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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU AESB2210)
Thursday April 16 2015, 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

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

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

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

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)

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

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

$$y_1' = y_2, \qquad y_2' = -4y_1 + 2t. \tag{16}$$

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

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 2t \end{pmatrix}, \tag{17}$$

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

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

$$A = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix}, \quad \underline{w}^1 = \begin{pmatrix} w_1^1 \\ w_2^1 \end{pmatrix}, \quad \underline{w}^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{18}$$

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

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

This gives

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

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

$$\begin{pmatrix} 1 & -\frac{1}{4} \\ 1 & 1 \end{pmatrix} \underline{w}^1 = \begin{pmatrix} 1 \\ -\frac{3}{4} \end{pmatrix}. \tag{21}$$

This system is solved by

$$\underline{w}^1 = \begin{pmatrix} 0.65 \\ -1.4 \end{pmatrix} \tag{22}$$

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

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

Subsequently, we perform the corrector step

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

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

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

[e] 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 $\lambda = \pm 2i$. Herewith, the following amplification factors are obtained:

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

This gives the following moduli

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

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. Furthermore the Modified Euler Method is instable.

2. (a) Consider $y(x) = x^2$, then $y'(x) = 2x$ and $y''(x) = 2$, substitution into the differential equation yields

$$-y''(x) + xy = -2 + xx^2 = x^3 - 2, \tag{29}$$

hence $y(x) = x^2$ satisfies the differential equation. Next, we check the boundary conditions: $y'(0) = 2 \cdot 0 = 0$ and $y(1) = 1^2 = 1$ which proves that the boundary conditions are also satisfied. Hence, $y(x) = x^2$ is a solution of the boundary value problem.

(b) Using central differences for the second order derivative at a node $x_j = jh$, gives

$$y''(x_j) \approx \frac{y_{j+1} - 2y_j + y_{j-1}}{h^2} =: Q(h). \tag{30}$$

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

$$\begin{aligned} y_{j+1} &= y(x_j + h) = y(x_j) + hy'(x_j) + \tfrac{h^2}{2}y''(x_j) + \tfrac{h^3}{3!}y'''(x_j) + \tfrac{h^4}{4!}y''''(\eta_+), \\ y_{j-1} &= y(x_j - h) = y(x_j) - hy'(x_j) + \tfrac{h^2}{2}y''(x_j) - \tfrac{h^3}{3!}y'''(x_j) + \tfrac{h^4}{4!}y''''(\eta_-). \end{aligned} \tag{31}$$

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 $|y''(x_j) - Q(h)| = O(h^2)$. This leads to the following discretization formula for internal grid nodes:

$$\frac{-w_{j-1} + 2w_j - w_{j+1}}{h^2} + x_j w_j = x_j^3 - 2. \tag{32}$$

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

$$0 = y'(0) \approx \frac{y_1 - y_{-1}}{2h} =: Q_b(h). \tag{33}$$

Using Taylor's Theorem, gives

$$\begin{aligned} &Q_b(h) = \\ &\frac{y(0) + hy'(0) + \frac{h^2}{2}y''(0) + \frac{h^3}{3!}y'''(\eta_+) - (y(0) - hy'(0) + \frac{h^2}{2}y''(0) - \frac{h^3}{3!}y'''(\eta_-))}{2h} = \\ &y'(0) + O(h^2). \end{aligned} \tag{34}$$

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

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

$$\frac{w_1 - w_{-1}}{2h} = 0 \Leftrightarrow w_{-1} = w_1. \tag{35}$$

The discretization at $x = 0$ is given by

$$\frac{-w_{-1} + 2w_0 - w_1}{h^2} = -2. \tag{36}$$

Substitution of equation (35) into the above equation, yields

$$\frac{2w_0 - 2w_1}{h^2} = -2. \tag{37}$$

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 (32), gives

$$\frac{-w_{n-2} + 2w_{n-1}}{h^2} + x_{n-1}w_{n-1} = x_{n-1}^3 - 2 + \frac{1}{h^2} = (1-h)^3 - 2 + \frac{1}{h^2}. \tag{38}$$

This concludes our discretization of the boundary conditions. In order to get a symmetric discretization matrix, one divides equation (37) by 2.

Next, we use $h = 1/3$, then, from equations (32, 37, 38), one obtains the following system

$$\begin{aligned} 9w_0 - 9w_1 &= -1 \\ -9w_0 + 18\tfrac{1}{3}w_1 - 9w_2 &= -\tfrac{53}{27} \\ -9w_1 + 18\tfrac{2}{3}w_2 &= \tfrac{197}{27}. \end{aligned} \tag{39}$$

(d) The truncation errors from the virtual grid point and internal points contain a third- and fourth order derivative, respectively (see part b). Since the exact solution is given by $y(x) = x^2$, the third and fourth order derivatives are zero. Hence, the error is zero at all grid points. Therefore, the numerical solution is given by $w_0 = y_0 = 0$, $w_1 = y_1 = 1/9$ and $w_2 = y_2 = 4/9$.

Remark: This numerical solution can also be obtained from the solution of system (39).

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

$$I_j^R = hf(x_{j-1}), \qquad h = x_j - x_{j-1}. \tag{40}$$

The composed integration rule is derived by

$$I^R = h(I_1^R + I_2^R + \ldots + I_n^R) = h(f(x_0) + \ldots + f(x_{n-1})), \tag{41}$$

which yields

$$I^R = \frac{1}{3} \cdot (0 + (\frac{1}{3})^2 + (\frac{2}{3})^2) = \frac{5}{27}. \tag{42}$$

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

$$I_j^T = \frac{h}{2}(f(x_{j-1}) + f(x_j)). \tag{43}$$

The composed integration rule is derived by

$$I^T = h(I_1^T + I_2^T + \ldots + I_n^T) = h(\frac{f(x_0)}{2} + f(x_1) + \ldots + f(x_{n-1}) + \frac{f(x_n)}{2}), \tag{44}$$

which leads to

$$I^T = \frac{1}{3} \cdot (0 + (\frac{1}{3})^2 + (\frac{2}{3})^2 + \frac{1}{2}) = \frac{19}{54}. \tag{45}$$

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

$$\begin{aligned} \varepsilon_R &\leq \frac{h}{2} \max_{x \in [0,1]} |y'(x)| \leq h = \frac{1}{n}, \\ \varepsilon_T &\leq \frac{h^2}{12} \max_{x \in [0,1]} |y''(x)| \leq \frac{h^2}{6} = \frac{1}{6n^2}. \end{aligned} \tag{46}$$

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

$$\lim_{n \to \infty} \frac{n+1}{n} = 1, \tag{47}$$

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