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

**ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU)**
**Tuesday April 1 2008, 14:00-17:00**

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 $t_n$:

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

From the Chain Rule of Differentiation, we know

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

Hence, one obtains

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

For $z_{n+1}$, we obtain, after substitution of the predictor step for $\bar{z}_{n+1}$ into the corrector step and after a Taylor expansion around $(t_n, y_n)$

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

Subsequently, 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) To compute the amplification factor one uses the test equation $y' = \lambda y$. Applying the Modified Euler method gives:

$$\begin{aligned} \text{predictor:} \quad \overline{w}_{n+1} &= w_n + hf(t_n, w_n), && (7) \\ \text{corrector:} \quad w_{n+1} &= w_n + \frac{h}{2}[f(t_n, w_n) + f(t_{n+1}, \overline{w}_{n+1})]. && (8) \end{aligned}$$

so

$$\begin{aligned} \text{predictor:} \quad \overline{w}_{n+1} &= w_n + h\lambda w_n, && (9) \\ \text{corrector:} \quad w_{n+1} &= w_n + \frac{h}{2}[\lambda w_n + \lambda(w_n + h\lambda w_n)]. && (10) \end{aligned}$$

Summarizing $w_{n+1} = (1 + h\lambda + \frac{1}{2}(h\lambda)^2)w_n$, which leads to the answer $Q(h\lambda) = 1 + h\lambda + \frac{1}{2}(h\lambda)^2$.

(c) Use the transformation:

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

This implies that

$$\begin{aligned} y_1' &= \Phi' = y_2\,, \\ y_2' &= \Phi'' = -\Phi' - \frac{1}{2}\Phi = -y_2 - \frac{1}{2}y_1 = -\frac{1}{2}y_1 - y_2;\,, \end{aligned}$$

So the matrix $A$ is given by $\begin{pmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \end{pmatrix}$.

(d) The eigenvalues of the matrix $A$ are $\lambda_1 = -\frac{1}{2} + \frac{i}{2}$ and $\lambda_2 = -\frac{1}{2} - \frac{i}{2}$. For stability it is needed that $|Q(h\lambda_1)| \leq 1$ and $|Q(h\lambda_2)| \leq 1$. Since $\lambda_2 = \bar{\lambda}_1$ it is sufficient to check the inequality $|Q(h\lambda_1)| \leq 1$. Using $h = 1$ we obtain $Q(h\lambda_1) = 1 + \lambda_1 + \frac{1}{2}\lambda_1^2 = \frac{1}{2} + \frac{i}{4}$. Note that $|Q(h\lambda_1)| = \sqrt{\frac{1}{4} + \frac{1}{16}} = 0.5590 \leq 1$, so the method is stable for $h = 1$.

(e) The Jacobian is defined by:

$$\begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} \end{pmatrix}$$

Using the definition it follows that

$$\begin{pmatrix} 0 & 1 \\ -\cos(y_1) & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\cos\frac{\pi}{4} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{\sqrt{2}}{2} & 0 \end{pmatrix}$$

(f) The eigenvalues of the Jacobian matrix are $\lambda_1 = i\sqrt{\cos(y_1)}$ and $\lambda_2 = -i\sqrt{\cos(y_1)}$ where we use that $\cos(y_1) \geq 0$ if $-\frac{\pi}{2} \leq \Phi \leq \frac{\pi}{2}$. Using the stability region of Modified Euler, it follows that the method is instable for all choices of $h$ (so also for $h = 1$) because $h\lambda_1$ and $h\lambda_2$ are both on the imaginary axis, which lies outside the stability region of Modified Euler.

Another way to show this is to compute $|Q(h\lambda_1)|$. For $h = 1$ this is equal to $|Q(\lambda_1)| = |1 + i\sqrt{\cos(y_1)} - \frac{1}{2}\cos(y_1)|$. Using the definition of the modulus of a complex number it follows that

$$|Q(\lambda_1)| = \sqrt{(1 - \frac{1}{2}\cos(y_1))^2 + \cos(y_1)} = \sqrt{1 + \frac{1}{4}(\cos(y_1))^2} > 1$$

so the method is instable for $h = 1$.

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] The central differences formula, $Q(h)$, gives

$$Q(h) = \frac{x(0.5) - x(0)}{0.5 - 0} = \frac{20 - 4}{0.5} = 32m/s = 115.2km/h. \tag{11}$$

[e] After application of Taylor's Theorem around $x$, one obtains for the truncation error

$$\begin{aligned} Q(h) - f'(x) &= \frac{f(x+h) - f(x-h)}{2h} - f'(x) = \\ &= \frac{f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{3!} f'''(x) + O(h^4)}{2h} + \\ &- \frac{f(x) - hf'(x) + \frac{h^2}{2} f''(x) - \frac{h^3}{3!} f'''(x) + O(h^4)}{2h} - f'(x) = \\ &= \frac{h^2}{3!} f'''(x) + O(h^3) = O(h^2). \end{aligned} \tag{12}$$

Hence, the truncation error has $O(h^2)$.

[f] Let $\hat{Q}(h)$ and $Q(h)$ be the central differences with the exact and measured data respectively, then the magnitude of the error from the measurements is bounded from above by

$$|\hat{Q}(h) - Q(h)| = |\frac{20 \pm \varepsilon - (4 \pm \varepsilon)}{0.5} - \frac{20 - 4}{0.5}| \leq \frac{2\varepsilon}{0.5} = 4\varepsilon\ m/s. \tag{13}$$

[g] The correction is 3 km/h = 0.8333 m/s. Further, the maximum error from the measurements equals the correction. Note that we should be consistent with the units. Hence, one gets

$$4\varepsilon = 0.8333\ m/s \Leftrightarrow \varepsilon = \frac{0.8333}{4} = 0.2083\ m. \tag{14}$$

The error from the measurements, corresponding to this correction, almost equals 21 cm.