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

ANSWERS OF THE TEST NUMERICAL METHODS FOR DIFFERENTIAL EQUATIONS (WI3097 TU / Minor AESB2210)
Thursday April 14th 2016, 18:30-21:30

1. (a) The **local truncation error** is defined by

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

where $y_n := y(t_n)$ represents the exact solution and

$$z_{n+1} = y_n + \Delta t f(t_{n+1}, z_{n+1}), \tag{2}$$

represents the approximation of the numerical solution at t_{n+1} upon using y_n for the previous time step. Since, we use the **test equation** $y' = \lambda y$, we express y_{n+1} in terms of y_n as follows

$$y_{n+1} = y_n e^{\lambda \Delta t} = y_n(1 + \lambda \Delta t + \frac{1}{2}(\lambda \Delta t)^2 + \mathcal{O}(\Delta t^3)). \tag{3}$$

From (2), we use the test equation and the **geometric series**

$$z_{n+1} = \frac{y_n}{1 - \lambda \Delta t} = y_n(1 + \lambda \Delta t + (\lambda \Delta t)^2 + \mathcal{O}(\Delta t^3)). \tag{4}$$

Substitution of equations (3) and (4) into the definition of the **local truncation error**, gives

$$\tau_{n+1}(\Delta t) = \frac{y_n}{\Delta t}\left(-\frac{(\lambda \Delta t)^2}{2} + \mathcal{O}(\Delta t^3)\right) = \mathcal{O}(\Delta t). \tag{5}$$

(b) Using the **test equation**, we get

$$w_{n+1} = w_n + \lambda \Delta t w_{n+1}, \tag{6}$$

where w_n denotes the numerical approximation of y_n. The above equation implies

$$w_{n+1} = \frac{w_n}{1 - \lambda \Delta t} =: Q(\lambda \Delta t) w_n. \tag{7}$$

Here $Q(\lambda \Delta t)$ represents the **amplification factor**. For numerical stability, we require the **modulus of the amplification factor** to satisfy

$$|Q(\lambda \Delta t)| \leq 1, \text{ hence } \left|\frac{1}{1 - \lambda \Delta t}\right| = \frac{1}{|1 - \lambda \Delta t|} \leq 1. \tag{8}$$

From the above equation, it is clear that

$$|1 - \lambda \Delta t| \geq 1, \tag{9}$$

and with $\lambda = \mu + i\nu$, we get

$$(1 - \mu \Delta t)^2 + (\nu \Delta t)^2 \geq 1. \tag{10}$$

This area is the whole **complex plane except the unit circle with center** $(1, 0)$, see Figure 1.

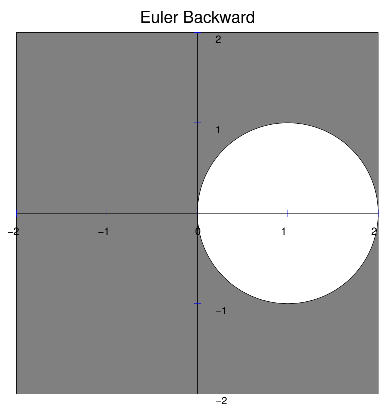

Figure 1: The region of stability of the backward Euler method (grey area).

(c) Consider the equations that we have to solve

$$\begin{aligned} y_1' &= -y_1^2 + 2y_1y_2 =: f_1(y_1, y_2), \\ y_2' &= -y_1y_2 - \tfrac{1}{2}y_2^2 =: f_2(y_1, y_2), \end{aligned} \tag{11}$$

Here, we introduced the functions $f_1(y_1, y_2)$ and $f_2(y_1, y_2)$. Then, the **Jacobian matrix** is given by

$$J(y_1, y_2) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(y_1, y_2) & \frac{\partial f_1}{\partial y_2}(y_1, y_2) \\ \\ \frac{\partial f_2}{\partial y_1}(y_1, y_2) & \frac{\partial f_2}{\partial y_2}(y_1, y_2) \end{pmatrix} = \begin{pmatrix} -2y_1 + 2y_2 & 2y_1 \\ \\ -y_2 & -y_1 - y_2 \end{pmatrix}. \tag{12}$$

For the equilibrium $(y_1(0), y_2(0)) = (1, 0)$, we have

$$J(y_1, y_2) := \begin{pmatrix} -2 & 2 \\ 0 & -1 \end{pmatrix}. \tag{13}$$

Hence the **eigenvalues** are given by $\lambda_1 = -2$ and $\lambda_2 = -1$.

(d) - We have $\lambda_1 = -2$ and $\lambda_2 = -1$, hence with $\Delta t > 0$, this implies that $\Delta t \lambda < 0$ (thus real-valued), then from Figure 1, it is clear that the **backward Euler is stable for any** $\Delta t > 0$.
  [The same argument holds for the backup choice $\lambda_1 = -4$ and $\lambda_2 = -3$.]
- Since the eigenvalues are real-valued and negative, we use

$$\Delta t \leq \frac{2}{|\lambda|}, \tag{14}$$

  as **stability bound for the forward Euler method**. With $\lambda_1 = -2$, and $\lambda_2 = -1$, we get $\Delta t \leq 1$ as the maximum allowable time step to warrant numerical stability, based on linear stability analysis around $(1, 0)$.
  [For the backup choice $\lambda_1 = -4$ and $\lambda_2 = -3$ we get $\Delta t \leq \frac{1}{2}$ as the maximum allowable time step to warrant numerical stability.]

(e) Applying the **forward Euler time integration method** to system (11), gives

$$\begin{aligned} u_{n+1} &= u_n + \Delta t(-u_n^2 + 2u_n v_n), \\ v_{n+1} &= v_n + \Delta t(-u_n v_n - \tfrac{1}{2}v_n^2). \end{aligned} \tag{15}$$

where $\mathbf{w}_n = (u_n, v_n)^\top$ denotes the numerical solution with components u_n and v_n. Using $\Delta t = 1$ and $u_0 = 1$ and $v_0 = 0$, gives

$$\begin{aligned} u_1 &= u_0 + \Delta t(-u_0^2 + 2u_0 v_0) = 1 - 1 = 0, \\ v_1 &= v_0 + \Delta t(-u_0 v_0 - \tfrac{1}{2}v_0^2) = 0. \end{aligned} \tag{16}$$

Hence $u_1 = 0$ and $v_1 = 0$.

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

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

Using $t = 10\,[\text{min}]$, and $h = 5\,[\text{min}]$ the approximation of the velocity is

$$\frac{d(10) - d(5)}{5}\left[\frac{\text{cm}}{\text{min}}\right] = \frac{550 - 250}{5}\left[\frac{\text{cm}}{\text{min}}\right] = 60\left[\frac{\text{cm}}{\text{min}}\right].$$

(b) **Taylor polynomials** are:

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

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

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

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

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

Using the new formula with $h = 5\,[\text{min}]$ we obtain the estimate:

$$\frac{d(0) - 4d(10) + 3d(20)}{20} \left[\frac{\text{cm}}{\text{min}}\right] = \frac{0 - 4 \times 250 + 3 \times 550}{10} \left[\frac{\text{cm}}{\text{min}}\right] = 65 \left[\frac{\text{cm}}{\text{min}}\right].$$

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

$$L_1(x) = \frac{x - x_1}{x_0 - x_1} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1). \tag{17}$$

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

$$\begin{aligned} L_2(x) &= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} f(x_0) && (18) \\ &+ \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} f(x_1) && (19) \\ &+ \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} f(x_2). && (20) \end{aligned}$$

This is also evident from application of the given formula.

(c) Obviously, $L_1(3) = 6$ and $L_2(3) = 6$ since the Lagrange interpolation polynomial satisfies $L_n(x_k) = f(x_k)$ for all points $x_0, x_1, \ldots, x_n$. Next, we compute $L_1(2)$ and $L_2(2)$ for both linear and quadratic Lagrangian interpolation as approximations at $x = 3$. For **linear interpolation**, we have

$$L_1(3) = \frac{2 - 3}{1 - 3} \cdot 3 + \frac{2 - 1}{3 - 1} \cdot 6 = \frac{9}{2}, \tag{21}$$

and for **quadratic interpolation**, one obtains

$$L_2(3) = \frac{(2-3)(2-4)}{(1-3)(1-4)} \cdot 3 + \frac{(2-1)(2-4)}{(3-1)(3-4)} \cdot 6 + \frac{(2-1)(2-3)}{(4-1)(4-3)} \cdot 5 = \frac{16}{3}. \quad (22)$$