DELFT UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING, MATHEMATICS AND COMPUTER SCIENCE

Examiner responsible: D. den Ouden-van der Horst
Examination reviewer: C. Vuik

TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS
(WI3197Minor & AESB2210-18)
February 3rd, 2023, 13:30 - 15:30

Number of questions: This is an exam with 9 open questions, subdivided in 3 main questions.
Answers All answers require arguments and/or shown calculation steps. Answers without arguments or calculation steps will give less or no points.
Electronic tools Only a non-graphical, non-programmable calculator is permitted. All other electronic tools are not permitted.
Notes, book and formula sheets Notes, books and formula sheets are not permitted.
Assessment In total 20 points can be earned. The final grade is given by $\max\{1, {}^{P}\!/_{2}\}$ rounded to one decimal, where P is the number of points earned.

1. Consider the following system of differential equations

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -4 & 1 \\ 1 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \cos(\pi t) \\ 0 \end{pmatrix}, \tag{A}$$

combined with the initial conditions $x_1(0) = 1$ and $x_2(0) = 0$ into an initial value problem.

For this initial value problem we use the following implicit numerical time integration method:

$$\begin{cases} k_1 = f(t_n + \frac{1}{2}\Delta t, w_n + \frac{1}{2}\Delta t k_1) \\ w_{n+1} = w_n + \Delta t k_1. \end{cases} \tag{B}$$

Here Δt denotes the time step and w_n represents the numerical approximation of $y(t_n)$ after n time steps.

(a) Show that the amplification factor $Q(\lambda\Delta t)$ of the above integration method (B) is given by: ($2\frac{1}{2}$ pt.)

$$Q(\lambda\Delta t) = \frac{1 + \frac{1}{2}\lambda\Delta t}{1 - \frac{1}{2}\lambda\Delta t}.$$

(b) Show that the local truncation error of the above time integration method (B) is of the order $\mathcal{O}(\Delta t^2)$ for the test equation $y' = \lambda y$. ($2\frac{1}{2}$ pt.)
Hint 1: $e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \mathcal{O}(x^4)$
Hint 2: $\frac{1}{1-x} = 1 + x + x^2 + x^3 + \mathcal{O}(x^4)$

(c) Determine for which time steps $\Delta t > 0$, the integration method (B), applied to the system (A), is stable. ($3\frac{1}{2}$ pt.)

(d) Calculate one step with the time integration method (B), in which $\Delta t = 1$ and $t_0 = 0$, applied to (A) and use the given initial conditions. (2 pt.)

Anwers to question 1

(a) The amplification factor is defined by

$$Q(\lambda\Delta t) = \frac{w_{n+1}}{w_n},$$

where w_{n+1} results from applying one step of the method to the test equation $y' = \lambda y$.
First we calculate k_1 and use $f(t, y) = \lambda y$:

$$\begin{aligned}
& & k_1 &= \lambda\left(w_n + \frac{1}{2}\Delta t k_1\right) \\
&\Rightarrow & k_1 &= \lambda w_n + \frac{1}{2}\lambda\Delta t k_1 \\
&\Rightarrow & k_1 - \frac{1}{2}\lambda\Delta t k_1 &= \lambda w_n \\
&\Rightarrow & \left(1 - \frac{1}{2}\lambda\Delta t\right) k_1 &= \lambda w_n \\
&\Rightarrow & k_1 &= \frac{\lambda}{1 - \frac{1}{2}\lambda\Delta t} w_n
\end{aligned}$$

Then we calculate w_{n+1}:

$$\begin{aligned}
w_{n+1} &= w_n + \Delta t k_1 \\
&= w_n + \frac{\lambda\Delta t}{1 - \frac{1}{2}\lambda\Delta t} w_n \\
&= \left(1 + \frac{\lambda\Delta t}{1 - \frac{1}{2}\lambda\Delta t}\right) w_n \\
&= \frac{1 + \frac{1}{2}\lambda\Delta t}{1 - \frac{1}{2}\lambda\Delta t} w_n
\end{aligned}$$

Finally division by w_n gives

$$Q(\lambda\Delta t) = \frac{1 + \frac{1}{2}\lambda\Delta t}{1 - \frac{1}{2}\lambda\Delta t}.$$

(b) The local truncation error for the test equation is given by

$$\tau_{n+1}(\Delta t) = \frac{e^{\lambda\Delta t} - Q(\lambda\Delta t)}{\Delta t} y_n., \tag{1}$$

$e^{\lambda\Delta t}$ can be expanded by the use of Taylor expansions around $\Delta t = 0$:

$$e^{\lambda\Delta t} = 1 + \lambda\Delta t + \frac{1}{2}(\lambda\Delta t)^2 + \mathcal{O}(\Delta t^3).$$

$\frac{1}{1 - \frac{1}{2}\lambda\Delta t}$ can be expanded by the use of Taylor expansions around $\Delta t = 0$:

$$\begin{aligned}
\frac{1}{1 - \frac{1}{2}\lambda\Delta t} &= 1 + \frac{1}{2}\lambda\Delta t + \left(\frac{1}{2}\lambda\Delta t\right)^2 + \mathcal{O}(\Delta t^3) \\
&= 1 + \frac{1}{2}\lambda\Delta t + \frac{1}{4}(\lambda\Delta t)^2 + \mathcal{O}(\Delta t^3).
\end{aligned}$$

This means the amplification factor can be rewritten to:

$$\begin{aligned}
Q(\lambda\Delta t) &= \left(1 + \frac{1}{2}\lambda\Delta t\right)\frac{1}{1 - \frac{1}{2}\lambda\Delta t} \\
&= \left(1 + \frac{1}{2}\lambda\Delta t\right)\left(1 + \frac{1}{2}\lambda\Delta t + \frac{1}{4}\left(\lambda\Delta t\right)^2 + \mathcal{O}(\Delta t^3)\right) \\
&= 1 + \lambda\Delta t + \frac{1}{2}\left(\lambda\Delta t\right)^2 + \mathcal{O}(\Delta t^3).
\end{aligned}$$

Substitution of the above in the local truncation error results in:

$$\begin{aligned}
\tau_{n+1} &= \frac{e^{\lambda\Delta t} - Q(\lambda\Delta t)}{\Delta t} y_n \\
&= \frac{\left(1 + \lambda\Delta t + \frac{1}{2}\left(\lambda\Delta t\right)^2 + \mathcal{O}(\Delta t^3)\right) - \left(1 + \lambda\Delta t + \frac{1}{2}\left(\lambda\Delta t\right)^2 + \mathcal{O}(\Delta t^3)\right)}{\Delta t} y_n \\
&= \frac{\mathcal{O}(\Delta t^3)}{\Delta t} y_n \\
&= \mathcal{O}(\Delta t^2).
\end{aligned}$$

(c) For stability,

$$|Q(\lambda\Delta t)| \leq 1,$$

must hold for all eigenvalues of the linear initial value problem, with Q the amplification factor of the given method.
First, we determine the eigenvalues of the matrix A. Subsequently, the eigenvalues are substituted into the amplification factor.
The eigenvalues of the matrix A are given by $\lambda_1 = -5$ and $\lambda_2 = -3$.
We first consider $\lambda_1 = -5$:

$$\begin{aligned}
Q(\lambda_1\Delta t) &= \frac{1 - \frac{5}{2}\Delta t}{1 + \frac{5}{2}\Delta t} \\
&= \frac{2 - 5\Delta t}{2 + 5\Delta t}.
\end{aligned}$$

Applying the stability criteria results in

$$-1 \leq \frac{2 - 5\Delta t}{2 + 5\Delta t} \leq 1,$$

and multiplying with the denominator gives

$$-2 - 5\Delta t \leq 2 - 5\Delta t \leq 2 + 5\Delta t.$$

First we solve the left inequality:

$$\begin{aligned}
& -2 - 5\Delta t \leq 2 - 5\Delta t \\
\Rightarrow \quad & -2 \leq 2
\end{aligned}$$

As this inequality is always true, we obtain no new information.
Then we solve the right inequality:

$$\begin{aligned}
& 2 - 5\Delta t \leq 2 + 5\Delta t \\
\Rightarrow \quad & 2 - 10\Delta t \leq 2 \\
\Rightarrow \quad & -10\Delta t \leq 0
\end{aligned}$$

As this inequality is always true for $\Delta t > 0$, we obtain no new information. Repeating this for $\lambda_2 = -3$ also results in no new information.

Therefor the time integration method applied to the initial value problem is stable for

$$\Delta t > 0.$$

(d) First we calculate $\mathbf{k}_1$, where we use $\Delta t = 1$:

$$\begin{aligned}
&& \mathbf{k}_1 &= \begin{pmatrix} -4 & 1 \\ 1 & -4 \end{pmatrix} \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{2}\mathbf{k}_1 \right) + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\Rightarrow && \mathbf{k}_1 &= \begin{pmatrix} -4 & 1 \\ 1 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} -4 & 1 \\ 1 & -4 \end{pmatrix} \mathbf{k}_1 \\
\Rightarrow && \mathbf{k}_1 - \frac{1}{2} \begin{pmatrix} -4 & 1 \\ 1 & -4 \end{pmatrix} \mathbf{k}_1 &= \begin{pmatrix} -4 & 1 \\ 1 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
\Rightarrow && \begin{pmatrix} 3 & -\frac{1}{2} \\ -\frac{1}{2} & 3 \end{pmatrix} \mathbf{k}_1 &= \begin{pmatrix} -4 \\ 1 \end{pmatrix} \\
\Rightarrow && \mathbf{k}_1 &= \begin{pmatrix} 3 & -\frac{1}{2} \\ -\frac{1}{2} & 3 \end{pmatrix}^{-1} \begin{pmatrix} -4 \\ 1 \end{pmatrix} \\
\Rightarrow && \mathbf{k}_1 &= \frac{1}{35} \begin{pmatrix} -46 \\ 4 \end{pmatrix}.
\end{aligned}$$

Then we calculate $\mathbf{w}_1$, again with $\Delta t = 1$:

$$\begin{aligned}
\mathbf{w}_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{35} \begin{pmatrix} -46 \\ 4 \end{pmatrix} \\
&= \frac{1}{35} \begin{pmatrix} -11 \\ 4 \end{pmatrix}.
\end{aligned}$$

2. We consider the following boundary-value problem:

$$\begin{cases} -y''(x) + 13y(x) &= \frac{1}{2+x}, \quad x \in [0,1], \\ y(0) &= 1, \\ y'(1) &= 0. \end{cases} \tag{C}$$

In this exercise we try to approximate the exact solution with a numerical method.

We solve the boundary value problem (C) using central finite differences with a local truncation error of $\mathcal{O}(\Delta x^2)$, upon setting $x_j = j\Delta x$, $(n+1)\Delta x = 1$, where Δx denotes the uniform step size. After discretization we obtain the following formulas:

$$\begin{aligned} -\frac{w_2 - 2w_1}{\Delta x^2} + 13w_1 &= \frac{1}{2+x_1} + \frac{1}{\Delta x^2}, \\ -\frac{w_{j+1} - 2w_j + w_{j-1}}{\Delta x^2} + 13w_j &= \frac{1}{2+x_j}, \qquad \text{for } j \in \{2, \dots, n\}, \\ -\frac{-w_{n+1} + w_n}{\Delta x^2} + \frac{13}{2}w_{n+1} &= \frac{1}{6}. \end{aligned}$$

(a) Give (with arguments) the derivation of this scheme. (3 pt.)

(b) Choose $\Delta x = 1/3$ and derive the system of equations resulting from this choice. Furthermore, rewrite this system to the form $A\mathbf{w} = \mathbf{b}$ with $\mathbf{w} = [w_1, \dots, w_{n+1}]^T$. Explicitly state A and $\mathbf{b}$ in your answer. (1 pt.)

Anwers to question 2

(a) Evaluation of the ode in $x = x_j$ and replacing $y''(x_j)$ with a finite difference of $\mathcal{O}\left(\Delta x^2\right)$ gives

$$-\frac{y(x_{j+1}) - 2y(x_j) + y(x_{j-1})}{\Delta x^2} + \mathcal{O}\left(\Delta x^2\right) + 13y(x_j) = \frac{1}{2 + x_j}.$$

Next, we neglect the truncation error and set $w_j \approx y(x_j)$ to obtain

$$-\frac{w_{j+1} - 2w_j + w_{j-1}}{\Delta x^2} + 13w_j = \frac{1}{2 + x_j}, \tag{2}$$

which is the second of the given equations.
At the left boundary, $x = 0$, we have $w_0 = 1$, which after substitution in (2) for $j = 1$ gives

$$-\frac{w_2 - 2w_1}{\Delta x^2} + 4w_1 = \frac{1}{2 + x_1} + \frac{1}{\Delta x^2},$$

which is the first of the given equations.
At the right boundary, $x = 1$, we approximate $y'(1)$ with a second-order central finite-difference, which transforms the boundary condition in:

$$\frac{y(x_{n+2}) - y(x_n)}{2\Delta x} + \mathcal{O}(\Delta x^2) = 0,$$

which after neglecting the errors results in

$$w_{n+2} = w_n.$$

Substitution of the above in (2) with $j = n + 1$ and division by two gives

$$-\frac{-w_{n+1} + w_n}{\Delta x^2} + \frac{13}{2}w_{n+1} = \frac{1}{6},$$

where we used $x_{n+1} = 1$ and which is the third of the given equations.

(b) We use $\Delta x = \frac{1}{3}$, so $n = 2$ and then, from the given equations, one obtains the following system:

$$\begin{aligned} 31w_1 - 9w_2 &= \frac{66}{7} \\ -9w_1 + 31w_2 - 9w_3 &= \frac{3}{8} \\ -9w_2 + \frac{31}{2}w_3 &= \frac{1}{6} \end{aligned}$$

This means with $\mathbf{w} = [w_1, w_2, w_3]^T$ that

$$A = \begin{bmatrix} 31 & -9 & 0 \\ -9 & 31 & -9 \\ 0 & -9 & \frac{31}{2} \end{bmatrix},$$

and

$$\mathbf{b} = \begin{bmatrix} \frac{66}{7} \\ \frac{3}{8} \\ \frac{1}{6} \end{bmatrix}.$$

3. Given is that the Trapezoidal rule satisfies

$$\left| \int_{x_L}^{x_R} f(x)\, \mathrm{d}x - \frac{x_R - x_L}{2} \left(f(x_L) + f(x_R) \right) \right| \leq \frac{1}{12} m_2 (x_R - x_L)^3,$$

where $m_2 = \max\limits_{x_L \leq x \leq x_R} |f''(x)|$.

We want to approximate the integral $\displaystyle\int_a^b f(x)\, \mathrm{d}x$ using the composite Trapezoidal rule I_T.

(a) Give the formula for the composite Trapezoidal I_T with stepsize $h = \frac{b-a}{n}$ that approximates $\displaystyle\int_a^b f(x)\, \mathrm{d}x$ and show that the composite Trapezoidal rule I_T satisfies

$$\left| \int_a^b f(x)\, \mathrm{d}x - I_T \right| \leq \frac{1}{12} M_2 (b-a) h^2,$$

where $M_2 = \max\limits_{a \leq x \leq b} |f''(x)|$. ($2\frac{1}{2}$ pt.)

(b) Approximate $\displaystyle\int_0^4 x^2\, \mathrm{d}x$ with the composite Trapezoidal rule with $h = 1$. (1 pt.)

(c) Give an appropriate upper bound for the absolute value of the error in the approximation in (b) and compare this error with the absolute value of the exact error. (2 pt.)

Anwers to question 3

(a) The composite Trapezoidal method is given by

$$I_T = \frac{h}{2}\sum_{k=1}^{n}\left(f(x_{k-1}) + f(x_k)\right).$$

Starting from the left of the inequality we can show:

$$\begin{aligned}
\left|\int_a^b f(x)\,\mathrm{d}x - I_T\right| &= \left|\sum_{k=1}^{n}\int_{x_{k-1}}^{x_k} f(x)\,\mathrm{d}x - \frac{h}{2}\sum_{k=1}^{n}\left(f(x_{k-1}) + f(x_k)\right)\right| \\
&= \left|\sum_{k=1}^{n}\left(\int_{x_{k-1}}^{x_k} f(x)\,\mathrm{d}x - \frac{h}{2}\left(f(x_{k-1}) + f(x_k)\right)\right)\right| \\
&\le \sum_{k=1}^{n}\left|\int_{x_{k-1}}^{x_k} f(x)\,\mathrm{d}x - \frac{h}{2}\left(f(x_{k-1}) + f(x_k)\right)\right| \\
&\le \sum_{k=1}^{n}\frac{1}{12}m_2^{(k)}h^3
\end{aligned}$$

where $m_2^{(k)} = \max_{x_{k-1}\le x\le x_k} |f''(x)|$

$$\begin{aligned}
&\le \sum_{k=1}^{n}\frac{1}{12}M_2h^3 \\
&= n\frac{1}{12}M_2h^3 \\
&= \frac{1}{12}M_2(b-a)h^2.
\end{aligned}$$

(b) Using $f(x) = x^2$ and $h = 1$ gives:

$$\begin{aligned}
\int_0^4 x^2\,\mathrm{d}x &\approx I_T \\
&= \frac{1}{2}\left(f(0) + 2f(1) + 2f(2) + 2f(3) + f(4)\right) \\
&= \frac{1}{2}\left(0 + 2 + 8 + 18 + 16\right) \\
&= 22
\end{aligned}$$

(c) We first calculate M_2:

$$\begin{aligned}
M_2 &= \max_{a\le x\le b} |f''(x)| \\
&= \max_{a\le x\le b} |2| \\
&= 2
\end{aligned}$$

This means

$$\begin{aligned}
\left|\int_0^4 x^2\,\mathrm{d}x - I_T\right| &\le \frac{1}{12}\cdot 2\cdot 4\cdot 1^2 \\
&= \frac{2}{3}
\end{aligned}$$

The exact integral can be calculated as

$$\int_0^4 x^2 \,\mathrm{d}x = \frac{64}{3}.$$

This means that the absolute value of the exact error is

$$\begin{aligned}\left|\int_0^4 x^2 \,\mathrm{d}x - I_T\right| &= \left|22 - \frac{64}{3}\right| \\ &= \frac{2}{3}\end{aligned}$$

This results in that the calculated upper bound for the error equals the exact error.