Scientific Computing

Ordinary Differential Equations: Numerical Methods

Exercise 13: Convergence of the Euler Method

Consider the ODE

\[ \frac{dy(t)}{dt} = Ay(t) + b \]  

with \( A \in \mathbb{R}^{N \times N} \), \( y(t) : \mathbb{R}^N \to \mathbb{R}^N \) and \( b \in \mathbb{R}^N \) (this could for example be the linear system arising from the two-species model). The explicit Euler method applied to this equation reads:

\[ y^{(n+1)} = y^{(n)} + \tau (Ay^{(n)} + b) \]  

with time step \( \tau \) and \( y^{(n)} := y(n \cdot \tau) \).

(a) Show the following statement: if the Euler method converges towards a vector \( y^* \), then \( y^* \) must be a critical point of the ODE.

(b) Under which conditions does the Euler discretisation from above (Eq. (2)) converge towards a critical point \( y^* \)?

Solution:

(a) Let the Euler method converge towards \( y^* \), that is \( y^{(n)} \xrightarrow{n \to \infty} y^* \). This implies that for the vector \( y^* \), it needs to hold:

\[ y^* = y^* + \tau (Ay^* + b) \iff 0 = \tau (Ay^* + b) \iff 0 = Ay^* + b \]

The latter equation exactly corresponds to the condition for a critical point of the ODE from Eq. (1).

(b) We first re-write the Euler update rule from Eq. (2) in a more compact form:

\[ y^{(n+1)} = y^{(n)} + \tau \left( Ay^{(n)} + b \right) = (I + \tau A)y^{(n)} + \tau b =: My^{(n)} + c, \]  

with \( M := I + \tau A \in \mathbb{R}^{N \times N} \) and \( c := \tau b \in \mathbb{R}^N \). Then, we can look at the sequence \( y^{(1)} \),
\[ y^{(2)} = My^{(0)} + c \]
\[ y^{(3)} = My^{(1)} + c = M(My^{(0)} + c) + c = M^2y^{(0)} + Mc + c \]
\[ \vdots \]
\[ y^{(n+1)} = M^{n+1}y^{(0)} + M^n c + \ldots + Mc + c = M^{n+1}y^{(0)} + \sum_{k=0}^{n} M^k c \]

We further assume that we can write the initial vector \( y^{(0)} \) and \( c \) as linear combinations of the eigenvectors \( v_i \) of \( M \):
\[ y^{(0)} = \sum_i \alpha_i v_i \]
\[ c = \sum_i \beta_i v_i \] (5)

with coefficients \( \alpha_i, \beta_i \). Inserting this into Eq. (4) yields:
\[ y^{(n+1)} = \sum_i \alpha_i \lambda_i^{n+1} v_i + \sum_{k=0}^{n} \beta_i \lambda_i^k v_i \]
\[ = \sum_i \left( \alpha_i \lambda_i^{n+1} v_i + \beta_i v_i \sum_{k=0}^{n} \lambda_i^k \right) \]
\[ = \sum_i \left( \alpha_i \lambda_i^{n+1} + \beta_i \sum_{k=0}^{n} \lambda_i^k \right) v_i \] (6)

Let’s consider the contributions of each eigenvector; therefore, we consider the cases \( \lambda_i = 1 \) and \( \lambda_i \neq 1 \):
\[ \lambda_i \neq 1 : \quad \alpha_i \lambda_i^{n+1} + \beta_i \frac{1 - \lambda_i^{n+1}}{1 - \lambda_i} \]
\[ \lambda_i = 1 : \quad \alpha_i + (n + 1) \beta_i \] (7)

For the first case in Eq. (7), we made use of the compact writing for geometric sums. In order to test for convergence towards a critical point, we can analyse the long-time behaviour of each eigenvector contribution:

- Assume that the matrix \( A \) has an eigenvalue \( \mu_i > 0 \) and a respective eigenvector \( w_i \). Then, we have:
\[ Mw_i = (I + \tau A)w_i = w_i + \tau \mu_i w_i = (1 + \tau \mu_i)w_i \]
\[ \lambda_i = 1 + \tau \mu_i \text{ is thus an eigenvalue of } M \text{ and } v_i = w_i \text{ is the corresponding eigenvector. Since } \mu_i > 0, \text{ the eigenvalue } \lambda_i \text{ must be bigger than one, } \lambda_i = 1 + \tau \mu_i > 1. \text{ As a consequence, the respective factor from Eq. (7) tends towards plus/minus infinity} \]
or remains zero:

\[
\alpha_i(1 + \tau \mu_i)^{n+1} + \beta_i \frac{1 - (1 + \tau \mu_i)^{n+1}}{-\tau \mu_i}
\]

\[
= \left( \alpha_i + \frac{\beta_i}{\tau \mu_i} \right)(1 + \tau \mu_i)^{n+1} - \frac{\beta_i}{\tau \mu_i} \rightarrow_{n \to \infty} \begin{cases} 
\infty & \text{if } \alpha_i + \frac{\beta_i}{\tau \mu_i} > 0 \\
-\infty & \text{if } \alpha_i + \frac{\beta_i}{\tau \mu_i} < 0 \\
-\frac{\beta_i}{\tau \mu_i} & \text{if } \alpha_i + \frac{\beta_i}{\tau \mu_i} = 0
\end{cases}
\]

(8)

- If we have an eigenvalue of \(A\) which fulfills \(\mu_i < -2/\tau\), then it follows \(\lambda_i = 1 + \tau \mu_i < 1 + \tau \cdot (-2/\tau) = -1\) for the corresponding eigenvalue of \(M\). For the case that \(\alpha_i + \frac{\beta_i}{\tau \mu_i} \neq 0\) in Eq. (8), no convergence is reached with the summand oscillating between \(\pm \infty\).

- If all eigenvalues of \(A\) fulfill \(-2/\tau < \mu_i < 0\), then all summands from Eq. (8) tend towards \(-\beta_i / (\tau \mu_i)\). We thus obtain:

\[
y^{(n+1)} = \sum_i \left( \alpha_i + \frac{\beta_i}{\tau \mu_i} \right)(1 + \tau \mu_i)^{n+1} - \frac{\beta_i}{\tau \mu_i} v_i \\
\rightarrow_{n \to \infty} \sum_i \left( -\frac{\beta_i}{\tau \mu_i} \right) v_i
\]

(9)

From (a), we know that the converged solution is also a critical point of our ODE. We further see that the critical point only depends on the coefficients \(\beta_i\), but not on \(\alpha_i\). This means that the critical point is solely defined via the vector \(c = \tau \cdot b\), but it does not depend on the initial condition \(y^{(0)}\). The initial condition which enters our update rule via \(M^{n+1}y^{(0)}\) decays over time and tends towards zero.

The condition \(-2/\tau < \mu_i < 0\) is typically interpreted as restriction for the time step: convergence towards a critical point can only be obtained if the time step \(\tau\) is chosen small enough such that it satisfies:

\[-\frac{2}{\tau} < \mu_i \iff \tau < -\frac{2}{\mu_i}\]

**Exercise 14: Analysis of Single-Step Methods**

Consider the ODE from last time

\[
\frac{d^2 y}{dt^2} = -y
\]

3
and its transform into a first-order system of ODEs

\[
\begin{pmatrix}
\frac{dy_0(t)}{dt} \\
\frac{dy_1(t)}{dt}
\end{pmatrix}
= \begin{pmatrix}
0 & 1 \\
-1 & 0
\end{pmatrix}
\cdot
\begin{pmatrix}
y_0(t) \\
y_1(t)
\end{pmatrix}
\tag{10}
\]

(a) Formulate the discrete update rule for the first-order system of Eq. (10) when applying the following single-step methods and using a time step \( \tau \):

- explicit Euler method
- implicit Euler method
- trapezoidal rule (Crank-Nicolson)

Write down the respective update scheme in matrix-vector form as

\[
\begin{pmatrix}
y_0^{n+1} \\
y_1^{n+1}
\end{pmatrix}
= A_{\text{method}}
\cdot
\begin{pmatrix}
y_0^n \\
y_1^n
\end{pmatrix}
\tag{11}
\]

where \( A_{\text{method}} \) denotes the method- and time step-dependent matrix for each of the single-step methods from above and \( y^n := y(n \cdot \tau) \). What can you say about the long-time behaviour of the system, that is for \((y_0^n, y_1^n)\) when \( n \to \infty \)?

(b) Write a maple sheet and check your analytical findings. You may consider solving the ODE from Eq. (10) for the initial values \( y(0) = 0, \frac{dy(0)}{dt} = 1. \)
Solution:

(a) Let’s start with the explicit Euler:

\[
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} =
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix} + \tau \cdot
\begin{pmatrix}
  0 & 1 \\
  -1 & 0
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix} =
\begin{pmatrix}
  1 & \tau \\
  -\tau & 1
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix}
\]

(12)

The implicit Euler update reads:

\[
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} =
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix} + \tau \cdot
\begin{pmatrix}
  0 & 1 \\
  -1 & 0
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} \iff
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} =
\begin{pmatrix}
  1 & -\tau \\
  \tau & 1
\end{pmatrix}^{-1}
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix}
\]

(13)

Note that for this particular scenario, an explicit update rule can be obtained for the implicit Euler with

\[
A_{\text{impl. Euler}} :=
\begin{pmatrix}
  1 & -\tau \\
  \tau & 1
\end{pmatrix}^{-1}
\]

In general, the matrix may not be known a priori, may be too complex or too huge. A simple and computationally cheap inversion as in this example is therefore often not possible. Then, a linear system of equations needs to be solved for each time step:

\[
\begin{pmatrix}
  1 & -\tau \\
  \tau & 1
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} =
\begin{pmatrix}
  1 & \tau \\
  -\tau & 1
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix}
\]

Solving this system delivers the solution \((y_{0}^{n+1}, y_{1}^{n+1})^\top\) of the next time step.

A similar arguing holds for the trapezoidal rule where both solutions of time step \(n\) and \(n+1\) are used to discretise the right hand side:

\[
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} =
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix} + \frac{\tau}{2} \cdot
\begin{pmatrix}
  0 & 1 \\
  -1 & 0
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix} + \frac{\tau}{2} \cdot
\begin{pmatrix}
  0 & 1 \\
  -1 & 0
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix}
\]

(14)

\[
\iff
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} =
\begin{pmatrix}
  \frac{1}{2} & -\frac{\tau}{2} \\
  \frac{\tau}{2} & 1
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} =
\begin{pmatrix}
  1 & \frac{\tau}{2} \\
  -\frac{\tau}{2} & 1
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix}
\]

\[
\iff
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} =
\begin{pmatrix}
  \frac{1}{2} & -\frac{\tau}{2} \\
  \frac{\tau}{2} & 1
\end{pmatrix}
\begin{pmatrix}
  \frac{1}{2} & -\frac{\tau}{2} \\
  \frac{\tau}{2} & 1
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix}
\]

\[
\iff
\begin{pmatrix}
  y_{0}^{n+1} \\
  y_{1}^{n+1}
\end{pmatrix} = \frac{1}{1 + \frac{\tau^2}{4}}
\begin{pmatrix}
  1 - \frac{\tau^2}{4} & \tau \\
  -\tau & 1 - \frac{\tau^2}{4}
\end{pmatrix}
\begin{pmatrix}
  y_{0}^{n} \\
  y_{1}^{n}
\end{pmatrix}
\]
In order to investigate the long-time behaviour, we can—similar to the previous investigations of discrete populations—analyse the eigenvalues of the update matrices

\[
A_{\text{expl. Euler}} := \begin{pmatrix} 1 & \tau \\ -\tau & 1 \end{pmatrix}
\]

\[
A_{\text{impl. Euler}} := \frac{1}{1+\tau^2} \begin{pmatrix} 1 & \tau \\ -\tau & 1 \end{pmatrix}
\]

\[
A_{\text{trapez.}} := \frac{1}{1+\tau^2} \begin{pmatrix} 1 - \frac{\tau^2}{4} & \tau \\ -\tau & 1 - \frac{\tau^2}{4} \end{pmatrix}
\]

We obtain

• for the explicit Euler method: \( \lambda_{1,2} = 1 \pm \tau i \)
  The magnitude of both eigenvalues is bigger than one. This means that the solution will increase over time. The amplitude of the sin-waves of the analytical solution will hence grow bigger and bigger over time.

• for the implicit Euler method: \( \lambda_{1,2} = \frac{1}{1 + \tau^2} \)
  These eigenvalues are exactly the inverse values of the eigenvalues from the explicit Euler method. Their magnitude is consequently < 1. This implies that the magnitude of the sin-waves will decay over time.

• for the trapezoidal rule: \( \lambda_{1,2} = \frac{4 - \tau^2 \pm 4\tau i}{\tau^2 + 4} \)
  Computing the magnitude of the eigenvalues delivers \( ||\lambda_{1,2}|| = 1 \). From this, we conclude that the magnitude of the solution is conserved over time. The amplitude of the sin-waves is thus expected to be conserved.

(b) See ws6_14b.mw.