CTB2400

Numerieke Wiskunde

Tentamenbundel Civiele Techniek
Het Gezelschap "Practische Studie"

LET OP! EEN REPRODUCERENDE LEERSTIJL IS SCHADELIJK VOOR DE ACADEMISCHE VORMING

Augustus 2017
Juli 2017
Augustus 2016
Juni 2016
Augustus 2015
Juli 2015

Vul dit schema in om een overzicht te hebben van welke onderwerpen van het vak al goed gaan en welke je nog moet oefenen

	Augustus 2017		Juli 2017		Augustus 2016		Juni 2016		Augustus 2015		Juli 2015	
Vraag	Onderwerp	Goed?	Onderwerp	Goed?	Onderwerp	Goed?	Onderwerp	Goed?	Onderwerp	Goed?	Onderwerp	Goed?
1												
2												
3												
4												
5												

TECHNISCHE UNIVERSITEIT DELFT
FACULTEIT ELEKTROTECHNIEK, WISKUNDE EN INFORMATICA

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (CTB2400)
Donderdag 17 Augustus 2017, 18:30-21:30

1. We beschouwen de volgende methode

$$\begin{cases} w_{n+1}^{*} = w_n + \Delta t f(t_n, w_n) \\ w_{n+1} = w_n + \Delta t \left(a_1 f(t_n, w_n) + a_2 f(t_{n+1}, w_{n+1}^{*})\right) \end{cases} \tag{1}$$

voor de integratie van het **beginwaardeprobleem** $y' = f(t, y)$, $y(t_0) = y_0$

(a) Toon aan dat de *locale afbreekfout* van de bovenstaande methode van de orde $O(\Delta t)$ is als $a_1 + a_2 = 1$. Voor welke waarde van a_1 en a_2 is de locale afbreekfout van de orde $O((\Delta t)^2)$? (3 pt.)

(b) Laat zien dat de *versterkingsfactor* voor algemene a_1 en a_2 gegeven wordt door

$$Q(\lambda \Delta t) = 1 + (a_1 + a_2)\lambda \Delta t + a_2(\lambda \Delta t)^2. \tag{2}$$

(2 pt.)

(c) Beschouw $\lambda < 0$ en $(a_1 + a_2)^2 - 8a_2 < 0$. Leid de *stabiliteitsvoorwaarde* af waar Δt aan moet voldoen. (2 pt.)

We passen de methode toe op het volgende *stelsel differentiaalvergelijkingen*

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 2t \end{pmatrix}, \tag{3}$$

met beginvoorwaarde $y_1(0) = 1$ en $y_2(0) = 0$.

(d) Gebruik $\Delta t = \frac{1}{2}$ om w^1 (één tijdsstap) te berekenen met de methode waarbij $a_1 = \frac{1}{2}$ en $a_2 = \frac{1}{2}$. (1 pt.)

(e) Is de methode met $a_1 = \frac{1}{2}$ en $a_2 = \frac{1}{2}$ toegepast op (3) stabiel voor de keuze $\Delta t = \frac{1}{2}$? (motiveer uw antwoord) (2 pt.)

2. We onderzoeken **Lagrange interpolatie**. Voor gegeven steunpunten x_0, x_1, …, x_n met bijbehorende functiewaarden $f(x_0)$, $f(x_1)$, …, $f(x_n)$, wordt het interpolatiepolynoom $L_n(x)$, gegeven door

$$\begin{aligned} L_n(x) &= \sum_{k=0}^{n} f(x_k) L_{kn}(x), \text{ met} \\ L_{kn}(x) &= \frac{(x - x_0)\cdots(x - x_{k-1})(x - x_{k+1})\cdots(x - x_n)}{(x_k - x_0)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n)}. \end{aligned} \tag{4}$$

voor vervolg z.o.z.

(a) Geef het *lineaire interpolatiepolynoom van Lagrange* $L_1(x)$ met steunpunten x_0 en x_1. (1 pt.)

(b) Geef het *kwadratische interpolatiepolynoom van Lagrange* $L_2(x)$ met steunpunten x_0, x_1 en x_2. (2 pt.)

(c) Bereken $L_n(2)$ en $L_n(3)$ eerst met lineaire interpolatie en dan met kwadratische interpolatie voor de volgende meetwaarden gegeven in tabelvorm:

k	x_k	$f(x_k)$
0	1	3
1	3	6
2	4	5

(2 pt.)

3. Vervolgens willen we de **integraal** $\int_0^1 y(x)dx$ met $y(x) = x^2$ **numeriek benaderen**.

(a) Geef de *Rechthoekregel* I^R. Geef ook de bijbehorende *samengestelde integratieregels* $I^R(h)$. Benader de integraal $\int_0^1 y(x)dx$ met behulp van de samengestelde Rechtshoekregel, met $h = 1/4$. (1 pt.)

(b) Herhaal deel (a) met de *Trapeziumregel* (I^T en $I^T(h)$), met $h = 1/4$. (1 pt.)

(c) Stel dat men $\int_0^1 y(x)dx$ benadert, dan is de grootte van de fout van de *samengestelde regels* (ε_R en ε_T voor de Rechthoek- en Trapeziumregel, respectievelijk) begrensd door

$$\varepsilon_R \leq \frac{h}{2} \max_{x\in[0,1]} |y'(x)|, \quad \varepsilon_T \leq \frac{h^2}{12} \max_{x\in[0,1]} |y''(x)|. \tag{5}$$

Geef *expliciete* bovengrenzen voor de fout aan met $y(x) = x^2$. Welke methode verdient de voorkeur als het aantal integratiepunten groot is? Motiveer uw voorkeur. (1 pt.)

(d) Gebruik de volgende formule voor de fout van de *Trapeziumregel*

$$\int_0^1 y(x)dx - I^T(h) = c_p h^p$$

en leid met behulp van de Richardson methode de volgende relatie af

$$\frac{I^T(2h) - I^T(4h)}{I^T(h) - I^T(2h)} = 2^p$$

Bereken daarmee de numerieke approximatie orde p voor $h = 1/4$. (1 pt.)

(e) Leid de volgende relatie af

$$\int_0^1 y(x)dx - I^T(h) = \frac{Q(h) - Q(2h)}{2^p - 1}$$

en gebruik deze om een schatting van de fout van de *Trapeziumregel* voor $h = 1/4$ te berekenen. (1 pt.)

Voor de uitwerkingen van dit tentamen zie:
http://ta.twi.tudelft.nl/nw/users/vuik/wi3097/tentamen.html

DELFT UNIVERSITY OF TECHNOLOGY
Faculty of Electrical Engineering, Mathematics and Computer Science

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (CTB2400)
Thursday August 17th 2017, 18:30-21:30

1. (a) The local truncation error is defined as

$$\tau_{n+1}(\Delta t) = \frac{y_{n+1} - z_{n+1}}{\Delta t}, \tag{1}$$

where z_{n+1} is given by

$$z_{n+1} = y_n + \Delta t \left(a_1 f(t_n, y_n) + a_2 f(t_n + \Delta t, y_n + \Delta t f(t_n, y_n)\right). \tag{2}$$

A Taylor expansion of f around (t_n, y_n) yields

$$f(t_n+\Delta t, y_n+\Delta t f(t_n, y_n)) = f(t_n, y_n)+\Delta t\frac{\partial f}{\partial t}(t_n, y_n)+\Delta t f(t_n, y_n)\frac{\partial f}{\partial y}(t_n, y_n)+O(\Delta t^2). \tag{3}$$

This is substituted into equation (2) to obtain

$$z_{n+1} = y_n+\Delta t \left(a_1 f(t_n, y_n) + a_2 \left[f(t_n, y_n) + \Delta t\frac{\partial f}{\partial t}(t_n, y_n) + \Delta t f(t_n, y_n)\frac{\partial f}{\partial y}(t_n, y_n)\right]\right)+O(\Delta t^3). \tag{4}$$

From the differential equation we know that:

$$y'(t_n) = f(t_n, y_n) \tag{5}$$

From the Chain Rule of Differentiation, we derive

$$y''(t_n) = \frac{df(t_n, y_n)}{dt} = \frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} y'(t_n) \tag{6}$$

after substitution of the differential equation one obtains:

$$y''(t_n) = \frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} f(t_n, y_n) \tag{7}$$

Substituting these expressions into (4) shows that

$$z_{n+1} = y_n + \Delta t(a_1 + a_2)y'(t_n) + \Delta t^2 a_2 y''(t_n) + O(\Delta t^3). \tag{8}$$

A Taylor series for $y(t)$ around t_n gives for y_{n+1}

$$y_{n+1} = y(t_n + \Delta t) = y_n + \Delta t y'(t_n) + \frac{\Delta t^2}{2} y''(t_n) + O(\Delta t^3). \tag{9}$$

Equations (9) and (8) are substituted into relation (1) to obtain

$$\tau_{n+1}(\Delta t) = y'(t_n)(1 - (a_1 + a_2)) + \Delta t y''(t_n)\left(\frac{1}{2} - a_2\right) + O(\Delta t^2) \tag{10}$$

Hence

(a) $a_1 + a_2 = 1$ implies $\tau_{n+1}(\Delta t) = O(\Delta t)$;

(b) $a_1 + a_2 = 1$ and $a_2 = \frac{1}{2}$, that is, $a_1 = a_2 = \frac{1}{2}$, gives $\tau_{n+1}(\Delta t) = O(\Delta t^2)$.

(b) The test equation is given by

$$y' = \lambda y. \tag{11}$$

Application of the predictor step to the test equation gives

$$w^*_{n+1} = w_n + \lambda \Delta t w_n = (1 + \lambda \Delta t) w_n. \tag{12}$$

The corrector step yields

$$w_{n+1} = w_n + \Delta t\,(a_1 \lambda w_n + a_2 \lambda (1 + \lambda \Delta t) w_n) = (1 + (a_1 + a_2)\lambda \Delta t + a_2 \lambda^2 \Delta t^2) w_n. \tag{13}$$

Hence the amplification factor is given by

$$Q(\Delta t \lambda) = 1 + (a_1 + a_2)\lambda \Delta t + a_2 (\lambda \Delta t)^2. \tag{14}$$

(c) Let $\lambda < 0$ (so λ is real), then, for stability, the amplification factor must satisfy

$$-1 \leq Q(\lambda \Delta t) \leq 1, \tag{15}$$

from the previous assignment, we have

$$-1 \leq 1 + (a_1 + a_2)\lambda \Delta t + a_2 (\lambda \Delta t)^2 \leq 1 \Leftrightarrow -2 \leq (a_1 + a_2)\lambda \Delta t + a_2 (\lambda \Delta t)^2 \leq 0. \tag{16}$$

First, we consider the left inequality:

$$a_2 (\lambda \Delta t)^2 + (a_1 + a_2)\lambda \Delta t + 2 \geq 0 \tag{17}$$

For $\lambda \Delta t = 0$, the above inequality is satisfied, further the discriminant is given by $(a_1 + a_2)^2 - 8a_2 < 0$. Here the last inequality follows from the given hypothesis. Hence the left inequality in relation (16) is always satisfied. Next we consider the right hand inequality of relation (16)

$$a_2 (\lambda \Delta t)^2 + (a_1 + a_2)\lambda \Delta t \leq 0. \tag{18}$$

This relation is rearranged into

$$a_2 (\lambda \Delta t)^2 \leq -(a_1 + a_2)\lambda \Delta t, \tag{19}$$

hence

$$a_2 |\lambda \Delta t|^2 \leq (a_1 + a_2)|\lambda \Delta t| \Leftrightarrow |\lambda \Delta t| \leq \frac{a_1 + a_2}{a_2}, \qquad a_2 \neq 0. \tag{20}$$

This results into the following condition for stability

$$\Delta t \leq \frac{a_1 + a_2}{a_2 |\lambda|}, \qquad a_2 \neq 0. \tag{21}$$

(d) We use the method with $a_1 = \frac{1}{2}$ and $a_2 = \frac{1}{2}$ and $\Delta t = \frac{1}{2}$. Let

$$A = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix}, \quad \underline{w}^1 = \begin{pmatrix} w_1^1 \\ w_2^1 \end{pmatrix}, \quad \underline{w}^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{22}$$

where the subscript stands for the component, whereas the superscript denotes the time–index. First, we carry out the prediction step

$$\underline{\hat{w}}^1 = \underline{w}^0 + \Delta t A \underline{w}^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ -4 \end{pmatrix} = \begin{pmatrix} 1 \\ \\ -2 \end{pmatrix}. \tag{23}$$

Subsequently, we perform the corrector step

$$\underline{w}^1 = \underline{w}^0 + \frac{\Delta t}{2}\left(A\underline{w}^0 + A\underline{\hat{w}}_1 + \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right). \tag{24}$$

Using $\Delta t = \frac{1}{2}$, gives

$$\underline{w}^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{4}\left(\begin{pmatrix} 0 \\ -4 \end{pmatrix} + \begin{pmatrix} -2 \\ -4 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} \frac{1}{2} \\ \\ -1\frac{3}{4} \end{pmatrix}. \tag{25}$$

(e) Before we can investigate the stability of the method we first have to determine the eigenvalues of the matrix

$$\begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix} \tag{26}$$

It is easy to see that this matrix has two complex eigenvalues $\lambda_1 = 2i$ and $\lambda_2 = -2i$. It is sufficient to investigate if $|Q(\lambda_1 \Delta t)| \leq 1$ because $|Q(\lambda_2 \Delta t)| = |Q(\lambda_1 \Delta t)|$. Using $a_1 = \frac{1}{2}$ and $a_2 = \frac{1}{2}$ and $\lambda_1 = 2i$ we obtain the following expression for $Q(\lambda_1 \Delta t)$

$$Q(\lambda_1 \Delta t) = 1 + \lambda_1 \Delta t + \frac{1}{2}(\lambda_1 \Delta t)^2 = 1 + 2\Delta t i - 2\Delta t^2$$

Substituting $\Delta t = \frac{1}{2}$ gives:

$$Q(\lambda_1 \Delta t) = 1 + i - \frac{1}{2}$$

Since $|Q(\lambda_1 \Delta t)| = \sqrt{(\frac{1}{2})^2 + 1^2)} = \sqrt{\frac{5}{4}} > 1$ the method is unstable for $\Delta t = \frac{1}{2}$.

2. (a) The **linear Lagrangian interpolatory polynomial**, with nodes x_0 and x_1, is given by

$$L_1(x) = \frac{x - x_1}{x_0 - x_1} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1). \tag{27}$$

This is evident from application of the given formula.

(b) The **quadratic Lagrangian interpolatory polynomial** with nodes x_0, x_1 and x_2 is given by

$$\begin{aligned} L_2(x) &= \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}f(x_0) & (28)\\ &+ \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}f(x_1) & (29)\\ &+ \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}f(x_2). & (30)\end{aligned}$$

This is also evident from application of the given formula.

(c) Obviously, $L_1(3) = 6$ and $L_2(3) = 6$ since the Lagrange interpolation polynomial satisfies $L_n(x_k) = f(x_k)$ for all points $x_0, x_1, \ldots, x_n$. Next, we compute $L_1(2)$ and $L_2(2)$ for both linear and quadratic Lagrangian interpolation as approximations at $x = 3$. For **linear interpolation**, we have

$$L_1(3) = \frac{2-3}{1-3}\cdot 3 + \frac{2-1}{3-1}\cdot 6 = \frac{9}{2}, \tag{31}$$

and for **quadratic interpolation**, one obtains

$$L_2(3) = \frac{(2-3)(2-4)}{(1-3)(1-4)}\cdot 3 + \frac{(2-1)(2-4)}{(3-1)(3-4)}\cdot 6 + \frac{(2-1)(2-3)}{(4-1)(4-3)}\cdot 5 = \frac{16}{3}. \tag{32}$$

3. (a) Consider an interval of integration $[x_{j-1}, x_j]$, then the **Rectangle Rule** reads

$$I_j^R = hf(x_{j-1}), \qquad h = x_j - x_{j-1}. \tag{33}$$

The *composed integration rule* is derived by

$$I^R = h(I_1^R + I_2^R + \ldots + I_n^R) = h(f(x_0) + \ldots + f(x_{n-1})), \tag{34}$$

which yields

$$I^R(h = 1/4) = \frac{1}{4}\cdot(0 + (\frac{1}{4})^2 + (\frac{2}{4})^2 + (\frac{3}{4})^2) = \frac{7}{32}. \tag{35}$$

(b) For the interval of integration $[x_{j-1}, x_j]$ the **Trapezoidal Rule** is

$$I_j^T = \frac{h}{2}(f(x_{j-1}) + f(x_j)). \tag{36}$$

The *composed integration rule* is derived by

$$I^T = h(I_1^T + I_2^T + \ldots + I_n^T) = h(\frac{f(x_0)}{2} + f(x_1) + \ldots + f(x_{n-1}) + \frac{f(x_n)}{2}), \tag{37}$$

which leads to

$$I^T(h = 1/4) = \frac{1}{4}\cdot(0 + (\frac{1}{4})^2 + (\frac{2}{4})^2 + (\frac{3}{4})^2 + \frac{1}{2}) = \frac{11}{32}. \tag{38}$$

(c) For a general number of subintervals, say n, the magnitude of the composed Rectangle- and Trapezoidal Rules, is bounded from above by

$$\begin{aligned} \varepsilon_R &\leq \frac{h}{2} \max_{x\in[0,1]} |y'(x)| \leq h = \frac{1}{n}, \\ \varepsilon_T &\leq \frac{h^2}{12} \max_{x\in[0,1]} |y''(x)| \leq \frac{h^2}{6} = \frac{1}{6n^2}. \end{aligned} \tag{39}$$

Here, the exact solution $y(x) = x^2$ was used. Hence, the error from the Trapezoidal Rule is much smaller. Furthermore, from the composed Rules, it is easy to see that the number of function evaluations for the composed Rectangle- and Trapezoidal Rules is given by n and $n+1$, respectively. Since

$$\lim_{n\to\infty} \frac{n+1}{n} = 1, \tag{40}$$

it follows that the amount of work for the Trapezoidal Rule is not significantly higher than it is for the Rectangle Rule. Hence, it is more attractive to use the Trapezoidal Rule.

(d) It follows from the given error formula that

$$\int_0^1 y(x)dx - I^T(h) = c_p h^p \tag{41}$$

$$\int_0^1 y(x)dx - I^T(2h) = c_p (2h)^p \tag{42}$$

$$\int_0^1 y(x)dx - I^T(4h) = c_p (4h)^p \tag{43}$$

By subtracting equation (42) from (43) and equation (41) from (42), the unknown exact integral value can be eliminated

$$I^T(2h) - I^T(4h) = c_p (2h)^p (2^p - 1) \tag{44}$$

$$I^T(h) - I^T(2h) = c_p (h)^p (2^p - 1) \tag{45}$$

By dividing these two expressions we obtain

$$\frac{I^T(2h) - I^T(4h)}{I^T(h) - I^T(2h)} = 2^p \tag{46}$$

From part (b) we know that $I^T(h = 1/4) = 11/32$. Moreover,

$$I^T(h = 1/2) = \frac{1}{2} \cdot (0 + (\frac{1}{2})^2 + \frac{1}{2}) = \frac{3}{8}. \tag{47}$$

and

$$I^T(h=1) = 1 \cdot (0 + \frac{1}{2}) = \frac{1}{2}. \tag{48}$$

Filling these three values into the above error formula yields

$$\frac{\frac{3}{8} - \frac{1}{2}}{\frac{11}{32} - \frac{3}{8}} = 4 = 2^p \tag{49}$$

from which it follows that $p = 2$.

(e) By dividing equation (45) by $(2^p - 1)$ we obtain

$$\frac{I^T(h) - I^T(2h)}{2^p - 1} = c_p(h)^p \tag{50}$$

which, according to the given error formula (41) equals

$$\frac{I^T(h) - I^T(2h)}{2^p - 1} = c_p(h)^p = \int_0^1 y(x)dx - I^T(h) \tag{51}$$

It follows that the error of the Trapozoidal Rule for $h = 1/4$ can be estimated as

$$\frac{\frac{11}{32} - \frac{3}{8}}{4 - 1} = -0.01042 \tag{52}$$

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 6 juli 2017, 18:30-21:30

1. We beschouwen de numerieke integratie van het volgende **beginwaardeprobleem**

$$y' = f(t, y), \quad y(t_0) = y_0. \tag{1}$$

We gebruiken de *voorwaartse methode van Euler* om de numerieke oplossing van dit beginwaardeprobleem (1) te bepalen. Deze methode is gegeven door

$$w_{n+1} = w_n + \Delta t f(t_n, w_n), \tag{2}$$

waarin Δt de tijdstap en w_n de numerieke oplossing op tijdstip t_n voorstelt.

(a) Bepaal de orde van de locale afbreekfout. (2 pt.)

(b) Geef de versterkingsfactor voor deze methode. Voor welke Δt is de methode stabiel als λ een negatief reëel getal is? (2 pt.)

(c) We beschouwen het beginwaarde probleem:

$$y'' = -y' - \frac{1}{2}y, \;\; y(0) = 1, \;\; y'(0) = 0.$$

Schrijf deze tweede orde differentiaalvergelijking als een stelsel eerste orde differentiaalvergelijkingen: $\mathbf{x}' = \mathbf{A}\mathbf{x}$. Toon aan dat de eigenwaarden van $\mathbf{A}$ gegeven worden door

$$\lambda_1 = -\frac{1}{2} + \frac{1}{2}i \text{ en } \lambda_2 = -\frac{1}{2} - \frac{1}{2}i.$$

(2 pt.)

(d) Doe 1 stap met de voorwaartse methode van Euler toegepast op het stelsel met $\Delta t = 1$. (2 pt.)

(e) Onderzoek de stabiliteit van de voorwaartse methode van Euler voor dit stelsel voor een algemene $\Delta t > 0$. (2 pt.)

[0] **voor vervolg z.o.z. Voor de uitwerkingen van dit tentamen zie: http://ta.twi.tudelft.nl/nw/users/vuik/wi3097/tentamen.html**

2. We beschouwen het volgende **iteratieproces** $x_{n+1} = g(x_n)$, met

$$g(x_n) = x_n + h(x_n)(x_n^3 - 27),$$

waarbij h een continue functie is met $h(x) \neq 0$ voor elke $x \neq 0$.

   (a) Als dit proces convergeert, naar welke limiet p convergeert het dan? (1 pt.)

   (b) Beschouw drie mogelijke keuzen voor $h(x)$:

      (i) $h_1(x) = -\frac{1}{x^4}$

      (ii) $h_2(x) = -\frac{1}{x^2}$

      (iii) $h_3(x) = -\frac{1}{3x^2}$

   Voor welke keuze kan het proces niet convergeren? Voor welke keuze convergeert het proces het snelst? Motiveer uw antwoord. (2 pt.)

   (c) Bepaal een functie $h_4(x)$ zodat de 'convergentiefaktor' één is. (1 pt.)

   (d) Laat p een nulpunt van een gegeven functie f zijn. $\hat{f}$ is de functie verstoord door meetfouten. Er is gegeven dat $|\hat{f}(x) - f(x)| \leq \epsilon_{max}$ voor alle x. Laat zien dat voor het nulpunt $\hat{p}$ van $\hat{f}$ geldt $|\hat{p} - p| \leq \frac{\epsilon_{max}}{|f'(p)|}$. (1 pt.)

3. We beschouwen we het volgende **randwaardeprobleem**

$$\begin{cases} -y''(x) + (x+1)y(x) = x^3 + x^2 - 2, & 0 < x < 1, \\ \\ y'(0) = 0, \quad y(1) = 1, \end{cases} \tag{3}$$

met $y' = \frac{dy}{dx}$ en $y'' = \frac{d^2y}{dx^2}$.

   (a) We willen het randwaardeprobleem (3) met behulp van de eindige differentie methode oplossen. Laat $x_j = j\Delta x$, $(n+1)\Delta x = 1$ waarin de stapgrootte Δx constant is. Geef een discretisatie (+ bewijs) met

   - locale afbreekfout van $\mathcal{O}((\Delta x)^2)$;
   - verwerking van de randvoorwaarden;
   - en een symmetrische discretisatie matrix.

   Gebruik een virtueel gridpunt voor de randvoorwaarde op $x = 0$. (2.5 pt.)

   (b) Gebruik een stapgrootte van $\Delta x = 1/3$ om het stelsel vergelijkingen $\mathbf{Aw} = \mathbf{f}$ voor de discretisatie uit deel (a) af te leiden. Verwerk de randvoorwaarden. Het afgeleide stelsel moet 3×3 zijn (drie onbekenden en drie vergelijkingen).

   Opmerking: U hoeft het stelsel **niet** op te lossen. (1 pt.)

   (c) Omdat de 3×3 discretisatie matrix $\mathbf{A}$ symmetrisch is, zijn alle eigenwaarden reëel. Maak gebruik van de Gershgorin-cirkel stelling om een schatting voor de kleinste eigenwaarde $|\lambda|_{\min}$ te berekenen. Concludeer hieruit dat het eindige differentie schema uit (a) stabiel is. Dat betekent dat $\mathbf{A}^{-1}$ bestaat en dat er een constante C bestaat zó dat $\|\mathbf{A}^{-1}\| \leq C$ als $\Delta x \to 0$. (1.5 pt.)

DELFT UNIVERSITY OF TECHNOLOGY
Faculty of Electrical Engineering, Mathematics and Computer Science

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Thursday July 6 2017, 18:30-21:30

1. (a) The local truncation error is defined by

$$\tau_h = \frac{y_{n+1} - z_{n+1}}{\Delta t}, \tag{1}$$

where

$$z_{n+1} = y_n + \Delta t f(t_n, y_n), \tag{2}$$

for the Forward Euler method. A Taylor expansion for y_{n+1} around t_n is given by

$$y_{n+1} = y_n + \Delta t y'(t_n) + \frac{\Delta t^2}{2} y''(\xi), \quad \exists\, \xi \in (t_n, t_{n+1}). \tag{3}$$

Since $y'(t_n) = f(t_n, y_n)$, we use equation (1), to get

$$\tau_h = \frac{\Delta t}{2} y''(\xi), \quad \exists\, \xi \in (t_n, t_{n+1}). \tag{4}$$

Hence, the truncation error is of first order.

(b) For the amplification factor we apply the method to the test equation: $y' = \lambda y$. Application of Forward Euler to this equation leads to:

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

so the amplification factor is $Q(\lambda \Delta t) = 1 + \lambda \Delta t$.

We have to check that $|Q(\lambda \Delta t)| \le 1$. For a negative real number λ this leads to the inequalities:

$$-1 \le 1 + \lambda \Delta t \le 1$$

The right hand inequality leads to $\lambda \Delta t \le 0$. Since $\Delta t > 0$ and $\lambda \le 0$ this inequality is always satisfied. The left hand inequality leads to $-1 \le 1 + \lambda \Delta t$ which is equavalent to $\lambda \Delta t \ge -2$. Dividing both sides by λ which is negative leads to:

$$\Delta t \le \frac{2}{-\lambda}.$$

(c) We use the following definition $x_1 = y$ and $x_2 = y'$. This implies that $x_1' = y' = x_2$ and $x_2' = y'' = -y' - \frac{1}{2}y = -x_2 - \frac{1}{2}x_1$. Writing this in vector notation shows that

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

so $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix}$. To compute the eigenvalues we look for values of λ such that

$$|\mathbf{A} - \lambda \mathbf{I}| = 0.$$

This implies that λ is a solution of

$$\lambda^2 + \lambda + \frac{1}{2} = 0,$$

which leads to the roots:

$$\lambda_1 = -\frac{1}{2} + \frac{1}{2}i \text{ and } \lambda_2 = -\frac{1}{2} - \frac{1}{2}i.$$

(d) We do one step with Forward Euler using $\Delta t = 1$.

$$\begin{bmatrix} w_{1,1} \\ w_{2,1} \end{bmatrix} = \begin{bmatrix} w_{1,0} \\ w_{2,0} \end{bmatrix} + \Delta t \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} w_{1,0} \\ w_{2,0} \end{bmatrix}$$

Substituting $\Delta t = 1$ and the initial conditions leads to:

$$\begin{bmatrix} w_{1,1} \\ w_{2,1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix}$$

(e) Since the eigenvalues are complex valued it is sufficient to check that the modulus: $|Q(\lambda_1 \Delta t)| \leq 1$. Substituting $\lambda_1 = -\frac{1}{2} + \frac{1}{2}i$ into $Q(\lambda_1 \Delta t)$ leads to the condition:

$$|1 + \Delta t(-\frac{1}{2} + \frac{1}{2}i)| \leq 1$$

This implies that

$$\sqrt{(1 - \frac{\Delta t}{2})^2 + (\frac{\Delta t}{2})^2} \leq 1$$

Rearranging the terms leads to

$$1 - \Delta t + \frac{1}{2}(\Delta t)^2 \leq 1$$

so

$$-\Delta t + \frac{1}{2}(\Delta t)^2 \leq 0$$

and thus

$$\Delta t \leq 2$$

.

2. (a) The iteration process is a fixed-point method. If the process converges we have: $\lim_{n\to\infty} x_n = p$. Using this in the iteration process yields:

$$\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} [x_n + h(x_n)(x_n^3 - 27)]$$

Since h is a continuous function one obtains:

$$p = p + h(p)(p^3 - 27)$$

so

$$h(p)(p^3 - 27) = 0.$$

Since $h(x) \neq 0$ for each $x \neq 0$ it follows that $p^3 - 27 = 0$ and thus $p = 27^{\frac{1}{3}} = 3$.

(b) The convergence of a fixed-point method $x_{n+1} = g(x_n)$ is determined by $g'(p)$. If $|g'(p)| < 1$ the method converges, whereas if $|g'(p)| > 1$ the method diverges. For all choices we compute the first derivative in p. For the first method we elaborate all steps. For the other methods we only give the final result. For h_1 we have $g_1(x) = x - \frac{x^3-27}{x^4}$. The first derivative is:

$$g_1'(x) = 1 - \frac{3x^2 \cdot x^4 - (x^3 - 27) \cdot 4x^3}{(x^4)^2}$$

Substitution of p yields:

$$g_1'(p) = 1 - \frac{3p^6 - (p^3 - 27) \cdot 4p^3}{p^8}.$$

Since $p = 3$ the final term cancels:

$$g_1'(p) = 1 - \frac{3p^6}{p^8} = 1 - 3^{-1} = \frac{2}{3}.$$

This implies that the method is convergent with convergence factor $\frac{2}{3}$.

For the second method we have:

$$g_2'(p) = 1 - \frac{3p^4 - (p^3 - 27) \cdot 2p}{p^4} = 1 - \frac{3p^4}{p^4} = -2$$

Thus the method diverges.

For the third method we have:

$$g_3'(p) = 1 - \frac{9p^4 - (p^3 - 27) \cdot 6p}{9p^4} = 1 - \frac{9p^4}{9p^4} = 0$$

Thus the method is convergent with convergence factor 0.

Concluding we note that the third method is the fastest.

(c) For a general function $h_4(x)$ the first derivative of $g_4(x) = x + h_4(x)(x^3 - 27)$ evaluated in p reads

$$g_4'(p) = 1 + h_4'(p)(p^3 - 27) + 3h_4(3)p^2$$

Since $p = 3$ we obtain $g_4'(3) = 1 + 27h_4(3)$. For $|g_4'(3)| = 1$ we need to find a differentiable function $h_4(x)$ that equals 0 in $p = 3$. A possible choice is

$$h_4(x) = x - 3.$$

(d) To estimate the error in p we first approximate the function f in the neighbourhood of p by the first order Taylor polynomial:

$$P_1(x) = f(p) + (x - p)f'(p) = (x - p)f'(p).$$

Due to the measurement errors we know that

$$(x - p)f'(p) - \epsilon_{max} \leq \hat{P}_1(x) \leq (x - p)f'(p) + \epsilon_{max}.$$

This implies that the perturbed root $\hat{p}$ is bounded by the roots of $(x-p)f'(p) - \epsilon_{max}$ and $(x - p)f'(p) + \epsilon_{max}$, which leads to

$$p - \frac{\epsilon_{max}}{|f'(p)|} \leq \hat{p} \leq p + \frac{\epsilon_{max}}{|f'(p)|}.$$

3. (a) Using central differences for the second order derivative at a node $x_j = j\Delta x$ gives

$$y''(x_j) \approx \frac{y_{j+1} - 2y_j + y_{j-1}}{\Delta x^2} =: Q(\Delta x). \tag{5}$$

Here, $y_j := y(x_j)$. Next, we will prove that this approximation is second order accurate, that is $|y''(x_j) - Q(\Delta x)| = \mathcal{O}(\Delta x^2)$.
Using Taylor's Theorem around $x = x_j$ gives

$$\begin{aligned} y_{j+1} &= y(x_j + \Delta x) = y(x_j) + \Delta x y'(x_j) + \tfrac{\Delta x^2}{2} y''(x_j) + \tfrac{\Delta x^3}{3!} y'''(x_j) + \tfrac{\Delta x^4}{4!} y''''(\eta_+), \\ y_{j-1} &= y(x_j - \Delta x) = y(x_j) - \Delta x y'(x_j) + \tfrac{\Delta x^2}{2} y''(x_j) - \tfrac{\Delta x^3}{3!} y'''(x_j) + \tfrac{\Delta x^4}{4!} y''''(\eta_-). \end{aligned} \tag{6}$$

Here, η_+ and η_- are numbers within the intervals (x_j, x_{j+1}) and (x_{j-1}, x_j), respectively. Substitution of these expressions into $Q(\Delta x)$ gives

$$|y''(x_j) - Q(\Delta x)| = \mathcal{O}(\Delta x^2).$$

This leads to the following discretisation formula for internal grid nodes:

$$\frac{-w_{j-1} + 2w_j - w_{j+1}}{\Delta x^2} + (x_j + 1)w_j = x_j^3 + x_j^2 - 2. \tag{7}$$

Here, w_j represents the numerical approximation of the solution y_j. To deal with the boundary $x = 0$, we use a virtual node at $x = -\Delta x$, and we define $y_{-1} := y(-\Delta x)$. Then, using central differences at $x = 0$ gives

$$0 = y'(0) \approx \frac{y_1 - y_{-1}}{2\Delta x} =: Q_b(\Delta x). \tag{8}$$

Using Taylor's Theorem, gives

$$\begin{aligned}
Q_b(\Delta x) &= \\
&= \frac{y(0) + \Delta x y'(0) + \frac{\Delta x^2}{2} y''(0) + \frac{\Delta x^3}{3!} y'''(\eta_+)}{2\Delta x} \\
&- \frac{y(0) - \Delta x y'(0) + \frac{\Delta x^2}{2} y''(0) - \frac{\Delta x^3}{3!} y'''(\eta_-)}{2\Delta x} \\
&= y'(0) + \mathcal{O}(\Delta x^2).
\end{aligned}$$

Again, we get an error of $\mathcal{O}(\Delta x^2)$.

(b) With respect to the numerical approximation at the virtual node, we get

$$\frac{w_1 - w_{-1}}{2\Delta x} = 0 \quad \Leftrightarrow \quad w_{-1} = w_1. \tag{9}$$

The discretisation at $x = 0$ is given by

$$\frac{-w_{-1} + 2w_0 - w_1}{\Delta x^2} + w_0 = -2. \tag{10}$$

Substitution of equation (9) into the above equation, yields

$$\frac{2w_0 - 2w_1}{\Delta x^2} + w_0 = -2. \tag{11}$$

Subsequently, we consider the boundary $x = 1$. To this extent, we consider its neighbouring point x_{n-1} and substitute the boundary condition $w_n = y(1) = y_n = 1$ into equation (7) to obtain

$$\begin{aligned}
&\frac{-w_{n-2} + 2w_{n-1}}{\Delta x^2} + (x_{n-1} + 1)w_{n-1} && (12) \\
&= x_{n-1}^3 + x_{n-1}^2 - 2 + \frac{1}{\Delta x^2} && (13) \\
&= (1 - \Delta x)^3 + (1 - \Delta x)^2 - 2 + \frac{1}{\Delta x^2}. && (14)
\end{aligned}$$

This concludes our discretisation of the boundary conditions. In order to get a symmetric discretisation matrix, one divides equation (11) by 2.

Next, we use $\Delta x = 1/3$. From equations (7, 11, 14) we obtain the following system

$$\begin{aligned} 9\frac{1}{2}w_0 - 9w_1 &= -1 \\ -9w_0 + 19\frac{1}{3}w_1 - 9w_2 &= -\frac{50}{27} \\ -9w_1 + 19\frac{2}{3}w_2 &= \frac{209}{27}. \end{aligned}$$

(c) The Gershgorin circle theorem states that the eigenvalues of a square matrix $\mathbf{A}$ are located in the complex plane in the union of circles

$$|z - a_{ii}| \leq \sum_{\substack{j \neq i \\ j=1}}^{n} |a_{ij}| \quad \text{where} \quad z \in \mathbb{C} \tag{15}$$

For the 3×3 matrix derived in part (b) we have

- For $i = 1$:
$$\left| z - 9\frac{1}{2} \right| \leq 9 \quad \Rightarrow \quad |\lambda_1|_{\min} \geq \frac{1}{2} \tag{16}$$
- For $i = 2$:
$$\left| z - 19\frac{1}{3} \right| \leq 18 \quad \Rightarrow \quad |\lambda_2|_{\min} \geq 1\frac{1}{3} \tag{17}$$
- For $i = 3$:
$$\left| z - 19\frac{2}{3} \right| \leq 9 \quad \Rightarrow \quad |\lambda_3|_{\min} \geq 10\frac{2}{3} \tag{18}$$

Hence, a lower bound for the smallest eigenvalue is $\frac{1}{2}$. For a symmetric matrix $\mathbf{A}$ we have

$$\|\mathbf{A}^{-1}\| = \frac{1}{|\lambda|_{\min}} \leq 2 \tag{19}$$

This proves that the finite-difference scheme is stable, e.g., with constant $C = 2$.

TECHNISCHE UNIVERSITEIT DELFT
FACULTEIT ELEKTROTECHNIEK, WISKUNDE EN INFORMATICA

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (CTB2400 WI3097TU)
Donderdag 11 Augustus 2016, 18:30-21:30

1. We beschouwen de volgende methode

$$\begin{cases} w_{n+1}^* = w_n + \Delta t f(t_n, w_n) \\ w_{n+1} = w_n + \Delta t \left(a_1 f(t_n, w_n) + a_2 f(t_{n+1}, w_{n+1}^*) \right) \end{cases} \tag{1}$$

voor de integratie van het **beginwaardeprobleem** $y' = f(t, y)$, $y(t_0) = y_0$.

(a) Toon aan dat de *locale afbreekfout* van de bovenstaande methode van de orde $O(\Delta t)$ is als $a_1 + a_2 = 1$. Voor welke waarde van a_1 en a_2 is de locale afbreekfout van de orde $O((\Delta t)^2)$? (3 pt.)

(b) Laat zien dat de *versterkingsfactor voor algemene* a_1 en a_2 gegeven wordt door

$$Q(\lambda \Delta t) = 1 + (a_1 + a_2)\lambda \Delta t + a_2 (\lambda \Delta t)^2. \tag{2}$$

(2 pt.)

(c) Beschouw $\lambda < 0$ en $(a_1 + a_2)^2 - 8a_2 < 0$, leid de *stabiliteitsvoorwaarde* af waar Δt aan moet voldoen. (2 pt.)

(d) We beschouwen het volgende *stelsel niet lineaire differentiaalvergelijkingen*:

$$\begin{aligned} x_1' &= -\sin x_1 + 2x_2 + t, \ \ x_1(0) = 0, \\ x_2' &= x_1 - x_2^2, \ \ x_2(0) = 1. \end{aligned} \tag{3}$$

Laat zien dat de Jacobiaan van het rechterlid van (3) op $t = 0$ gegeven wordt door:

$$\begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix}.$$

(1.5 pt.)

(e) Kies $a_1 = a_2 = \frac{1}{2}$. Voor welke waarden van Δt is de methode toegepast op (3) stabiel op $t = 0$? (1.5 pt.)

2. We zoeken een **differentie formule** van de vorm:

$$Q(h) = \frac{\alpha_0}{h^2} f(0) + \frac{\alpha_{-1}}{h^2} f(-h) + \frac{\alpha_{-2}}{h^2} f(-2h),$$

zodat

$$f''(0) - Q(h) = \mathcal{O}(h).$$

(a) Geef het *lineaire stelsel vergelijkingen* waar α_0, α_{-1} en α_{-2} aan moeten voldoen. (2 pt.)

voor vervolg z.o.z.

x	$f(x)$
0	0
$-\frac{1}{4}$	0.0156
$-\frac{1}{2}$	0.1250
$-\frac{3}{4}$	0.4219
- 1	1.0000

Tabel 1: De gebruikte waarden

(b) De oplossing van het in het vorige onderdeel afgeleide stelsel wordt gegeven door $\alpha_0 = 1$, $\alpha_{-1} = -2$ en $\alpha_{-2} = 1$. Geef voor deze waarden een uitdrukking voor de *afbreekfout* $f''(0) - Q(h)$. (1 pt.)

(c) Geef met behulp van de *Richardson methode* een schatting van de fout $f''(0) - Q(\frac{1}{4})$. Gebruik daarvoor de getallen gegeven in Tabel 1. (2 pt.)

3. We beschouwen de **convectie-diffusie vergelijking** met Dirichlet randvoorwaarden:

$$\begin{cases} -\epsilon u'' + u' = 1, \qquad 0 < x < 1, \\ \\ u(0) = 0, \qquad u(1) = 0, \end{cases} \tag{4}$$

waarin $u = u(x)$, $u' = \frac{du}{dx}$ en $u'' = \frac{d^2u}{dx^2}$.

(a) Laat zien dat

$$u(x) = x - \frac{1 - e^{x/\epsilon}}{1 - e^{1/\epsilon}}, \tag{5}$$

de *exacte oplossing* is van het randwaardeprobleem (4). (1 pt.)

(b) We lossen het randwaardeprobleem (4) op met behulp van *centrale eindige differenties* voor de diffusieve term en *upwind eindige differenties* voor de convectieve term.

Voor alle *inwendige punten* x_j geldt de discretisatievergelijking

$$-\epsilon \frac{w_{j+1} - 2w_j + w_{j-1}}{(\Delta x)^2} + \frac{w_j - w_{j-1}}{\Delta x} = 1, \text{ for } j \in \{1, \ldots, n\}. \tag{6}$$

met $x_j = j\Delta x$, $(n+1)\Delta x = 1$, waarbij Δx de uniforme stapgrootte is.

Geef de *discretisatievergelijkingen* voor de twee randpunten x_1 en x_n. (1 pt.)

(c) Gebruik een stapgrootte van $\Delta x = 1/4$ om het stelsel vergelijkingen $\mathbf{Aw} = \mathbf{f}$ af te leiden. Verwerk de randvoorwaarden. Het afgeleide stelsel heeft drie vergelijkingen met drie onbekenden, dat betekent dat $\mathbf{A}$ een 3×3 matrix is en $\mathbf{w}$ en $\mathbf{f}$ 1×3 kolomvectoren zijn.

Dit stelsel vergelijkingen hoeft **niet** opgelost te worden. (2 pt.)

(d) Zal de discretisatiemethode (6) oscillerende oplossingen geven? Motiveer je antwoord. (1 pt.)

Voor de uitwerkingen van dit tentamen zie:
http://ta.twi.tudelft.nl/nw/users/vuik/wi3097/tentamen.html

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (CTB2400 WI3097TU)
Thursday August 11th 2016, 18:30-21:30

1. (a) The local truncation error is defined as

$$\tau_{n+1}(\Delta t) = \frac{y_{n+1} - z_{n+1}}{\Delta t}, \tag{1}$$

where z_{n+1} is given by

$$z_{n+1} = y_n + \Delta t \left(a_1 f(t_n, y_n) + a_2 f(t_n + \Delta t, y_n + \Delta t f(t_n, y_n)) \right). \tag{2}$$

A Taylor expansion of f around (t_n, y_n) yields

$$f(t_n+\Delta t, y_n+\Delta t f(t_n, y_n)) = f(t_n, y_n)+\Delta t \frac{\partial f}{\partial t}(t_n, y_n)+\Delta t f(t_n, y_n)\frac{\partial f}{\partial y}(t_n, y_n)+O((\Delta t)^2). \tag{3}$$

This is substituted into equation (2) to obtain

$$z_{n+1} = y_n+\Delta t \left(a_1 f(t_n, y_n) + a_2 \left[f(t_n, y_n) + \Delta t \frac{\partial f}{\partial t}(t_n, y_n) + \Delta t f(t_n, y_n) \frac{\partial f}{\partial y}(t_n, y_n) \right] \right)+O((\Delta t)^3) \tag{4}$$

A Taylor series for $y(t)$ around t_n gives for y_{n+1}

$$y_{n+1} = y(t_n + \Delta t) = y_n + \Delta t y'(t_n) + \frac{(\Delta t)^2}{2} y''(t_n) + O((\Delta t)^3). \tag{5}$$

From the differential equation we know that:

$$y'(t_n) = f(t_n, y_n) \tag{6}$$

From the Chain Rule of Differentiation, we derive

$$y''(t_n) = \frac{df(t_n, y_n)}{dt} = \frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} y'(t_n) \tag{7}$$

after substitution of the differential equation one obtains:

$$y''(t_n) = \frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} f(t_n, y_n) \tag{8}$$

Equations (5) and (4) are substituted into relation (1) to obtain

$$\tau_{n+1}(\Delta t) = f(t_n, y_n)(1-(a_1+a_2))+\Delta t \left(\frac{\partial f}{\partial t} + f \frac{\partial f}{\partial y} \right) \left(\frac{1}{2} - a_2 \right)+O((\Delta t)^2) \tag{9}$$

Hence

i. $a_1 + a_2 = 1$ implies $\tau_{n+1}(\Delta t) = O(\Delta t)$;

ii. $a_1+a_2 = 1$ and $a_2 = 1/2$, that is, $a_1 = a_2 = 1/2$, gives $\tau_{n+1}(\Delta t) = O((\Delta t)^2)$.

(b) The test equation is given by

$$y' = \lambda y. \tag{10}$$

Application of the predictor step to the test equation gives

$$w^*_{n+1} = w_n + \lambda \Delta t w_n = (1 + \lambda\Delta t)w_n. \tag{11}$$

The corrector step yields

$$w_{n+1} = w_n + \Delta t\,(a_1\lambda w_n + a_2\lambda(1 + \lambda\Delta t)w_n) = (1 + (a_1 + a_2)\lambda\Delta t + a_2(\lambda\Delta t)^2)w_n. \tag{12}$$

Hence the amplification factor is given by

$$Q(\lambda\Delta t) = 1 + (a_1 + a_2)\lambda\Delta t + a_2(\lambda\Delta t)^2. \tag{13}$$

(c) Let $\lambda < 0$ (so λ is real), then, for stability, the amplification factor must satisfy

$$-1 \le Q(\lambda\Delta t) \le 1, \tag{14}$$

from the previous assignment, we have

$$-1 \le 1 + (a_1 + a_2)\lambda\Delta t + a_2(\lambda\Delta t)^2 \le 1 \Leftrightarrow -2 \le (a_1 + a_2)\lambda\Delta t + a_2(\lambda\Delta t)^2 \le 0. \tag{15}$$

First, we consider the left inequality:

$$a_2(\lambda\Delta t)^2 + (a_1 + a_2)\lambda\Delta t + 2 \ge 0 \tag{16}$$

For $\lambda\Delta t = 0$, the above inequality is satisfied. The discriminant of the quadratic equation is given by $(a_1 + a_2)^2 - 8a_2$. From the given assumption it follows that $(a_1 + a_2)^2 - 8a_2 < 0$ so the quadratic equation does not have real roots. Hence the left inequality in relation (15) is always satisfied. Next we consider the right hand inequality of relation (15)

$$a_2(\lambda\Delta t)^2 + (a_1 + a_2)\lambda\Delta t \le 0. \tag{17}$$

This relation is rearranged into

$$a_2(\lambda\Delta t)^2 \le -(a_1 + a_2)\lambda\Delta t, \tag{18}$$

hence

$$a_2|\lambda\Delta t|^2 \le (a_1 + a_2)|\lambda\Delta t| \Leftrightarrow |\lambda\Delta t| \le \frac{a_1 + a_2}{a_2}, \qquad a_2 \ne 0. \tag{19}$$

This results into the following condition for stability

$$\Delta t \le \frac{a_1 + a_2}{a_2|\lambda|}, \qquad a_2 \ne 0. \tag{20}$$

(d) In order to compute the Jacobian, we note that the right-hand side of the non linear system is given by:

$$f_1(x_1, x_2) = -\sin x_1 + 2x_2 + t$$

$$f_2(x_1, x_2) = x_1 - x_2^2$$

From the definition of the Jacobian it follows that:

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} -\cos x_1 & 2 \\ 1 & -2x_2 \end{pmatrix}.$$

Substitution of $\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ shows that

$$J = \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix}.$$

(e) For the stability it is sufficient to check that $|Q(\lambda_i \Delta t)| \leq 1$ for all the eigenvalues of the Jacobian matrix. It is easy to see that the eigenvalues of the Jacobian matrix are $\lambda_1 = -3$ and $\lambda_2 = 0$.

For the choice $a_1 = a_2 = \frac{1}{2}$ we use the expression

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

For $\lambda_2 = 0$ it appears that $Q(\lambda_2 \Delta t) = 1$ so the inequality is satisfied for all Δt. For $\lambda_1 = -3$ we have to check the following inequalities:

$$-1 \leq 1 - 3\Delta t + \frac{9}{2}(\Delta t)^2 \leq 1$$

For the left-hand inequality we arrive at

$$0 \leq \frac{9}{2}(\Delta t)^2 - 3\Delta t + 2$$

It appears that the discriminant $9 - 4 \cdot \frac{9}{2} \cdot 2$ is negative, so there are no real roots which implies that the inequality is satisfied for all Δt.

For the right-hand inequality we get

$$-3\Delta t + \frac{9}{2}(\Delta t)^2 \leq 0$$

$$\frac{9}{2}(\Delta t)^2 \leq 3\Delta t$$

so

$$\Delta t \leq \frac{2}{3}$$

(another option is to see that for $a_1 = a_2 = \frac{1}{2}$ the method is equal to the modified Euler method, and remember that this method is stable for real eigenvalues if $\Delta t \leq \frac{-2}{\lambda}$)

2. (a) The *Taylor polynomials* around 0 are given by:

$$\begin{aligned} f(0) &= f(0)\ , \\ f(-h) &= f(0) - hf'(0) + \frac{h^2}{2}f''(0) - \frac{h^3}{6}f'''(\xi_1)\ , \\ f(-2h) &= f(0) - 2hf'(0) + 2h^2 f''(0) - \frac{(2h)^3}{6}f'''(\xi_2)\ . \end{aligned}$$

Here $\xi_1 \in (-h, 0)$, $\xi_2 \in (-2h, 0)$. We know that $Q(h) = \frac{\alpha_0}{h^2}f(0) + \frac{\alpha_{-1}}{h^2}f(-h) + \frac{\alpha_{-2}}{h^2}f(-2h)$, which should be equal to $f''(0) + \mathcal{O}(h)$. This leads to the following conditions:

$$\begin{array}{lcccccl} f(0): & \frac{\alpha_0}{h^2} & + & \frac{\alpha_{-1}}{h^2} & + & \frac{\alpha_{-2}}{h^2} & = 0\ , \\ f'(0): & & & -\frac{h\alpha_{-1}}{h^2} & - & \frac{2h\alpha_{-2}}{h^2} & = 0\ , \\ f''(0): & & & \frac{h^2}{2h^2}\alpha_{-1} & + & \frac{2h^2\alpha_{-2}}{h^2} & = 1\ . \end{array}$$

This can also be written as

$$\begin{array}{lcccccl} f(0): & \alpha_0 & + & \alpha_{-1} & + & \alpha_{-2} & = 0\ , \\ f'(0): & & & -\alpha_{-1} & - & 2\alpha_{-2} & = 0\ , \\ f''(0): & & & \frac{\alpha_{-1}}{2} & + & 2\alpha_{-2} & = 1\ . \end{array}$$

(b) The *truncation error* follows from the Taylor polynomials:

$$f''(0) - Q(h) = f''(0) - \frac{f(0) - 2f(-h) + f(-2h)}{h^2} = -\left(\frac{\frac{2h^3}{6}f'''(\xi_1) - \frac{8h^3}{6}f'''(\xi_2)}{h^2}\right)$$

$$= hf'''(\xi).$$

(c) Note that

$$f''(0) - Q(2h) = K2h \tag{21}$$

$$f''(0) - Q(h) = Kh \tag{22}$$

Subtraction gives:

$$Q(h) - Q(2h) = K2h - Kh = Kh. \tag{23}$$

We choose $h = \frac{1}{4}$. Then $Q(2h) = Q(\frac{1}{2}) = \frac{0-2\times 0.1250+1}{0.25} = 3$ and $Q(h) = Q(\frac{1}{4}) = \frac{0-2\times 0.0156+0.1250}{(\frac{1}{4})^2} = 1.5008$. Combining (22) and (23) shows that

$$f''(0) - Q(\frac{1}{4}) = Q(\frac{1}{4}) - Q(\frac{1}{2}) = -1.4992$$

3. (a) First, we check the boundary conditions:

$$u(0) = 0 - \frac{1-e^{0/\epsilon}}{1-e^{1/\epsilon}} = \frac{1-1}{1-e^{1/\epsilon}} = 0, \quad u(1) = 1 - \frac{1-e^{1/\epsilon}}{1-e^{1/\epsilon}} = 0. \tag{24}$$

Further, we have

$$u'(x) = 1 + \frac{e^{x/\epsilon}}{\epsilon(1-e^{1/\epsilon})}, \tag{25}$$

$$u''(x) = \frac{e^{x/\epsilon}}{\epsilon^2(1-e^{1/\epsilon})}. \tag{26}$$

Hence, we immediately see

$$-\epsilon u''(x) + u'(x) = -\frac{\epsilon e^{x/\epsilon}}{\epsilon^2(1-e^{1/\epsilon})} + 1 + \frac{e^{x/\epsilon}}{\epsilon(1-e^{1/\epsilon})} = 1. \tag{27}$$

Hence, the solution $u(x) = 1 - \frac{1-e^{x/\epsilon}}{1-e^{1/\epsilon}}$ satisfies the differential equation and the Dirichlet boundary conditions, and therewith $u(x)$ is the solution to the boundary value problem (uniqueness can be demonstrated in a straightforward way, but this was not asked for).

(b) The domain of computation, being $(0,1)$, is divided into subintervals with mesh points, we set $x_j = j\Delta x$, where we use n unknowns, such that $x_{n+1} = (n+1)\Delta x = 1$. The discretization method for an interior node is given by

$$-\epsilon\frac{w_{j+1} - 2w_j + w_{j-1}}{(\Delta x)^2} + \frac{w_j - w_{j-1}}{\Delta x} = 1, \text{ for } j \in \{1, \ldots, n\}. \tag{28}$$

At the boundaries, we see for $j = 1$ and $j = n$, upon substituting $w_0 = 0$ and $w_{n+1} = 0$, respectively:

$$\begin{aligned} -\epsilon\frac{w_2 - 2w_1 + 0}{(\Delta x)^2} + \frac{w_1 - 0}{\Delta x} &= 1, \\ -\epsilon\frac{0 - 2w_n + w_{n-1}}{(\Delta x)^2} + \frac{w_n - w_{n-1}}{\Delta x} &= 1. \end{aligned} \tag{29}$$

This can be rewritten more neatly as follows:

$$\begin{aligned} \epsilon\frac{-w_2 + 2w_1}{(\Delta x)^2} + \frac{w_1}{\Delta x} &= 1, \\ \epsilon\frac{2w_n - w_{n-1}}{(\Delta x)^2} + \frac{w_n - w_{n-1}}{\Delta x} &= 1. \end{aligned} \tag{30}$$

(c) Next, we use $\Delta x = 1/4$ and thus $n = 3$. Then, from equations (28) and (30), one obtains the following 3×3 linear system of equations

$$\begin{bmatrix} 32\epsilon + 4 & -16\epsilon & 0 \\ -16\epsilon - 4 & 32\epsilon + 4 & -16\epsilon \\ 0 & -16\epsilon - 4 & 16\epsilon + 4 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{31}$$

(d) No, the upwind difference method for the convective terms is designed to not produce oscillatory solutions independent of the size of the diffusion coefficient ϵ and the velocity v which is equal to one in this case.

TECHNISCHE UNIVERSITEIT DELFT
FACULTEIT ELEKTROTECHNIEK, WISKUNDE EN INFORMATICA

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (CTB2400 WI3097 TU)
Donderdag 30 Juni 2016, 18:30-21:30

1. (a) De **vierde orde Runge-Kutta methode** (RK_4) voor de differentiaalvergelijking $y' = f(t, y)$ wordt gegeven door de volgende formules:

$$\begin{aligned} k_1 &= \Delta t f(t_n, w_n) \\ k_2 &= \Delta t f(t_n + \tfrac{1}{2}\Delta t, w_n + \tfrac{1}{2}k_1) \\ k_3 &= \Delta t f(t_n + \tfrac{1}{2}\Delta t, w_n + \tfrac{1}{2}k_2) \\ k_4 &= \Delta t f(t_n + \Delta t, w_n + k_3) \end{aligned}$$

$$w_{n+1} = w_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4).$$

Bepaal de versterkingsfactor $Q(\lambda\Delta t)$ van RK_4 door de methode toe te passen op de *homogene testvergelijking* $y' = \lambda y$. (2 pt.)

(b) Gebruik het feit dat $y(t_{n+1}) = e^{\lambda\Delta t}y(t_n)$ geldt voor de exacte oplossing van $y' = \lambda y$ om te laten zien dat RK_4 toegepast op de *homogene testvergelijking* een lokale afbreekfout heeft van $O(\Delta t^4)$. (Hint: $e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + ...$) (2 pt.)

In de volgende onderdelen mag je aannemen dat dit resultaat ook geldt voor stelsels, zodat de globale fout van RK_4 toegepast op (1) gelijk is aan $O(\Delta t^4)$.

We beschouwen het volgende **tweede orde beginwaarde probleem**:

$$y'' + py' + qy = \sin t, \quad y(0) = y_0, \; y'(0) = y'_0. \tag{1}$$

(c) Schrijf (1) als een stelsel eerste orde differentiaalvergelijkingen van het type

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g}(t).$$

Geef $\mathbf{A}$ en $\mathbf{g}$ en bepaal de eigenwaarden van $\mathbf{A}$, voor willekeurige p en q. (2 pt.)

(d) We nemen nu p=1000 en q=250001. Geef een benadering van de stabiliteitsvoorwaarde voor dit geval.

Hint: gebruik het figuur van het stabiliteitsgebied van RK_4 op pagina 3. (2 pt.)

(e) Stel dat je een keuze moet maken tussen de Trapeziumregel en RK_4 om het probleem gegeven in (d) op te lossen. Motiveer je keuze zo goed mogelijk. Hierbij moeten aan de orde komen: de stabiliteitsvoorwaarde en de orde van grootte van de globale fout. (2 pt.)

voor vervolg z.o.z.

2. Vervolgens willen we de **integraal** $\int_0^1 y(x)dx$ met $y(x) = x^2$ **numeriek benaderen**.

   (a) Geef de *Rechthoekregel* I^R. Geef ook de bijbehorende *samengestelde integratieregels* $I^R(h)$. Benader de integraal $\int_0^1 y(x)dx$ met behulp van de samengestelde Rechtshoekregel, met $h = 1/3$. (1 pt.)

   (b) Herhal deel (a) met de *Trapeziumregel* (I^T en $I^T(h)$), met $h = 1/3$. (1 pt.)

   (c) Stel dat men $\int_0^1 y(x)dx$ benadert, dan is de grootte van de fout van de *samengestelde regels* (ε_R en ε_T voor de Rechthoek- en Trapeziumregel, respectievelijk) begrensd door

   $$\varepsilon_R \leq \frac{h}{2} \max_{x\in[0,1]} |y'(x)|, \quad \varepsilon_T \leq \frac{h^2}{12} \max_{x\in[0,1]} |y''(x)|. \tag{2}$$

   Welke methode verdient de voorkeur als het aantal integratiepunten groot is? Motiveer uw voorkeur. (2 pt.)

3. Vervolgens leiden we de **Newton-Raphson methode** af en gebruiken we deze om een niet lineair probleem op te lossen.

   (a) Gegeven is het *scalaire* niet lineaire probleem:

   $$\text{Bepaal } p \in \mathbb{R} \text{ zodat } f(p) = 0. \tag{3}$$

   Leid de formule van Newton-Raphson

   $$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \text{ voor } n \geq 1 \tag{4}$$

   met beginschatting p_0 af om het probleem op te lossen. (2 pt.)

   (b) Om te laten zien dat de Newton-Raphson methode convergeert, schrijven we de methode als een vaste punt methode:

   $$g(x) = x - \frac{f(x)}{f'(x)}. \tag{5}$$

   Eén van de eisen van het convergentiebewijs is dat $g'(x)$ bestaat en voldoet aan:

   $$|g'(x)| \leq k < 1. \tag{6}$$

   Laat zien dat de Newton-Raphson methode toegepast op $f(x) = \sin(x)$ naar de oplossing $p = 0$ convergeert als de beginschatting p_0 gekozen wordt in het interval $(-\sqrt{2}/2, \sqrt{2}/2)$. (1 pt.)

   (c) Leid Newton-Raphson's methode voor het volgende *algemeene* niet lineare problem af:

   $$\text{Bepaal } \mathbf{p} \in \mathbb{R}^m \text{ zodat } \mathbf{f}(\mathbf{p}) = \mathbf{0}. \tag{7}$$

   (1 pt.)

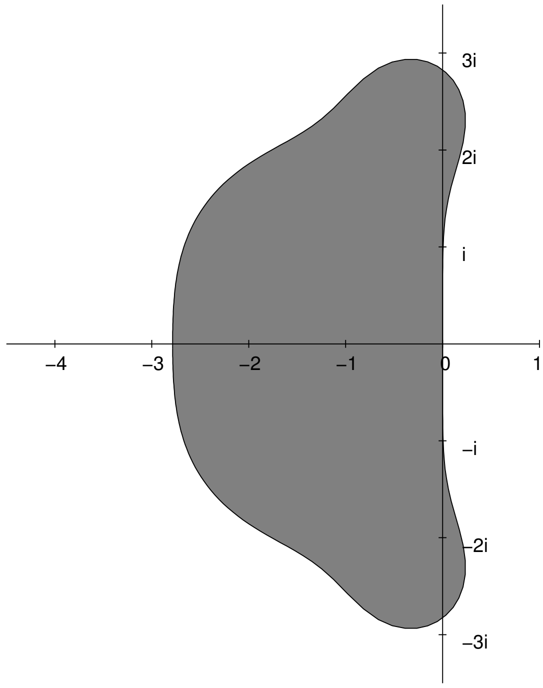
   (d) Voer **één** Newton-Raphson stap uit op het volgende niet lineaire problem voor w_1 en w_2:

   $$\begin{cases} 18w_1 - 9w_2 + w_1^2 = 0, \\ -9w_1 + 18w_2 + w_2^2 = 9. \end{cases} \tag{8}$$

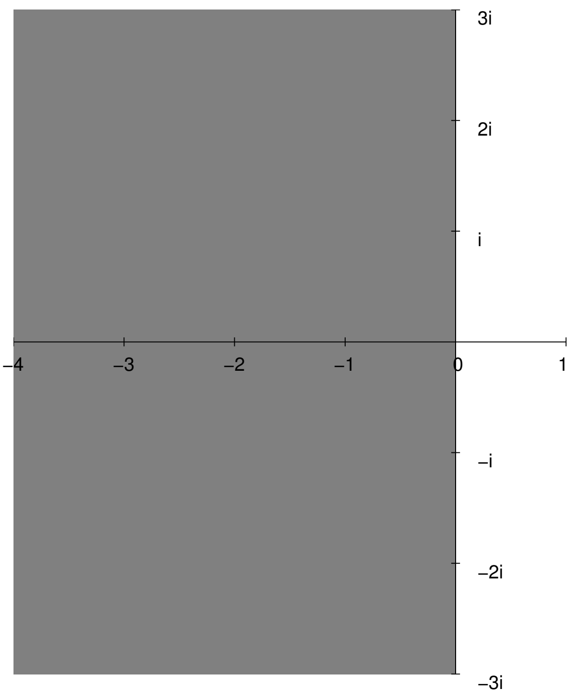
   Gebruik $w_1 = w_2 = 0$ als de beginschatting. (2 pt.)

Voor de uitwerkingen van dit tentamen zie:
http://ta.twi.tudelft.nl/nw/users/vuik/wi3097/tentamen.html

Lijst van figuren



Figuur 1: Stabiliteitsgebied voor RK_4



Figuur 2: Stabiliteitsgebied voor de Trapeziumregel

DELFT UNIVERSITY OF TECHNOLOGY
Faculty of Electrical Engineering, Mathematics and Computer Science

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (CTB2400 WI3097 TU)
Thursday June 30th 2016, 18:30-21:30

1. (a) Replace $f(t, y)$ by λy in the RK$_4$ formulas:

$$\begin{aligned}
k_1 &= \lambda \Delta t w_n \\
k_2 &= \lambda \Delta t(w_n + \tfrac{1}{2}k_1) = \lambda \Delta t(1 + \tfrac{1}{2}\lambda \Delta t)w_n \\
k_3 &= \lambda \Delta t(w_n + \tfrac{1}{2}k_2) = \lambda \Delta t(1 + \tfrac{1}{2}\lambda \Delta t(1 + \tfrac{1}{2}\lambda \Delta t))w_n \\
k_4 &= \lambda \Delta t(w_n + k_3) = \lambda \Delta t(1 + \lambda \Delta t(1 + \tfrac{1}{2}\lambda \Delta t(1 + \tfrac{1}{2}\lambda \Delta t))w_n
\end{aligned}$$

Substitution of these expressions into:

$$w_{n+1} = w_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

and collecting like powers of $\lambda \Delta t$ yields:

$$w_{n+1} = [1 + \lambda \Delta t + \frac{1}{2}(\lambda \Delta t)^2 + \frac{1}{6}(\lambda \Delta t)^3 + \frac{1}{24}(\lambda \Delta t)^4]w_n.$$

The amplification factor is therefore:

$$Q(\lambda \Delta t) = 1 + \lambda \Delta t + \frac{1}{2}(\lambda \Delta t)^2 + \frac{1}{6}(\lambda \Delta t)^3 + \frac{1}{24}(\lambda \Delta t)^4.$$

(b) The local truncation error is defined as

$$\tau_{n+1} = \frac{y(t_{n+1}) - z_{n+1}}{\Delta t}, \tag{1}$$

where z_{n+1} is the numerical solution at t_{n+1}, obtained by starting from the exact value $y(t_n)$ in stead of w_n. Repeating the derivation under (a), with w_n replaced by $y(t_n)$, gives:

$$z_{n+1} = Q(\lambda \Delta t)y(t_n).$$

Using furthermore $y(t_{n+1}) = e^{\lambda \Delta t}y(t_n)$ in (1) it follows that

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

Canceling the first five terms of the expansion of $e^{\lambda \Delta t}$ against $Q(\lambda \Delta t)$, the required order of magnitude of τ_{n+1} follows.

(c) Use the transformation:

$$\begin{aligned} y_1 &= y\ , \\ y_2 &= y'\ , \end{aligned}$$

This implies that

$$\begin{aligned} y_1' &= y' = y_2\ , \\ y_2' &= y'' = -qy_1 - py_2 + \sin t\ , \end{aligned}$$

So the matrix $\mathbf{A}$ and vector $\mathbf{g}$ are:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix};\ \ \mathbf{g}(t) = \begin{pmatrix} 0 \\ \sin t \end{pmatrix}.$$

Characteristic equation: $\lambda^2 + p\lambda + q = 0$. $\lambda_{1,2} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$.

(d) Substitution of the values of p and q into the matrix $\mathbf{A}$ yields the eigenvalues $\lambda_{1,2} = -500 \pm i$. From the given drawing of the stability region the following can be inferred. Because the imaginary part is much smaller than the real part, an approximate stability condition can be obtained by simply neglecting the imaginary part. Then $\Delta t \leq 2.8/500 = 0.0056$ follows as the stability condition.

(e)

$$y'' + py' + qy = \sin t,\quad y(0) = y_0,\ y'(0) = y_0'. \tag{2}$$

After a short time the solution is close to a linear combination of $\sin t$ and $\cos t$, which is called a smooth solution.

The smooth solution can be integrated accurately by RK_4 with a 'large' step size: a step size of 0.1, let us say, would give an error of order 10^{-4} which is sufficient for most engineering purposes. However stability, governed by the eigenvalues, requires that the step size be restricted (see part (d)) to 0.0056. So the stability requirement forces us to choose a step size yielding an unnecessarily accurate solution, which is inefficient.

The Trapezoidal rule, on the other hand, is stable for all step sizes. So the step size is restricted by accuracy requirements only. The Trapezoidal rule has a global error of order Δt^2 such that a good accuracy may be expected for step sizes of about 0.01, which is much larger than the step size for RK4: 0.0056. An efficiency gain may be obtained in spite of the extra work connected with the implicitness of the method.

2. (a) Consider an interval of integration $[x_{j-1}, x_j]$, then the **Rectangle Rule** reads

$$I_j^R = hf(x_{j-1}), \qquad h = x_j - x_{j-1}. \tag{3}$$

The *composed integration rule* is derived by

$$I^R = h(I_1^R + I_2^R + \ldots + I_n^R) = h(f(x_0) + \ldots + f(x_{n-1})), \tag{4}$$

which yields

$$I^R = \frac{1}{3} \cdot (0 + (\frac{1}{3})^2 + (\frac{2}{3})^2) = \frac{5}{27}. \tag{5}$$

(b) For the interval of integration $[x_{j-1}, x_j]$ the **Trapezoidal Rule** is

$$I_j^T = \frac{h}{2}(f(x_{j-1}) + f(x_j)). \tag{6}$$

The *composed integration rule* is derived by

$$I^T = h(I_1^T + I_2^T + \ldots + I_n^T) = h(\frac{f(x_0)}{2} + f(x_1) + \ldots + f(x_{n-1}) + \frac{f(x_n)}{2}), \tag{7}$$

which leads to

$$I^T = \frac{1}{3} \cdot (0 + (\frac{1}{3})^2 + (\frac{2}{3})^2 + \frac{1}{2}) = \frac{19}{54}. \tag{8}$$

(c) For a general number of subintervals, say n, the magnitude of the composed Rectangle- and Trapezoidal Rules, is bounded from above by

$$\begin{aligned} \varepsilon_R &\leq \frac{h}{2} \max_{x \in [0,1]} |y'(x)| \leq h = \frac{1}{n}, \\ \varepsilon_T &\leq \frac{h^2}{12} \max_{x \in [0,1]} |y''(x)| \leq \frac{h^2}{6} = \frac{1}{6n^2}. \end{aligned} \tag{9}$$

Here, the exact solution $y(x) = x^2$ was used. Hence, the error from the Trapezoidal Rule is much smaller. Furthermore, from the composed Rules, it is easy to see that the number of function evaluations for the composed Rectangle- and Trapezoidal Rules is given by n and $n+1$, respectively. Since

$$\lim_{n \to \infty} \frac{n+1}{n} = 1, \tag{10}$$

it follows that the amount of work for the Trapezoidal Rule is not significantly higher than it is for the Rectangle Rule. Hence, it is more attractive to use the Trapezoidal Rule.

3. (a) **Newton-Raphson's method** is an iterative method to find $p \in \mathbb{R}$ such that $f(p) = 0$. Suppose $f \in C^2[a,b]$. Let $\bar{x} \in [a,b]$ be an approximation of the root p such that $f'(\bar{x}) \neq 0$, and suppose that $|p - \bar{x}|$ is small. Consider the first-degree Taylor polynomial about $\bar{x}$:

$$f(x) = f(\bar{x}) + (x - \bar{x})f'(\bar{x}) + \frac{(x - \bar{x})^2}{2} f''(\xi(x)), \tag{11}$$

in which $\xi(x)$ between x and $\bar{x}$. Using that $f(p) = 0$, equation (11) yields

$$0 = f(\bar{x}) + (p - \bar{x})f'(\bar{x}) + \frac{(p - \bar{x})^2}{2} f''(\xi(x)).$$

Because $|p - \bar{x}|$ is small, $(p - \bar{x})^2$ can be neglected, such that

$$0 \approx f(\bar{x}) + (p - \bar{x})f'(\bar{x}).$$

Note that the right-hand side is the formula for the tangent in $(\bar{x}, f(\bar{x}))$. Solving for p yields

$$p \approx \bar{x} - \frac{f(\bar{x})}{f'(\bar{x})}.$$

This motivates the Newton-Raphson method, that starts with an approximation p_0 and generates a sequence $\{p_n\}$ by

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \quad \text{for} \quad n \geq 1.$$

Remark 1 *One can also give a graphical derivation following Figure 4.2 from the book.*

(b) The first derivative of g equals

$$g'(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}.$$

Substitution of $f(x) = \sin(x)$, $f'(x) = \cos(x)$ and $f''(x) = -\sin(x)$ yields

$$g'(x) = -\frac{\sin^2(x)}{\cos^2(x)} = -\tan^2(x).$$

Since $\tan(-\sqrt{2}/2) = -1$, $\tan(\sqrt{2}/2) = 1$ and the tangent function is monotonically increasing on the interval $[-1,1]$ any initial guess inside the interval $(-1,1)$ will lead to a convergent iteration process.

(c) It follows from the linearization of the function $\mathbf{f}$ about the iterate $\mathbf{x}_{n-1}$ that

$$f_1(\mathbf{p}) \approx f_1(\mathbf{p}^{(n-1)}) + \frac{\partial f_1}{\partial p_1}(\mathbf{p}^{(n-1)})(p_1 - p_1^{(n-1)}) + \ldots + \frac{\partial f_1}{\partial p_m}(\mathbf{p}^{(n-1)})(p_m - p_m^{(n-1)}),$$
$$\vdots$$
$$f_m(\mathbf{p}) \approx f_m(\mathbf{p}^{(n-1)}) + \frac{\partial f_m}{\partial p_1}(\mathbf{p}^{(n-1)})(p_1 - p_1^{(n-1)}) + \ldots + \frac{\partial f_m}{\partial p_m}(\mathbf{p}^{(n-1)})(p_m - p_m^{(n-1)}).$$

Defining the Jacobian matrix of $\mathbf{f}(\mathbf{x})$ by

$$\mathbf{J}(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \ldots & \frac{\partial f_1}{\partial x_m}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \ldots & \frac{\partial f_m}{\partial x_m}(\mathbf{x}) \end{pmatrix},$$

the linearization can be written in the more compact form

$$\mathbf{f}(\mathbf{p}) \approx \mathbf{f}(\mathbf{p}^{(n-1)}) + \mathbf{J}(\mathbf{p}^{(n-1)})(\mathbf{p} - \mathbf{p}^{(n-1)}).$$

The next iterate, $\mathbf{p}^{(n)}$, is obtained by setting the linearization equal to zero:

$$\mathbf{f}(\mathbf{p}^{(n-1)}) + \mathbf{J}(\mathbf{p}^{(n-1)})(\mathbf{p}^{(n)} - \mathbf{p}^{(n-1)}) = 0, \tag{12}$$

which can be rewritten as

$$\mathbf{J}(\mathbf{p}^{(n-1)})\mathbf{s}^{(n)} = -\mathbf{f}(\mathbf{p}^{(n-1)}), \tag{13}$$

where $\mathbf{s}^{(n)} = \mathbf{p}^{(n)} - \mathbf{p}^{(n-1)}$. The new approximation equals $\mathbf{p}^{(n)} = \mathbf{p}^{(n-1)} + \mathbf{s}^{(n)}$. Finally, Newton-Raphson's formula for general nonlinear problems reads:

$$\mathbf{p}^{(n)} = \mathbf{p}^{(n-1)} - \mathbf{J}^{-1}(\mathbf{p}^{(n-1)})\mathbf{f}(\mathbf{p}^{(n-1)}). \tag{14}$$

(d) First, we rewrite the system into the form

$$\begin{aligned} f_1(w_1, w_2) &= 0, \\ f_2(w_1, w_2) &= 0, \end{aligned} \tag{15}$$

by setting

$$\begin{aligned} f_1(w_1, w_2) &:= 18w_1 - 9w_2 + (w_1)^2, \\ f_2(w_1, w_2) &:= -9w_1 + 18w_2 + (w_2)^2 - 9. \end{aligned} \tag{16}$$

We denote the Jacobi-matrix by $J(w_1, w_2)$. At the first step we compute

$$\underline{w}^{(1)} = \underline{w}^{(0)} - J(\underline{w}^{(0)})^{-1}F(\underline{w}^{(0)}), \tag{17}$$

where $\underline{w} = [w_1\ w_2]^T$. Note that

$$J(\underline{w}^{(0)}) = \begin{pmatrix} 18 + 2w_1^{(0)} & -9 \\ -9 & 18 + 2w_2^{(0)} \end{pmatrix}. \tag{18}$$

Using $w_1^{(0)} = w_2^{(0)} = 0$ we obtain:

$$J(\underline{w}^{(0)}) = \begin{pmatrix} 18 & -9 \\ -9 & 18 \end{pmatrix}. \tag{19}$$

This implies that

$$J(\underline{w}^{(0)})^{-1} = \frac{1}{18^2 - 81} \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix}. \tag{20}$$

Furthermore

$$F(\underline{w}^{(0)}) = \begin{pmatrix} 0 \\ -9 \end{pmatrix}, \tag{21}$$

so

$$\underline{w}^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{18^2 - 81} \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix} \begin{pmatrix} 0 \\ -9 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}. \tag{22}$$

TECHNISCHE UNIVERSITEIT DELFT
FACULTEIT ELEKTROTECHNIEK, WISKUNDE EN INFORMATICA

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU CTB2400)
Donderdag 13 Augustus 2015, 18:30-21:30

1. In deze opgave maken we gebruik van de trapeziummethode voor de integratie van het beginwaardeprobleem $y' = f(t, y)$ met $y(t_0) = y_0$:

$$w_{n+1} = w_n + \frac{\Delta t}{2}\left(f(t_n, w_n) + f(t_{n+1}, w_{n+1})\right) \quad (1)$$

   (a) Laat zien dat de versterkingsfactor van de trapeziummethode gegeven wordt door

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

   (1 pt.)

   (b) Geef de orde (+ bewijs) van de lokale afbreekfout van de trapeziummethode voor de testvergelijking (hint de volgende reeksen kunnen gebruikt worden: $e^x = \sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$ en $\frac{1}{1-x} = \sum\limits_{k=0}^{\infty} x^k$). (3 pt.)

   (c) Doe één stap met de trapeziummethode voor het volgende beginwaardeprobleem

$$y' = -2y + e^t, \text{ met } y(0) = 2,$$

   en stapgrootte $\Delta t = 1$. (2 pt.)

   (d) We beschouwen het beginwaarde probleem:

$$y'' = -y' - \frac{1}{2}y, \;\; y(0) = 1, \;\; y'(0) = 0.$$

   Schrijf deze tweede orde differentiaalvergelijking als een stelsel eerste orde differentiaalvergelijkingen: $\mathbf{x}' = \mathbf{A}\mathbf{x}$. Toon aan dat de eigenwaarden van $\mathbf{A}$ gegeven worden door

$$\lambda_1 = -\frac{1}{2} + \frac{1}{2}i \text{ en } \lambda_2 = -\frac{1}{2} - \frac{1}{2}i.$$

   (2 pt.)

   (e) Onderzoek de stabiliteit van de trapeziummethode toegepast op het stelsel gegeven in (d). (2 pt.)

[0] **voor vervolg z.o.z. Voor de uitwerkingen van dit tentamen zie: http://ta.twi.tudelft.nl/nw/users/vuik/wi3097/tentamen.html**

2. We onderzoeken Lagrange interpolatie. Voor gegeven steunpunten x_0, x_1, ..., x_n met bijbehorende functiewaarden $f(x_0)$, $f(x_1)$, ..., $f(x_n)$, wordt het interpolatiepolynoom $p_n(x)$, gegeven door

$$
\begin{aligned}
p_n(x) &= \sum_{i=0}^{n} f(x_i) L_i(x), \text{ met} \\
L_i(x) &= \frac{(x-x_0)(\ldots)(x-x_{i-1})(x-x_{i+1})(\ldots)(x-x_n)}{(x_i-x_0)(\ldots)(x_i-x_{i-1})(x_i-x_{i+1})(\ldots)(x_i-x_n)}.
\end{aligned}
\tag{2}
$$

Verder zijn de volgende meetwaarden gegeven in tabelvorm:

i	x_i	$f(x_i)$
0	-1	3
1	0	2
2	1	5

(a) Geef het lineaire interpolatiepolynoom van Lagrange met steunpunten x_0 en x_1. (1pt.)

(b) Geef de kwadratische interpolatieformule van Lagrange met steunpunten x_0, x_1 en x_2. (2 pt.)

(c) Benader $f(0)$ en $f(0.5)$ eerst met lineaire interpolatie en dan met kwadratische interpolatie. (2 pt.)

Gegeven is de Newton-Raphson methode

$$p_{n+1} = p_n - \frac{f(p_n)}{f'(p_n)}.$$

(d) Leid de bovenstaande Newton-Raphson methode af. (1.5 pt.)

(e) We zoeken het positieve nulpunt van $f(x) = e^x - x^3$. Neem als startwaarde $p_0 = 2$ en bepaal p_1 en p_2 met de Newton-Raphson methode. (1.5 pt.)

(f) Laat p de oplossing van $f(p) = 0$ zijn. Toon aan dat dan geldt

$$|p - p_{n+1}| = K|p - p_n|^2, \text{ voor } n \to \infty \tag{3}$$

en bepaal de waarde van de constante K. (2 pt.)

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU CTB2400)
Thursday August 13th 2015, 18:30-21:30

1. (a) The amplification factor can be derived as follows. Consider the test equation $y' = \lambda y$. Application of the trapezoidal rule to this equation gives:

$$w_{j+1} = w_j + \frac{\Delta t}{2}(\lambda w_j + \lambda w_{j+1}) \tag{1}$$

Rearranging of w_{j+1} and w_j in (1) yields

$$\left(1 - \frac{\Delta t}{2}\lambda\right) w_{j+1} = \left(1 + \frac{\Delta t}{2}\lambda\right) w_j.$$

It now follows that

$$w_{j+1} = \frac{1 + \frac{\Delta t}{2}\lambda}{1 - \frac{\Delta t}{2}\lambda} w_j,$$

and thus

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

(b) The definition of the local truncation error is

$$\tau_{j+1} = \frac{y_{j+1} - Q(\Delta t\lambda)y_j}{\Delta t}.$$

The exact solution of the test equation is given by

$$y_{j+1} = e^{\Delta t\lambda} y_j.$$

Combination of these results shows that the local truncation error of the test equation is determined by the difference between the exponential function and the amplification factor $Q(\Delta t\lambda)$

$$\tau_{j+1} = \frac{e^{\Delta t\lambda} - Q(\Delta t\lambda)}{\Delta t} y_j. \tag{2}$$

The difference between the exponential function and amplification factor can be computed as follows. The Taylor series of $e^{\Delta t\lambda}$ with known point 0 is:

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

The Taylor series of $\frac{1}{1-\frac{\Delta t}{2}\lambda}$ with known point 0 is:

$$\frac{1}{1-\frac{\Delta t}{2}\lambda} = 1 + \frac{1}{2}\Delta t\lambda + \frac{1}{4}\Delta t^2\lambda^2 + \mathcal{O}(\Delta t^3). \tag{4}$$

With (4) it follows that $\frac{1+\frac{\Delta t}{2}\lambda}{1-\frac{\Delta t}{2}\lambda}$ is equal to

$$\frac{1+\frac{\Delta t}{2}\lambda}{1-\frac{\Delta t}{2}\lambda} = 1 + \Delta t\lambda + \frac{1}{2}(\Delta t\lambda)^2 + \mathcal{O}(\Delta t^3). \tag{5}$$

In order to determine $e^{\Delta t\lambda} - Q(\Delta t\lambda)$, we subtract (5) from (3). Now it follows that

$$e^{\Delta t\lambda} - Q(\Delta t\lambda) = \mathcal{O}(\Delta t^3). \tag{6}$$

The local truncation error can be found by substituting (6) into (2), which leads to

$$\tau_{j+1} = \mathcal{O}(\Delta t^2).$$

(c) Application of the trapezoidal rule to

$$y' = -2y + e^t, \text{ with } y(0) = 2,$$

and step size $\Delta t = 1$ gives:

$$w_1 = w_0 + \frac{\Delta t}{2}[-2w_0 + e^0 - 2w_1 + e].$$

Using the initial value $w_0 = y(0) = 2$ and step size $\Delta t = 1$ gives:

$$w_1 = 2 + \frac{1}{2}[-4 - 2w_1 + 1 + e].$$

This leads to

$$2w_1 = 2 + \frac{-3+e}{2} = \frac{1}{2} + \frac{e}{2}, \text{ so } w_1 = \frac{1}{4} + \frac{e}{4}.$$

(d) We use the following definition $x_1 = y$ and $x_2 = y'$. This implies that $x_1' = y' = x_2$ and $x_2' = y'' = -y' - \frac{1}{2}y = -x_2 - \frac{1}{2}x_1$. Writing this in vector notation shows that

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

so $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix}$. To compute the eigenvalues we look for values of λ such that

$$|\mathbf{A} - \lambda\mathbf{I}| = 0.$$

This implies that λ is a solution of

$$\lambda^2 + \lambda + \frac{1}{2} = 0,$$

which leads to the roots:

$$\lambda_1 = -\frac{1}{2} + \frac{1}{2}i \text{ and } \lambda_2 = -\frac{1}{2} - \frac{1}{2}i.$$

(e) To investigate the stability it is sufficient that

$$|Q(\Delta t \lambda_1)| \leq 1 \text{ and } |Q(\Delta t \lambda_2)| \leq 1.$$

Since λ_1 and λ_2 are complex valued, it is sufficient to check only the first inequality. This leads to

$$\left| \frac{1 + \frac{\Delta t(-\frac{1}{2}+\frac{1}{2}i)}{2}}{1 - \frac{\Delta t(-\frac{1}{2}+\frac{1}{2}i)}{2}} \right| \leq 1,$$

which is equivalent to

$$\frac{|1 - \frac{\Delta t}{4} + \frac{\Delta t i}{4}|}{|1 + \frac{\Delta t}{4} - \frac{\Delta t i}{4}|} \leq 1.$$

Using the definition of the absolute value we arrive at the inequality

$$\frac{\sqrt{(1 - \frac{\Delta t}{4})^2 + (\frac{\Delta t}{4})^2}}{\sqrt{(1 + \frac{\Delta t}{4})^2 + (\frac{\delta t}{4})^2}} \leq 1.$$

This equality is valid for all values of Δt because

$$\sqrt{(1 - \frac{\Delta t}{4})^2 + (\frac{\Delta t}{4})^2} \leq \sqrt{(1 + \frac{\Delta t}{4})^2 + (\frac{\Delta t}{4})^2},$$

for all $\Delta t > 0$.

2. (a) The linear Lagrangian interpolatory polynomial, with nodes x_0 and x_1, is given by

$$p_1(x) = \frac{x - x_1}{x_0 - x_1} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1). \tag{7}$$

This is evident from application of the given formula.

(b) The quadratic Lagrangian interpolatory polynomial with nodes x_0, x_1 and x_2 is given by

$$p_2(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} f(x_0) + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} f(x_1) + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} f(x_2). \tag{8}$$

This is also evident from application of the given formula.

(c) Obviously, $p_1(0) = 2$ and $p_2(0) = 2$ since the Lagrange interpolation polynomial satisfies $p_n(x_i) = f(x_i)$ for all points $x_0, x_1, \ldots, x_n$. Next, we compute $p_1(0.5)$ and $p_2(0.5)$ for both linear and quadratic Lagrangian interpolation as approximations at $x = 0.5$. For linear interpolation, we have

$$p_1(0.5) = \frac{0.5 - 0}{-1 - 0} \cdot 3 + \frac{0.5 + 1}{0 + 1} \cdot 2 = \frac{3}{2}, \tag{9}$$

and for quadratic interpolation, one obtains

$$p_2(0.5) = \frac{(0.5-0)(0.5-1)}{(-1)\cdot(-2)} \cdot 3 + \frac{(0.5+1)(0.5-1)}{1\cdot(-1)} \cdot 2 + \frac{(0.5+1)(0.5-0)}{2\cdot 1} \cdot 5 = 3. \tag{10}$$

(d) The method of Newton-Raphson is based on linearization around the iterate p_n. This is given by

$$L(x) = f(p_n) + (x - p_n)f'(p_n). \tag{11}$$

Next, we determine p_{n+1} such that $L(p_{n+1}) = 0$, that is

$$f(p_n) + (p_{n+1} - p_n)f'(p_n) = 0 \Leftrightarrow p_{n+1} = p_n - \frac{f(p_n)}{f'(p_n)}, \qquad f'(p_n) \neq 0. \tag{12}$$

This result can also be proved graphically, see book, chapter 4.

(e) We have $f(x) = e^x - x^3$, so $f'(x) = e^x - 3x^2$ and hence

$$p_{n+1} = p_n - \frac{e^{p_n} - p_n^3}{e^{p_n} - 3p_n^2}.$$

With the initial value $p_0 = 3$, this gives

$$p_1 = 3 - \frac{e^3 - 3^3}{e^3 - 3 \cdot 3^2} = 2$$

and consequently

$$p_2 = 2 - \frac{e^2 - 2^3}{e^2 - 3 \cdot 2^2} = \frac{e^2 - 16}{e^2 - 12} \approx 1.8675$$

(f) We consider a Taylor polynomial around p_n, to express p

$$0 = f(p) = f(p_n) + (p - p_n)f'(p_n) + \frac{(p - p_n)^2}{2} f''(\xi_n), \tag{13}$$

for some ξ_n between p and p_n. Note that this form gives the exact representation. Subsequently, we consider the Newton-Raphson approximation

$$0 = L(p_{n+1}) = f(p_n) + (p_{n+1} - p_n)f'(p_n). \tag{14}$$

Subtraction of these two above equations gives

$$p_{n+1} - p = \frac{(p_n - p)^2}{2} \frac{f''(\xi_n)}{f'(p_n)}, \text{ provided that } f'(p_n) \neq 0, \tag{15}$$

and hence

$$|p_{n+1} - p| = \frac{(p_n - p)^2}{2} |\frac{f''(\xi_n)}{f'(p_n)}|, \text{ provided that } f'(p_n) \neq 0, \tag{16}$$

Using $p_n \to p$, $\xi_n \to p$ as $n \to \infty$ and continuity of $f(x)$ up to at least the second derivative, we arrive at $K = |\frac{f''(p)}{f'(p)}|$. $\square$

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU AESB2210 CTB2400)
Thursday Juli 2 2015, 18:30-21:30

1. We beschouwen de volgende predictor-corrector methode voor de integratie van het beginwaardeprobleem $y' = f(t, y)$, $y(t_0) = y_0$:

$$\begin{aligned} w_{n+1}^* &= w_n + \Delta t f(t_n, w_n), \\ w_{n+1} &= w_n + \Delta t \left((1-\mu) f(t_n, w_n) + \mu f(t_{n+1}, w_{n+1}^*)\right), \end{aligned} \tag{1}$$

waarin Δt de tijdstap, μ een reëel getal ($0 \leq \mu \leq 1$) en w_n de numerieke oplossing op tijdstip t_n voorstelt.

(a) Toon aan dat de locale afbreekfout van de bovenstaande methode voor $0 \leq \mu \leq 1$ van de orde $O(\Delta t)$ is en voor $\mu = \frac{1}{2}$ is de orde $O((\Delta t)^2)$. (*N.B. Dit moet afgeleid worden voor de algemene differentiaalvergelijking $y' = f(t, y)$*). (3 pt)

(b) Leid af dat de versterkingsfactor van deze methode gegeven wordt door

$$Q(\lambda \Delta t) = 1 + \lambda \Delta t + \mu (\lambda \Delta t)^2.$$

(2 pt)

(c) We beschouwen het volgende stelsel niet lineaire differentiaalvergelijkingen:

$$\begin{aligned} x_1' &= -x_1 + \cos x_1 + 2x_2 + t, \ x_1(0) = 0, \\ x_2' &= x_1 - x_2^2, \ x_2(0) = 1. \end{aligned} \tag{2}$$

Voer één stap uit met de methode gegeven in (1) met $\Delta t = \frac{1}{2}$ en $\mu = \frac{1}{2}$. (1 pt)

(d) Laat zien dat de Jacobiaan van het rechterlid van (2) op $t = 0$ gegeven wordt door:

$$\begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix}.$$

(1 pt)

(e) Kies $\mu = 0$. Voor welke waarden van Δt is de methode toegepast op (2) stabiel op $t = 0$? Beantwoord dezelfde vraag voor $\mu = \frac{1}{2}$. (3 pt)

[0] **voor vervolg z.o.z. Voor de uitwerkingen van dit tentamen zie: http://ta.twi.tudelft.nl/nw/users/vuik/wi3097/tentamen.html**

2. We beschouwen de convectie–diffusie vergelijking met Dirichlet randvoorwaarden:

$$\begin{cases} -u'' + u' = 1, & 0 < x < 1, \\ u(0) = 0, & u(1) = 0, \end{cases} \tag{3}$$

waarin $u = u(x)$, $u' = \frac{du}{dx}$ en $u'' = \frac{d^2u}{dx^2}$.

(a) Laat zien dat

$$u(x) = x - \frac{1 - e^x}{1 - e}, \tag{4}$$

de exacte oplossing is van randwaardeprobleem (3). (1 pt.)

(b) We lossen randwaardeprobleem (3) op met eindige differenties, waarin $x_j = j\Delta x$, $(n+1)\Delta x = 1$, met Δx als uniforme stapgrootte. Geef een discretisatiemethode (+bewijs) waarvoor de afbreekfout van orde $O((\Delta x)^2)$ is. Behandel ook de randvoorwaarden. (2 pt.)

(c) Leg uit waarom oscillerende numerieke oplossingen voor (3) als onbetrouwbaar gezien moeten worden. (1 pt.)

(d) Gebruik een stapgrootte van $\Delta x = 1/4$ om het stelsel vergelijkingen $Ay = b$ af te leiden. Verwerk de randvoorwaarden. Het afgeleide stelsel heeft drie vergelijkingen met drie onbekenden, dat betekent dat A een 3×3 matrix is en y en b 1×3 kolomvectoren zijn. Dit stelsel vergelijkingen hoeft **niet** opgelost te worden. (2 pt.)

(e) Gegeven is het iteratieproces $x_{n+1} = g(x_n)$, met

$$g(x_n) = x_n + h(x_n)(x_n^3 - 27),$$

waarbij h een continue functie is met $h(x) \neq 0$ voor elke $x \neq 0$. Als dit proces convergeert, naar welke limiet p convergeert het dan? (1 pt.)

(f) Beschouw drie mogelijke keuzen voor $h(x)$:

i. $h_1(x) = -\frac{1}{x^4}$
ii. $h_2(x) = -\frac{1}{x^2}$
iii. $h_3(x) = -\frac{1}{3x^2}$

Voor welke keuze kan het proces niet convergeren? Voor welke keuze convergeert het proces het snelst? Motiveer uw antwoord. (2 pt.)

(g) p is een nulpunt van een gegeven functie f. $\hat{f}$ is de functie verstoord door meetfouten. Er is gegeven dat $|\hat{f}(x) - f(x)| \leq \epsilon_{max}$ voor alle x. Laat zien dat voor het nulpunt $\hat{p}$ van $\hat{f}$ geldt $|\hat{p} - p| \leq \frac{\epsilon_{max}}{|f'(p)|}$. (1 pt.)

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU AESB2210 CTB2400)
Thursday July 2 2015, 18:30-21:30

1. (a) The local truncation error is given by

$$\tau_{n+1}(\Delta t) = \frac{y_{n+1} - z_{n+1}}{\Delta t} \tag{1}$$

where z_{n+1} is the result of applying the method once with starting solution y_n. Here we obtain y_{n+1} by a Taylor expansion around t_n:

$$y_{n+1} = y_n + \Delta t y'(t_n) + \frac{(\Delta t)^2}{2} y''(t_n) + O((\Delta t)^3). \tag{2}$$

For z_{n+1}, we obtain, after substitution of the predictor step for z^*_{n+1} into the corrector step

$$z_{n+1} = y_n + \Delta t \left((1-\mu) f(t_n, y_n) + \mu f(t_n + \Delta t, y_n + \Delta t f(t_n, y_n))\right) \tag{3}$$

After a Taylor expansion of $f(t_n + \Delta t, y_n + \Delta t f(t_n, y_n))$ around (t_n, y_n) one obtains:

$$z_{n+1} = y_n + \Delta t \left((1-\mu) f(t_n, y_n) + \mu (f(t_n, y_n) + \Delta t (\frac{\partial f(t_n, y_n)}{\partial t} + f(t_n, y_n) \frac{\partial f(t_n, y_n)}{\partial y})) + O((\Delta t)^2) \right). \tag{4}$$

From the differential equation we know that:

$$y'(t_n) = f(t_n, y_n) \tag{5}$$

From the Chain Rule of Differentiation, we derive

$$y''(t_n) = \frac{df(t_n, y_n)}{dt} = \frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} y'(t_n) \tag{6}$$

after substitution of the differential equation one obtains:

$$y''(t_n) = \frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} f(t_n, y_n) \tag{7}$$

This implies that $z_{n+1} = y_n + \Delta t y'(t_n) + \mu (\Delta t)^2 y''(t_n) + O((\Delta t)^3)$. Subsequently, it follows that

$$y_{n+1} - z_{n+1} = O((\Delta t)^2), \text{ and, hence } \tau_{n+1}(\Delta t) = \frac{O((\Delta t)^2)}{\Delta t} = O(\Delta t) \text{ for } 0 \le \mu \le 1, \tag{8}$$

$$y_{n+1} - z_{n+1} = O((\Delta t)^3), \text{ and, hence } \tau_{n+1}(\Delta t) = \frac{O((\Delta t)^3)}{\Delta t} = O((\Delta t)^2) \text{ for } \mu = \frac{1}{2}. \tag{9}$$

(b) Consider the test equation $y' = \lambda y$, then, herewith, one obtains

$$\begin{aligned} w^*_{n+1} &= w_n + \lambda\Delta t w_n = (1+\lambda\Delta t)w_n, \\ w_{n+1} &= w_n + ((1-\mu)\lambda\Delta t w_n + \mu\lambda\Delta t w^*_{n+1}) = \\ &= w_n + ((1-\mu)\lambda\Delta t w_n + \mu\lambda\Delta t(w_n + \lambda\Delta t w_n)) = (1+\lambda\Delta t + \mu(\lambda\Delta t)^2)w_n. \end{aligned} \tag{10}$$

Hence the amplification factor is given by

$$Q(\lambda\Delta t) = 1 + \lambda\Delta t + \mu(\lambda\Delta t)^2. \tag{11}$$

(c) Doing one step with the given method with $\Delta t = \frac{1}{2}$ and $\mu = \frac{1}{2}$ leads to the following steps:
Predictor:

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} -0 + \cos(0) + 2 + 0 \\ 0 - 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \end{pmatrix}$$

Corrector:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2}\left(\frac{1}{2}\begin{pmatrix} 3 \\ -1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} -\frac{3}{2} + \cos(\frac{3}{2}) + 2\cdot\frac{1}{2} + \frac{1}{2} \\ \frac{3}{2} - (\frac{1}{2})^2 \end{pmatrix}\right)$$

which can be written as:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 + \frac{3}{4} - \frac{3}{8} + \frac{1}{4}\cos(\frac{3}{2}) + \frac{3}{8} \\ 1 - \frac{1}{4} + \frac{3}{8} - \frac{1}{16} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} + \frac{1}{4}\cos(\frac{3}{2}) \\ \frac{17}{16} \end{pmatrix} = \begin{pmatrix} 0.7677 \\ 1.0625 \end{pmatrix}$$

(d) In order to compute the Jacobian, we note that the right-hand side of the non linear system can be noted by:

$$f_1(x_1, x_2) = -x_1 + \cos x_1 + 2x_2 + t$$

$$f_2(x_1, x_2) = x_1 - x_2^2$$

From the definition of the Jacobian it follows that:

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} -1 - \sin x_1 & 2 \\ 1 & -2x_2 \end{pmatrix}.$$

Substitution of $\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ shows that

$$J = \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix}.$$

(e) For the stability it is sufficient to check that $|Q(\lambda_i \Delta t)| \leq 1$ for all the eigenvalues of the Jacobian matrix. It is easy to see that the eigenvalues of the Jacobian matrix are $\lambda_1 = -3$ and $\lambda_2 = 0$.

For the choice $\mu = 0$ we note that the method is equal to the Euler Forward method. For real eigenvalues the Euler Forward method is stable if $\Delta t \leq \frac{-2}{\lambda}$. Since $\lambda_1 = -3$ and $\lambda_2 = 0$ we know that the method is stable if $\Delta t \leq \frac{-2}{-3} = \frac{2}{3}$ (another option is to derive the values of Δt such that $|Q(\lambda_i \Delta t)| \leq 1$ by using the description of $Q(\lambda \Delta t)$)

For the choice $\mu = \frac{1}{2}$ we use the expression

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

For $\lambda_2 = 0$ it appears that $Q(\lambda_2 \Delta t) = 1$ so the inequality is satisfied for all Δt. For $\lambda_1 = -3$ we have to check the following inequalities:

$$-1 \leq 1 - 3\Delta t + \frac{9}{2}(\Delta t)^2 \leq 1$$

For the left-hand inequality we arrive at

$$0 \leq \frac{9}{2}(\Delta t)^2 - 3\Delta t + 2$$

It appears that the discriminant $9 - 4 \cdot \frac{9}{2} \cdot 2$ is negative, so there are no real roots which implies that the inequality is satisfied for all Δt.

For the right-hand inequality we get

$$-3\Delta t + \frac{9}{2}(\Delta t)^2 \leq 0$$

$$\frac{9}{2}(\Delta t)^2 \leq 3\Delta t$$

so

$$\Delta t \leq \frac{2}{3}$$

(another option is to see that for $\mu = \frac{1}{2}$ the method is equal to the modified Euler method, and remember that this method is stable for real eigenvalues if $\Delta t \leq \frac{-2}{\lambda}$)

2. (a) First, we check the boundary conditions:

$$u(0) = 0 - \frac{1 - e^0}{1 - e} = \frac{1 - 1}{1 - e} = 0, \quad u(1) = 1 - \frac{1 - e^1}{1 - e} = 0. \tag{12}$$

Further, we have

$$u'(x) = 1 + \frac{e^x}{1-e}, \tag{13}$$

$$u''(x) = \frac{e^x}{1-e}. \tag{14}$$

Hence, we immediately see

$$-u''(x) + u'(x) = -\frac{e^x}{1-e} + 1 + \frac{e^x}{1-e} = 1. \tag{15}$$

Hence, the solution $u(x) = 1 - \frac{1-e^x}{1-e}$ satisfies the differential and the boundary conditions, and therewith $u(x)$ is the solution to the boundary value problem (uniqueness can be demonstrated in a straightforward way, but this was not asked for).

(b) The domain of computation, being $(0,1)$, is divided into subintervals with mesh points, we set $x_j = j\Delta x$, where we use n unknowns, such that $x_{n+1} = (n+1)\Delta x = 1$. We are looking for a discretization with an error of second order, $O((\Delta x)^2)$. To this extent, we use the following central differences approximation at x_j:

$$u'(x_j) \approx \frac{u(x_{j+1}) - u(x_{j-1})}{2\Delta x}, \text{ for } j \in \{1, \dots, n\}. \tag{16}$$

We note that the above formula can be derived formally by writing the derivative as

$$u'(x_j) = \frac{\alpha_0 u(x_{j-1}) + \alpha_1 u(x_j) + \alpha_2 u(x_{j+1})}{\Delta x}, \tag{17}$$

and solve α_0, α_1 and α_2 from checking the zeroth, first and second order derivatives of $u(x)$. Further, the second order derivative is approximated by

$$u''(x_j) \approx \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{(\Delta x)^2}. \tag{18}$$

Since we approximate the derivatives at the point x_j, we use Taylor series expansion about x_j, to obtain:

$$\begin{aligned} u(x_{j+1}) &= u(x_j + \Delta x) = u(x_j) + \Delta x u'(x_j) + \frac{(\Delta x)^2}{2} u''(x_j) + \frac{(\Delta x)^3}{6} u'''(x_j) + O((\Delta x)^4), \\ u(x_{j-1}) &= u(x_j - \Delta x) = u(x_j) - \Delta x u'(x_j) + \frac{(\Delta x)^2}{2} u''(x_j) - \frac{(\Delta x)^3}{6} u'''(x_j) + O((\Delta x)^4), \end{aligned} \tag{19}$$

This gives

$$\begin{aligned} &-\frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{(\Delta x)^2} + \frac{u(x_{j+1}) - u(x_{j-1})}{2\Delta x} = -u''(x_j) + u'(x_j) \\ &+\frac{O((\Delta x)^3)}{2\Delta x} + \frac{O((\Delta x)^4)}{(\Delta x)^2} = -u''(x_j) + u'(x_j) + O((\Delta x)^2). \end{aligned} \tag{20}$$

Hence the error is second order, that is $O((\Delta x)^2)$. Next, we neglect the truncation error, and set $w_j := u(x_j)$ to get

$$-\frac{w_{j+1} - 2w_j + w_{j-1}}{(\Delta x)^2} + \frac{w_{j+1} - w_{j-1}}{2\Delta x} = 1, \text{ for } j \in \{1, \ldots, n\}. \tag{21}$$

At the boundaries, we see for $j = 1$ and $j = n$, upon substituting $w_0 = 0$ and $w_{n+1} = 0$, respectively:

$$\begin{aligned} -\frac{w_2 - 2w_1 + 0}{(\Delta x)^2} + \frac{w_2 - 0}{2\Delta x} &= 1, \\ -\frac{0 - 2w_n + w_{n-1}}{(\Delta x)^2} + \frac{0 - w_{n-1}}{2\Delta x} &= 1. \end{aligned} \tag{22}$$

This can be rewritten more neatly as follows:

$$\begin{aligned} \frac{-w_2 + 2w_1}{(\Delta x)^2} + \frac{w_2}{2\Delta x} &= 1, \\ \frac{2w_n - w_{n-1}}{(\Delta x)^2} - \frac{w_{n-1}}{2\Delta x} &= 1. \end{aligned} \tag{23}$$

(c) The real-valued exact solution and its first and second derivative are given by

$$u(x) = x - \frac{1 - e^x}{1 - e}, \tag{24}$$

$$u'(x) = 1 + \frac{e^x}{1 - e}, \tag{25}$$

$$u''(x) = \frac{e^x}{1 - e}. \tag{26}$$

First, we calculate the point $x^* = \ln(1/(e - 1))$, where $u'(x^*) = 0$ and verify that $u(x)$ attains its maximum value at x^* (since $u''(x^*) = -1/(e - 1)^2 < 0$). Since $u(0) = u(1) = 0$ we can conclude that the exact solution is monotonically increasing on $[0, x^*]$ and monotonically decreasing on $[x^*, 1]$. Since the numerical solution should have the same characteristics as the exact solution, oscillatory solutions should be considered as not reflecting the analytic solution.

(d) Next, we use $\Delta x = 1/4$, then, from equations (21) and (23), one obtains the following system

$$32w_1 - 14w_2 = 1 \tag{27}$$

$$-18w_1 + 32w_2 - 14w_3 = 1 \tag{28}$$

$$-18w_2 + 32w_3 = 1 \tag{29}$$

(e) The iteration process is a fixed point method. If the process converges we have: $\lim_{n\to\infty} x_n = p$. Using this in the iteration process yields:

$$\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} [x_n + h(x_n)(x_n^3 - 27)]$$

Since h is a continuous function one obtains:

$$p = p + h(p)(p^3 - 27)$$

so

$$h(p)(p^3 - 27) = 0.$$

Since $h(x) \neq 0$ for each $x \neq 0$ it follows that $p^3 - 27 = 0$ and thus $p = 27^{\frac{1}{3}} = 3$.

(f) The convergence of a fixed point method $x_{n+1} = g(x_n)$ is determined by $g'(p)$. If $|g'(p)| < 1$ the method converges, whereas if $|g'(p)| > 1$ the method diverges. For all choices we compute the first derivative in p. For the first method we elaborate all steps. For the other methods we only give the final result. For h_1 we have $g_1(x) = x - \frac{x^3-27}{x^4}$. The first derivative is:

$$g_1'(x) = 1 - \frac{3x^2 \cdot x^4 - (x^3 - 27) \cdot 4x^3}{(x^4)^2}$$

Substitution of p yields:

$$g_1'(p) = 1 - \frac{3p^6 - (p^3 - 27) \cdot 4p^3}{p^8}.$$

Since $p = 3$ the final term cancels:

$$g_1'(p) = 1 - \frac{3p^6}{p^8} = 1 - 3^{-1} = \frac{2}{3}.$$

This implies that the method is convergent with convergence factor $\frac{2}{3}$.

For the second method we have:

$$g_2'(p) = 1 - \frac{3p^4 - (p^3 - 27) \cdot 2p}{p^4} = 1 - \frac{3p^4}{p^4} = -2$$

Thus the method diverges.

For the third method we have:

$$g_3'(p) = 1 - \frac{9p^4 - (p^3 - 27) \cdot 6p}{9p^4} = 1 - \frac{9p^4}{9p^4} = 0$$

Thus the method is convergent with convergence factor 0.

Concluding we note that the third method is the fastest.

(g) To estimate the error in p we first approximate the function f in the neighboorhood of p by the first order Taylor polynomial:

$$P_1(x) = f(p) + (x - p)f'(p) = (x - p)f'(p).$$

Due to the measurement errors we know that

$$(x - p)f'(p) - \epsilon_{max} \leq \hat{P}_1(x) \leq (x - p)f'(p) + \epsilon_{max}.$$

This implies that the perturbed root $\hat{p}$ is bounded by the roots of $(x-p)f'(p) - \epsilon_{max}$ and $(x - p)f'(p) + \epsilon_{max}$, which leads to

$$p - \frac{\epsilon_{max}}{|f'(p)|} \leq \hat{p} \leq p + \frac{\epsilon_{max}}{|f'(p)|}.$$