## DELFT UNIVERSITY OF TECHNOLOGY

Faculty of Electrical Engineering, Mathematics and Computer Science

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

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

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

where $z_{n+1}$ is given by

$$
\begin{equation*}
z_{n+1}=y_{n}+\Delta t\left(a_{1} f\left(t_{n}, y_{n}\right)+a_{2} f\left(t_{n}+\Delta t, y_{n}+\Delta t f\left(t_{n}, y_{n}\right)\right)\right. \tag{2}
\end{equation*}
$$

A Taylor expansion of $f$ around $\left(t_{n}, y_{n}\right)$ yields

$$
\begin{equation*}
f\left(t_{n}+\Delta t, y_{n}+\Delta t f\left(t_{n}, y_{n}\right)\right)=f\left(t_{n}, y_{n}\right)+\Delta t \frac{\partial f}{\partial t}\left(t_{n}, y_{n}\right)+\Delta t f\left(t_{n}, y_{n}\right) \frac{\partial f}{\partial y}\left(t_{n}, y_{n}\right)+O\left((\Delta t)^{2}\right) \tag{3}
\end{equation*}
$$

This is substituted into equation (2) to obtain

$$
\begin{equation*}
z_{n+1}=y_{n}+\Delta t\left(a_{1} f\left(t_{n}, y_{n}\right)+a_{2}\left[f\left(t_{n}, y_{n}\right)+\Delta t \frac{\partial f}{\partial t}\left(t_{n}, y_{n}\right)+\Delta t f\left(t_{n}, y_{n}\right) \frac{\partial f}{\partial y}\left(t_{n}, y_{n}\right)\right]\right)+O\left((\Delta t)^{3}\right) \tag{4}
\end{equation*}
$$

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

$$
\begin{equation*}
y_{n+1}=y\left(t_{n}+\Delta t\right)=y_{n}+\Delta t y^{\prime}\left(t_{n}\right)+\frac{(\Delta t)^{2}}{2} y^{\prime \prime}\left(t_{n}\right)+O\left((\Delta t)^{3}\right) . \tag{5}
\end{equation*}
$$

From the differential equation we know that:

$$
\begin{equation*}
y^{\prime}\left(t_{n}\right)=f\left(t_{n}, y_{n}\right) \tag{6}
\end{equation*}
$$

From the Chain Rule of Differentiation, we derive

$$
\begin{equation*}
y^{\prime \prime}\left(t_{n}\right)=\frac{d f\left(t_{n}, y_{n}\right)}{d t}=\frac{\partial f\left(t_{n}, y_{n}\right)}{\partial t}+\frac{\partial f\left(t_{n}, y_{n}\right)}{\partial y} y^{\prime}\left(t_{n}\right) \tag{7}
\end{equation*}
$$

after substitution of the differential equation one obtains:

$$
\begin{equation*}
y^{\prime \prime}\left(t_{n}\right)=\frac{\partial f\left(t_{n}, y_{n}\right)}{\partial t}+\frac{\partial f\left(t_{n}, y_{n}\right)}{\partial y} f\left(t_{n}, y_{n}\right) \tag{8}
\end{equation*}
$$

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

$$
\begin{equation*}
\tau_{n+1}(\Delta t)=f\left(t_{n}, y_{n}\right)\left(1-\left(a_{1}+a_{2}\right)\right)+\Delta t\left(\frac{\partial f}{\partial t}+f \frac{\partial f}{\partial y}\right)\left(\frac{1}{2}-a_{2}\right)+O\left((\Delta t)^{2}\right) \tag{9}
\end{equation*}
$$

Hence
i. $a_{1}+a_{2}=1$ implies $\tau_{n+1}(\Delta t)=O(\Delta t)$;
ii. $a_{1}+a_{2}=1$ and $a_{2}=1 / 2$, that is, $a_{1}=a_{2}=1 / 2$, gives $\tau_{n+1}(\Delta t)=O\left((\Delta t)^{2}\right)$.
(b) The test equation is given by

$$
\begin{equation*}
y^{\prime}=\lambda y \tag{10}
\end{equation*}
$$

Application of the predictor step to the test equation gives

$$
\begin{equation*}
w_{n+1}^{*}=w_{n}+\lambda \Delta t w_{n}=(1+\lambda \Delta t) w_{n} . \tag{11}
\end{equation*}
$$

The corrector step yields

$$
\begin{equation*}
w_{n+1}=w_{n}+\Delta t\left(a_{1} \lambda w_{n}+a_{2} \lambda(1+\lambda \Delta t) w_{n}\right)=\left(1+\left(a_{1}+a_{2}\right) \lambda \Delta t+a_{2}(\lambda \Delta t)^{2}\right) w_{n} . \tag{12}
\end{equation*}
$$

Hence the amplification factor is given by

$$
\begin{equation*}
Q(\lambda \Delta t)=1+\left(a_{1}+a_{2}\right) \lambda \Delta t+a_{2}(\lambda \Delta t)^{2} . \tag{13}
\end{equation*}
$$

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

$$
\begin{equation*}
-1 \leq Q(\lambda \Delta t) \leq 1 \tag{14}
\end{equation*}
$$

from the previous assignment, we have

$$
\begin{equation*}
-1 \leq 1+\left(a_{1}+a_{2}\right) \lambda \Delta t+a_{2}(\lambda \Delta t)^{2} \leq 1 \Leftrightarrow-2 \leq\left(a_{1}+a_{2}\right) \lambda \Delta t+a_{2}(\lambda \Delta t)^{2} \leq 0 \tag{15}
\end{equation*}
$$

First, we consider the left inequality:

$$
\begin{equation*}
a_{2}(\lambda \Delta t)^{2}+\left(a_{1}+a_{2}\right) \lambda \Delta t+2 \geq 0 \tag{16}
\end{equation*}
$$

For $\lambda \Delta t=0$, the above inequality is satisfied. The discriminant of the quadratic equation is given by $\left(a_{1}+a_{2}\right)^{2}-8 a_{2}$. From the given assumption it follows that $\left(a_{1}+a_{2}\right)^{2}-8 a_{2}<0$ so the quadratic equation does not have real roots. Hence the left inequality in relation (15) is always satisfied. Next we consider the right hand inequality of relation (15)

$$
\begin{equation*}
a_{2}(\lambda \Delta t)^{2}+\left(a_{1}+a_{2}\right) \lambda \Delta t \leq 0 \tag{17}
\end{equation*}
$$

This relation is rearranged into

$$
\begin{equation*}
a_{2}(\lambda \Delta t)^{2} \leq-\left(a_{1}+a_{2}\right) \lambda \Delta t \tag{18}
\end{equation*}
$$

hence

$$
\begin{equation*}
a_{2}|\lambda \Delta t|^{2} \leq\left(a_{1}+a_{2}\right)|\lambda \Delta t| \Leftrightarrow|\lambda \Delta t| \leq \frac{a_{1}+a_{2}}{a_{2}}, \quad a_{2} \neq 0 \tag{19}
\end{equation*}
$$

This results into the following condition for stability

$$
\begin{equation*}
\Delta t \leq \frac{a_{1}+a_{2}}{a_{2}|\lambda|}, \quad a_{2} \neq 0 \tag{20}
\end{equation*}
$$

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

$$
\begin{gathered}
f_{1}\left(x_{1}, x_{2}\right)=-\sin x_{1}+2 x_{2}+t \\
f_{2}\left(x_{1}, x_{2}\right)=x_{1}-x_{2}^{2}
\end{gathered}
$$

From the definition of the Jacobian it follows that:

$$
\left(\begin{array}{ll}
\frac{\partial f_{1}}{\partial x_{1}} & \frac{\partial f_{1}}{\partial x_{2}} \\
\frac{\partial f_{2}}{\partial x_{1}} & \frac{\partial f_{2}}{\partial x_{2}}
\end{array}\right)=\left(\begin{array}{cc}
-\cos x_{1} & 2 \\
1 & -2 x_{2}
\end{array}\right) .
$$

Substitution of $\binom{x_{1}(0)}{x_{2}(0)}=\binom{0}{1}$ shows that

$$
J=\left(\begin{array}{cc}
-1 & 2 \\
1 & -2
\end{array}\right)
$$

(e) For the stability it is sufficient to check that $\left|Q\left(\lambda_{i} \Delta t\right)\right| \leq 1$ for all the eigenvalues of the Jacobian matrix. It is easy to see that the eigenvalues of the Jacobian matrix are $\lambda_{1}=-3$ and $\lambda_{2}=0$.

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

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

For $\lambda_{2}=0$ it appears that $Q\left(\lambda_{2} \Delta t\right)=1$ so the inequality is satisfied for all $\Delta t$. For $\lambda_{1}=-3$ we have to check the following inequalities:

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

For the left-hand inequality we arrive at

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

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

For the right-hand inequality we get

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

so

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

(another option is to see that for $a_{1}=a_{2}=\frac{1}{2}$ the method is equal to the modified Euler method, and remember that this method is stable for real eigenvalues if $\Delta t \leq \frac{-2}{\lambda}$ )
2. (a) The Taylor polynomials around 0 are given by:

$$
\begin{aligned}
f(0) & =f(0) \\
f(-h) & =f(0)-h f^{\prime}(0)+\frac{h^{2}}{2} f^{\prime \prime}(0)-\frac{h^{3}}{6} f^{\prime \prime \prime}\left(\xi_{1}\right), \\
f(-2 h) & =f(0)-2 h f^{\prime}(0)+2 h^{2} f^{\prime \prime}(0)-\frac{(2 h)^{3}}{6} f^{\prime \prime \prime}\left(\xi_{2}\right) .
\end{aligned}
$$

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

$$
\begin{aligned}
f(0): & \frac{\alpha_{0}}{h^{2}}+\frac{\alpha-1}{h^{2}}+\frac{\alpha-2}{h^{2}} & =0, \\
f^{\prime}(0): & -\frac{h \alpha-1}{h^{2}}-\frac{2 h-2}{h^{2}} & =0, \\
f^{\prime \prime}(0): & \frac{h^{2}}{2 h^{2}} \alpha_{-1}+\frac{2 h^{2} \alpha-2}{h^{2}} & =1 .
\end{aligned}
$$

This can also be written as

$$
\begin{array}{rlrl}
f(0): & \alpha_{0}+\alpha_{-1}+\alpha_{-2} & =0, \\
f^{\prime}(0): & & -\alpha_{-1}-2 \alpha_{-2} & =0, \\
f^{\prime \prime}(0): & \frac{\alpha_{-1}}{2}+2 \alpha_{-2} & =1 .
\end{array}
$$

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

$$
\begin{gathered}
f^{\prime \prime}(0)-Q(h)=f^{\prime \prime}(0)-\frac{f(0)-2 f(-h)+f(-2 h)}{h^{2}}=-\left(\frac{\frac{2 h^{3}}{6} f^{\prime \prime \prime}\left(\xi_{1}\right)-\frac{8 h^{3}}{6} f^{\prime \prime \prime}\left(\xi_{2}\right)}{h^{2}}\right) \\
=h f^{\prime \prime \prime}(\xi) .
\end{gathered}
$$

(c) Note that

$$
\begin{align*}
f^{\prime \prime}(0)-Q(2 h) & =K 2 h  \tag{21}\\
f^{\prime \prime}(0)-Q(h) & =K h \tag{22}
\end{align*}
$$

Subtraction gives:

$$
\begin{equation*}
Q(h)-Q(2 h)=K 2 h-K h=K h . \tag{23}
\end{equation*}
$$

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

$$
f^{\prime \prime}(0)-Q\left(\frac{1}{4}\right)=Q\left(\frac{1}{4}\right)-Q\left(\frac{1}{2}\right)=-1.4992
$$

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

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

Further, we have

$$
\begin{align*}
u^{\prime}(x) & =1+\frac{e^{x / \epsilon}}{\epsilon\left(1-e^{1 / \epsilon}\right)}  \tag{25}\\
u^{\prime \prime}(x) & =\frac{e^{x / \epsilon}}{\epsilon^{2}\left(1-e^{1 / \epsilon}\right)} \tag{26}
\end{align*}
$$

Hence, we immediately see

$$
\begin{equation*}
-\epsilon u^{\prime \prime}(x)+u^{\prime}(x)=-\frac{\epsilon e^{x / \epsilon}}{\epsilon^{2}\left(1-e^{1 / \epsilon}\right)}+1+\frac{e^{x / \epsilon}}{\epsilon\left(1-e^{1 / \epsilon}\right)}=1 . \tag{27}
\end{equation*}
$$

Hence, the solution $u(x)=1-\frac{1-e^{x / \epsilon}}{1-e^{1 / \epsilon}}$ satisfies the differential equation and the Dirichlet boundary conditions, and therewith $u(x)$ is the solution to the boundary value problem (uniqueness can be demonstrated in a straightforward way, but this was not asked for).
(b) The domain of computation, being $(0,1)$, is divided into subintervals with mesh points, we set $x_{j}=j \Delta x$, where we use $n$ unknowns, such that $x_{n+1}=(n+$ 1) $\Delta x=1$. The discretization method for an interior node is given by

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

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

$$
\begin{align*}
& -\epsilon \frac{w_{2}-2 w_{1}+0}{(\Delta x)^{2}}+\frac{w_{1}-0}{\Delta x}=1 \\
& -\epsilon \frac{0-2 w_{n}+w_{n-1}}{(\Delta x)^{2}}+\frac{w_{n}-w_{n-1}}{\Delta x}=1 \tag{29}
\end{align*}
$$

This can be rewritten more neatly as follows:

$$
\begin{align*}
& \epsilon \frac{-w_{2}+2 w_{1}}{(\Delta x)^{2}}+\frac{w_{1}}{\Delta x}=1  \tag{30}\\
& \epsilon \frac{2 w_{n}-w_{n-1}}{(\Delta x)^{2}}+\frac{w_{n}-w_{n-1}}{\Delta x}=1
\end{align*}
$$

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

$$
\left[\begin{array}{ccc}
32 \epsilon+4 & -16 \epsilon & 0  \tag{31}\\
-16 \epsilon-4 & 32 \epsilon+4 & -16 \epsilon \\
0 & -16 \epsilon-4 & 16 \epsilon+4
\end{array}\right]\left[\begin{array}{l}
w_{1} \\
w_{2} \\
w_{3}
\end{array}\right]=\left[\begin{array}{l}
1 \\
1 \\
1
\end{array}\right]
$$

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

