a good approximation, in other words that the local error $\|F(h)f - T(h)f\|$ is small. The condition above is in applications hard to verify, since T is a priori unknown. Motivated by the finite dimensional ODE case, we relate the consistency condition to the derivative of F at $t = 0$.

Proposition 4.5. *Let $Y \subseteq D(A)$ be a Banach space that is dense in X and is continuously embedded in the Banach space $D(A)$. Suppose that it is invariant under the semigroup T so that the restriction*[1] *is a strongly continuous semigroup again. Then a finite difference scheme F is consistent with* (ACP) *on Y if and only if*

$$Af = \lim_{h \searrow 0} \frac{F(h)f - f}{h} = F'(0)f \tag{4.1}$$

holds for all $f \in Y$.

Proof. Suppose first that F is consistent. Since $Y \subseteq D(A)$, we have by definition

$$\lim_{h \searrow 0} \frac{T(h)f - f}{h} = Af \quad \text{for } f \in Y.$$

By specialising $t = 0$ in the definition of consistency we obtain

$$0 = \lim_{h \searrow 0} \frac{F(h)f - T(h)f}{h} = \lim_{h \searrow 0} \frac{F(h)f - f + f - T(h)}{h}.$$

This yields the convergence in (4.1).

For the other direction, suppose that (4.1) holds. Then the function $G : [0,1] \to \mathscr{L}(Y, X)$ defined by

$$G(h) := \begin{cases} F'(0) & \text{for } h = 0, \\ \dfrac{F(h) - I}{h} & \text{for } h \in (0,1] \end{cases}$$

[1] We restrict, of course, the semigroup operators: $T|_Y(t) = T(t)|_Y$.

is strongly continuous. First of all note that for $h > 0$ indeed $G(h) \in \mathscr{L}(Y, X)$, since Y is continuously embedded in $D(A)$ hence in X. On the other hand we have $G(0) = A \in \mathscr{L}(Y, X)$, again by the continuous embedding $Y \subseteq D(A)$. The strong continuity on $(0, 1]$ is obvious, whereas at 0 it follows from the assumption (4.1). In particular, we obtain from Proposition 2.2 that

$$\Big\|\frac{F(h) - I}{h}\Big\|_{\mathscr{L}(Y,X)} \leq M \quad \text{for all } h \in [0, 1] \text{ and for some } M \geq 0.$$

The same arguments yield

$$\Big\|\frac{T(h) - I}{h}\Big\|_{\mathscr{L}(Y,X)} \leq M \quad \text{for all } h \in [0, 1] \text{ and for some } M \geq 0.$$

We now prove consistency on Y. Take $t_0 > 0$. By assumption, T is strongly continuous on the Banach space Y, whence for $f \in Y$ fixed the set

$$C := \big\{T(s)f \,:\, t \in [0, t_0]\big\} \subseteq Y$$

is compact (being the continuous image of a closed interval). By using Proposition 2.30, we conclude

$$\frac{F(h) - I}{h} g \to Ag \quad \text{and} \quad \frac{T(h) - I}{h} g \to Ag$$

uniformly for $g \in C$ as $h \to 0$. This implies the uniform convergence

$$\frac{F(h) - T(h)}{h} T(t)f \to 0$$

for $t \in [0, t_0]$ as $h \to 0$, i.e., consistency. □

4.1 The Lax Equivalence Theorem

It turned out quite early that for partial differential equations finite difference schemes do not always converge. Famous examples are due to Richardson[2] or Courant, Friedrichs and Lewy[3].

Using the notation and the notions above, we can formulate the following fundamental result.

Theorem 4.6 (Lax Equivalence Theorem[4])**.** *For a consistent finite difference scheme, stability is equivalent to convergence.*

Proof. Suppose first that the consistent finite different scheme F is convergent but not stable. Fix $t_0 > 0$ such that

$$\sup\big\{\|F(h)^n\| : h \geq 0,\ n \in \mathbb{N},\ nh \in [0, t_0]\big\} = \infty,$$

and take a sequence $h_k \to 0$ with $n_k h_k \in [0, t_0]$ and $\|F(h_k)^{n_k}\| \to \infty$. By passing to the subsequence we may suppose that $n_k h_k \to t$ for some $t \in [0, t_0]$. Convergence and the uniform boundedness principle, Theorem 2.28, now imply boundedness of $\|F(h_k)^{n_k}\|$, hence a contradiction. So F is stable.

[2] L. F. Richardson, Weather Prediction by Numerical Process. Cambridge University Press, London 1922.

[3] R. Courant, K. Friedrichs, H. Lewy, "Über die partiellen Differenzengleichungen der mathematischen Physik," Math. Annalen **100** (1928), 32–74.

[4] P. D. Lax and R. D. Richtmyer, "Survey of the stability of linear finite difference equations", Comm. Pure Appl. Math. **9** (1956), 267–293.

Fix $t > 0$ and take sequences $h_k \to 0$, $n_k \to \infty$ with $h_k n_k \in [0, t]$ and $h_k n_k \to t$. Notice first of all that, by the strong continuity of T, it suffices to prove

$$F(h_k)^{n_k} f - T(n_k h_k) f \to 0.$$

Now, one can use the well-known algebraic identity on the difference of two n^{th} powers to obtain the "telescopic sum"

$$F(h_k)^{n_k} f - T(n_k h_k) f = F(h_k)^{n_k} f - T(h_k)^{n_k} f = \sum_{j=0}^{n_k - 1} F(h_k)^{n_k - 1 - j} \big(F(h_k) - T(h_k)\big) T(h_k)^j f. \quad (4.2)$$

From this point on, the proof is a standard epsilon-argument. Taking $f \in D$ and fixing $\varepsilon > 0$, it follows by the consistency assumption in Definition 4.1 a) that there is $N \in \mathbb{N}$ so that

$$\|F(h_k)T(s)f - T(h_k)T(s)f\| \le \varepsilon h_k$$

holds $s \in [0, t]$ for all $k \ge N$ and $t \in [0, t_0]$. For $k \ge N$ we can now estimate (4.2) as follows

$$\|F(h_k)^{n_k} f - T(n_k h_k) f\| \le \sum_{j=0}^{n_k - 1} \|F(h_k)^{n_k - 1 - j}\| \cdot \varepsilon h_k \le \sum_{j=0}^{n_k - 1} M \cdot \varepsilon h_k \ \le M t \varepsilon.$$

This proves the convergence on the space D. The claim follows then from the stability condition and the denseness of the set D in X. □

Applications of the Lax equivalence theorem are numerous. We list here some of them.

Corollary 4.7 (Implicit Euler Scheme). *Assume that the operator A generates the strongly continuous semigroup T of type $(M, 0)$ where $M \ge 1$. Then for all $f \in X$ we have*

$$T(t)f = \lim_{n \to \infty} \left(\tfrac{n}{t} R(\tfrac{n}{t}, A)\right)^n f = \lim_{n \to \infty} \left(I - \tfrac{t}{n} A\right)^{-n} f.$$

Proof. Stability follows from the properties of the generator discussed in Lecture 2, especially from equation (2.2). Let us introduce the function

$$F(t) := \begin{cases} I & \text{for } t = 0, \\ \frac{1}{t} R(\frac{1}{t}, A) & \text{for } t > 0. \end{cases}$$

Then, from the identity $\lambda R(\lambda, A) - I = AR(\lambda, A)$ we see that for $f \in D(A)$

$$\frac{F(h)f - f}{h} = \tfrac{1}{h} R(\tfrac{1}{h}, A) A f.$$

In order to apply the Lax equivalence theorem 4.6, we need to check consistency, i.e., the convergence of this expression to Af as $h \to 0$ (use Proposition 4.5). The proof can be finished by applying the following result, which we state separately because of its importance. □

Proposition 4.8. *Let A be a closed, densely defined operator. Assume that there are $M \ge 1$ and $\omega \in \mathbb{R}$ such that for all $\lambda > \omega$, we have $\lambda \in \rho(A)$ and $\|\lambda R(\lambda, A)\| \le M$. Then*

a) $\lambda R(\lambda, A) f \to f$ for all $f \in X$ as $\lambda \to \infty$, and

b) $\lambda R(\lambda, A) A f \to Af$ for all $f \in D(A)$ as $\lambda \to \infty$.

Proof. Taking $g \in D(A)$, we see that $\lambda R(\lambda, A)g = R(\lambda, A)Ag + g$. By assumption,

$$\|R(\lambda, A)Ag\| \leq \frac{M}{\lambda}\|Ag\|,$$

and hence $\lambda R(\lambda, A)g \to g$ as $\lambda \to \infty$. By the denseness of $D(A)$ and the boundedness, the convergence follows for all $g \in X$. The second statement is an immediate consequence of the first one. □

The operators $\lambda R(\lambda, A)$, $\lambda > 0$, are called **Yosida approximants**.

Remark 4.9. It can be proved that although the Crank–Nicolson scheme is consistent it is not stable for every generator. The left shift semigroup on $C_0(\mathbb{R})$ or on $L^1(\mathbb{R})$ provides a notable counterexample. We will come back to this problem in later lectures.

The following approximation formula is usually called **Lie-Trotter product formula**[5] in functional analysis and has deep applications. We will come back to it in later lectures in more detail.

Corollary 4.10. *Suppose that the operators A, B, and C are generators of strongly continuous semigroups T, S, and U, respectively. Suppose further that*

$$D(A) \cap D(B) = D(C) \text{ and for } f \in D(A) \cap D(B) \text{ we have } Cf = Af + Bf,$$

and that there is $M \geq 1$, $\omega \in \mathbb{R}$ such that

$$\left\|\left(S(\tfrac{t}{n})T(\tfrac{t}{n})\right)^n\right\| \leq M\mathrm{e}^{\omega t}.$$

Then

$$U(t)f = \lim_{n\to\infty} \left(S(\tfrac{t}{n})T(\tfrac{t}{n})\right)^n f$$

for all $f \in X$, locally uniformly in $t \geq 0$.

Proof. We introduce $F(t) = S(t)T(t)$ and check the consistency. For $f \in D(A)\cap D(B)$ we conclude that

$$\frac{F(t)f - f}{t} = S(t)\frac{T(t)f - f}{t} + \frac{S(t)f - f}{t}.$$

Clearly, we have for $f \in D(A) \cap D(B)$ by definition that

$$\frac{S(t)f - f}{t} \to Bf,$$
$$\frac{T(t)f - f}{t} \to Af,$$
$$S(t)g \to g \quad \text{for all } g \in X$$

as $t \to 0$. Since the set $\left\{\frac{T(t)f-f}{t} : t \in (0, 1]\right\} \cup \{0\}$ is compact, we can apply Proposition 2.30 to infer that

$$\lim_{t\searrow 0} \frac{F(t)f - f}{t} = Bf + Af = Cf,$$

which proves the assertion. □

[5] H. F. Trotter, "On the product of semi-groups of operators," Proc. Amer. Math. Soc. **10** (1959), 545–551.

4.2 Order of convergence

The quantitative version of the Lax equivalence theorem is quite immediate. Before stating it, we need the following definitions.

Definition 4.11. Let A generate the semigroup T on X and let F be a finite difference scheme. Suppose that there is a dense subspace $Y \subset X$ invariant under the semigroup which is a Banach space and let $p > 0$.

a) The finite difference scheme F is called **consistent of order** $p > 0$ **on** Y, if there is a subspace $Y \subset D(A)$ dense in X and invariant under the semigroup operators T, so that there is a $C > 0$ such that for all $f \in Y$ we have
$$\|F(h)f - T(h)f\| \leq Ch^{p+1}\|f\|_Y. \tag{4.3}$$

b) The finite difference scheme F is called **convergent of order** p **on** Y, if for all $t_0 > 0$ there is $K > 0$ such that for all $g \in Y$ we have
$$\|F(h)^n g - T(nh)g\| \leq K t_0 h^p \|g\|_Y$$
for all $n \in \mathbb{N}$, $h \geq 0$ with $nh \in [0, t_0]$.

We will occasionally say that the finite difference scheme has consistency/convergence order $p > 0$ on the subspace Y. Let us stress that the order p may depend on the subspace Y. This will be extremely important in applications later on.

Proposition 4.12. *Suppose that there is a subspace $Y \subset D(A)$ dense and invariant under the semigroup operators $T(t)$, which is a Banach space with its norm, satisfying*
$$\|T(t)\|_Y \leq M\mathrm{e}^{\omega t}.$$
Let F be a stable finite difference scheme which is consistent of order $p > 0$ on Y. Then the finite difference scheme is convergent of order p on Y.

Proof. The proof goes along the same lines as the one for the Lax equivalence theorem 4.6, the only difference is that we have some bound for the local error $\|F(h) - T(h)\|$. For simplicity, we may take $\omega \geq 0$. Let $t_0 > 0$ be fixed. For $g \in Y$, $n \in \mathbb{N}$ and $h \geq 0$ with $nh \in [0, t_0]$ we can write
$$\begin{aligned}\|F(h)^n g - T(nh)g\| &\leq \sum_{j=0}^{n-1} \left\|F(h)^{n-1-j}\right\| \left\|(F(h) - T(h))\, T\big(jh\big)g\right\| \\ &\leq \sum_{j=0}^{n-1} MCh^{p+1} \left\|T\big(jh\big)g\right\|_Y \leq \sum_{j=0}^{n-1} MCh^{p+1} M\mathrm{e}^{\omega jh}\|g\|_Y \\ &\leq M^2 C\mathrm{e}^{\omega t_0} t_0 h^p \|g\|_Y = K t_0 h^p \|g\|_Y. \qquad \square\end{aligned}$$

Remark 4.13. We will elaborate later on the choice of the subspace Y. For example, if $X = \mathrm{L}^2(0, \pi)$, you may think of some Sobolev space $Y \subset X$. It is a dense subset, but a Hilbert space with its own norm. Actually, we will spend quite a lot of time later on with the following two topics: Having a large scale of possible invariant subspaces Y at hand, and developing technical tools to verify the consistency estimate (4.3) for various methods.

It is possible to make the convergence of the implicit Euler scheme quantitative along these lines. Going in this direction, we note first that domains of powers of the generator are good candidates to be such subspaces Y.

Proposition 4.14. *Let A generate the semigroup T of type (M,ω), where $M \geq 1$ and $\omega \in \mathbb{R}$. For $n \in \mathbb{N}$, define*

$$X_n := D(A^n), \quad \|f\|_n := \|f\| + \|A^n f\|,$$

the domain of A^n with its graph norm. Then X_n is a Banach space, the restriction $T_n(t) := T(t)|_{X_n}$ defines a strongly continuous semigroup T_n of the same type (M,ω) in X_n.

Proof. We leave the proof of the fact that this space is a Banach space as Exercise 1. By Proposition 2.18, X_n is invariant and dense in the space X. Notice that for $f \in X_n$,

$$\|T(h)f - f\|_n = \|T(h)f - f\| + \|A^n(T(h)f - f)\| = \|T(h)f - f\| + \|T(h)A^n f - A^n f\| \to 0$$

as $h \to 0$ by the strong continuity of T in X. Finally,

$$\|T(t)f\|_n = \|T(t)f\| + \|A^n T(t)f\| \leq M\mathrm{e}^{\omega t}\|f\| + \|T(t)A^n f\| \leq M\mathrm{e}^{\omega t}\left(\|f\| + \|A^n f\|\right) = M\mathrm{e}^{\omega t}\|f\|_n$$

shows that T_n is of type (M,ω). □

The following can be considered as a very simple special case of a celebrated result by Brenner and Thomée[6] on the convergence of rational approximation schemes.

Corollary 4.15. *Let A generate the semigroup T of type $(M,0)$ where $M \geq 1$. Consider the implicit Euler scheme of Corollary 4.7. Then there is $C > 0$ such that for all $f \in D(A^2)$*

$$\left\|(I - hA)^{-n}f - T(nh)f\right\| \leq Kt_0 h\|f\|_2,$$

holds for all $n \in \mathbb{N}$, $h \geq 0$ such that $nh \in [0,t_0]$.

Proof. Stability follows from (2.2), because A is a generator. We have to deal with consistency. We prove consistency on $D(A^2)$. So take $f \in D(A^2)$, and note that

$$F(h) = (I - hA)^{-1} = \tfrac{1}{h}R(\tfrac{1}{h},A) = AR(\tfrac{1}{h},A) + I \tag{4.4}$$

holds. Hence, by Proposition 2.9, and since $f \in D(A)$ we have

$$\begin{aligned} F(h)f - T(h)f &= AR(\tfrac{1}{h},A)f - A\int_0^h T(s)f\,\mathrm{d}s. \\ &= R(\tfrac{1}{h},A)Af - \int_0^h T(s)Af\,\mathrm{d}s = \int_0^h \left(\tfrac{1}{h}R(\tfrac{1}{h},A) - T(s)\right)Af\,\mathrm{d}s. \end{aligned}$$

In a similar manner as before, we analyse the integrand. Since $g = Af \in D(A)$ we obtain

$$\left(\tfrac{1}{h}R(\tfrac{1}{h},A) - T(s)\right)g = AR(\tfrac{1}{h},A)g - A\int_0^h T(s)g = R(\tfrac{1}{h},A)Ag - \int_0^h T(s)Ag\,\mathrm{d}s.$$

[6] P. Brenner and V. Thomée, "On rational approximations of semigroups," SIAM J. Numer. Anal. **16** (1979), 683-694.

By (2.2), the inequality

$$\big\|R(\tfrac{1}{h},A)Ag\big\| \le hM\|Ag\|$$

follows. Since we are integrating a bounded function, we have

$$\left\|\int_0^h T(s)Ag\,\mathrm{d}s\right\| \le sM\|Ag\| \le hM\|Ag\|.$$

Summarizing, for $f \in D(A^2)$, the estimate

$$\|F(h)f - T(h)f\| \le h^2(2M\|A^2 f\|)$$

holds. From Proposition 4.12 the assertion follows. □

4.3 Exercises

1. Let A be the generator of a semigroup, and consider the space $X_n = D(A^n)$ with the graph norm.

a) For $n \in \mathbb{N}$ and $x \in D(A^n)$ define $|||x||| := \|x\| + \|Ax\| + \cdots + \|A^n x\|$. Prove that $|||\cdot|||$ and $\|\cdot\|_n$ are equivalent norms.

b) Prove that X_n is a Banach space.

2. Let $X = \ell^2$ and $m = (m_n)$ be a sequence with $\operatorname{Re} m_n \le 0$. Consider the semigroup T generated by the multiplication operator $A = M_m$ and define the Crank–Nicolson method as

$$F(h) = (I + \tfrac{h}{2}A)(I - \tfrac{h}{2}A)^{-1}.$$

a) Show that it is stable.

b) Show that it is consistent.

3. Consider the heat equation of Section 1.1 and show that the implicit Euler scheme converges in the operator norm, and has first order convergence.

4. Solve the exercises in the appendix.