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 $K = |\frac{f''(z)}{2 * f'(z)}| = |\frac{12z^2}{2(4z^3-3)}| \approx 1.3867$. □