CTB2400 – Numerieke methoden voor differentiaalvergelijkingen

April 2013
Januari 2013
Augustus 2012
Juli 2012
April 2012
Januari 2012
Augustus 2011
Juni 2011
April 2011
Januari 2011

Tentamenbundel Civiele Techniek
Het Gezelschap "Practische Studie"

EEN REPRODUCERENDE
LEERSTIJL
IS SCHADELIJK VOOR
DE ACADEMISCHE
VORMING

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 18 april 2013, 18:30-21:30

1. We beschouwen de numerieke integratie van het volgende beginwaardeprobleem $y' = f(t, y)$, $y(t_0) = y_0$ met de voorwaartse methode van Euler

$$w_{n+1} = w_n + hf(t_n, w_n). \tag{1}$$

   a Bepaal de orde van de locale afbreekfout. (2.5pt.)

   b We beschouwen het volgende tweede orde beginwaardeprobleem

$$\begin{cases} y'' + \varepsilon y' + y = sin(t), \\ y(0) = 1, y'(0) = 0. \end{cases} \tag{2}$$

   Herschrijf dit beginwaardeprobleem in de vorm van een stelsel eerste orde differentiaalvergelijkingen. Neem ook de beginvoorwaarden mee. (1pt.)

   We gaan verder met het volgende stelsel beginwaardeproblemen

$$\begin{cases} y_1' = -y_2, \\ y_2' = y_1 + \varepsilon y_2, \end{cases} \tag{3}$$

   met beginvoorwaarden $y_1(0) = 1$ en $y_2(0) = 2$ en $\varepsilon \in \mathbb{R}$ een gegeven constante.

   c Wat is de maximaal toelaatbare waarde van h voor numerieke stabiliteit als $\varepsilon = 0$? Geef een gedegen toelichting. (2.5pt.)

   d Voor welke waarden van ε is het gegeven stelsel (analytisch) stabiel? (2pt.)

   e Wat is de maximaal toelaatbare waarde van h voor numerieke stabiliteit indien $-2 \leq \varepsilon < 0$? Licht het antwoord toe. (2pt.)

[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 randwaardeprobleem:

$$\begin{cases} -\dfrac{d^2y}{dx^2} + \dfrac{dy}{dx} + y = x^2 + 2x - 2, \quad x \in (0,1), \\ y(0) = 0, \qquad y'(1) = 2. \end{cases} \tag{4}$$

a Toon aan dat $y(x) = x^2$ aan dit randwaardeprobleem voldoet. (1pt.)

We gebruiken een eindige differentiemethode om de oplossing van bovenstaand randwaardeprobleem te benaderen. Laat de gridpunten gegeven worden door $x_j = jh$, met h als stapgrootte. Laat $x_n = nh = 1$.

b Geef een eindige differentieschema (+ motivatie) waarvan de locale afbreekfout van $O(h^2)$ is. *Hint:* Gebruik een virtueel roosterpunt voor de randvoorwaarde op $x = 1$. (3pt.)

c Beredeneer, gebruikmakend van de schatting van de afbreekfout in vorig onderdeel en de exacte oplossing van dit randwaardeprobleem, dat het verschil tussen de numerieke benadering van de oplossing en de exacte oplossing gelijk is aan nul. (2pt.)

Gegeven zijn de volgende tabelwaarden voor de benadering van de functie $y(x) = x^2$.

Tabel 1: Functiewaarden voor $\tilde{y}(x)$ (afgerond op drie decimalen).

x	$\tilde{y}(x)$
0	0
0.25	0.063
0.5	0.25

d Schat $y'(0)$ met behulp van voorwaartse differenties gebruikmakend van de waarden uit Tabel 1 met $h = 0.25$ en $h = 0.5$. (1pt.)

e We bekijken de nauwkeurigheid van de berekening.

i Stel dat de tabelwaarden een (afrond)fout van maximale grootte $\varepsilon = 0.0005$ bevatten, zeg $|\tilde{y}(x_j) - y(x_j)| \leq \varepsilon$ ($y(0) = 0$ is exact), wat is de invloed van deze afrondfout op de fout van de voorwaartse differenties? (1pt.)

ii Toon aan dat de afbreekfout van $O(h)$ is. (1pt.)

iii Gebruik de methode van Richardson om de fout te schatten. (1pt.)

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 14 april 2011, 18:30-21:30

1. We beschouwen de volgende numerieke tijdsintegratiemethode

$$y_{n+1} = y_n + h\left(\alpha f(t_n, y_n) + \beta f(t_{n-1}, y_{n-1})\right). \tag{1}$$

   (a) Laat zien dat voor $\alpha = \frac{3}{2}$ en $\beta = -\frac{1}{2}$ de locale afbreekfout $\mathcal{O}(h^2)$ is. *Hint: Gebruik $y'_{n-1} = f(t_{n-1}, y_{n-1})$, waarin $f(t_{n-1}, y_{n-1})$ verkregen kan worden door een Taylorpolynoom van y' rond t_n.* (3pt.)

   (b) Gebruik de testvergelijking om de versterkingsfactor af te leiden. *Hint:* $y_j = [Q(h\lambda)]y_{j-1}$. (2pt.)

   (c) Laat zien dat de methode stabiel is voor $h \leq -\frac{1}{\lambda}$ als λ een reëel en negatief getal is. (2pt.)

   Beschouw het stelsel

$$\mathbf{y}' = \begin{bmatrix} -4 & 1 \\ 1 & -4 \end{bmatrix} \mathbf{y} + \begin{bmatrix} 0 \\ -\cos(t) \end{bmatrix}. \tag{2}$$

   (d) Bereken de eigenwaarden van de matrix in (2). Bepaal de waarden van h waarvoor het schema stabiel is als we dit toepassen op (2). (2pt.)

   (e) Voor welke waarden van h convergeert het schema in (1)? (1pt.)

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

$$p_n(x) = \sum_{i=0}^{n} f(x_i)L_i(x), \text{ met}$$
$$L_i(x) = \frac{(x - x_0)(\ldots)(x - x_{i-1})(x - x_{i+1})(\ldots)(x - x_n)}{(x_i - x_0)(\ldots)(x_i - x_{i-1})(x_i - x_{i+1})(\ldots)(x_i - x_n)}. \tag{3}$$

   Verder zijn de volgende meetwaarden gegeven in tabelvorm:

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

i	x_i	$f(x_i)$
0	0	1
1	1	2
2	2	4

(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. (2pt.)

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

(d) Stel dat de functiewaarden in de tabel een meetfout bevatten met grootte van ten hoogste ε.

i Laat zien dat de fout, ten gevolge van de onnauwkeurigheid van de meetdata, voor lineaire interpolatie binnen steunpunten x_0 en x_1 begrensd is. (1pt.)

ii Hoe zit dit voor lineaire extrapolatie buiten de steunpunten x_0 en x_1? Geef een motivatie (met name voor het geval dat x ver buiten het interval van de steunpunten ligt). (1pt.)

(e) We beschouwen de trapeziumregel voor numerieke integratie.

i Leid met behulp van het lineaire interpolatiepolynoom de trapeziumregel om $\int_{x_0}^{x_1} f(x)dx$ te benaderen af. (1.5pt.)

ii Leid af dat de afbreekfout van de enkelvoudige trapeziumregel over het interval $[x_0, x_1]$ gegeven is door

$$\frac{1}{12}(x_1 - x_0)^3 \max_{x \in [x_0, x_1]} |f''(x)|, \tag{4}$$

indien de tweede orde afgeleide van $f(x)$ continu is op $[x_0, x_1]$. *Hint: De fout voor lineaire interpolatie over steunpunten x_0 en x_1 wordt gegeven door*

$$f(x) - p_1(x) = \frac{1}{2}(x - x_0)(x - x_1)f''(\chi), \text{ voor zekere } \chi \in (x_0, x_1),$$

waarin $p_1(x)$ het lineaire interpolatiepolynoom voorstelt. (1.5pt.)

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 19 april 2012, 18:30-21:30

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

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

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

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

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

   (2 pt.)

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

   d Beschouw het volgende stelsel

$$\begin{cases} y_1' = -y_1 y_2, \\ y_2' = y_1 y_2 - y_2, \end{cases} \tag{3}$$

   Laat zien dat de Jacobiaan van het rechterlid (die gebruikt wordt voor linearisatie van het stelsel) voor beginvoorwaarde $y_1(0) = 1$ en $y_2(0) = 2$ gegeven wordt door

$$\begin{pmatrix} -2 & -1 \\ 2 & 0 \end{pmatrix}.$$

   (1.5 pt.)

   e Beschouw nu de numerieke methode in vergelijking (1) voor het geval dat $a_1 = a_2 = 1/2$ toegepast op stelsel (3). Is de methode stabiel rond de beginvoorwaarde $y_1(0) = 1$ en $y_2(0) = 2$ en stapgrootte $h = 1$ (+ motivatie)? (1.5 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 randwaardeprobleem:

$$(P_1)\begin{cases} -v'' + v = 2 + x(2-x), & x \in (0,1), \\ v(0) = 0, & v'(1) = 0. \end{cases}$$

a Laat zien dat $v(x) = x(2-x)$ de oplossing is van randwaardeprobleem (P_1). (1 pt.)

b Geef de eindige differentie discretisatie met fout van $O(h^2)$ (+ bewijs), waarin h de afstand tussen gridpunten voorstelt (*Hint: gebruik een virtueel gridpunt bij* $x = 1$). De discretisatie moet symmetrisch zijn. (3 pt.)

c Geef het stelsel vergelijkingen dat verkregen wordt na eindige differentie discretisatie met drie (na verwerking van het virtuele gridpunt) onbekenden ($h = 1/3$). (2 pt.)

d Bereken de fout van de numerieke oplossing, en verklaar uw antwoord. (1 pt.)

e Vervolgens, beschouwen we het volgende niet–lineaire stelsel vergelijkingen

$$\begin{cases} 18v_1 - 9v_2 + v_1^2 = \frac{20}{9}, \\ -9v_1 + 18v_2 + v_2^2 = \frac{20}{9}. \end{cases}$$

Voer een stap met de methode van Newton uit op bovenstaand stelsel waarin u $v_1 = v_2 = 0$ gebruikt als beginschatting. (3 pt.)

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 25 augustus 2011, 18:30-21:30

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

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

waarin h de tijdstap, θ een reëel getal ($0 \le \theta \le 1$) en w_n de numerieke oplossing op tijdstip t_n voorstelt.

   (a) Toon aan dat de locale afbreekfout van de bovenstaande methode voor $0 \le \theta \le 1$ van de orde $O(h)$ en voor $\theta = \frac{1}{2}$ van de orde $O(h^2)$ is (*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

$$Q(h\lambda) = 1 + h\lambda + \theta(h\lambda)^2.$$

(2 pt)

   (c) Gegeven is het tweede orde beginwaardeprobleem:

$$y'' + 4y' + 8y = t^2 - 1, y(0) = 0 \text{ en } y'(0) = 1.$$

   Schrijf dit als een stelsel eerste orde differentiaalvergelijkingen

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} f(t) \\ g(t) \end{pmatrix}.$$

   Geef f en g en laat zien dat $A = \begin{pmatrix} 0 & 1 \\ -8 & -4 \end{pmatrix}$. ($1\frac{1}{2}$ pt)

   (d) Doe één stap met de methode gegeven in (1) met $h = 1$ en $\theta = \frac{1}{2}$. ($1\frac{1}{2}$ pt)

   (e) Voor welke $\theta \in [0, 1]$ is de methode gegeven in (1) met $h = 1$ stabiel bij het toepassen op het stelsel gegeven in onderdeel (c). Geef een duidelijke motivatie. (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. (a) We zoeken een 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

$$f''(0) - Q(h) = O(h).$$

Geef het lineaire stelsel vergelijkingen waar α_0, α_1 en α_2 aan moeten voldoen. (2 pt)

(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 $f''(0) - Q(h)$. (2 pt)

(c) Gebruik de getallen gegeven in Tabel 1. Geef met behulp van de Richardson

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

methode een schatting van de fout: $f''(0) - Q(\frac{1}{4})$. (2 pt)

(d) Gegeven is dat de tabelwaarden een maximale afrondfout hebben van ϵ: $|f(x) - \hat{f}(x)| \leq \epsilon$. Laat zien, dat voor de afrondfout in de benadering geldt:

$$|Q(h) - \hat{Q}(h)| \leq \frac{C_1 \epsilon}{h^2}$$

en geef C_1 en ϵ. (2 pt)

(e) Als gegeven is dat $f''(0) - Q(h) = 6h$, geef dan de optimale waarde van h zodat de totale fout $|f''(0) - \hat{Q}(h)|$ minimaal is. (2 pt)

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 30 augustus 2012, 18:30-21:30

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

$$\begin{aligned} w^*_{n+1} &= w_n + hf(t_n, w_n), \\ w_{n+1} &= w_n + h\left((1-\mu)f(t_n, w_n) + \mu f(t_{n+1}, w^*_{n+1})\right), \end{aligned} \qquad (1)$$

waarin h 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 $O(h)$ is en voor $\mu = \frac{1}{2}$ is de orde $O(h^2)$. (*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

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

(2 pt)

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

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

Voer één stap uit met de methode gegeven in (1) met $h = \frac{1}{2}$ 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 h 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 benaderen de integraal $\int_a^b f(x)dx$ met steunpunten $x_j = a + (j-1)h$, waarin $x_{n+1} = b$. Voor een interval tussen twee naburige steunpunten, (x_j, x_{j+1}), geeft de rechthoekregel de benadering $hf(x_j)$ en herhaalde toepassing geeft $\int_a^b f(x)dx \approx T_0 = h\sum_{j=1}^n f(x_j)$.

   (a) Laat zien dat de locale afbreekfout, $|E_0^I| := |\int_{x_j}^{x_{j+1}} f(x)dx - hf(x_j)|$, en globale fout, $|E_0| := |\int_a^b f(x)dx - T_0|$ achtereenvolgens gegeven worden door

$$|E_0^I| \le \frac{h^2}{2} \max_{x\in[x_j,x_{j+1}]} |f'(x)|, \text{ en } |E_0| \le \frac{(b-a)h}{2} \max_{x\in[a,b]} |f'(x)|. \tag{3}$$

   Hint: U kunt de stelling van Taylor gebruiken en $\max_{x\in[a,b]}|f(x)|$ *staat voor het maximum van* $|f(x)|$ *over het interval* $[a,b]$. (2 pt.)

   (b) Nu nemen we ook de eerste afgeleide van f in de steunpunten $\{x_j\}$ mee. Gebruik de stelling van Taylor om af te leiden dat de integraal m.b.v. de steunpunten benaderd kan worden door T_1 met globale fout E_1, waarin

$$\int_a^b f(x)dx \approx T_1 = h\sum_{j=1}^n \left[f(x_j) + \frac{h}{2}f'(x_j)\right], \; |E_1| \le \frac{(b-a)h^2}{6} \max_{x\in[a,b]} |f''(x)|. \tag{4}$$

   (2pt.)

   (c) Gebruik de methode uit vergelijking (4) met $h = \frac{1}{2}$ om $\int_0^1 x^2 dx$ te benaderen en vergelijk de fout met de schatting voor $|E_1|$. (2pt.)

   (d) Nu gebruiken we de afgeleiden van f tot en met de 2-de orde. Verder zijn T_2 en E_2 achtereenvolgens de benadering van $\int_a^b f(x)dx$ met deze afgeleiden en de bijbehorende globale fout. Toon aan dat

$$T_2 = T_1 + \frac{h^3}{3!}\sum_{j=1}^n f''(x_j), \text{ en } |E_2| \le \frac{(b-a)h^3}{4!} \max_{x\in[a,b]} |f'''(x)|. \tag{5}$$

   (2pt.)

   (e) Stel dat alle functiewaarden en hun afgeleiden een meet-of afrondfout van ten hoogste ε bevatten, d.w.z. $|\tilde{f}^{(k)}(x_j) - f^{(k)}(x_j)| \le \varepsilon$, voor alle j en k–de afgeleiden ($k = 0$ geeft f zelf). Laat T_2 en $\tilde{T}_2$ achtereenvolgens berekend zijn met de exacte (f) en beschikbare waarden ($\tilde{f}$) van f en zijn afgeleiden, toon aan dat de invloed van deze fout afgeschat kan worden door:

$$|\tilde{T}_2 - T_2| \le (b-a)\varepsilon\left(1 + \frac{h}{2} + \frac{h^2}{3!}\right). \tag{6}$$

   (2pt.)

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 20 januari 2011, 18:30-21:30

1. De modified Euler methode voor de integratie van beginwaarde probleem $y' = f(t, y)$, $y(t_0) = y_0$, is gegeven door

$$\begin{cases} w_{n+1}^* = w_n + hf(t_n, w_n) \\ w_{n+1} = w_n + \dfrac{h}{2}\left(f(t_n, w_n) + f(t_{n+1}, w_{n+1}^*)\right), \end{cases} \tag{1}$$

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

(a) Toon aan dat de locale afbreekfout van de modified Euler methode van de orde $O(h^2)$ is. (U mag hier niet de testvergelijking gebruiken.) (3pt.)

Gegeven het beginwaardeprobleem

$$\begin{cases} \dfrac{d^2y}{dt^2} + 4\dfrac{dy}{dt} + 3y = \cos t, \\ y(0) = 1, \quad \dfrac{dy}{dt}(0) = 2. \end{cases} \tag{2}$$

(b) Laat zien dat bovenstaand beginwaarde probleem geschreven kan worden als

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -3 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \cos t \end{pmatrix}. \tag{3}$$

Geef ook de beginvoorwaarden voor $x_1(0)$ en $x_2(0)$. (1pt.)

(c) Bereken één stap met de modified Euler methode, waarbij $h = 0.1$ en $t_0 = 0$ met de gegeven beginvoorwaarden. (2pt.)

(d) Leid de versterkingsfactor voor de modified Euler methode af. (2pt.)

(e) Bepaal voor welke stapgrootte $h > 0$ de modified Euler methode toegepast op beginwaarde probleem (2), stabiel is. (2pt.)

[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. In deze opgave beschouwen we twee numerieke methoden voor het oplossen van niet-lineaire vergelijkingen.

   (a) We bepalen het nulpunt van een algemene gegeven functie $f(x)$ die een continue afgeleide heeft. We gebruiken een vaste punts methode van Picard, met

   $$p_{k+1} = g(p_k) = p_k - \frac{f(p_k)}{\alpha}, \text{ waarin } \alpha \in \mathbb{R}.$$

   Stel dat p een vast punt is, laat zien dat als $0 < f'(p) < \alpha$ de bovenstaande keuze voor $g(x)$ altijd convergentie oplevert voor een beginschatting p_0 gekozen voldoende dicht bij p. (3pt.)

   Gegeven is de Newton-Raphson methode

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

   (b) Leid de bovenstaande Newton-Raphson methode af. (2pt.)

   (c) We zoeken het positieve nulpunt van $f(x) = x^2 - 2x - 2$. Neem als startwaarde $p_0 = 2$ en bepaal p_1 met de Newton-Raphson methode. (1pt.)

   (d) Motiveer waarom de startwaarde $p_0 = 1$ geen logische keuze is voor de Newton-Raphson methode. (2pt.)

   (e) We voeren nu een interpolatie uit op een functie $y = y(x)$ met steunpunten $(\frac{2}{3}, \frac{1}{3})$ en (1,1).

   - Geef de formule voor het lineaire interpolatiepolynoom $P(x)$ met steunpunten $y(\frac{2}{3}) = \frac{1}{3}$ en $y(1) = 1$. (1pt.)
   - Bepaal het punt $\tilde{x}$ waar $P(\tilde{x}) = \frac{1}{2}$ met behulp van steunpunten in het eerste onderdeel van deze vraag. (Dit is inverse lineaire interpolatie ofwel Regula-Falsi.) (1pt.)

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 26 januari 2012, 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 $y(t_0) = y_0$:

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

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

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

   (2 pt.)

   (b) Geef de orde (+ bewijs) van de lokale afbreekfout van de Trapeziummethode voor de testvergelijking. *Hint:* $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots,\quad \frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots.$ (3 pt.)

   (c) Toon aan dat voor een algemene complexe $\lambda = \mu + i\nu$ de methode stabiel is voor elke stapgroote $h > 0$ als $\mu \leq 0$. (2 pt.)

   (d) Doe één stap met de Trapeziummethode voor het volgende beginwaardeprobleem

$$y' = -(1 + t)y + t, \text{ met } y(0) = 1,$$

   en stapgrootte $h = 1$. (1.5 pt.)

   (e) Maak voor dit probleem (gegeven in onderdeel d) een vergelijking van de Trapeziummethode en de Euler Voorwaarts methode. Aan welke methode geeft u de voorkeur (+ motivatie)? (1.5 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. In de eerste drie onderdelen van deze opgave maken we gebruik van een hypothetische computer die met floating point (decimale) getallen kan rekenen. Deze computer heeft de volgende specificaties:

- Ieder reëel getal wordt voorgesteld als floating point number met vier cijfers achter de komma;
- De floating point weergave vindt plaats door *afronding*.

Dus als voorbeeld: $fl(5/7) = fl(0.714285714\ldots) = 0.7143 \cdot 10^0$. In de opgave beschouwen we volgende twee gegeven getallen $x = 2/3 = 0.666666666\ldots$ en $y = 1999/3000 = 0.666333333\ldots$.

[a] Bereken $x + y$, $x - y$, $fl(fl(x) + fl(y))$ en $fl(fl(x) - fl(y))$, met de hierboven gegeven waarden voor x en y, als exacte uitkomsten en computerweergaven van deze uitkomsten. (1.5 pt)

[b] Geef de relatieve fout die optreedt als gevolg van de afronding in de berekeningen door onze computer voor $x + y$ en $x - y$. (1.5 pt)

[c] Geef een motivatie waarom de relatieve fout in het algemeen als $x \approx y$ voor $x - y$ dramatisch hoger ligt dan voor $x + y$ onder aanname dat $x, y > 0$. (2 pt)

In the tweede deel van deze som, beschouwen we het volgende randwaardeprobleem (differentiaalvergelijking met randvoorwaarden):

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

[d] Laat h de stapgrootte zijn. Geef een discretisatie met een fout van $O(h^2)$ (+ bewijs) zo dat $x_n = 1$. Gebruik een virtueel gridpunt bij $x = 0$. (3pt.)

[e] Gebruik een stapgrootte van $h = 1/3$ om het stelsel vergelijkingen af te leiden. Verwerk de randvoorwaarden. Het afgeleide stelsel moet 3×3 zijn (drie onbekenden en drie vergelijkingen). (2pt.)

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
maandag 28 januari 2013, 18:30-21:30

1. De modified Euler methode voor de integratie van beginwaarde probleem $y' = f(t, y)$, $y(t_0) = y_0$, is gegeven door

$$\begin{cases} w_{n+1}^* = w_n + hf(t_n, w_n) \\ \\ w_{n+1} = w_n + \frac{h}{2}\left(f(t_n, w_n) + f(t_{n+1}, w_{n+1}^*)\right). \end{cases} \quad (1)$$

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

[a] Toon aan dat de locale afbreekfout van deze methode $O(h^2)$ is. (3 pt)

De versterkingsfactor wordt gegeven door

$$Q(h\lambda) = 1 + h\lambda + \frac{(h\lambda)^2}{2}.$$

[b] Leid deze versterkingsfactor voor de modified Euler methode af. (1 pt)

Gegeven het beginwaardeprobleem

$$\begin{cases} \frac{d^2y}{dt^2} + 12\frac{dy}{dt} + 72y = \sin t, \\ y(0) = 1, \quad \frac{dy}{dt}(0) = 2. \end{cases} \quad (2)$$

[c] Laat zien dat bovenstaand beginwaarde probleem geschreven kan worden als

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -72 & -12 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \sin t \end{pmatrix}. \quad (3)$$

Geef ook de beginvoorwaarden voor $x_1(0)$ en $x_2(0)$. (2 pt)

[d] Bereken één stap met de modified Euler methode, waarbij $h = 0.1$ en $t_0 = 0$ met de gegeven beginvoorwaarden uit (2). (2 pt)

[e] Ga na of de modified Euler methode, toegepast op het gegeven beginwaarde probleem (2), stabiel is voor $h = 0.25$. (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. Van een voertuig wordt de snelheid geschat. De toegestane maximumsnelheid is 40 m/s. De gemeten posities van het voertuig staan in de onderstaande tabel.

t (s)	0	1	2
$f(t)$ (m)	200	215	250

(a) Geef de 1^e orde achterwaartse differentieformule en bepaal hiermee een schatting van de snelheid op $t = 2$ ($f'(2)$). (1 pt.)

(b) We zoeken een differentie formule voor de eerste afgeleide van f in het punt $2h$ van de vorm: $Q(h) = \frac{\alpha_0}{h} f(0) + \frac{\alpha_1}{h} f(h) + \frac{\alpha_2}{h} f(2h)$, zodat $f'(2h) - Q(h) = O(h^2)$. Laat zien dat de coefficienten α_0, α_1 en α_2 moeten voldoen aan het volgende stelsel:

$$\begin{array}{rcrcrcl} \frac{\alpha_0}{h} & + & \frac{\alpha_1}{h} & + & \frac{\alpha_2}{h} & = & 0\,, \\ -2\alpha_0 & - & \alpha_1 & & & = & 1\,, \\ 2\alpha_0 h & + & \frac{1}{2}\alpha_1 h & & & = & 0\,. \end{array}$$

(2 pt.)

(c) De oplossing van dit stelsel wordt gegeven door $\alpha_0 = \frac{1}{2}$, $\alpha_1 = -2$ en $\alpha_2 = \frac{3}{2}$. Geef een uitdrukking voor de afbreekfout $f'(2h) - Q(h)$ en een schatting van de snelheid. (2 pt.)

(d) De gemeten posities hebben een maximale meetfout van ϵ: $|f(t) - \hat{f}(t)| \leq \epsilon$. Laat zien, dat voor de meetfout in de benadering geldt: $|Q(h) - \hat{Q}(h)| \leq \frac{C_1 \epsilon}{h}$ en geef C_1. (1.5 pt.)

(e) Leid met behulp van het lineaire interpolatiepolynoom de trapeziumregel om $\int_{x_0}^{x_1} f(x)dx$ te benaderen af. (1.5pt.)

(f) Leid af dat de afbreekfout van de enkelvoudige trapeziumregel over het interval $[x_0, x_1]$ gegeven is door $\frac{1}{12}(x_1 - x_0)^3 \max_{x\in[x_0,x_1]} |f''(x)|$, indien de tweede orde afgeleide van $f(x)$ continu is op $[x_0, x_1]$. *Hint: De fout voor lineaire interpolatie over steunpunten x_0 en x_1 wordt gegeven door*

$$f(x) - p_1(x) = \frac{1}{2}(x - x_0)(x - x_1)f''(\chi), \text{ voor zekere } \chi \in (x_0, x_1),$$

waarin $p_1(x)$ het lineaire interpolatiepolynoom voorstelt. (2pt.)

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 5 juli 2012, 18:30-21:30

1. Om het beginwaardeprobleem, gegeven door

$$y' = f(t, y(t)), \qquad y(t^0) = y^0, \tag{1}$$

te integreren, beschouwen we de Trapezium Regel

$$w^{n+1} = w^n + \frac{h}{2}(f(t^n, w^n) + f(t^{n+1}, w^{n+1})), \tag{2}$$

en de Modified Euler Methode

$$\begin{cases} \hat{w}^{n+1} = w^n + hf(t^n, w^n), & \text{predictie–stap,} \\ w^{n+1} = w^n + \frac{h}{2}(f(t^n, w^n) + f(t^{n+1}, \hat{w}^{n+1})), & \text{correctie–stap.} \end{cases} \tag{3}$$

Hier staat w^n voor de numerieke benadering op tijdstip $t^n = t^0 + nh$, en geeft h de tijdsstap weer.

[a] Gebruik de test–vergelijking om aan te tonen dat de versterkingsfactoren van beide methoden gegeven worden door

$$\begin{aligned} Q_T(h\lambda) &= \frac{1 + \frac{h\lambda}{2}}{1 - \frac{h\lambda}{2}}, && \text{Trapezium Regel,} \\ Q_{ME}(h\lambda) &= 1 + h\lambda + \tfrac{(h\lambda)^2}{2}, && \text{Modified Euler Methode.} \end{aligned} \tag{4}$$

(2pt.)

[b] Laat zien dat de locale afbreekfout van orde $O(h^2)$ is.

Hint: U mag de test–vergelijking voor zowel de Trapezium Regel als de Modified Euler Methode gebruiken. Verder geldt $e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + O(x^4)$ *en voor* $|x| < 1$ *geldt* $\frac{1}{1-x} = 1 + x + x^2 + x^3 + O(x^4)$. (3pt.)

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

We passen beide methoden toe op het beginwaardeprobleem

$$y'' + y = t(1-t), \quad y(0) = 0, \quad y'(0) = 1. \tag{5}$$

[c] Laat zien dat, door gebruik te maken van $y_1(t) = y(t)$ en $y_2(t) = y'(t)$, dat dit beginwaardeprobleem herschreven kan worden als het volgende stelsel vergelijkingen

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ t(1-t) \end{pmatrix}, \tag{6}$$

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

[d] Gebruik $h = \frac{1}{2}$ om w_1 (één tijdsstap) te berekenen met zowel de Trapezium Regel als de Modified Euler Methode. (2pt.)

[e] Welke van de twee methoden toegepast op het huidige beginwaardeprobleem (zie opgaven [c–d]) heeft volgens u de voorkeur? Licht uw keuze toe in termen van nauwkeurigheid, stabiliteit en hoeveelheid werk. (2pt.)

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

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

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

(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. (2pt.)

(c) 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 $|\hat{p} - p| \leq \frac{\epsilon_{max}}{|f'(p)|}$. (1pt.)

(d) We gebruiken vervolgens het Newton-Raphson schema, gegeven door

$$z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}.$$

We nemen nu $f(x) = x^4 - 3x$. Voer één stap uit met dit Newton-Raphson schema met beginschatting $z_0 = 1$. (2 pt.)

(e) Leid de Newton-Raphson methode af. (2 pt.)

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

$$|z - z_{k+1}| = K|z - z_k|^2, \text{ voor } k \to \infty \tag{7}$$

en bepaal de waarde van de constante K (voor $k \to \infty$). (2 pt.)

TECHNISCHE UNIVERSITEIT DELFT
Faculteit Elektrotechniek, Wiskunde en Informatica

TENTAMEN NUMERIEKE METHODEN VOOR DIFFERENTIAALVERGELIJKINGEN (WI3097 TU)
donderdag 30 juni 2011, 18:30-21:30

1. We beschouwen de numerieke integratie van het volgende beginwaardeprobleem $y' = f(t, y)$, $y(t_0) = y_0$. We gebruiken de voorwaartse methode van Euler om de numerieke oplossing van dit beginwaardeprobleem te bepalen. Deze methode is gegeven door
$$w_{n+1} = w_n + hf(t_n, w_n), \quad (1)$$
waarin h de tijdstap en w_n de numerieke oplossing op tijdstip t_n voorstelt.

   a Bepaal, met gedegen toelichting, de orde van de locale afbreekfout. (2.5pt.)

   b We beschouwen het volgende tweede orde beginwaardeprobleem
$$\begin{cases} y'' + \varepsilon y' + y = sin(t), \\ y(0) = 1, y'(0) = 0. \end{cases} \quad (2)$$
   Herschrijf, met gedegen toelichting, dit beginwaardeprobleem in de vorm van een stelsel eerste orde differentiaalvergelijkingen. Neem ook de beginvoorwaarden mee. (1pt.)

   We gaan verder met het volgende stelsel beginwaardeproblemen
$$\begin{cases} y_1' = -y_2, \\ y_2' = y_1 + \varepsilon y_2, \end{cases} \quad (3)$$
   met beginvoorwaarden $y_1(0) = 1$ en $y_2(0) = 2$, en verder is $\varepsilon \in \mathbb{R}$ een gegeven constante.

   c Wat is de maximaal toelaatbare waarde van h voor numerieke stabiliteit als $\varepsilon = 0$? Geef een gedegen toelichting. (2.5pt.)

   d Voor welke waarden van ε is het gegeven stelsel (analytisch) stabiel? Geef een goede toelichting. (2pt.)

   We onderzoeken de numerieke stabiliteit met de voorwaartse methode van Euler voor het gegeven stelsel beginwaardeproblemen voor algemene waarden van ε.

   e Wat is de maximaal toelaatbare waarde van h voor numerieke stabiliteit indien $-2 \leq \varepsilon < 0$? Licht het antwoord toe. (2pt.)

[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. Gegeven het volgende randwaardeprobleem:

$$\begin{cases} -y'' + x^2 y = x, \text{ voor } x \in (0,1) \\ y(0) = 0, y(1) = 1. \end{cases} \tag{4}$$

We benaderen de oplossing van dit probleem met behulp van eindige differenties. De roosterpunten worden gegeven door $x_j = jh$, $j \in \{0, \ldots, n+1\}$ met $h = \dfrac{1}{n+1}$.

a Geef een discretisatie (+bewijs) van $-y'' + x^2 y = x$ waarbij de lokale afbreekfout van de orde $O(h^2)$ is. (3 pt)

b Geef voor $n = 3$ het stelsel $A\mathbf{w} = \mathbf{b}$ waaraan de numerieke oplossing $\mathbf{w}$ moet voldoen. (2 pt)

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

$$g(x_n) = x_n + h(x_n)(x_n^2 - 4),$$

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

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

i. $h_1(x) = -\frac{1}{2}x$

ii. $h_2(x) = -\frac{1}{3}$

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

We beperken ons tot de limiet $p > 0$. Voor welke keuze kan het proces niet convergeren? Voor welke keuze convergeert het proces het snelst? Motiveer uw antwoord. (2pt.)

e Doe 3 iteraties met de keuze $h_2(x) = \frac{1}{3}$ met startwaarde $x_0 = 3$. (1pt.)

f We beschouwen nu het geval waarin we het nulpunt p van een gegeven functie f bepalen. $\hat{f}$ 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 $|\hat{p} - p| \leq \frac{\epsilon_{max}}{|f'(p)|}$. (1pt.)

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Thursday April 18 2013, 18:30-21:30

1. a The local truncation error is defined by

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

where

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

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

$$y_{n+1} = y_n + hy'(t_n) + \frac{h^2}{2}y''(\xi), \quad \exists\, \xi \in (t_n, t_{n+1}). \tag{3}$$

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

$$\tau_h = \frac{h}{2}y''(\xi), \quad \exists\, \xi \in (t_n, t_{n+1}). \tag{4}$$

Hence, the truncation error is of first order.

b We define $y_1 := y$ and $y_2 := y'$, hence $y_1' = y_2$. Further, we use the differential equation to obtain

$$y'' + \varepsilon y' + y = y_1'' + \varepsilon y_1' + y_1 = y_2' + \varepsilon y_2 + y_1. \tag{5}$$

Hence, we obtain

$$y_2' = -y_1 - \varepsilon y_2 + sin(t). \tag{6}$$

Hence the system is given by

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= -y_1 - \varepsilon y_2 + sin(t). \end{aligned} \tag{7}$$

The initial conditions are given by

$$\begin{aligned} 1 &= y(0) = y_1(0), \\ 0 &= y'(0) = y_1'(0) = y_2(0). \end{aligned} \tag{8}$$

c First, we use the test equation, $y' = \lambda y$, to analyze numerical stability. For forward Euler, we obtain

$$w_{n+1} = w_n + h\lambda w_n = Q(h\lambda) w_n, \tag{9}$$

hence the amplification factor becomes

$$Q(h\lambda) = 1 + h\lambda. \tag{10}$$

The numerical solution is stable if and only if $|Q(h\lambda)| \leq 1$. Next, we deal with the case $\varepsilon = 0$, to obtain the following system

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{11}$$

This system gives the following eigenvalues $\lambda_{1,2} = \pm i$, where i is the imaginary unit. Hence, the amplification factor is given by

$$Q(h\lambda) = 1 \pm hi. \tag{12}$$

Then, it is immediately clear that $|Q(h\lambda)| > 1$ for all $h > 0$. Hence, we conclude that the forward Euler method is never stable if $\varepsilon = 0$.

d From Assignment 1.c., we know that if $\varepsilon = 0$, the eigenvalues of the system are purely imaginary. This implies that the system is analytically (zero) stable if $\varepsilon = 0$.

Nonzero values of ε give the following system

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & \varepsilon \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{13}$$

then we get the following eigenvalues $\lambda_{1,2} = \frac{\varepsilon}{2} \pm \frac{1}{2}\sqrt{\varepsilon^2 - 4}$ (real-valued), if $\varepsilon^2 - 4 \geq 0$ and $\lambda = \frac{\varepsilon}{2} \pm \frac{i}{2}\sqrt{4 - \varepsilon^2}$ (nonreal-valued) if $\varepsilon^2 - 4 < 0$. Hence, we consider two cases: real-valued and nonreal-valued eigenvalues.

Real-valued eigenvalues
In this case $|\varepsilon| \geq 2$, and $0 \leq \varepsilon^2 - 4 < \varepsilon^2$, and hence the real-valued eigenvalues have the same sign, which is determined by the sign of ε. Hence, if $\varepsilon \leq -2$, then, the system is stable. Furthermore, if $\varepsilon \geq 2$, then, the system is unstable.

Nonreal-valued eigenvalues
In this case $|\varepsilon| < 2$. The system is analytically unstable if and only if the real part of the eigenvalues is positive. Further, the real part of the eigenvalues is positive if and only if $\varepsilon > 0$. Hence, the system is analytically unstable if and only if $\varepsilon > 0$. Hence, the system is stable if and only if $(-2 <)\varepsilon \leq 0$.

From these arguments, it follows that the system is stable if and only if $\varepsilon \leq 0$.

e Since currently the discriminant, ε^2-4, is negative, the eigenvalues are nonreal. Substitution into the amplification factor yields

$$Q(h\lambda) = 1 + \frac{\varepsilon}{2}h \pm \frac{ih}{2}\sqrt{4-\varepsilon^2}. \tag{14}$$

Hence, numerical stability is warranted if

$$|Q(h\lambda)|^2 = (1+\frac{\varepsilon}{2}h)^2 + \frac{h^2}{4}(4-\varepsilon^2) \leq 1. \tag{15}$$

Hence for stability, we have

$$1 + \varepsilon h + \frac{\varepsilon^2 h^2}{4} + h^2 - \frac{\varepsilon^2 h^2}{4} = 1 + h\varepsilon + h^2 \leq 1. \tag{16}$$

Since $h > 0$, we obtain the following stability criterion

$$h \leq -\varepsilon = |\varepsilon|. \tag{17}$$

If $\varepsilon = -2$, then both eigenvalues are real-valued and given by $\lambda_{1,2} = -1$. For this case, we obtain $Q(\lambda h) = 1 - h$, and stability is warranted if and only if $-1 \leq Q(h\lambda) \leq 1$, hence $h \leq 2(= |\varepsilon|)$.

We conclude that for $-2 \leq \varepsilon < 0$, we have a numerically stable solution if and only if $h \leq |\varepsilon|$.

2. a First we check that $y(x) = x^2$ satisfies the boundary conditions. It immediately follows that $y(0) = 0$ and using $y'(x) = 2x$, gives $y'(1) = 2$, and hence the boundary conditions are satisfied. Further, substitution of $y = x^2$, using $y''(x) = 2$, gives

$$-y'' + y' + y = -2 + 2x + x^2, \tag{18}$$

which is equal to the right-hand side of the differential equation and hence $y(x) = x^2$ satisfies the boundary value problem (the differential equation and the boundary conditions).

b Let $x_j = jh$, $x_n = 1$, hence $h = \frac{1}{n}$. We use a Taylor Series to express the relation between the differences formulae and the derivatives. Using the convention that

$y_j = y(x_j)$, gives

$$
\begin{aligned}
&-\frac{y_{j+1}-2y_j+y_{j-1}}{h^2}+\frac{y_{j+1}-y_{j-1}}{2h}+y_j=\\
&-\frac{y_j+hy'(x_j)+\frac{h^2}{2}y''(x_j)+\frac{h^3}{3!}y'''(x_j)+\frac{h^4}{4!}y''''(x_j)+O(h^5)-2y_j}{h^2}-\\
&\frac{y_j-hy'(x_j)+\frac{h^2}{2}y''(x_j)-\frac{h^3}{3!}y'''(x_j)+\frac{h^4}{4!}y''''(x_j)+O(h^5)}{h^2}+\\
&\frac{y_j+hy'(x_j)+\frac{h^2}{2}y''(x_j)+\frac{h^3}{3!}y'''(x_j)+O(h^4)}{2h}-\\
&\frac{y_j-hy'(x_j)+\frac{h^2}{2}y''(x_j)-\frac{h^3}{3!}y'''(x_j)+O(h^4)}{2h}+y_j=\\
&-y''(x_j)+y'(x_j)+y(x_j)+\frac{h^2}{12}\left(y''''(x_j)+2y'''(x_j)\right)+O(h^3).
\end{aligned}
\tag{19}
$$

Hence the local trunction error for the discretization in the interior gives a order $O(h^2)$, where minimal third–order derivatives are involved. Further, using a virtual gridnode at $x_{n+1} = 1 + h$, gives

$$
\begin{aligned}
&\frac{y_{n+1}-y_{n-1}}{2h}=\frac{y(1)+hy'(1)+\frac{h^2}{2}y''(1)+\frac{h^3}{3!}y'''(1)+O(h^4)}{2h}-\\
&\frac{y(1)-hy'(1)+\frac{h^2}{2}y''(1)-\frac{h^3}{3!}y'''(1)+O(h^4)}{2h}=y'(1)+\frac{h^2}{6}y'''(1)+O(h^3)=\\
&2+\frac{h^2}{6}y'''(1)+O(h^3).
\end{aligned}
\tag{20}
$$

Hence, also for the differencing at $x = 1$, a local truncation error of $O(h^2)$ is obtained with derivatives of minimal third order. Hence all difference formulae give a (local) truncation error of order $O(h^2)$. Neglecting the truncation errors, and setting $f(x) = x^2 + 2x - 2$, gives the following finite difference approach for the numerical approximation w_j:

$$\frac{-w_{j-1}+2w_j-w_{j+1}}{h^2}+\frac{w_{j+1}-w_{j-1}}{2h}+w_j=f(x_j), \qquad j=1\dots n. \tag{21}$$

The above equation can be simplified to

$$-(\frac{1}{h^2}+\frac{1}{2h})w_{j-1}+(1+\frac{2}{h^2})w_j+(-\frac{1}{h^2}+\frac{1}{2h})w_{j+1}=f(x_j), \qquad j=1\dots n. \tag{22}$$

Using the boundary condition $w_0 = 0$, gives for $j = 0$:

$$(1 + \frac{2}{h^2})w_1 + (-\frac{1}{h^2} + \frac{1}{2h})w_2 = f(x_1). \tag{23}$$

For $j = n$, we substitute

$$\frac{w_{n+1} - w_{n-1}}{2h} = 2 \Leftrightarrow w_{n+1} = w_{n-1} + 4h, \tag{24}$$

to obtain for $j = n$

$$-\frac{2}{h^2}w_{n-1} + (1 + \frac{2}{h^2})w_n = f(x_j) + \frac{4}{h} - 2. \tag{25}$$

Herewith, we got a discretization with local truncation errors of $O(h^2)$.

c In the previous assignment, we saw that all truncation errors are of order $O(h^2)$ with derivatives of minimal third order. Since $y(x) = x^2$ is the (only) solution to the boundary value problem considered currently, we see that all p–th order derivatives $y^{(p)}(x) = 0$, for $p \geq 3$, and hence all truncation errors are zero. Therefore, for the present boundary value problem, the current finite differences approach gives the exact solution to the boundary value problem (hence the difference between the exact solution and the numerical approximation vanishes).

d The forward difference formula, $Q(h)$, to approximate $y'(0)$ is given by

$$Q(h) = \frac{y(h) - y(0)}{h}. \tag{26}$$

For $h = 0.25$ and $h = 0.5$ from the tabular values, we, respectively, get $\tilde{Q}(0.25) = 0.252$ and $\tilde{Q}(0.5) = 0.5$. Note that the tildes indicate that we used the approximate values for $y(x)$ from Table 1.

e i Let $\tilde{y}(x_j)$, and $y(x_j)$, respectively, represent the approximate values and exact values, and let $\tilde{Q}(h)$ denote the differencing executed with the approximate values for y, then

$$|Q(h) - \tilde{Q}(h)| = |\frac{y(h) - y(0)}{h} - \frac{\tilde{y}(h) - y(0)}{h}| = \frac{|y(h) - \tilde{y}(h)|}{h} \leq \frac{\varepsilon}{h} = \frac{0.0005}{h}. \tag{27}$$

(Note that this gives an upperbound $|Q(h) - \tilde{Q}(h)| \leq 0.002$.)

ii The truncation error is given by

$$y'(0) - \frac{y(h) - y(0)}{h} = y'(0) - \frac{y(0) + hy'(0) + \frac{h^2}{2}y''(0) + O(h^3) - y(0)}{h} = -\frac{h}{2}y''(0) + O(h^2). \tag{28}$$

Hence the truncation error is of order $O(h)$.

iii The truncation error is of first order, hence for h sufficiently small, we have

$$y'(0) \approx Q(h) + Kh, \tag{29}$$

where Kh is an estimate of the error, and for $2h$, we get

$$y'(0) \approx Q(2h) + 2Kh, \tag{30}$$

Subtraction of these two equations and using the values computed earlier, gives the following estimate of the error

$$Kh \approx Q(h) - Q(2h) = 0.252 - 0.5 = -0.248. \tag{31}$$

(Not asked for: This estimate can be used to update the originally computed approximation:

$$y'(0) = Q(h) + Kh = 0.25 - 0.248 = 0.002. \tag{32}$$

It is possible to show that the discrepance with zero follows from the influence of rounding.)

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Thursday April 14 2011, 18:30-21:30

1. We consider the following method:

$$y_{n+1} = y_n + h\left(\alpha f(t_n, y_n) + \beta f(t_{n-1}, y_{n-1})\right). \tag{1}$$

   (a) The local truncation error is given by

$$\tau_{n+1} = \frac{y_{n+1} - z_{n+1}}{h},$$

   where y_{n+1} is the exact solution at time t_{n+1} and z_{n+1} is the numerical method (1) applied to y_{n-1} and y_n. We will need the following Taylor expansions:

$$y_{n+1} = y_n + hy'_n + \frac{h^2}{2}y''_n + \mathcal{O}(h^3),$$

$$y'_{n-1} = f(t_{n-1}, y_{n-1}) = y'_n - hy''_n + \mathcal{O}(h^2).$$

   We then have for z_{n+1} :

$$z_{n+1} = y_n + h(\alpha+\beta)y'_n - h^2\beta y''_n + \mathcal{O}(h^3).$$

   Subtracting this from y_{n+1} and dividing by h we have the local truncation error is:

$$\tau_{n+1} = (1-(\alpha+\beta))y'_n + h\left(\frac{1}{2}+\beta\right)y''_n + \mathcal{O}(h^2).$$

   Since

$$\alpha = \frac{3}{2}, \qquad \beta = -\frac{1}{2}$$

   we obtain $\tau_{n+1} = \mathcal{O}(h^2)$.

   (b) Substituting the relation $y_j = [Q(h\lambda)]y_{j-1}$ and the test equation $y' = \lambda y = f(t, y)$ in (1) gives the following:

$$[Q(h\lambda)]^2 y_{n-1} = Q(h\lambda)y_{n-1} + h\left(\frac{3}{2}\lambda Q(h\lambda)y_{n-1} - \frac{1}{2}\lambda y_{n-1}\right).$$

   This can be rewritten as

$$[Q(h\lambda)]^2 - Q(h\lambda)\left(1 + \frac{3}{2}h\lambda\right) + \frac{1}{2}h\lambda = 0.$$

We therefore have that

$$Q_1(h\lambda) = \frac{1}{2}\left(1 + \frac{3}{2}h\lambda + \sqrt{\left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1}\right) \tag{2}$$

$$Q_2(h\lambda) = \frac{1}{2}\left(1 + \frac{3}{2}h\lambda - \sqrt{\left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1}\right) \tag{3}$$

(c) Since the discriminant of $\left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1$ is negative the value of $\left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1$ is always positive, so both $Q_1(h\lambda)$ and $Q_2(h\lambda)$ are real numbers. This implies that we must have $-1 \leq Q_2(h\lambda) < Q_1(h\lambda) \leq 1$. Note that $Q_1(h\lambda) \leq 1$ is satisfied for all h because it simplifies to

$$\frac{1}{2}\left(1 + \frac{3}{2}h\lambda + \sqrt{\left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1}\right) \leq 1$$

$$\sqrt{\left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1} \leq 2 - 1 - \frac{3}{2}h\lambda$$

$$\sqrt{\left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1} \leq 1 - \frac{3}{2}h\lambda$$

Squaring both sides gives

$$\left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1 \leq 1 - 3h\lambda + \left(\frac{3}{2}h\lambda\right)^2,$$

which implies

$$0 \leq -4h\lambda$$

which is always true for negative real values of λ.

For $-1 \leq Q_2(h\lambda)$, we can write this as

$$\begin{aligned} -2 &\leq 1 + \frac{3}{2}h\lambda - \sqrt{\left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1} \\ (3 + \frac{3}{2}h\lambda)^2 &\geq \left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1 \\ \left(\frac{3}{2}h\lambda\right)^2 + 9h\lambda + 9 &\geq \left(\frac{3}{2}h\lambda\right)^2 + h\lambda + 1 \end{aligned} \tag{4}$$

which simplifies to

$$h \leq -\frac{1}{\lambda}.$$

Consider the system

$$\mathbf{y}' = \begin{bmatrix} -4 & 1 \\ 1 & -4 \end{bmatrix} \mathbf{y} + \begin{bmatrix} 0 \\ -\cos(t) \end{bmatrix}. \tag{5}$$

(d) The eigenvalues of the matrix in (5) are given by

$$det(A - \lambda I) = (-4-\lambda)^2 - 1 = \lambda^2 + 8\lambda + 15 = 0.$$

This gives the values $\lambda_1 = -3$ and $\lambda_2 = -5$. Therefore, in order to apply our method to the system (5), we have the stability criteria

$$h \leq \frac{1}{3} \text{ and } h \leq \frac{1}{5}.$$

Since the strongest condition should be satisfied the method is stable for

$$h \leq \frac{1}{5}.$$

(e) Method (1) converges as long as $h < \frac{1}{\max_\lambda |\lambda|}$ because a stable and consistent scheme converges (Lax equivalence theorem).

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

$$p_1(x) = \frac{x - x_1}{x_0 - x_1} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1). \tag{6}$$

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

$$p_2(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} f(x_0) + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} f(x_1) + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} f(x_2). \tag{7}$$

This is also evident from application of the given formula.

(c) To this extent, 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

$$p_1(0.5) = 0.5 + \frac{1}{2} \cdot 2 = \frac{3}{2}, \tag{8}$$

and for quadratic interpolation, one obtains

$$p_2(0.5) = \frac{(0.5-1)(0.5-2)}{1 \cdot (-2)} \cdot 1 + \frac{(0.5-0)(0.5-2)}{1 \cdot (-1)} \cdot 2 + \frac{(0.5-0)(0.5-1)}{2 \cdot 1} \cdot 4 = \frac{11}{8} = 1.375. \tag{9}$$

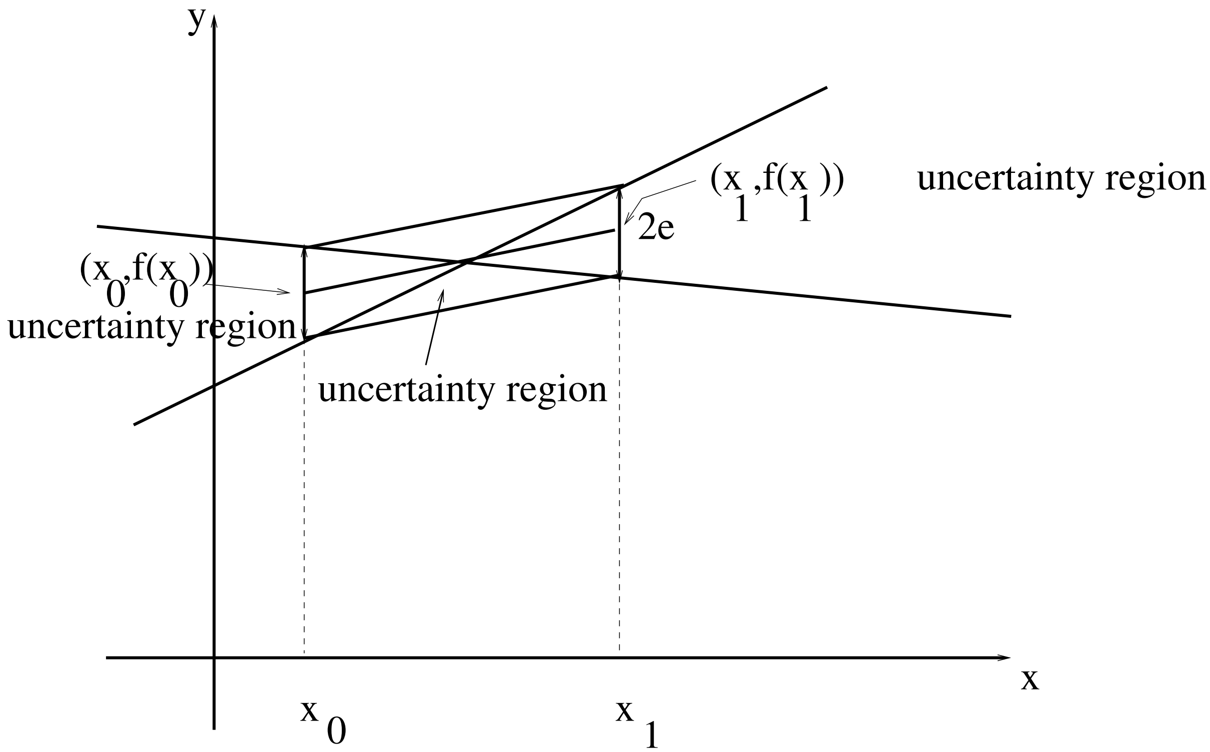

Figure 1: The measured values and the error using linear interpolation.

(d) Consider Figure 1. For interpolation, the error is bounded and for extrapolation, the error may become arbitrarily large as we move more and more outside the interval of the measured values. Of course, also a more algebraic motivation is allowed. We note that this effect may become worse if a higher order interpolatory formula is used.

(e) i We integrate $f(x)$, in which we approximate $f(x)$ by $p_1(x)$, then it follows:

$$\int_{x_0}^{x_1} f(x)dx \approx \int_{x_0}^{x_1} p_1(x)dx = \int_{x_0}^{x_1} \left\{ f(x_0)\frac{x-x_1}{x_0-x_1} + f(x_1)\frac{x-x_0}{x_1-x_0} \right\} dx =$$

$$= \left[\frac{1}{2}\frac{(x-x_0)^2}{x_1-x_0}f(x_1)\right]_{x_0}^{x_1} + \left[\frac{1}{2}\frac{(x-x_1)^2}{x_0-x_1}f(x_0)\right]_{x_0}^{x_1} = \frac{1}{2}(x_1-x_0)(f(x_0)+f(x_1)). \tag{10}$$

This is the Trapezoidal Rule.

ii The magnitude of the error of the numerical integration over interval $[x_0, x_1]$

is given by

$$
|\int_{x_0}^{x_1} f(x)dx - \int_{x_0}^{x_1} p_1(x)dx| = |\int_{x_0}^{x_1} \left(f(x) - p_1(x)\right) dx| =
$$

$$
|\int_{x_0}^{x_1} \frac{1}{2}(x-x_0)(x-x_1)f''(\chi(x))dx| \leq \frac{1}{2} \max_{x\in[x_0,x_1]} |f''(x)| \int_{x_0}^{x_1} (x-x_0)(x-x_1)dx =
$$

$$
\frac{1}{12}(x_1-x_0)^3 \max_{x\in[x_0,x_1]} |f''(x)|. \tag{11}
$$

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Thursday April 19 2012, 18:30-21:30

1.

a The local truncation error is defined as

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

where z_{n+1} is given by

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

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

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

This is substituted into equation (2) to obtain

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

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

$$y_{n+1} = y(t_n + h) = y_n + hy'(t_n) + \frac{h^2}{2}y''(t_n) + O(h^3). \tag{5}$$

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

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

Hence

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

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

b The test equation is given by

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

Application of the predictor step to the test equation gives

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

The corrector step yields

$$w_{n+1} = w_n + h\left(a_1\lambda w_n + a_2\lambda(1 + h\lambda)w_n\right) = (1 + (a_1 + a_2)h\lambda + a_2h^2\lambda^2)w_n. \tag{12}$$

Hence the amplification factor is given by

$$Q(h\lambda) = 1 + (a_1 + a_2)h\lambda + a_2h^2\lambda^2. \tag{13}$$

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

$$-1 \leq Q(h\lambda) \leq 1, \tag{14}$$

from the previous assignment, we have

$$-1 \leq 1 + (a_1 + a_2)h\lambda + a_2(h\lambda)^2 \leq 1 \Leftrightarrow -2 \leq (a_1 + a_2)h\lambda + a_2(h\lambda)^2 \leq 0. \tag{15}$$

First, we consider the left inequality:

$$a_2(h\lambda)^2 + (a_1 + a_2)h\lambda + 2 \geq 0 \tag{16}$$

For $h\lambda = 0$, 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 (15) is always satisfied. Next we consider the right hand inequality of relation (15)

$$a_2(h\lambda)^2 + (a_1 + a_2)h\lambda \leq 0. \tag{17}$$

This relation is rearranged into

$$a_2(h\lambda)^2 \leq -(a_1 + a_2)h\lambda, \tag{18}$$

hence

$$a_2|h\lambda|^2 \leq (a_1 + a_2)|h\lambda| \Leftrightarrow |h\lambda| \leq \frac{a_1 + a_2}{a_2}, \qquad a_2 \neq 0. \tag{19}$$

This results into the following condition for stability

$$h \leq \frac{a_1 + a_2}{a_2|\lambda|}, \qquad a_2 \neq 0. \tag{20}$$

d The Jacobian, J, is given by

$$J = \begin{pmatrix} \dfrac{\partial f_1}{\partial y_1} & \dfrac{\partial f_1}{\partial y_2} \\ \\ \dfrac{\partial f_2}{\partial y_1} & \dfrac{\partial f_2}{\partial y_2} \end{pmatrix}. \tag{21}$$

Since $f_1(y_1, y_2) = -y_1 y_2$ and $f_2(y_1, y_2) = y_1 y_2 - y_2$, we obtain

$$J = \begin{pmatrix} -y_2 & -y_1 \\ y_2 & y_1 - 1 \end{pmatrix}. \tag{22}$$

Substitution of the initial values $y_1(0) = 1$ and $y_2(0) = 2$, gives

$$J = \begin{pmatrix} -2 & -1 \\ 2 & 0 \end{pmatrix}. \tag{23}$$

e The eigenvalues of the Jacobian at $y_1(0) = y_2(0) = 1$ are given by $\lambda_{1,2} = 1 \pm i$. For our case, we have

$$Q(h\lambda) = -1 + h\lambda + 1/2(h\lambda)^2. \tag{24}$$

Since our eigenvalues are not real valued, it is required for stability that

$$|Q(h\lambda)| \leq 1. \tag{25}$$

Since the eigenvalues are complex conjugates, we can proceed with one of the eigenvalues, say $\lambda = -1 + i$ with $\lambda^2 = -2i$ to obtain

$$Q(h\lambda) \quad = 1 + h(-1 + i) + 1/2h^2(-2i) \tag{26}$$

Substitution of $h = 1$ shows that $Q(h\lambda) = 0$. This implies that $|Q(h\lambda)| = 0 \leq 1$ so the method is stable.

2. a Given $v(x) = x(2 - x)$, then $v''(x) = -2$, and hence $-v'' + v = 2 + x(2 - x)$ follows by simple addition. Further, $v(0) = 0$ and $v'(x) = 2 - 2x$ and hence $v'(1) = 0$. Hence the differential equation, as well as the boundary conditions are satisfied.

   b Let $v_j = v(x_j)$, and let $x_n = 1$, hence $h = 1/n$, then

$$v_{j-1} = v(x_j - h) = v_j - hv'(x_j) + h^2/2v''(x_j) - h^3/3!v'''(x_j) + h^4/4!v''''(x_j) + O(h^5);$$

$$v_{j+1} = v(x_j + h) = v_j + hv'(x_j) + h^2/2v''(x_j) + h^3/3!v'''(x_j) + h^4/4!v''''(x_j) + O(h^5). \tag{27}$$

From the above expression, it can be seen that

$$v''(x_j) = \frac{v_{j-1} - 2v_j + v_{j+1}}{h^2} + \frac{h^2}{12}v''''(x_j) + O(h^3), \tag{28}$$

and hence the error is $O(h^2)$. This gives the following discretization

$$\frac{-w_{j-1} + 2w_j - w_{j+1}}{h^2} + w_j = 2 + x_j(2 - x_j), \qquad \text{for } j = 1 \dots n, \tag{29}$$

where $x_j = jh$ and $w_j \approx v_j$ as the numerical (finite difference) solution under neglecting the error. Further, we use a virtual gridnode near $x = 1$, $x_{n+1} = 1+h$, with

$$0 = v'(1) = \frac{v_{n+1} - v_{n-1}}{2h} - \frac{h^2}{3}v'''(1) + O(h^3), \tag{30}$$

hence the error is $O(h^2)$. Neglecting the error, and substitution into the discretization equation $j = n$, gives

$$\frac{-2w_{n-1} + 2w_n}{h^2} + w_n = 3. \tag{31}$$

Division by 2 to make the discretization symmetric, gives

$$\frac{-w_{n-1} + w_n}{h^2} + \frac{1}{2}w_n = \frac{3}{2}. \tag{32}$$

The boundary condition at $x = 0$, gives

$$\frac{2w_1 - w_2}{h^2} + w_1 = 2 + h(2 - h).. \tag{33}$$

c For $j = 1$, we get, using $h = 1/3$,

$$18w_1 - 9w_2 + w_1 = 2 + 1/3 * 5/3 = 23/9. \tag{34}$$

For $j = 2$, we obtain

$$-9w_1 + 18w_2 - 9w_3 + w_2 = 26/9. \tag{35}$$

For $j = 3 = n$, we use $w_4 = w_2$, which gives

$$-9w_2 + 9w_3 + 1/2w_3 = 3/2. \tag{36}$$

Hence, the system of equations is

$$\begin{cases} 19w_1 - 9w_2 = 23/9, \\ \\ -9w_1 + 19w_2 - 9w_3 = 26/9, \\ \\ -9w_2 + 19/2w_3 = 3/2. \end{cases} \tag{37}$$

d The exact solution is given by $v(x) = x(2-x)$, and hence all derivatives of order three and larger are zero. Further, the error is determined by the derivatives of third order and larger. This implies that the error is zero.

e To this extent, we consider the determination of the zeros of the following system of equations

$$\begin{cases} F_1(v_1, v_2) = 18v_1 - 9v_2 + v_1^2 - \frac{20}{9}, \\ \\ F_2(v_1, v_2) = -9v_1 + 18v_2 + v_2^2 - \frac{20}{9}. \end{cases}$$

We consider (v_1^k, v_2^k) as the kth estimate of the successive approximations. Linearization around the estimate (v_1^k, v_2^k) gives the following Newton method:

$$\frac{\partial(F_1, F_2)}{\partial(v_1, v_2)}(v_1^k, v_2^k) \begin{pmatrix} v_1^{k+1} - v_1^k \\ \\ v_2^{k+1} - v_2^k \end{pmatrix} = -\underline{F}(v_1^k, v_2^k), \tag{38}$$

where $\underline{F}(v_1, v_2) = [F_1(v_1, v_2) \; F_2(v_1, v_2)]^T$, and

$$\frac{\partial(F_1, F_2)}{\partial(v_1, v_2)}(v_1^k, v_2^k) = \begin{pmatrix} \frac{\partial F_1}{\partial v_1}(v_1^k, v_2^k) & \frac{\partial F_1}{\partial v_2}(v_1^k, v_2^k) \\ \\ \frac{\partial F_2}{\partial v_1}(v_1^k, v_2^k) & \frac{\partial F_2}{\partial v_2}(v_1^k, v_2^k) \end{pmatrix} = \begin{pmatrix} 18 + 2v_1^k & -9 \\ \\ -9 & 18 + 2v_2^k \end{pmatrix}, \tag{39}$$

is the Jacobian matrix. Using $v_1^0 = v_2^0 = 0$, we get

$$\begin{pmatrix} 18 & -9 \\ \\ -9 & 18 \end{pmatrix} \begin{pmatrix} v_1^1 - v_1^0 \\ \\ v_2^1 - v_2^0 \end{pmatrix} = \begin{pmatrix} 20/9 \\ \\ 20/9 \end{pmatrix}. \tag{40}$$

The solution is given by $v_1^1 - v_1^0 = 20/81 = v_2^1 - v_2^0$, and hence $v_1^1 = v_2^1 = 20/81$.

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Thursday August 25 2011, 18:30-21:30

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

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

Here we obtain y_{n+1} by a Taylor expansion around t_n:

$$y_{n+1} = y_n + hy'(t_n) + \frac{h^2}{2}y''(t_n) + O(h^3). \tag{2}$$

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

$$z_{n+1} = y_n + h\left((1-\theta)f(t_n, y_n) + \theta f(t_n + h, y_n + hf(t_n, y_n))\right) \tag{3}$$

After a Taylor expansion of $f(t_n+h, y_n+hf(t_n, y_n))$ around (t_n, y_n) one obtains:

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

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 $z_{n+1} = y_n + hy'(t_n) + \theta h^2 y''(t_n)$. Subsequently, it follows that

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

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

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

$$\begin{aligned} w_{n+1}^* &= w_n + h\lambda w_n = (1+h\lambda)w_n, \\ w_{n+1} &= w_n + h((1-\theta)\lambda w_n + \theta\lambda w_{n+1}^*) = \\ &= w_n + h((1-\theta)\lambda w_n + \theta\lambda(w_n + h\lambda w_n)) = (1 + h\lambda + \theta(h\lambda)^2)w_n. \end{aligned} \tag{10}$$

Hence the amplification factor is given by

$$Q(h\lambda) = 1 + h\lambda + \theta(h\lambda)^2. \tag{11}$$

(c) We start this exercise by using the following vector:

$$x_1 = y$$

$$x_2 = y'$$

From this it follows that

$$x_1' = y' = x_2$$
$$x_2' = y'' = -4y' - 8y + t^2 - 1 = -4x_2 - 8x_1 + t^2 - 1$$

where we have used the second order differential equation. We can write this as follows in matrix-vector notation:

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -8 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ t^2 - 1 \end{pmatrix}$$

So it follows that $A = \begin{pmatrix} 0 & 1 \\ -8 & -4 \end{pmatrix}$ and $f(t) = 0$ and $g(t) = t^2 - 1$.

(d) In order to do one step we first note that

$$\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} y(0) \\ y'(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The predictor step with $h = 1$ now gives:

$$w_1^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 1\left(\begin{pmatrix} 0 & 1 \\ -8 & -4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -5 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$$

Finally the correction step with $\theta = \frac{1}{2}$ leads to

$$w_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ -5 \end{pmatrix} + \frac{1}{2}\left(\begin{pmatrix} 0 & 1 \\ -8 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ -4 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} -1.5 \\ 2.5 \end{pmatrix}$$

(e) Compute the eigenvalues of matrix $\begin{pmatrix} 0 & 1 \\ -8 & -4 \end{pmatrix}$. To do this we compute the determinant of $\begin{pmatrix} -\lambda & 1 \\ -8 & -4-\lambda \end{pmatrix}$, which is equal to $\lambda^2+4\lambda+8$. The roots of this polynomial are equal to $\lambda_1 = -2+2i$ and $\lambda_2 = -2-2i$. Since $\lambda_2 = \bar{\lambda_1}$ it is sufficient to consider λ_1 only. For $h = 1$ we obtain $h\lambda_1 = -2+2i$. This implies that

$$Q(h\lambda_1) = 1 + h\lambda_1 + \theta(h\lambda_1)^2$$

$$Q(h\lambda_1) = 1 + (-2+2i) + \theta(-2+2i)^2$$

$$Q(h\lambda_1) = 1 - 2 + 2i + \theta(4 - 8i - 4) = -1 + i(2-8\theta)$$

In order to check that $|Q(h\lambda_1)| \leq 1$, we compute the modulus of $Q(h\lambda_1)$, which is equal to

$$\sqrt{1^2 + (2-8\theta)^2}$$

It is easy to see that this is only less than or equal to 1 if $\theta = \frac{1}{4}$.

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

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

Here $\xi_1 \in (0,h)$, $\xi_2 \in (0,2h)$. 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 $f''(0) + O(h)$. This leads to the following conditions:

$$\begin{array}{lccccccc} f(0): & \frac{\alpha_0}{h^2} & + & \frac{\alpha_1}{h^2} & + & \frac{\alpha_2}{h^2} & = & 0\,, \\ f'(0): & & & \frac{h\alpha_1}{h^2} & + & \frac{2h\alpha_2}{h^2} & = & 0\,, \\ f''(0): & & & \frac{h^2}{2h^2}\alpha_1 & + & \frac{2h^2\alpha_2}{h^2} & = & 1\,. \end{array}$$

This can also be written as

$$\begin{array}{lccccccc} f(0): & \alpha_0 & + & \alpha_1 & + & \alpha_2 & = & 0\,, \\ f'(0): & & & \alpha_1 & + & 2\alpha_2 & = & 0\,, \\ f''(0): & & & \frac{\alpha_1}{2} & + & 2\alpha_2 & = & 1\,. \end{array}$$

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

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

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

(c) Note that

$$f''(0) - Q(h) = Kh \tag{12}$$

$$f''(0) - Q(\frac{h}{2}) = K(\frac{h}{2}) \tag{13}$$

Subtraction gives:

$$Q(\frac{h}{2}) - Q(h) = Kh - K\frac{h}{2} = K(\frac{h}{2}). \tag{14}$$

We choose $h = \frac{1}{2}$. Then $Q(h) = Q(\frac{1}{2}) = \frac{0-2\times 0.1250+1}{0.25} = 3$ and $Q(\frac{h}{2}) = Q(\frac{1}{4}) = \frac{0-2\times 0.0156+0.1250}{(\frac{1}{4})^2} = 1.5008$. Combining (13) and (14) shows that

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

(d) To estimate the rounding error we note that

$$|Q(h) - \hat{Q}(h)| = |\frac{(f(0) - \hat{f}(0)) - 2(f(h) - \hat{f}(h)) + (f(2h) - \hat{f}(2h))}{h^2}|$$

$$\leq \frac{|f(0) - \hat{f}(0)| + 2|f(h) - \hat{f}(h)| + |f(2h) - \hat{f}(2h)|}{h^2} \leq \frac{4\epsilon}{h^2},$$

so $C_1 = 4$. Since only 4 digits are given the rounding error is: $\epsilon = 0.00005$.

(e) The total error is bounded by

$$|f''(0) - \hat{Q}(h)| = |f''(0) - Q(h) + Q(h) - \hat{Q}(h)|$$

$$\leq |f''(0) - Q(h)| + |Q(h) - \hat{Q}(h)|$$

$$\leq 6h + \frac{4\epsilon}{h^2} = g(h)$$

This is minimal for h_{opt}, for which $g'(h_{opt}) = 0$. Note that $g'(h) = 6 - \frac{8\epsilon}{h^3}$. This implies that $h_{opt}^3 = \frac{4\epsilon}{3}$, so $h_{opt} = (\frac{4\epsilon}{3})^{\frac{1}{3}} \approx 0.0405$.

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Thursday August 30 2012, 18:30-21:30

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

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

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:

$$y_{n+1} = y_n + hy'(t_n) + \frac{h^2}{2}y''(t_n) + O(h^3). \tag{2}$$

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

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

After a Taylor expansion of $f(t_n+h, y_n+hf(t_n, y_n))$ around (t_n, y_n) one obtains:

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

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 $z_{n+1} = y_n + hy'(t_n) + \mu h^2 y''(t_n)$. Subsequently, it follows that

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

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

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

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

Hence the amplification factor is given by

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

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

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} -\sin(0) + 2 + 0 \\ 0 - 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}$$

Corrector:

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

which can be written as:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 + \frac{1}{2} - \frac{1}{4}\sin(1) + \frac{3}{8} \\ 1 - \frac{1}{4} + \frac{1}{4} - \frac{1}{16} \end{pmatrix} = \begin{pmatrix} \frac{7}{8} - \frac{1}{4}\sin(1) \\ \frac{15}{16} \end{pmatrix} = \begin{pmatrix} 0.6646 \\ 0.9375 \end{pmatrix}$$

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

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

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

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 $|Q(h\lambda_i)| \leq 1$ 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 $h \leq \frac{-2}{\lambda}$. Since $\lambda_1 = -3$ and $\lambda_2 = 0$ we know that the method is stable if $h \leq \frac{-2}{-3} = \frac{2}{3}$ (another option is to derive the values of h such that $|Q(h\lambda_i)| \leq 1$ by using the description of $Q(h\lambda)$)

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

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

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

$$-1 \leq 1 - 3h + \frac{9}{2}h^2 \leq 1$$

For the left-hand inequality we arrive at

$$0 \leq \frac{9}{2}h^2 - 3h + 2$$

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 h.

For the right-hand inequality we get

$$-3h + \frac{9}{2}h^2 \leq 0$$

$$\frac{9}{2}h^2 \leq 3h$$

so

$$h \leq \frac{2}{3}$$

(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 $h \leq \frac{-2}{\lambda}$)

2. (a) Taylor's Theorem (or here the Mean Value Theorem) gives for a zeroth order approximation around x_j:

$$f(x) = f(x_j) + (x - x_j)f'(\xi(x)), \tag{12}$$

for a $\xi(x) \in (x_j, x)$ if $x > x_j$. Then we consider the interval $[x_j, x_{j+1})$ and use Taylor's Theorem around x_j in the integration to get

$$\int_{x_j}^{x_{j+1}} f(x)dx = \int_{x_j}^{x_{j+1}} f(x_j) + (x - x_j)f'(\xi(x))dx = hf(x_j) + \int_{x_j}^{x_{j+1}} (x - x_j)f'(\xi(x))dx. \tag{13}$$

Hence we get

$$|\int_{x_j}^{x_{j+1}} f(x)dx - \int_{x_j}^{x_{j+1}} f(x_j)| = |\int_{x_j}^{x_{j+1}} (x - x_j)f'(\xi(x))dx|. \tag{14}$$

Taking the maximum value of f' over the interval $[x_j, x_{j+1}]$, yields

$$|\int_{x_j}^{x_{j+1}} (x-x_j)f'(\xi(x))dx| \leq \max_{x\in[x_j,x_{j+1}]} |f'(x)| \int_{x_j}^{x_{j+1}} (x-x_j)dx = \frac{h^2}{2} \max_{x\in[x_j,x_{j+1}]} |f'(x)|. \tag{15}$$

By combining relations (14) and (15), we proved that

$$|\int_{x_j}^{x_{j+1}} f(x)dx - \int_{x_j}^{x_{j+1}} f(x_j)| \leq \frac{h^2}{2} \max_{x\in[x_j,x_{j+1}]} |f'(x)|. \tag{16}$$

Next, we deal with the entire interval $[a, b]$, then

$$|\int_a^b f(x)dx - h\sum_{j=1}^{n} f(x_j)| = |\sum_{j=1}^{n} \left(\int_{x_j}^{x_{j+1}} f(x)dx - hf(x_j)\right)|. \tag{17}$$

We use the Triangle Inequality to get

$$|\sum_{j=1}^{n} \left(\int_{x_j}^{x_{j+1}} f(x)dx - hf(x_j)\right)| \leq \sum_{j=1}^{n} |\int_{x_j}^{x_{j+1}} f(x)dx - hf(x_j)|. \tag{18}$$

From relation (16), it follows that

$$\sum_{j=1}^{n} |\int_{x_j}^{x_{j+1}} f(x)dx - hf(x_j)|. \leq \frac{h^2}{2} \sum_{j=1}^{n} \max_{x\in[x_j,x_{j+1}]} |f'(x)|. \tag{19}$$

Since $\max_{x\in([a,b]} |f'(x)| \geq \max_{x\in([x_j,x_{j+1}]} |f'(x)|,\ \forall j \in \{1, \ldots, n\}$, we get

$$\frac{h^2}{2} \sum_{j=1}^{n} \max_{x\in[x_j,x_{j+1}]} |f'(x)| \leq \frac{h^2}{2} \cdot n \cdot \max_{x\in[a,b]} |f'(x)|. \tag{20}$$

Since $x_{n+1} = a + nh = b$, we have $nh = b - a$ and hence the above inequality gives

$$\frac{h^2}{2} \sum_{j=1}^{n} \max_{x\in[x_j,x_{j+1}]} |f'(x)| \leq \frac{h^2}{2} \cdot n \cdot \max_{x\in[a,b]} |f'(x)| = \frac{h}{2}(b-a) \max_{x\in[a,b]} |f'(x)|. \tag{21}$$

Hence the global error can be estimated from above by

$$|\int_a^b f(x)dx - h\sum_{j=1}^{n} f(x_j)| \leq \frac{h}{2}(b-a) \max_{x\in[a,b]} |f'(x)|. \tag{22}$$

(b) Incorporating the first–order derivative in Taylor's Theorem (linearization) gives

$$f(x) = f(x_j) + (x - x_j)f'(x_j) + \frac{(x - x_j)^2}{2} f''(\xi(x)), \tag{23}$$

for a $\xi(x) \in (x_j, x)$ if $x > x_j$. We start integrating over the interval $[x_j, x_{j+1}]$ to get

$$\int_{x_j}^{x_{j+1}} f(x)dx = \int_{x_j}^{x_{j+1}} f(x_j) + (x - x_j)f'(x_j) + \frac{(x - x_j)^2}{2} f''(\xi(x))dx = \\ hf(x_j) + \frac{h^2}{2} f'(x_j) + \int_{x_j}^{x_{j+1}} \frac{(x - x_j)^2}{2} f''(\xi(x))dx. \tag{24}$$

Hence, we obtain

$$|\int_{x_j}^{x_{j+1}} f(x)dx - (hf(x_j) + \frac{h^2}{2} f'(x_j))| = |\int_{x_j}^{x_{j+1}} \frac{(x - x_j)^2}{2} f''(\xi(x))dx| \leq \\ \max_{x \in [x_j, x_{j+1}]} |f''(x)| \int_{x_j}^{x_{j+1}} \frac{(x - x_j)^2}{2} = \frac{h^3}{6} \max_{x \in [x_j, x_{j+1}]} |f''(x)|. \tag{25}$$

Analogously to the previous assignment, we get

$$|E_1| = |\int_a^b f(x)dx - \sum_{j=1}^{n} \left(hf(x_j) + \frac{h^2}{2} f'(x_j) \right)| = |\sum_{j=1}^{n} \left(\int_{x_j}^{x_{j+1}} f(x)dx - \left(hf(x_j) + \frac{h^2}{2} f'(x_j) \right) \right. \\ \sum_{j=1}^{n} |\left(\int_{x_j}^{x_{j+1}} f(x)dx - \left(hf(x_j) + \frac{h^2}{2} f'(x_j) \right) \right)| \leq \frac{h^3}{6} \sum_{j=1}^{n} \max_{x \in [x_j, x_{j+1}]} |f''(x)| \leq \\ \frac{h^3}{6} \cdot n \cdot \max_{x \in [a,b]} |f''(x)| = \frac{h^2}{6}(b - a) \max_{x \in [a,b]} |f''(x)|. \tag{26}$$

Hence $\int_a^b f(x)dx \approx \sum_{j=1}^{n} h(f(x_j) + \frac{h}{2} f'(x_j)) = T_1$ where the global error is estimated from above by the above expression.

(c) Upon considering the interval $(0, 1)$ with $h = \frac{1}{2}$, we use $x_1 = 0$ and $x_2 = \frac{1}{2}$ $(n = 2)$. Then, we get

$$\int_0^1 x^2 dx \approx h(f(x_1)+f(x_2)+\frac{h}{2}(f'(x_1)+f'(x_2)) = \frac{1}{2}(0+(\frac{1}{2})^2+\frac{1}{4}(0+2\cdot\frac{1}{2})) = \frac{1}{4}. \tag{27}$$

The exact answer is given by $\frac{1}{3}$, hence the error is $\frac{1}{12}$. To check our result, we use the upper bound of the error given in relation (26):

$$\frac{h^2}{6}(b - a) \max_{x \in [a,b]} |f''(x)| = \frac{1}{6} \cdot (\frac{1}{2})^2 \cdot 1 \cdot 2 = \frac{1}{12}. \tag{28}$$

Note that here it was used that the second–order derivative of x^2 is given by 2. Hence our the error that we found using the exact solution does not exceed the upper bound from relation (26), and hence our result makes sense.

(d) T_1 is the approximation of the integral obtained by the use the first order derivatives, hence T_2 is the analogon with the first and second order derivatives, hence

$$
\begin{aligned}
&T_2 = \\
&\sum_{j=1}^{n}\left(\int_{x_j}^{x_{j+1}} f(x_j) + (x-x_j)f'(x_j) + \frac{(x-x_j)^2}{2}f''(x_j)dx\right) = \\
&\sum_{j=1}^{n}\left(\int_{x_j}^{x_{j+1}} f(x_j) + (x-x_j)f'(x_j)dx\right) + \sum_{j=1}^{n}\int_{x_j}^{x_{j+1}} \frac{(x-x_j)^2}{2}f''(x_j)dx \\
&= T_1 + \sum_{j=1}^{n}\int_{x_j}^{x_{j+1}} \frac{(x-x_j)^2}{2}f''(x_j)dx = T_1 + \frac{h^3}{3!}\sum_{j=1}^{n} f''(x_j).
\end{aligned}
\tag{29}
$$

The last step follows from evaluation of the integral. Hence we demonstrated that

$$T_2 = T_1 + \frac{h^3}{3!}\sum_{j=1}^{n} f''(x_j). \tag{30}$$

Further, the local error is found by using Taylor's Theorem over the interval $[x_j, x_{j+1}]$ to get

$$
\begin{aligned}
&|\int_{x_j}^{x_{j+1}} f(x)dx - \left(\int_{x_j}^{x_{j+1}} f(x_j) + \ldots + \frac{(x-x_j)^2}{2!}f''(x_j)dx\right)| = \\
&|\int_{x_j}^{x_{j+1}} \frac{(x-x_j)^3}{3!}f'''(\xi(x))dx| \le \max_{x\in[x_j,x_{j+1}]}|f'''(x)|\int_{x_j}^{x_{j+1}} \frac{(x-x_j)^3}{3!}dx = \\
&\frac{h^4}{4!}\max_{x\in[x_j,x_{j+1}]}|f'''(x)|.
\end{aligned}
\tag{31}
$$

Here, the last step follows from evaluation of the integral. A summation procedure over all intervals, similar to assignment 2.a., gives the global error bound:

$$
\begin{aligned}
&|E_2| = |\int_a^b f(x)dx - T_2| \le \frac{h^4}{4!}\sum_{j=1}^{n}\max_{x\in[x_j,x_{j+1}]}|f'''(x)| \le \\
&\frac{h^4}{4!}\cdot n \cdot \max_{x\in[a,b]}|f'''(x)| = \frac{h^3(b-a)}{4!}\max_{x\in[a,b]}|f'''(x)|.
\end{aligned}
\tag{32}
$$

(e) Let T_2 and $\tilde{T}_2$, respectively, be the approximation of $\int_a^b f(x)dx$ using the exact and available values of f and its derivatives. Then, we have

$$T_2 = \sum_{j=1}^{n} \int_{x_j}^{x_{j+1}} f(x_j) + \ldots + \frac{(x-x_j)^2}{2!} f''(x_j) dx =$$

$$\sum_{j=1}^{n} \left(hf(x_j) + \frac{h^2}{2} f'(x_j) + \frac{h^3}{3!} f''(x_j) \right) = \tag{33}$$

$$h \sum_{j=1}^{n} f(x_j) + \frac{h^2}{2} \sum_{j=1}^{n} f'(x_j) + \frac{h^3}{3!} \sum_{j=1}^{n} f''(x_j).$$

For $\tilde{T}_2$, we similarly have

$$\tilde{T}_2 = h \sum_{j=1}^{n} \tilde{f}(x_j) + \frac{h^2}{2} \sum_{j=1}^{n} \tilde{f}'(x_j) + \frac{h^3}{3!} \sum_{j=1}^{n} \tilde{f}''(x_j). \tag{34}$$

Subtraction of the above two equations, taking the absolute value, and using the Triangle Inequality, gives

$$|T_2 - \tilde{T}_2| \leq$$

$$h \sum_{j=1}^{n} |f(x_j) - \tilde{f}(x_j)| + \frac{h^2}{2} \sum_{j=1}^{n} |f'(x_j) - \tilde{f}'(x_j)| + \frac{h^3}{3!} \sum_{j=1}^{n} |f''(x_j) - \tilde{f}''(x_j)|, \tag{35}$$

Using $|f^{(k)}(x_j) - \tilde{f}^{(k)}(x_j)| \leq \varepsilon$ for all k and j, and $nh = b - a$, gives

$$|T_2 - \tilde{T}_2| \leq h \cdot n \cdot \varepsilon + \frac{h^2}{2} \cdot n \cdot \varepsilon + \frac{h^3}{3!} \cdot n \cdot \varepsilon =$$

$$(b-a)\varepsilon \left(1 + \frac{h}{2} + \frac{h^2}{3!} \right) = (b-a)\varepsilon \sum_{k=1}^{3} \frac{h^{k-1}}{k!}. \tag{36}$$

.

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Thursday January 20 2011, 18:30-21:30

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

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

in which we determine y_{n+1} by the use of Taylor expansions around t_n:

$$y_{n+1} = y_n + hy'(t_n) + \frac{h^2}{2}y''(t_n) + O(h^3). \tag{2}$$

We bear in mind that

$$\begin{aligned} y'(t_n) &= f(t_n, y_n) \\ y''(t_n) &= \frac{df(t_n, y_n)}{dt} = \frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} y'(t_n) = \\ &\frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} f(t_n, y_n). \end{aligned} \tag{3}$$

Hence

$$y_{n+1} = y_n + hy'(t_n) + \frac{h^2}{2}\left(\frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y} f(t_n, y_n)\right) + O(h^3). \tag{4}$$

After substitution of the predictor $z^*_{n+1} = y_n + hf(t_n, y_n)$ into the corrector, and after using a Taylor expansion around (t_n, y_n), we obtain for z_{n+1}

$$\begin{aligned} z_{n+1} &= y_n + \tfrac{h}{2}\left(f(t_n, y_n) + f(t_n + h, y_n + hf(t_n, y_n))\right) = \\ y_n &+ \frac{h}{2}\left(f(t_n, y_n) + f(t_n, y_n) + h(\frac{\partial f(t_n, y_n)}{\partial t} + f(t_n, y_n)\frac{\partial f(t_n, y_n)}{\partial y}) + O(h^2)\right). \end{aligned} \tag{5}$$

Herewith, one obtains

$$y_{n+1} - z_{n+1} = O(h^3), \text{ and hence } \tau_{n+1}(h) = \frac{O(h^3)}{h} = O(h^2). \tag{6}$$

(b) Let $x_1 = y$ and $x_2 = y'$, then $y'' = x_2'$, and hence

$$\begin{aligned} x_2' + 4x_2 + 3x_1 &= \cos(t), \\ x_2 &= x_1'. \end{aligned} \tag{7}$$

We write this as

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -3x_1 - 4x_2 + \cos(t). \end{aligned} \tag{8}$$

Finally, this is represented in the following matrix-vector form:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -3 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \cos(t) \end{pmatrix}. \tag{9}$$

In which, we have the following matrix $A = \begin{pmatrix} 0 & 1 \\ -3 & -4 \end{pmatrix}$ and $f = \begin{pmatrix} 0 \\ \cos(t) \end{pmatrix}$.

The initial conditions are defined by $\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

(c) Application of the Modified Euler method to the system $\underline{x}' = A\underline{x} + \underline{f}$, gives

$$\begin{aligned} \underline{w}_1^* &= \underline{w}_0 + h\left(A\underline{w}_0 + \underline{f}_0\right), \\ \underline{w}_1 &= \underline{w}_0 + \tfrac{h}{2}\left(A\underline{w}_0 + f_0 + A\underline{w}_1^* + \underline{f}_1\right). \end{aligned} \tag{10}$$

With the initial condition $\underline{w}_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $h = 0.1$, this gives the following result for the predictor

$$\underline{w}_1^* = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{1}{10}\left(\begin{pmatrix} 0 & 1 \\ -3 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 6/5 \\ 1 \end{pmatrix}. \tag{11}$$

The corrector is calculated as follows

$$\begin{aligned} \underline{w}_1 &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \tfrac{1}{20}\left(\begin{pmatrix} 0 & 1 \\ -3 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -3 & -4 \end{pmatrix}\begin{pmatrix} 6/5 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ \cos(\frac{1}{10}) \end{pmatrix}\right) = \\ &= \begin{pmatrix} 1.1500 \\ 1.1698 \end{pmatrix} \end{aligned} \tag{12}$$

(d) Consider the test equation $y' = \lambda y$, then one gets

$$\begin{aligned} w_{n+1}^* &= w_n + h\lambda w_n = (1 + h\lambda)w_n, \\ w_{n+1} &= w_n + \frac{h}{2}(\lambda w_n + \lambda w_{n+1}^*) = \\ &= w_n + \frac{h}{2}(\lambda w_n + \lambda(w_n + h\lambda w_n)) = (1 + h\lambda + \frac{(h\lambda)^2}{2})w_n. \end{aligned} \tag{13}$$

Hence the amplification factor is given by

$$Q(h\lambda) = 1 + h\lambda + \frac{(h\lambda)^2}{2}. \tag{14}$$

(e) 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 $\lambda_1 = -1$ and $\lambda_2 = -3$. We first check the amplification factor of $\lambda_1 = -1$:

$$-1 \leq 1 - h + \frac{1}{2}h^2 \leq 1 \tag{15}$$

The first inequality leads to

$$0 \leq 2 - h + \frac{1}{2}h^2$$

Since the discriminant of this equation is equal to $1-4*\frac{1}{2}*2 = -3$ the inequality always holds. The second inequality leads to

$$-h + \frac{1}{2}h^2 \leq 0$$

so

$$\frac{1}{2}h^2 \leq h$$

which implies

$$h \leq 2$$

Now we check the amplification factor of $\lambda_2 = -3$:

$$-1 \leq 1 - 3h + \frac{1}{2}9h^2 \leq 1 \tag{16}$$

The first inequality leads to

$$0 \leq 2 - 3h + \frac{1}{2}9h^2$$

Since the discriminant of this equation is equal to $9 - 4 * \frac{9}{2} * 2 = -27$ the inequality always holds. The second inequality leads to

$$-3h + \frac{9}{2}h^2 \leq 0$$

so

$$\frac{3}{2}h^2 \leq h$$

which implies

$$h \leq \frac{2}{3}$$

So the modified Euler method is stable if $h \leq \frac{2}{3}$.

2. (a) We have to check whether the requirements for the Convergence Theorem (see Theorem 4.3.2 in Vuik et al.) on convergence are satisfied. We have to remark that these requirements give a sufficient condition for convergence to the fixed point if we choose an initial value in a neighborhood around the fixed point p. The theorem is formulated as follows:

Theorem: *If there exists a $\delta > 0$ such that $g(x) \in C^1[p-\delta, p+\delta]$ (the first order derivative of $g(x)$ is continuous), then, the fixed point method converges for each initial value $p_0 \in [p-\delta, p+\delta]$ if the following hypotheses are satisfied:*

i. $g : [p-\delta, p+\delta] \longrightarrow [p-\delta, p+\delta]$;

ii. There exists a $r > 0$ such that

$$|g'(x)| \leq r < 1, \text{ for } x \in [p-\delta, p+\delta].$$

Firstly, the derivative of $g(x)$ is given by

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

Further, we have

$$g'(p) = 1 - \frac{f'(p)}{\alpha} > 1 - \frac{f'(p)}{f'(p)} = 0,$$

since $0 < f'(p) < \alpha$. From this, it also follows that

$$g'(p) = 1 - \frac{f'(p)}{\alpha} < 1,$$

since $f'(p) > 0$ and $\alpha > 0$. When we combine these bounds for the derivative of g' with continuity of $f'(x)$, and hence also with continuity of $g'(x)$ around p, it follows that there is a neighborhood around p for which we have $0 < g'(x) < 1$. In other words, mathematically speaking: There exists a $\delta > 0$ for which $0 < g'(x) < 1$ for each $x \in [p-\delta, p+\delta]$. Hence the first hypothesis of the convergence theorem is satisfied.

Further, via the Mean Value Theorem, we get

$$g(p+\delta) = g(p) + \delta g'(\xi_1), \text{ for a } \xi_1 \in (p-\delta, p+\delta) \text{ and,}$$

$$g(p-\delta) = g(p) - \delta g'(\xi_2), \text{ for a } \xi_2 \in (p-\delta, p+\delta).$$

This yields with $0 < g'(\xi) < 1$ and $g(p) = p$:

$$g(p-\delta) = g(p) - \delta g'(\xi_1) > p - \delta, \text{ and } g(p+\delta) = g(p) + \delta g'(\xi_2) < p + \delta.$$

Hence, we have $g(p \pm \delta) \in [p-\delta, p+\delta]$. Since $g(x)$ is monotonical on $[p-\delta, p+\delta]$, $g(x)$ cannot assume any values outside the range $[p-\delta, p+\delta]$ if $x \in [p-\delta, p+\delta]$. Hence, we have

$$g(x) \in [p-\delta, p+\delta], \text{ for } x \in [p-\delta, p+\delta],$$

which is equivalent to the second hypothesis. This all sustains convergence if the initial guess is chosen within a neighborhood around the fixed point p.

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

$$L(x) = f(p_n) + (x - p_n)f'(p_n). \tag{17}$$

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

$$f(p_n) + (p_{n+1} - p_n)f'(p_n) = 0 \Leftrightarrow p_{n+1} = p_n - \frac{f(p_n)}{f'(p_n)}, \qquad f'(p_n) \neq 0. \tag{18}$$

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

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

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

With the initial value $p_0 = 2$, this gives

$$p_1 = 2 - \frac{4 - 4 - 2}{4 - 2} = 3.$$

(d) We have $f'(x) = 2x - 2$ and hence $f'(1) = 0$ with starting value $p_0 = 1$. In the recursion, one divides by zero. Division by zero does not make any sense, so $p_0 = 1$ is not a suitable starting value. Geometrically, one may remark that the tangent is horizontal on $p_0 = 1$.

(e) We answer both questions sequentially:

- The linear interpolation polynomial with points x_0 en x_1 is given by:

$$P_1(x) = y(x_0)\frac{x - x_1}{x_0 - x_1} + y(x_1)\frac{x - x_0}{x_1 - x_0} = -(x-1) + 3(x-2/3) = 2x-1. \tag{19}$$

- We have $P_1(x) = 1/2 \Leftrightarrow 2x - 1 = 1/2$. Solution of this equation in x gives $x = \frac{3}{4}$.

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Thursday January 26 2012, 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:

$$w_{j+1} = w_j + \frac{h}{2}\left(\lambda w_j + \lambda w_{j+1}\right) \tag{1}$$

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

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

It now follows that

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

and thus

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

(b) The definition of the local truncation error is

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

The exact solution of the test equation is given by

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

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 $Q(h\lambda)$

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

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

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

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

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

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

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

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

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

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

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

(c) The trapezoidal rule is stable if

$$\frac{|1+\frac{h}{2}\lambda|}{|1-\frac{h}{2}\lambda|} \leq 1.$$

Using the complex valued $\lambda = \mu + i\nu$ it appears that the condition is equal to:

$$\frac{|1+\frac{h}{2}(\mu+i\nu)|}{|1-\frac{h}{2}(\mu+i\nu)|} \leq 1$$

This is equivalent with

$$\frac{\sqrt{(1+\frac{h}{2}\mu)^2 + (\frac{h}{2}\nu)^2}}{\sqrt{(1-\frac{h}{2}\mu)^2 + (\frac{h}{2}\nu)^2}} \leq 1$$

Since $\mu \leq 0$ it easily follows that

$$\sqrt{(1+\frac{h}{2}\mu)^2 + (\frac{h}{2}\nu)^2} \leq \sqrt{(1-\frac{h}{2}\mu)^2 + (\frac{h}{2}\nu)^2}$$

which implies that

$$\frac{|1+\frac{h}{2}\lambda|}{|1-\frac{h}{2}\lambda|} \leq 1.$$

and the method is stable.

(d) Application of the trapezoidal rule to

$$y' = -(1+t)y + t, \text{ with } y(0) = 1,$$

and step size $h = 1$ gives:

$$w_1 = w_0 + \frac{h}{2}[-w_0 + 0 - 2w_1 + 1].$$

Using the initial value $w_0 = y(0) = 1$ and step size $h = 1$ gives:

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

This leads to

$$2w_1 = 1, \text{ so } w_1 = \frac{1}{2}.$$

(e) For the comparison we use the following items: accuracy, stability, and amount of work. Below we make the comparison:

- Accuracy: since the error of Euler Forward is $O(h)$ and that of the trapezoidal rule is $O(h^2)$, the error is less for the trapezoidal rule.
- Stability: since the value of $-(1+t)$ is always negative the trapezoidal rule is stable for all step sizes, whereas for Euler Forward the step size should satisfy the inequality $h \leq \frac{1+t}{2}$.
- Amount of work: since the differential equation is linear the amount of work for the implicit trapezoidal rule is comparable to the work of the explicit Euler Forward method.

From the above comparisons we conclude that for this problem the trapezoidal rule is preferred.

2. [a] We compute

$$x + y = 2/3 + 1999/3000 = 1.333,$$

and

$$x - y = 2/3 - 1999/3000 = 1/3000 = 0.3333\ldots \cdot 10^{-3}.$$

Further, we have $fl(x) = 0.6667$, $fl(y) = 0.6663$, and

$$fl(x) + fl(y) = 0.1333 \cdot 10^1,$$

hence $fl(fl(x) + fl(y)) = 0.1333 \cdot 10^1$.

For the subtraction, one obtains

$$fl(x) - fl(y) = 0.4 \cdot 10^{-3},$$

and hence

$$fl(fl(x) - fl(y)) = fl(0.4 \cdot 10^{-3}) = 0.4000 \cdot 10^{-3}.$$

[b] After the addition, the relative error is given by

$$|\frac{0.1333 \cdot 10^1 - 1.333}{0.1333 \cdot 10^1}| = 0,$$

and after the subtraction, one gets

$$|\frac{0.4000 \cdot 10^{-3} - 0.3333 \ldots \cdot 10^{-3}}{0.3333 \ldots \cdot 10^{-3}}| = 0.2.$$

[c] The relative error due to subtraction of two positive numbers is divided by the difference between these numbers. If this difference gets arbitrarily small, then the relative error gets arbitrarily large for a given absolute error.

[d] Using central differences for the second order derivative at a node $x_j = jh$, gives

$$y''(x_j) \approx \frac{y_{j+1} - 2y_j + y_{j-1}}{h^2} =: Q(h). \tag{7}$$

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

$$\begin{aligned} y_{j+1} &= y(x_j + h) = y(x_j) + hy'(x_j) + \tfrac{h^2}{2} y''(x_j) + \tfrac{h^3}{3!} y'''(x_j) + \tfrac{h^4}{4!} y''''(\eta_+), \\ y_{j-1} &= y(x_j - h) = y(x_j) - hy'(x_j) + \tfrac{h^2}{2} y''(x_j) - \tfrac{h^3}{3!} y'''(x_j) + \tfrac{h^4}{4!} y''''(\eta_-). \end{aligned} \tag{8}$$

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 $Q(h)$ gives $|y''(x_j) - Q(h)| = O(h^2)$. Therewith, we obtain the following discretization formula for the internal grid nodes:

$$\frac{-w_{j-1} + 2w_j - w_{j+1}}{h^2} + x_j w_j = x_j^3 - 2. \tag{9}$$

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 $x = -h$, and we define $y_{-1} := y(-h)$. Then, using central differences at $x = 0$ gives

$$0 = y'(0) \approx \frac{y_1 - y_{-1}}{2h} =: Q_b(h). \tag{10}$$

Using Taylor's Theorem, gives

$$Q_b(h) =$$

$$\frac{y(0)+hy'(0)+\frac{h^2}{2}y''(0)+\frac{h^3}{3!}y'''(\eta_+)-(y(0)-hy'(0)+\frac{h^2}{2}y''(0)-\frac{h^3}{3!}y'''(\eta_-))}{2h} =$$

$$y'(0)+O(h^2). \tag{11}$$

Again, we get an error of $O(h^2)$.

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

$$\frac{w_1 - w_{-1}}{2h} = 0 \Leftrightarrow w_{-1} = w_1. \tag{12}$$

The discretization at $x = 0$ is given by

$$\frac{-w_{-1} + 2w_0 - w_1}{h^2} = -2. \tag{13}$$

Substitution of equation (12) into the above equation, yields

$$\frac{2w_0 - 2w_1}{h^2} = -2. \tag{14}$$

Subsequently, we consider the boundary $x = 1$. To this extent, we consider its neighboring point x_{n-1}, here substitution of the boundary condition $w_n = y(1) = y_n = 1$ into equation (9), gives

$$\frac{-w_{n-2} + 2w_{n-1}}{h^2} + x_{n-1}w_{n-1} = x_{n-1}^3 - 2 + \frac{1}{h^2} = (1-h)^3 - 2 + \frac{1}{h^2}. \tag{15}$$

This concludes our discretization of the boundary conditions. In order to get a symmetric discretization matrix, one divides equation (14) by 2.

[e] Next, we use $h = 1/3$, then, from equations (9, 14, 15), one obtains the following system

$$\begin{aligned} 9w_0 - 9w_1 &= -1 \\ -9w_0 + 18\tfrac{1}{3}w_1 - 9w_2 &= -\tfrac{53}{27} \\ -9w_1 + 18\tfrac{2}{3}w_2 &= \tfrac{197}{27}. \end{aligned} \tag{16}$$

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Monday January 28 2013, 18:30-21:30

1. [a] The local truncation error is given by

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

where z_{n+1} is computed by one step of the method starting from y_n, and we determine y_{n+1} by the use of a Taylor Series around t_n:

$$y_{n+1} = y_n + hy'(t_n) + \frac{h^2}{2}y''(t_n) + O(h^3). \tag{2}$$

We realize that

$$\begin{aligned} y'(t_n) &= f(t_n, y_n) \\ y''(t_n) &= \frac{df(t_n, y_n)}{dt} = \frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y}y'(t_n) = \\ &= \frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y}f(t_n, y_n). \end{aligned} \tag{3}$$

Hence, this gives

$$y_{n+1} = y_n + hy'(t_n) + \frac{h^2}{2}\left(\frac{\partial f(t_n, y_n)}{\partial t} + \frac{\partial f(t_n, y_n)}{\partial y}f(t_n, y_n)\right) + O(h^3). \tag{4}$$

For z_{n+1}, after substitution of the predictor–step for z^*_{n+1} into the corrector–step, and using the Taylor Series around (t_n, y_n)

$$\begin{aligned} &z_{n+1} = y_n + \tfrac{h}{2}\left(f(t_n, y_n) + f(t_n + h, y_n + hf(t_n, y_n))\right) = \\ &y_n + \frac{h}{2}\left(f(t_n, y_n) + f(t_n, y_n) + h(\frac{\partial f(t_n, y_n)}{\partial t} + f(t_n, y_n)\frac{\partial f(t_n, y_n)}{\partial y}) + O(h^2)\right). \end{aligned} \tag{5}$$

Then, it follows that

$$y_{n+1} - z_{n+1} = O(h^3), \text{ and hence } \tau_{n+1}(h) = \frac{O(h^3)}{h} = O(h^2). \tag{6}$$

[b] Consider the test–equation $y' = \lambda y$, then it follows that

$$\begin{aligned} w^*_{n+1} &= w_n + h\lambda w_n = (1+h\lambda)w_n, \\ w_{n+1} &= w_n + \frac{h}{2}(\lambda w_n + \lambda w^*_{n+1}) = \\ &= w_n + \frac{h}{2}(\lambda w_n + \lambda(w_n + h\lambda w_n)) = (1 + h\lambda + \frac{(h\lambda)^2}{2})w_n. \end{aligned} \tag{7}$$

Hence the amplification factor is given by

$$Q(h\lambda) = 1 + h\lambda + \frac{(h\lambda)^2}{2}. \tag{8}$$

[c] Let $x_1 = y$ and $x_2 = y'$, then it follows that $y'' = x_2'$, and hence we get

$$\begin{aligned} &x_2' + 12x_2 + 72x_1 = \sin(t), \\ &x_2 = x_1'. \end{aligned} \tag{9}$$

This expression is written as

$$\begin{aligned} &x_1' = x_2, \\ &x_2' = -72x_1 - 12x_2 + \sin(t). \end{aligned} \tag{10}$$

Finally, we get the following matrix–form:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -72 & -12 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \sin(t) \end{pmatrix}. \tag{11}$$

Here, we have $A = \begin{pmatrix} 0 & 1 \\ -72 & -12 \end{pmatrix}$ and $f = \begin{pmatrix} 0 \\ \sin(t) \end{pmatrix}$. The initial conditions are given by $\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

[d] The Modified Euler Method, applied to the system $\underline{x}' = A\underline{x} + \underline{f}$, gives

$$\begin{aligned} \underline{w}^*_1 &= \underline{w}_0 + h\left(A\underline{w}_0 + \underline{f}_0\right), \\ \underline{w}_1 &= \underline{w}_0 + \tfrac{h}{2}\left(A\underline{w}_0 + \underline{f}_0 + A\underline{w}^*_1 + \underline{f}_1\right). \end{aligned} \tag{12}$$

With the initial condition and $h = 0.1$, this gives

$$\underline{w}^*_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{1}{10}\left(\begin{pmatrix} 0 & 1 \\ -72 & -12 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1.2 \\ -7.6 \end{pmatrix}. \tag{13}$$

Then, the correction–step is given by

$$\underline{w}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \tfrac{1}{20}\left(\begin{pmatrix} 0 & 1 \\ -72 & -12 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -72 & -12 \end{pmatrix}\begin{pmatrix} 1.2 \\ -7.6 \end{pmatrix} + \begin{pmatrix} 0 \\ \sin(\frac{1}{10}) \end{pmatrix}\right) =$$

$$= \begin{pmatrix} 0.72 \\ -2.55501 \end{pmatrix} \tag{14}$$

[e] To this extent, we determine the eigenvalues of the matrix A. Subsequently, these eigenvalues are substituted into the amplification factor. The eigenvalues of A are given by $-6 \pm 6i$. Using $h = 0.25$, it follows that

$$Q(h\lambda) = 1+h\lambda+\frac{1}{2}h^2\lambda^2 = 1+\frac{1}{4}(-6+6i)+\frac{1}{32}(-6+6i)^2 = 1-\frac{3}{2}+\frac{3}{2}i-\frac{72}{32}i = -\frac{1}{2}-\frac{3}{4}i. \quad (15)$$

Herewith, it follows that $|Q(h\lambda)|^2 = \frac{1}{4} + \frac{9}{16} = \frac{13}{16} < 1$. Hence for $h = 0.25$, it follows that the method applied to the given system is stable. Note that this conclusion holds for both the eigenvalues of A since they are complex conjugates.

2. (a) The first order backward difference formula for the first derivative is given by

$$f'(t) \approx \frac{f(t) - f(t-h)}{h}.$$

Using $t = 2$, and $h = 1$ the approximation of the velocity is

$$\frac{f(2) - f(1)}{1} = 250 - 215 = 35 \text{ (m/s)}.$$

(b) Taylor polynomials are:

$$\begin{aligned} f(0) &= f(2h) - 2hf'(2h) + 2h^2 f''(2h) - \frac{(2h)^3}{6} f'''(\xi_0)\,, \\ f(h) &= f(2h) - hf'(2h) + \frac{h^2}{2} f''(2h) - \frac{h^3}{6} f'''(\xi_1)\,, \\ f(2h) &= f(2h). \end{aligned}$$

We know that $Q(h) = \frac{\alpha_0}{h} f(0) + \frac{\alpha_1}{h} f(h) + \frac{\alpha_2}{h} f(2h)$, which should be equal to $f'(2h) + O(h^2)$. This leads to the following conditions:

$$\begin{aligned} \tfrac{\alpha_0}{h} + \tfrac{\alpha_1}{h} + \tfrac{\alpha_2}{h} &= 0\,, \\ -2\alpha_0 - \alpha_1 &= 1\,, \\ 2\alpha_0 h + \tfrac{1}{2}\alpha_1 h &= 0\,. \end{aligned}$$

(c) The truncation error follows from the Taylor polynomials:

$$f'(2h) - Q(h) = f'(2h) - \frac{f(0) - 4f(h) + 3f(2h)}{2h} = \frac{\frac{8h^3}{6} f'''(\xi_0) - 4(\frac{h^3}{6} f'''(\xi_1))}{2h} = \frac{1}{3}h^2 f'''(\xi).$$

Using the new formula with $h = 1$ we obtain the estimate:

$$\frac{f(0) - 4f(1) + 3f(2)}{2} = \frac{200 - 4 \times 215 + 3 \times 250}{2} = 45 \text{ (m/s)}.$$

Note that the estimated velocity of the vehicle is larger than the maximum speed of 40 (m/s).

(d) To estimate the measuring error we note that

$$|Q(h) - \hat{Q}(h)| = |\frac{(f(0) - \hat{f}(0)) - 4(f(h) - \hat{f}(h)) + 3(f(2h) - \hat{f}(2h))}{2h}|$$

$$\leq \frac{|f(0) - \hat{f}(0)| + 4|f(h) - \hat{f}(h)| + 3|f(2h) - \hat{f}(2h)|}{2h} \leq \frac{4\epsilon}{h},$$

so $C_1 = 4$.

(e) We integrate $f(x)$, in which we approximate $f(x)$ by $p_1(x)$, then it follows:

$$\int_{x_0}^{x_1} f(x)dx \approx \int_{x_0}^{x_1} p_1(x)dx = \int_{x_0}^{x_1} \left\{ f(x_0)\frac{x - x_1}{x_0 - x_1} + f(x_1)\frac{x - x_0}{x_1 - x_0} \right\} dx =$$

$$= \left[\frac{1}{2}\frac{(x - x_0)^2}{x_1 - x_0}f(x_1)\right]_{x_0}^{x_1} + \left[\frac{1}{2}\frac{(x - x_1)^2}{x_0 - x_1}f(x_0)\right]_{x_0}^{x_1} = \frac{1}{2}(x_1 - x_0)(f(x_0) + f(x_1)). \tag{16}$$

This is the Trapezoidal Rule.

(f) The magnitude of the error of the numerical integration over interval $[x_0, x_1]$ is given by

$$|\int_{x_0}^{x_1} f(x)dx - \int_{x_0}^{x_1} p_1(x)dx| = |\int_{x_0}^{x_1} (f(x) - p_1(x))\, dx| =$$

$$|\int_{x_0}^{x_1} \frac{1}{2}(x - x_0)(x - x_1)f''(\chi(x))dx| \leq \frac{1}{2}\max_{x\in[x_0,x_1]} |f''(x)| \int_{x_0}^{x_1} (x - x_0)(x_1 - x)dx =$$

$$\frac{1}{12}(x_1 - x_0)^3 \max_{x\in[x_0,x_1]} |f''(x)|. \tag{17}$$

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 5 2012, 18:30-21:30

1. [a] The test–equation is given by $y' = \lambda y$, and we bear in mind that the amplification factor is defined by
$$Q(h\lambda) = \frac{w^{n+1}}{w^n}. \tag{1}$$
Then for the Trapezoidal Rule, we get
$$w^{n+1} = w^n + \frac{h}{2}(\lambda w^n + \lambda w^{n+1}) = w^n + \frac{h\lambda}{2}(w^n + w^{n+1}). \tag{2}$$
The above equation is rewritten as
$$w^{n+1}(1 - \frac{h\lambda}{2}) = w^n(1 + \frac{h\lambda}{2}). \tag{3}$$
Then, using the definition of the amplification factor, we immediately have
$$Q_T(h\lambda) = \frac{1 + \frac{h\lambda}{2}}{1 - \frac{h\lambda}{2}}. \tag{4}$$
The Modified Euler Method is treated analogously, to get
$$\begin{aligned} \hat{w}^{n+1} &= w^n + h\lambda w^n, && \text{predictor} \\ w^{n+1} &= w^n + \tfrac{h}{2}(\lambda w^n + \lambda \hat{w}^{n+1}), && \text{corrector.} \end{aligned} \tag{5}$$
Combining the predictor and corrector, gives
$$w^{n+1} = w^n + \frac{h\lambda}{2}(w^n + w^n + h\lambda w^n) = w^n(1 + h\lambda + \frac{(h\lambda)^2}{2}). \tag{6}$$
Finally, the definition of the amplification factor implies that
$$Q_{ME}(h\lambda) = 1 + h\lambda + \frac{(h\lambda)^2}{2}. \tag{7}$$
[b] The local truncation error is defined by
$$\tau_{n+1}(h) = \frac{y^{n+1} - z^{n+1}}{h}, \tag{8}$$

where y^{n+1} and z^{n+1}, respectively, denote the exact solution and the numerical approximation at time t^{n+1} under using y^n. Since, we use the test–equation to estimate the local truncation error, we get

$$z^{n+1} = Q(h\lambda)y^n. \tag{9}$$

The exact solution to the test–equation at time t^{n+1} is expressed in terms of y^n by

$$y^{n+1} = y^n e^{\lambda h}. \tag{10}$$

Substitution into the definition of the local truncation error, gives

$$\tau_{n+1}(h) = \frac{y^n}{h}(e^{h\lambda} - Q(h\lambda)) = \frac{y^n}{h}(1 + h\lambda + \frac{(h\lambda)^2}{2} + \frac{(h\lambda)^3}{3!} + O(h^4) - Q(h\lambda)), \tag{11}$$

where we used the Taylor expansion of the exponential around 0. For the Trapezoidal Rule, we have

$$\begin{aligned} Q_T(h\lambda) = \frac{1+\frac{h\lambda}{2}}{1-\frac{h\lambda}{2}} = (1+\frac{h\lambda}{2})(1+\frac{h\lambda}{2}+(\frac{h\lambda}{2})^2+(\frac{h\lambda}{2})^3+O(h^4)) = \\ 1 + h\lambda + \frac{(h\lambda)^2}{2} + \frac{(h\lambda)^3}{4} + O(h^4). \end{aligned} \tag{12}$$

Using equation (11), we get after some rearrangements

$$\tau_{n+1}(h) = -\frac{y^n\lambda^3 h^2}{12} + O(h^3) = O(h^2). \tag{13}$$

The Modified Euler Method is treated similarly with

$$Q_{ME}(h\lambda) = 1 + h\lambda + \frac{(h\lambda)^2}{2}, \tag{14}$$

to give via equation (11)

$$\tau_{n+1}(h) = \frac{y^n\lambda^3 h^2}{6} + O(h^3) = O(h^2). \tag{15}$$

[c] Let $y_1 = y$ and let $y_2 = y_1'$, then $y_2' = y_1'' = y''$. Hence we have

$$y_1' = y_2, \qquad y_2' = -y_1 + t(1-t). \tag{16}$$

The two equations are linear and therewith, one can rewrite this system using a matrix representation:

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ t(1-t) \end{pmatrix}, \tag{17}$$

Further, we have $y_1(0) = y(0) = 0$ and $y_2(0) = y'(0) = 1$.

[d] We use $h = \frac{1}{2}$, and let

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \underline{w}^1 = \begin{pmatrix} w_1^1 \\ w_2^1 \end{pmatrix}, \quad \underline{y}^0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{18}$$

where the subscript stands for the component, whereas the superscript denotes the time–index. The Trapezoidal Rule gives

$$\underline{w}^1 = \underline{y}^0 + \frac{h}{2}(A\underline{y}^0 + A\underline{w}^1 + \begin{pmatrix} 0 \\ \frac{1}{4} \end{pmatrix}). \tag{19}$$

This gives

$$(I - \frac{h}{2}A)\underline{w}^1 = (I + \frac{h}{2}A)\underline{y}^0 + \frac{h}{2}\begin{pmatrix} 0 \\ \frac{1}{4} \end{pmatrix}. \tag{20}$$

Substitution of $h = \frac{1}{2}$, gives the following linear system

$$\begin{pmatrix} 1 & -\frac{1}{4} \\ \frac{1}{4} & 1 \end{pmatrix} \underline{w}^1 = \begin{pmatrix} \frac{1}{4} \\ \frac{17}{16} \end{pmatrix}. \tag{21}$$

This system is solved by

$$\underline{w}^1 = \begin{pmatrix} \frac{33}{68} \\ \\ \frac{16}{17} \end{pmatrix} \tag{22}$$

Next, we treat the Modified Euler Method. First, we carry out the prediction step

$$\underline{\hat{w}}^1 = \underline{y}^0 + hA\underline{y}^0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \\ 1 \end{pmatrix}. \tag{23}$$

Subsequently, we perform the corrector step

$$\underline{w}^1 = \underline{y}^0 + \frac{h}{2}\left(A\underline{y}^0 + A\underline{\hat{w}}_1 + \begin{pmatrix} 0 \\ \frac{1}{4} \end{pmatrix}\right). \tag{24}$$

Using $h = \frac{1}{2}$, gives

$$\underline{w}^1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{4}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{4} \end{pmatrix}\right) = \begin{pmatrix} \frac{1}{2} \\ \\ \frac{15}{16} \end{pmatrix}. \tag{25}$$

[e] The local truncation errors for both methods are approximated by

$$\tau_{n+1}^T(h) = -\frac{y^n\lambda^3h^2}{12}, \qquad \tau_{n+1}^{EM}(h) = \frac{y^n\lambda^3h^2}{6}. \tag{26}$$

From these equations, it can be seen that the errors have the same order, although the error from the Trapezoidal Rule is about twice as small as the one from the Modified Euler Method in the limit for $h \to 0$.

With regard to stability, the eigenvalues of A in the present initial value problem, are given by $\lambda = \pm i$. Herewith, the following amplification factors are obtained:

$$Q_T(h) = \frac{1 + \frac{\pm ih}{2}}{1 - \frac{\pm ih}{2}}, \qquad Q_{ME}(h) = 1 \pm ih - \frac{1}{2}h^2. \tag{27}$$

This gives the following moduli

$$|Q_T(h)| = 1, \qquad |Q_{ME}(h)| = \sqrt{(1 - \frac{h^2}{2})^2 + h^2} = \sqrt{1 + \frac{h^4}{4}} > 1. \tag{28}$$

Hence the Trapezoidal Rule is neutrally stable, whereas the Modified Euler Method is unstable.

The workload is smaller for the Modified Euler Method, since no linear system needs to be solved. Although the solution of the linear system may require considerable computation time if A is a very large matrix, the issue is not very important for the present case.

Therefore, the Trapezoidal Rule is to be preferred for the present system since the system is just a two-by-two set of equations.

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:

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

Since h is a continuous function one obtains:

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

so

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

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

(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 $g_1(x) = x - \frac{x^3-3}{x^4}$. The first derivative is:

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

Substitution of p yields:

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

Since $p = 3^{\frac{1}{3}}$ the final term cancels:

$$g_1'(p) = 1 - \frac{3p^6}{p^8} = 1 - 3^{\frac{1}{3}} = -0.4422.$$

This implies that the method is convergent with convergence factor 0.4422.

For the second method we have:

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

Thus the method diverges.

For the third method we have:

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

Thus the method is convergent with convergence factor 0.

Concluding we note that the third method is the fastest.

(c) 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

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

(d) Using the Newton-Raphson iteration method

$$z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}$$

for $f(x) = x^4 - 3x$ we have to compute $f'(x)$. It easily follows that $f'(x) = 4x^3 - 3$. Substituting this together with the initial guess $z_0 = 1$ into the definition of the Newton-Raphson method leads to:

$$z_1 = z_0 - \frac{f(z_0)}{f'(z_0)} = z_0 - \frac{z_0^4 - 3z_0}{4z_0^3 - 3} = 1 - \frac{1 - 3}{4 - 3} = 3.$$

(e) The Newton-Raphson iteration method can be derived using a graph of a function, in which the zero of the tangent at z_k on $f(x)$ defines z_{k+1}. We consider a linearization of $f(x)$ around z_k:

$$L(x) := f(z_k) + (x - z_k)f'(z_k),$$

and determine its zero, that is $L(z_{k+1}) = 0$, this gives

$$z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}, \text{ provided that } f'(z_k) \neq 0,$$

□

(f) We consider a Taylor polynomial around z_k, to express z

$$0 = f(z) = f(z_k) + (z - z_k)f'(z_k) + \frac{(z - z_k)^2}{2} f''(\xi_k), \tag{29}$$

for some ξ_k between z and z_k. Note that this form gives the exact representation. Subsequently, we consider the Newton-Raphson approximation

$$0 = L(z_{k+1}) = f(z_k) + (z_{k+1} - z_k)f'(z_k). \tag{30}$$

Subtraction of these two above equations gives

$$z_{k+1} - z = \frac{(z_k - z)^2}{2} \frac{f''(\xi_k)}{f'(z_k)}, \text{ provided that } f'(z_k) \neq 0, \tag{31}$$

and hence

$$|z_{k+1} - z| = \frac{(z_k - z)^2}{2} \left|\frac{f''(\xi_k)}{f'(z_k)}\right|, \text{ provided that } f'(z_k) \neq 0, \tag{32}$$

Note that $f''(x) = 12x^2$ and $z = 3^{\frac{1}{3}}$. Using $z_k \to z$, $\xi_k \to z$ as $k \to \infty$ and continuity of $f(x)$ up to at least the second derivative, we arrive at $K = |\frac{f''(z)}{2*f'(z)}| = |\frac{12z^2}{2(4z^3-3)}| \approx 1.3867$. □

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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)
Thursday June 30 2011, 18:30-21:30

1. a The local truncation error is defined by

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

where

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

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

$$y_{n+1} = y_n + hy'(t_n) + \frac{h^2}{2}y''(\xi), \quad \exists\, \xi \in (t_n, t_{n+1}). \tag{3}$$

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

$$\tau_h = \frac{h}{2}y''(\xi), \quad \exists\, \xi \in (t_n, t_{n+1}). \tag{4}$$

Hence, the truncation error is of first order.

b We define $y_1 := y$ and $y_2 := y'$, hence $y_1' = y_2$. Further, we use the differential equation to obtain

$$y'' + \varepsilon y' + y = y_1'' + \varepsilon y_1' + y_1 = y_2' + \varepsilon y_2 + y_1. \tag{5}$$

Hence, we obtain

$$y_2' = -y_1 - \varepsilon y_2 + sin(t). \tag{6}$$

Hence the system is given by

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= -y_1 - \varepsilon y_2 + sin(t). \end{aligned} \tag{7}$$

The initial conditions are given by

$$\begin{aligned} 1 &= y(0) = y_1(0), \\ 0 &= y'(0) = y_1'(0) = y_2(0). \end{aligned} \tag{8}$$

c First, we use the test equation, $y' = \lambda y$, to analyze numerical stability. For forward Euler, we obtain

$$w_{n+1} = w_n + h\lambda w_n = Q(h\lambda)w_n, \tag{9}$$

hence the amplification factor becomes

$$Q(h\lambda) = 1 + h\lambda. \tag{10}$$

The numerical solution is stable if and only if $|Q(h\lambda)| \leq 1$. Next, we deal with the case $\varepsilon = 0$, to obtain the following system

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{11}$$

This system gives the following eigenvalues $\lambda_{1,2} = \pm i$, where i is the imaginary unit. Hence, the amplification factor is given by

$$Q(h\lambda) = 1 \pm hi. \tag{12}$$

Then, it is immediately clear that $|Q(h\lambda)| > 1$ for all $h > 0$. Hence, we conclude that the forward Euler method is never stable if $\varepsilon = 0$.

d From Assignment 1.c., we know that if $\varepsilon = 0$, the eigenvalues of the system are purely imaginary. This implies that the system is analytically (zero) stable if $\varepsilon = 0$.

Nonzero values of ε give the following system

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & \varepsilon \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{13}$$

then we get the following eigenvalues $\lambda_{1,2} = \frac{\varepsilon}{2} \pm \frac{1}{2}\sqrt{\varepsilon^2 - 4}$ (real-valued), if $\varepsilon^2 - 4 \geq 0$ and $\lambda = \frac{\varepsilon}{2} \pm \frac{i}{2}\sqrt{4 - \varepsilon^2}$ (nonreal-valued) if $\varepsilon^2 - 4 < 0$. Hence, we consider two cases: real-valued and nonreal-valued eigenvalues.

Real-valued eigenvalues
In this case $|\varepsilon| \geq 2$, and $0 \leq \varepsilon^2 - 4 < \varepsilon^2$, and hence the real-valued eigenvalues have the same sign, which is determined by the sign of ε. Hence, if $\varepsilon \leq -2$, then, the system is stable. Furthermore, if $\varepsilon \geq 2$, then, the system is unstable.

Nonreal-valued eigenvalues
In this case $|\varepsilon| < 2$. The system is analytically unstable if and only if the real part of the eigenvalues is positive. Further, the real part of the eigenvalues is positive if and only if $\varepsilon > 0$. Hence, the system is analytically unstable if and only if $\varepsilon > 0$. Hence, the system is stable if and only if $(-2 <)\varepsilon \leq 0$.

From these arguments, it follows that the system is stable if and only if $\varepsilon \leq 0$.

e Since currently the discriminant, $\varepsilon^2 - 4$, is negative, the eigenvalues are nonreal. Substitution into the amplification factor yields

$$Q(h\lambda) = 1 + \frac{\varepsilon}{2}h \pm \frac{ih}{2}\sqrt{4-\varepsilon^2}. \tag{14}$$

Hence, numerical stability is warranted if

$$|Q(h\lambda)|^2 = (1 + \frac{\varepsilon}{2}h)^2 + \frac{h^2}{4}(4-\varepsilon^2) \leq 1. \tag{15}$$

Hence for stability, we have

$$1 + \varepsilon h + \frac{\varepsilon^2 h^2}{4} + h^2 - \frac{\varepsilon^2 h^2}{4} = 1 + h\varepsilon + h^2 \leq 1. \tag{16}$$

Since $h > 0$, we obtain the following stability criterion

$$h \leq -\varepsilon = |\varepsilon|. \tag{17}$$

If $\varepsilon = -2$, then both eigenvalues are real-valued and given by $\lambda_{1,2} = -1$. For this case, we obtain $Q(\lambda h) = 1 - h$, and stability is warranted if and only if $-1 \leq Q(h\lambda) \leq 1$, hence $h \leq 2(= |\varepsilon|$.

We conclude that for $-2 \leq \varepsilon < 0$, we have a numerically stable solution if and only if $h \leq |\varepsilon|$.

2. a After discretization by the use of finite differences one obtains

$$\frac{-w_{i-1} + 2w_i - w_{i+1}}{h^2} + x_i^2 w_i = x_i. \tag{18}$$

The truncation error is defined by

$$e_i = \frac{-y_{i-1} + 2y_i - y_{i+1}}{h^2} + x_i^2 y_i - x_i. \tag{19}$$

Taylor series of y_{i-1} and y_{i+1} around x_i, gives

$$\begin{aligned} y_{i+1} &= y_i + hy'(x_i) + \frac{h^2}{2!}y''(x_i) + \frac{h^3}{3!}y'''(x_i) + \frac{h^4}{4!}y''''(x_i) + O(h^5), \\ y_{i-1} &= y_i - hy'(x_i) + \frac{h^2}{2!}y''(x_i) - \frac{h^3}{3!}y'''(x_i) + \frac{h^4}{4!}y''''(x_i) - O(h^5), \end{aligned} \tag{20}$$

Substitution of the above expressions into the definition of the truncation error gives

$$\varepsilon_i = -y''(x_i) + O(h^2) + x_i^2 y(x_i) - x_i. \tag{21}$$

Using the differential equation $-y'' + x^2 y = x$ finally gives

$$\varepsilon_i = O(h^2). \tag{22}$$

b For this case we have $h = 0.25$, for the points $j \in \{1,2,3\}$, the discretization with $w_0 = 0$ and $w_4 = 1$:

$$\begin{aligned} &32w_1 - 16w_2 + \frac{1}{16}w_1 = \frac{1}{4}, \\ &-16w_1 + 32w_2 - 16w_3 + \frac{1}{4}w_2 = \frac{1}{2}, \\ &-16w_2 + 32w_3 + \frac{9}{16}w_3 = \frac{3}{4} + 16. \end{aligned} \tag{23}$$

Hence in matrix-vector form:

$$\begin{pmatrix} 32.0625 & -16 & 0 \\ -16 & 32.25 & -16 \\ 0 & -16 & 32.5625 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} 0.25 \\ 0.5 \\ 16.75 \end{pmatrix} \tag{24}$$

c 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:

$$\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} [x_n + h(x_n)(x_n^2 - 4)]$$

Since h is a continuous function one obtains:

$$p = p + h(p)(p^2 - 4)$$

so

$$h(p)(p^2 - 4) = 0.$$

Since $h(x) \neq 0$ for each $x \neq 0$ it follows that $p^2 - 4 = 0$ and thus there are two limits $p = -2$ and $p = 2$.

d 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 if $p_0 \neq p$. 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 $g_1(x) = x - \frac{1}{2}x(x^2 - 4) = 3x - \frac{1}{2}x^3$. The first derivative is:

$$g_1'(x) = 3 - \frac{3}{2}x^2$$

Substitution of $p = 2$ yields:

$$g_1'(2) = 3 - \frac{3}{2}4 = 3 - 6 = -3.$$

Since $|g_1'(2)| = 3 > 1$ this method is divergent.

For the second method we have:

$$g_2(x) = x - \frac{1}{3}(x^2 - 4)$$

$$g_2'(x) = 1 - \frac{2}{3}x$$

Since $|g_2'(2)| = |-\frac{1}{3}| = \frac{1}{3} < 1$ the method converges with convergence factor $\frac{1}{3}$.

For the third method we have:

$$g_3(x) = x - \frac{1}{2x}(x^2 - 4) = \frac{x}{2} + \frac{2}{x}$$

$$g_3'(x) = \frac{1}{2} - \frac{2}{x^2}$$

Note that $g_3'(2) = \frac{1}{2} - \frac{2}{4} = 0$ the method is convergent with convergence factor 0.

Concluding we note that the third method is the fastest since $|g_3'(2)| < |g_2'(2)|$.

e We use the iteration process:

$$x_{n+1} = x_n - \frac{1}{3}(x_n^2 - 4)$$

Starting from $x_0 = 3$ we obtain the following iterates:

$$x_1 = 1.3333$$

$$x_2 = 2.0741$$

$$x_3 = 1.9735$$

Note that the method indeed converges and that the convergence is alternating.

f 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

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