# Solutions to the exercises in Lecture 4
# The Lax equivalence theorem

## Exercise 1

(a) For $n \in \mathbb{N}$ we show the equivalence of the two norms $|||\cdot|||$ and $\|\cdot\|_n$:

Let us assume the operator $A$ has a non-empty resolvent set $\rho(A)$. Then there exists a $\lambda \in \rho(A)$ such that the resolvent $R(\lambda, A)$ is well-defined and bounded. In particular, if $A$ is the generator of a $\mathcal{C}_0$-semigroup, this assumption is fulfilled. Therefore, we have

$$\begin{aligned}
R(\lambda, A)A^n x &= R(\lambda, A)\left(\lambda - (\lambda - A)\right) A^{n-1}x = \\
&= \lambda R(\lambda, A)A^{n-1}x - A^{n-1}x = \\
&= \lambda R(\lambda, A)\left(\lambda - (\lambda - A)\right) A^{n-2}x - A^{n-1}x = \\
&= \lambda^2 R(\lambda, A)A^{n-2}x - \lambda A^{n-2}x - A^{n-1}x = \\
&= \ldots = \\
&= \lambda^n R(\lambda, A)x - \lambda^{n-1}Ax - \ldots - A^{n-1}x
\end{aligned}$$

for $x \in D(A^n)$. As an immediate consequence we get

$$\|A^{n-1}x\| \le c_1\|x\| + \ldots + c_{n-2}\|A^{n-2}\| + c\|A^n\|$$

for some $c_1, \ldots, c_{n-2}, c > 0$. Thus,

$$|||x||| \le C\|x\|_n\,, \quad C > 0\,,$$

holds for all $x \in D(A^n)$. Finally, the converse inequality

$$\|x\|_n \le \widetilde{C}|||x|||\,, \quad \widetilde{C} > 0\,,$$

follows trivially from the definition of the two norms and this yields the equivalence of $|||\cdot|||$ and $\|\cdot\|_n$.

(b) $X_n = D(A^n)$ furnished with the norm $|||\cdot|||$ is a Banach space:

Let $\{x_m\}_{m\in\mathbb{N}}$ be a Cauchy sequence in $D(A^n)$. Since

$$|||x_m - x_k||| = \|x_m - x_k\| + \|A(x_m - x_k)\| + \ldots + \|A^n(x_m - x_k)\| \to 0$$

for $m, k \to \infty$, the sequence $\{x_m\}_{m\in\mathbb{N}}$ is Cauchy in $D(A^j)$ for all $0 \le j \le n$. Note that the domains of the powers of $A$ are nested, i.e.

$$X = D(A^0) \supseteq D(A) \supseteq \ldots \supseteq D(A^n).$$

The completeness of $X$ implies the existence of an $x \in X$ with

$$\lim_{m\to\infty} \|x_m - x\| = 0.$$

As $A$ is closed, we inductively conclude

$$\lim_{m\to\infty} A^j x_m = A^j x \quad \text{in } D(A^j)$$

by using (a) for all $1 \le j \le n$. Therefore,

$$|||x_m - x||| = \|x_m - x\| + \|A(x_m - x)\| + \ldots + \|A^n(x_m - x)\| \to 0$$

for $m \to \infty$, which proves the assertion. We point out that assuming $A$ to be closed and having a non-empty resolvent set is sufficient to conclude the completeness of $(X_n, |||\cdot|||)$.

# Exercise 2

Let $X = \ell^2$, $m = (m_n)$ a sequence with Re $m_n \le 0$ and $A : X \to X$ the multiplication operator $A = M_m$ defined for $x \in X$ by

$$(M_m x)_n = m_n x_n.$$

Consider the semigroup $T : [0, \infty) \to \mathcal{L}(X, X)$ generated by $A$, i.e. for $x \in \ell^2$,

$$(T(t)x)_n = \mathrm{e}^{m_n t} x_n.$$

We define further the Crank Nicolson method

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

elementwise for the operator $A$ and $x \in X$ by

$$(F(h)x)_n = \frac{1 + \frac{h}{2}m_n}{1 - \frac{h}{2}m_n} x_n.$$

1. Now for $x \in X$

$$(F(h)^k x)_n = \left(\frac{1+\frac{h}{2}m_n}{1-\frac{h}{2}m_n}\right)^k x_n$$

so that

$$\|F(h)^k\| = \sup_n \left|\left(\frac{1+\frac{h}{2}m_n}{1-\frac{h}{2}m_n}\right)^k\right|.$$

We see that

$$\left|\left(\frac{1+\frac{h}{2}m_n}{1-\frac{h}{2}m_n}\right)\right| = \frac{1+h\text{Re } m_n + \frac{h^2}{4}|m_n|^2}{1-h\text{Re } m_n + \frac{h^2}{4}|m_n|^2} \leq 1 \quad \forall n \in \mathbb{N},$$

since Re $m_n \leq 0$, and thus the method is stable.

2. The consistency we see from the following. Let $x \in X$. Then

$$\lim_{h\to 0}\left(\frac{F(h)T(t)x - T(t+h)x}{h}\right)_n = \frac{\frac{1+\frac{h}{2}m_n}{1-\frac{h}{2}m_n}\mathrm{e}^{m_n t} - \mathrm{e}^{m_n(t+h)}}{h}x_n$$

$$= \lim_{h\to 0}\frac{\left(\left(1+\frac{hm_n}{1-\frac{h}{2}m_n}\right) - \mathrm{e}^{m_n h}\right)\mathrm{e}^{m_n t}}{h}x_n = \lim_{h\to 0}\frac{\left(\left(1+\frac{hm_n}{1-\frac{h}{2}m_n}\right) - (1+m_n h\varphi_1(m_n h)\right)\mathrm{e}^{m_n t}}{h}x_n$$

$$= \lim_{h\to 0}\left(\frac{m_n}{1-\frac{h}{2}m_n} - m_n\varphi_1(m_n h)\right)\mathrm{e}^{m_n t}x_n,$$

where $\varphi_1$ is the entire function $\varphi_1(z) = (\mathrm{e}^z - 1)/z$, for which $\lim_{|z|\to 0}\varphi_1(z) = 1$. Thus

$$\lim_{h\to 0}\left(\frac{F(h)T(t)x - T(t+h)x}{h}\right)_n = (m_n - m_n)\,\mathrm{e}^{m_n t}x_n = 0 \quad \forall n \geq 1,$$

and the method is consistent.

# Exercise 3

We want to show that $|(I-hA)^{-n} - T(nh)| \leq \frac{C}{n}$ or equivalent that $|n(I-hA)^{-n} - nT(nh)|$ is bounded for all $n \in \mathbb{N}$. Therefore we use the representation of Proposition 1.1, i.e. we consider the operator acting on the spectral coefficients and have to bound the term

$$\left|n(1+hk^2)^{-n} - n\mathrm{e}^{-nhk^2}\right|. \tag{1}$$

In the following we use the abbreviation $\alpha := hk^2$ and distinguish the cases $\alpha \geq 1$ and $\alpha < 1$.

(i) For $\alpha \geq 1$ we easily see that

$$n(1+\alpha)^{-n} \leq \frac{n}{1+n\alpha} \leq \frac{1}{\alpha} \leq 1,$$

since $(1+\alpha)^n \geq 1+n\alpha$. Furthermore the second summand in (1) can be estimated by

$$n\mathrm{e}^{-n\alpha} \leq \frac{1}{\alpha \mathrm{e}} \leq 1.$$

Application of the triangular inequality yields the bound of (1) in this case.

(ii) For $\alpha < 1$ we represent $1+\alpha$ by

$$1+\alpha = \mathrm{e}^{\alpha+g(\alpha)}.$$

We conclude that $-1 \leq g(\alpha) \leq 0$ and that $\alpha + g(\alpha) \geq 0$, since $1 \leq 1+\alpha \leq \mathrm{e}^{\alpha}$ for $0 \leq \alpha \leq 1$. Using these bounds for $g$, we get the estimate

$$\left|n(1+\alpha)^{-n} - n\mathrm{e}^{-n\alpha}\right| = \left|n\mathrm{e}^{-n\alpha-ng(\alpha)}\left(1-\mathrm{e}^{ng(\alpha)}\right)\right| = \left|n\mathrm{e}^{-n\alpha-ng(\alpha)}\frac{1}{n}\int_{g(\alpha)}^{0} \mathrm{e}^{n\tau}\mathrm{d}\tau\right| \leq$$
$$\leq \mathrm{e}^{-n(\alpha+g(\alpha))}|g(\alpha)| \sup_{g(\alpha)\leq\tau\leq 0} \mathrm{e}^{n\tau} \leq 1$$

and proved the bound of (1) for $\alpha < 1$.

# The exercises in Appendix B.

1. As in Appendix B, we consider here a continuous function $f : [t_0, t_{\max}] \times \mathbb{R}^m \to \mathbb{R}^m$ and the initial value problem on $[t_0, t_{\max}]$:

$$\begin{cases} y'(t) &= f(t, y(t)) \\ y(t_0) &= y_0. \end{cases}$$

We consider the order of the error $e_0 = \|y_1 - y(t_0+h)\|$, where $y_1$ is the solution obtained after one step by the numerical method and $y(t_0+h)$ is the exact solution

$$y(t_0+h) = y_0 + \int_{t_0}^{t_0+h} f(t, y(t))\mathrm{d}t.$$

Expanding the integrand into Taylor series at $t = t_0$ and integrating w.r.t. $t$, we see that

$$y(t_0 + h) = y_0 + hf(t_0, y_0) + \frac{h^2}{2}(f_t(t_0, y_0) + f_y(t_0, y_0)f(t_0, y_0)) + \mathcal{O}(h^3), \quad (2)$$

where $f_t$ and $f_y$ denote the corresponding derivatives. From this we straightforwardly see that the explicit Euler method

$$y_1 = y_0 + hf(t_0, y_0)$$

is consistent of order 1. For Runge's method

$$y_1 = y_0 + hf\left(t_0 + \tfrac{h}{2}, y_0 + \tfrac{h}{2}f(t_0, y_0)\right)$$

we see by expanding $f\left(t_0 + \frac{h}{2}, y_0 + \frac{h}{2}f(t_0, y_0)\right)$ into Taylor series at $t = t_0$, that

$$y_1 = y_0 + hf(t_0, y_0) + \tfrac{h^2}{2}(f_t(t_0, y_0) + f_y(t_0, y_0)f(t_0, y_0)) + \mathcal{O}(h^3),$$

and by comparing to (2) we see that $e_0 = \mathcal{O}(h^3)$ and thus Runge's method is consistent of order 2.

2. The exponential function $\mathrm{e}^t$ is given by the solution $y(t)$ for the ordinary differential equation

$$y'(t) = y(t), \quad y(0) = 1.$$

We obtain the solution by applying the Euler method with step size $h = t/n$, $n \in \mathbb{N}$, for some fixed $t > 0$ and by letting the time step $h \to 0$ ($n \to \infty$). Applying the method, we find that

$$\begin{aligned} y_1 &= y_0 + hy_0 \\ y_2 &= y_1 + hy_1 = (1+h)^2\, y_0 \\ &= \ldots \\ y_n &= (1+h)^n y_0 = \left(1 + \tfrac{t}{n}\right)^n y_0. \end{aligned}$$

Since Euler's method is convergent, the exact solution at time $t = 1$, i.e. $y(1) = \mathrm{e}$, is the limit of the numerical solutions $y_n$ computed with step size $1/h$ as $n \to \infty$, that is

$$\mathrm{e} = \lim_{n\to\infty} \left(1 + \tfrac{1}{n}\right)^n.$$

3. The collocation polynomial $u$, defined by the $s$ distinct real numbers $c_1, \ldots, c_s$ between 0 and 1, coincides with the function $f$ at the collocation points:

$$\begin{aligned} u'(t_n + c_j h_n) &= f(t_n + c_j h_n, u(t_n + c_j h_n)), \quad \text{for } j = 1, \ldots, s \\ u(t_n) &= y_n, \end{aligned}$$

and the numerical solution $y_{n+1}$ at time $t_{n+1} = t_n + h_n$ is then defined as

$$y_{n+1} = u(t_n + h_n).$$

For the Radau IIA rule with $s = 1$ and $c_1 = 1$ we see that

$$u'(t_n + h_n) = f(t_n + h_n, u(t_n + h_n)),$$

which gives the implicit Euler method, and for which the Butcher tableau is given by

$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}$$

For the Gauss method with $s = 1$ and $c_1 = \frac{1}{2}$, we see that

$$u'(t_n + \frac{1}{2}h_n) = f\left(t_n + \tfrac{1}{2}h_n, u(t_n + \tfrac{1}{2}h_n)\right),$$

and since $u$ is first order polynomial with steepness $f\left(t_n + \frac{1}{2}h_n, u(t_n + \frac{1}{2}h_n)\right)$, we have

$$u'(t_n + h_n) = y_0 + h_n f\left(t_n + \tfrac{1}{2}h_n, u(t_n + \tfrac{1}{2}h_n)\right),$$

and the Butcher tableau is given by

$$\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array}$$

For $s = 2$ we define the coefficients using the Largrange interpolation polynomials (as discussed in Appendix B), which are defined for the collocation points $c_1, ..., c_s$ by

$$l_i(\tau) = \prod_{\substack{i=1 \\ i \neq m}}^{s} \frac{\tau - c_m}{c_i - c_m}.$$

The coefficients $b_i$ and $a_{ij}$ can be then determined by

$$b_i = \int_0^1 l_i(\tau) d\tau, \quad \text{and} \quad a_{ij} = \int_0^{c_i} l_j(\tau) d\tau$$

for $i, j = 1, ..., s$. For Radau IIA with $c_1 = \frac{1}{3}$ and $c_2 = 1$, we find that

$$l_1(\tau) = -\frac{3}{2}\tau + \frac{3}{2} \quad \text{and} \quad l_2(\tau) = \frac{3}{2}\tau - \frac{1}{2},$$

and further

$$b_1 = \int_0^1 l_1(\tau)d\tau = \left[-\frac{3}{4}\tau^2 + \frac{3}{2}\tau\right]_0^1 = \frac{3}{4}$$

$$a_{11} = \int_0^{\frac{1}{3}} l_1(\tau)d\tau = \left[-\frac{3}{4}\tau^2 + \frac{3}{2}\tau\right]_0^{\frac{1}{3}} = \frac{5}{12}$$

and so on. After computing all the 6 integrals, we get the Butcher tableau

$$\begin{array}{c|cc} \frac{1}{3} & \frac{5}{12} & -\frac{1}{12} \\ 1 & \frac{3}{4} & \frac{1}{4} \\ \hline & \frac{3}{4} & \frac{1}{4} \end{array}$$

For the Gauss method with $c_1 = \frac{1}{2} - \frac{\sqrt{3}}{6}$ and $c_2 = \frac{1}{2} + \frac{\sqrt{3}}{6}$ we similarly compute the 6 integrals for the Lagrange interpolation polynomials

$$l_1(\tau) = \frac{\tau - c_2}{c_1 - c_2} \quad \text{and} \quad l_2(\tau) = \frac{\tau - c_1}{c_2 - c_1}$$

and get the Butcher tableau

$$\begin{array}{c|cc} \frac{1}{2} - \frac{\sqrt{3}}{6} & \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{2} + \frac{\sqrt{3}}{6} & \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

4. We analyse here the stability of the above collocation methods, when applied to the problem $y' = -k^2 y$.

   Applying the Radau IIA method with $s = 1$, we get

   $$y_{n+1} = y_n - h_n k^2 y_{n+1},$$

   from which we see that

   $$\|y_{n+1}\| \leq \frac{\|y_n\|}{|1 + h_n k^2|} \leq \|y_n\|,$$

   and thus the method is stable for all $k \in \mathbb{R}$.
   Applying the Gauss method with $s = 1$, we get

   $$k_1 = y_n - \frac{h_n}{2} k^2 k_1,$$

from which we get

$$k_1 = -k^2 \left(1 + \frac{h_n}{2} k^2\right)^{-1} y_n,$$

and furthermore

$$y_{n+1} = y_n - \frac{h_n k^2}{1 + \frac{h_n}{2} k^2} y_n = \frac{1 - \frac{h_n}{2} k^2}{1 + \frac{h_n}{2} k^2} y_n,$$

and thus also for this method we find that

$$\|y_{n+1}\| \leq \|y_n\| \quad \forall k \in \mathbb{R}.$$

To analyse the stability of the $s = 2$ methods, we may write as before

$$y_{n+1} = R(h_n k^2) y_n$$

with a suitable function $R(z)$, for $z \in \mathbb{C}$. In theory of numerical ordinary differential equations these functions are called stability functions and for example for the $s$-stage Radau IIA method they are given by the $(s-1, s)$ subdiagonal Padé approximations of the exponential function (Hairer and Wanner: Solving Ordinary Differential Equations II). For $s = 2$ the stability function is given by

$$R_{1,2}(z) = \frac{1 + \frac{1}{3} z}{1 - \frac{2}{3} z + \frac{1}{3} \frac{z^2}{2!}}.$$

Thus for the above problem we get

$$y_{n+1} = \frac{1 - \frac{1}{3} h_n k^2}{1 + \frac{2}{3} h_n k^2 + \frac{1}{6} h_n^2 k^4} y_n,$$

which implies stability for all $k \in \mathbb{R}$. Respectively for $s$-stage Gauss method, the stability function is given by the $(s, s)$ diagonal Padé approximation. Thus when applying the 2-stage Gauss method to the above problem, we find that

$$y_{n+1} = \frac{1 - \frac{1}{2} h_n k^2 + \frac{1}{12} h_n^2 k^4}{1 + \frac{1}{2} h_n k^2 + \frac{1}{12} h_n^2 k^4} y_n,$$

which implies stability for all $k \in \mathbb{R}$.

5. The Crank Nicolson scheme

$$y_{n+1} = y_n + \tfrac{h_n}{2} \left(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\right) \tag{3}$$

has the form

$$y_{n+1} = y_n + h_n \left(\tfrac{1}{2} k_1 + \tfrac{1}{2} k_2\right),$$

where $k_1 = f(t_n, y_n)$ and $k_2 = f(t_{n+1}, y_{n+1})$. Thus from (3) it follows that

$$k_2 = f(t_n + h_n, y_n + h_n \left(\tfrac{1}{2}k_1 + \tfrac{1}{2}k_2\right))$$

and the Butcher tableau for the Crank-Nicholson scheme is then given by

$$\begin{array}{c|cc} & & \\ 1 & \frac{1}{2} & \frac{1}{2} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$