

Queens College, CUNY, Department of Computer Science
Numerical Methods
CSCI 361 / 761
Summer 2018
Instructor: Dr. Sateesh Mane

© Sateesh R. Mane 2018

due Friday, August 3, 2018, 11.59 pm

16 Homework lecture 16

- As experience has demonstrated, if you do not understand the above expressions/questions, **THEN ASK**.
- If you do not understand the words/sentences in the lectures, **THEN ASK**.
- Send me an email, explain what you do not understand.
- Do not just keep quiet and produce nonsense in exams.

16.1 Forward and reverse integration: higher order methods

- The following differential equation is a special case of **Riccati's equation**

$$\frac{dy}{dx} = x - y^2. \quad (16.1.1)$$

- We integrate eq. (16.1.1) starting at $x_0 = 0$, with the initial condition $y_0 = 0$.
- Suppose we integrate eq. (16.1.1) forwards to $x = 1$ and then backwards to $x = 0$.
State the value of the exact solution $y(x)$ after integrating back to $x = 0$.
- In this question we shall numerically integrate eq. (16.1.1) forward for n steps to $x_n = 1$, and then backwards for n steps to $x_{2n} = 0$. Employ a uniform stepsize for the integration.
 - Use (i) explicit Euler, (ii) midpoint, (iii) trapezoid, and (iv) Runge–Kutta fourth order RK4.**
 - Compute the value of the numerical solution y_n at $x_n = 1$ and fill the first table of values below.**
 - In all cases, compute the value of y_n to 3 decimal places.**

n	y_n^{Euler}	y_n^{midpoint}	$y_n^{\text{trapezoid}}$	y_n^{RK4}
10	3 d.p.	3 d.p.	3 d.p.	3 d.p.
100	3 d.p.	3 d.p.	3 d.p.	3 d.p.
1000	3 d.p.	3 d.p.	3 d.p.	3 d.p.

- Compute the value of the numerical solution y_{2n} at $x_{2n} = 0$ and fill the table of values below.**
- Tabulate the values of (i) ny_{2n}^{Euler} , (ii) $n^3y_{2n}^{\text{midpoint}}$, (iii) $n^3y_{2n}^{\text{trapezoid}}$ and (iv) $n^5y_{2n}^{\text{RK4}}$.**
- *** note the powers of n in each case. *****
- In all cases, state your answers to 3 decimal places.**

$2n$	ny_{2n}^{Euler}	$n^3y_{2n}^{\text{midpoint}}$	$n^3y_{2n}^{\text{trapezoid}}$	$n^5y_{2n}^{\text{RK4}}$
20	3 d.p.	3 d.p.	3 d.p.	3 d.p.
200	3 d.p.	3 d.p.	3 d.p.	3 d.p.
2000	3 d.p.	3 d.p.	3 d.p.	3 d.p.

- If you have done your work correctly, the results in each column should be approximately independent of n .
- Can you explain why?* You don't have to. See next page for an explanation.

- The explicit Euler method is first order, hence its local truncation error is $O(h^2) = O(1/n^2)$.
 1. The numerical errors add up (accumulate in magnitude) on both the forward and reverse integrations.
 2. Hence the overall numerical error after $2n$ integration steps is $O(nh^2) = O(1/n)$.
- The midpoint and trapezoid methods are second order, hence the local truncation error in both cases is $O(h^3) = O(1/n^3)$.
 1. However notice that the overall error in the value of y_{2n} is also $O(h^3) = O(1/n^3)$.
 2. The overall error has NOT grown to $O(nh^3) = O(1/n^2)$.
 3. This is because both integrators have some symmetry, if the direction of integration is reversed, and the numerical errors on the reverse integration (partially) cancel the errors accumulated on the forward integration.
 4. Runge–Kutta RK4 also displays the same (good) cancellation property. The local truncation error is $O(h^5) = O(1/n^5)$ and the overall error in the value of y_{2n} is also $O(h^5) = O(1/n^5)$.
- This is a good feature for an iterator to have.
 1. note that the forward–backward symmetry is not exact, for the midpoint, trapezoid and RK4 methods.
 2. It is an approximate symmetry.
 3. Hence the cancellation is partial, not complete.
 4. nevertheless, it is a good feature.
 5. There are other integrators where the forward–backward symmetry is exact.
 6. Typically, for such integrators, the differential equations must have a special structure.

16.2 Area enclosed by curve

- Let $\alpha > 0$ and $\beta > 0$ be positive constant real numbers.
- You are given the following differential equation:

$$\frac{dA}{dx} = 4(1 - x^\alpha)^{1/\beta}. \quad (16.2.1)$$

- The initial condition is $A = 0$ at $x = 0$.
- **Show that the value of $A(1)$ is given by the following integral:**

$$A(1) = 4 \int_0^1 (1 - x^\alpha)^{1/\beta} dx. \quad (16.2.2)$$

- **Show that the value of $A(1)$ is the area enclosed by the following closed curve:**

$$|x|^\alpha + |y|^\beta = 1. \quad (16.2.3)$$

- Set $\alpha = 0.7$ and $\beta = 1.2$.
- **Integrate eq. (16.2.1) using n steps from $x_0 = 0$ to $x_n = 1$.**
 1. **Employ a uniform stepsize $h = 1/n$.**
 2. **Use (i) explicit Euler, (ii) midpoint, (iii) trapezoid, and (iv) Runge–Kutta fourth order RK4.**
 3. **Define A_n as the numerical value of $A(1)$.**
 4. **Compute the value of the numerical solution A_n at $x_n = 1$ and fill the table of values below.**
 5. **In all cases, compute the values of A_n to 4 decimal places.**

n	A_n^{Euler}	A_n^{midpoint}	$A_n^{\text{trapezoid}}$	A_n^{RK4}
10	4 d.p.	4 d.p.	4 d.p.	4 d.p.
100	4 d.p.	4 d.p.	4 d.p.	4 d.p.
1000	4 d.p.	4 d.p.	4 d.p.	4 d.p.

16.3 Auxiliary variables

- The **pendulum equation** can be written in the form

$$\frac{d^2\theta}{dt^2} + k \sin \theta = 0. \quad (16.3.1)$$

- The independent variable is the time t and $k > 0$ is a real positive constant.
- Introduce an auxiliary variable v , and write

$$\frac{d\theta}{dt} = v, \quad (16.3.2a)$$

$$\frac{dv}{dt} = -k \sin \theta. \quad (16.3.2b)$$

- Define a parameter E as follows

$$E = \frac{v^2}{2} + k(1 - \cos \theta). \quad (16.3.3)$$

- **Use eqs. (16.3.2a) and (16.3.2b) to show the following:**

$$\frac{dE}{dt} = 0. \quad (16.3.4)$$

- We call E an **invariant**. Its value does not change along a trajectory in the (θ, v) space.
- In fact, E is the energy of the pendulum.
- Define a two-component vector \mathbf{y} via

$$\mathbf{y} = (\theta, v). \quad (16.3.5)$$

- Define a two-component vector $\mathbf{f} = (f_1, f_2)$ via

$$f_1(t, \theta, v) = v, \quad f_2(t, \theta, v) = -k \sin \theta. \quad (16.3.6)$$

- The formal equation is $d\mathbf{y}/dt = \mathbf{f}$. **Use eqs. (16.3.2a) and (16.3.2b) to show the following:**

$$\frac{d\mathbf{y}}{dt} = \frac{d}{dt} \begin{pmatrix} \theta \\ v \end{pmatrix} = \begin{pmatrix} f_1(t, \theta, v) \\ f_2(t, \theta, v) \end{pmatrix} = \mathbf{f}. \quad (16.3.7)$$

- The initial conditions are $\theta = 0$ and $v = \sqrt{2}$ at $t = t_0 = 0$.
- Set $k = 1$.
- Calculate the value of E in eq. (16.3.3) for these initial conditions.
- Integrate eq. (16.3.7) forward in time using a stepsize of $h = 0.01$.
 1. Use (i) explicit Euler, (ii) midpoint, (iii) trapezoid, and (iv) Runge–Kutta fourth order RK4.
 2. The C++ function code is shown on the next page.
 3. For each method, compute the numerical value of the following quantity:

$$D_i = \frac{v_i^2}{2} + k(1 - \cos \theta_i) - E. \quad (16.3.8)$$

4. Fill the table below with the values of (i) D_i^{Euler} , (ii) $D_i^{\text{midpoint}} \times 10^5$, (iii) $D_i^{\text{trapezoid}} \times 10^5$ and (iv) $D_i^{\text{RK4}} \times 10^{10}$.
5. *** Note the powers of 10. ***
6. In all cases, state the result to 2 decimal places.

i	D_i^{Euler}	$D_i^{\text{midpoint}} \times 10^5$	$D_i^{\text{trapezoid}} \times 10^5$	$D_i^{\text{RK4}} \times 10^{10}$
0	2 d.p.	2 d.p.	2 d.p.	2 d.p.
10	2 d.p.	2 d.p.	2 d.p.	2 d.p.
100	2 d.p.	2 d.p.	2 d.p.	2 d.p.
1000	2 d.p.	2 d.p.	2 d.p.	2 d.p.

- (Optional)
 1. For $n = 1000$, plot (separate) graphs of D_i against θ_i , $i = 0, 1, \dots, n$, computed using (i) explicit Euler, (ii) midpoint, (iii) trapezoid, and (iv) Runge–Kutta fourth order RK4.
 2. For $n = 1000$, plot (separate) graphs of the solution in the (θ, v) plane, i.e. plot the points (θ_i, v_i) , $i = 0, 1, \dots, n$, computed using (i) explicit Euler, (ii) midpoint, (iii) trapezoid, and (iv) Runge–Kutta fourth order RK4.
- Let us analyze the details of various integration schemes.
- See next page(s).

16.3.1 C++ function code

All the integration schemes call the following C++ function.
(It says “ x ” instead of “ t ”):

```
int f(int m, double x, const std::vector<double> & y, std::vector<double> & g)
{
    // value of k is hard-wired for simplicity
    const double k = 1.0;

    // first component "f1(t,theta,v) = v"
    g[0] = y[1];

    // second component "f2(t,theta,v) = -k*sin(theta)"
    g[1] = -k*sin(y[0]);

    return 0;
}
```

All the integration schemes can be called with the following inputs:

```
// m = 2
// x = t_i
// h = step size
// vector (array) y_in  = (theta, v)_i
// vector (array) y_out = (theta, v)_{i+1}

int Euler_explicit(int m, double x, double h,
                  std::vector<double> & y_in,
                  std::vector<double> & y_out);

int midpoint(int m, double x, double h,
            std::vector<double> & y_in,
            std::vector<double> & y_out);

int trapezoid(int m, double x, double h,
             std::vector<double> & y_in,
             std::vector<double> & y_out);

int RK4(int m, double x, double h,
       std::vector<double> & y_in,
       std::vector<double> & y_out);
```

16.3.2 Explicit Euler method

- In this scheme, we have

$$\mathbf{y}_{i+1} = \mathbf{y}_i + h \mathbf{f}(t_i, \mathbf{y}_i). \quad (16.3.1)$$

- In matrix notation this means

$$\begin{pmatrix} \theta \\ v \end{pmatrix}_{i+1} = \begin{pmatrix} \theta \\ v \end{pmatrix}_i + h \begin{pmatrix} f_1(t_i, \theta_i, v_i) \\ f_2(t_i, \theta_i, v_i) \end{pmatrix}. \quad (16.3.2)$$

- In vector component notation this means

$$\theta_{i+1} = \theta_i + h f_1(t_i, \theta_i, v_i) = \theta_i + h v_i, \quad (16.3.3a)$$

$$v_{i+1} = v_i + h f_2(t_i, \theta_i, v_i) = v_i - h k \sin \theta_i. \quad (16.3.3b)$$

- First step $i = 0$.

1. Then $t_0 = 0$, $\theta_0 = 0$ and $v_0 = \sqrt{2}$.

2. Hence

$$\theta_1 = \theta_0 + h f_1(t_0, \theta_0, v_0) = \theta_0 + h v_0 = \sqrt{2} h, \quad (16.3.4a)$$

$$v_1 = v_0 + h f_2(t_0, \theta_0, v_0) = v_0 - h k \sin \theta_0 = \sqrt{2}. \quad (16.3.4b)$$

- Second step $i = 1$.

1. Then $t_1 = h$, $\theta_1 = \sqrt{2} h$ and $v_1 = \sqrt{2}$.

2. Hence

$$\theta_2 = \theta_1 + h f_1(t_1, \theta_1, v_1) = \theta_1 + h v_1 = 2\sqrt{2} h, \quad (16.3.5a)$$

$$v_2 = v_1 + h f_2(t_1, \theta_1, v_1) = v_1 - h k \sin \theta_1 = \sqrt{2} - h k \sin(\sqrt{2} h). \quad (16.3.5b)$$

16.3.3 Midpoint method

- In this scheme, we have

$$\mathbf{g}_1 = \mathbf{f}(t_i, \mathbf{y}_i), \quad (16.3.1a)$$

$$\mathbf{g}_2 = \mathbf{f}(t_i + \frac{1}{2}h, \mathbf{y}_i + \frac{1}{2}h\mathbf{g}_1), \quad (16.3.1b)$$

$$\mathbf{y}_{i+1} = \mathbf{y}_i + h\mathbf{g}_2. \quad (16.3.1c)$$

- In matrix notation this means

$$\begin{pmatrix} g_\theta \\ g_v \end{pmatrix}_1 = \begin{pmatrix} f_1(t_i, \theta_i, v_i) \\ f_2(t_i, \theta_i, v_i) \end{pmatrix} = \begin{pmatrix} v_i \\ -k \sin \theta_i \end{pmatrix}. \quad (16.3.2)$$

- Next define

$$\mathbf{y}_i + \frac{1}{2}h\mathbf{g}_1 = \begin{pmatrix} \theta_{\text{tmp}} \\ v_{\text{tmp}} \end{pmatrix} = \begin{pmatrix} \theta_i \\ v_i \end{pmatrix} + \frac{h}{2} \begin{pmatrix} g_\theta \\ g_v \end{pmatrix}_1 = \begin{pmatrix} \theta_i + \frac{1}{2}hv_i \\ v_i - \frac{1}{2}hk \sin \theta_i \end{pmatrix}. \quad (16.3.3)$$

- Next

$$\begin{pmatrix} g_\theta \\ g_v \end{pmatrix}_2 = \begin{pmatrix} f_1(t_i + \frac{1}{2}h, \theta_{\text{tmp}}, v_{\text{tmp}}) \\ f_2(t_i + \frac{1}{2}h, \theta_{\text{tmp}}, v_{\text{tmp}}) \end{pmatrix} = \begin{pmatrix} v_{\text{tmp}} \\ -k \sin(\theta_{\text{tmp}}) \end{pmatrix}. \quad (16.3.4)$$

- First step $i = 0$.

1. Then $t_0 = 0$, $\theta_0 = 0$ and $v_0 = \sqrt{2}$.

2. Hence

$$(g_\theta)_1 = f_1(t_0, \theta_0, v_0) = v_0 = \sqrt{2}, \quad (16.3.5a)$$

$$(g_v)_1 = f_2(t_0, \theta_0, v_0) = -k \sin \theta_0 = 0. \quad (16.3.5b)$$

3. Then

$$\theta_{\text{tmp}} = \theta_0 + \frac{1}{2}h(g_\theta)_1 = h/\sqrt{2}, \quad (16.3.6a)$$

$$v_{\text{tmp}} = v_0 + \frac{1}{2}h(g_v)_1 = \sqrt{2}. \quad (16.3.6b)$$

4. Next

$$(g_\theta)_2 = f_1(t_0 + \frac{1}{2}h, \theta_{\text{tmp}}, v_{\text{tmp}}) = v_{\text{tmp}} = \sqrt{2}, \quad (16.3.7a)$$

$$(g_v)_2 = f_2(t_0 + \frac{1}{2}h, \theta_{\text{tmp}}, v_{\text{tmp}}) = -k \sin(\theta_{\text{tmp}}) = -k \sin(h/\sqrt{2}). \quad (16.3.7b)$$

5. Then

$$\theta_1 = \theta_0 + h(g_\theta)_2 = \sqrt{2}h, \quad (16.3.8a)$$

$$v_1 = v_0 + h(g_v)_2 = \sqrt{2} - hk \sin(h/\sqrt{2}). \quad (16.3.8b)$$

- Second step $i = 1$.

1. Yuck.

16.3.4 Trapezoid method

- In this scheme, we have

$$\mathbf{g}_1 = \mathbf{f}(t_i, \mathbf{y}_i), \quad (16.3.1a)$$

$$\mathbf{g}_2 = \mathbf{f}(t_i + h, \mathbf{y}_i + h\mathbf{g}_1), \quad (16.3.1b)$$

$$\mathbf{y}_{i+1} = \mathbf{y}_i + \frac{h}{2}(\mathbf{g}_1 + \mathbf{g}_2). \quad (16.3.1c)$$

- In matrix notation this means

$$\begin{pmatrix} g_\theta \\ g_v \end{pmatrix}_1 = \begin{pmatrix} f_1(t_i, \theta_i, v_i) \\ f_2(t_i, \theta_i, v_i) \end{pmatrix} = \begin{pmatrix} v_i \\ -k \sin \theta_i \end{pmatrix}. \quad (16.3.2)$$

- Next define

$$\mathbf{y}_i + h\mathbf{g}_1 = \begin{pmatrix} \theta_{\text{tmp}} \\ v_{\text{tmp}} \end{pmatrix} = \begin{pmatrix} \theta_i \\ v_i \end{pmatrix} + h \begin{pmatrix} g_\theta \\ g_v \end{pmatrix}_1 = \begin{pmatrix} \theta_i + hv_i \\ v_i - hk \sin \theta_i \end{pmatrix}. \quad (16.3.3)$$

- Next

$$\begin{pmatrix} g_\theta \\ g_v \end{pmatrix}_2 = \begin{pmatrix} f_1(t_i + h, \theta_{\text{tmp}}, v_{\text{tmp}}) \\ f_2(t_i + h, \theta_{\text{tmp}}, v_{\text{tmp}}) \end{pmatrix} = \begin{pmatrix} v_{\text{tmp}} \\ -k \sin(\theta_{\text{tmp}}) \end{pmatrix}. \quad (16.3.4)$$

- First step $i = 0$.

1. Then $t_0 = 0$, $\theta_0 = 0$ and $v_0 = \sqrt{2}$.

2. Hence

$$(g_\theta)_1 = f_1(t_0, \theta_0, v_0) = v_0 = \sqrt{2}, \quad (16.3.5a)$$

$$(g_v)_1 = f_2(t_0, \theta_0, v_0) = -k \sin \theta_0 = 0. \quad (16.3.5b)$$

3. Then

$$\theta_{\text{tmp}} = \theta_0 + h(g_\theta)_1 = \sqrt{2}h, \quad (16.3.6a)$$

$$v_{\text{tmp}} = v_0 + h(g_v)_1 = \sqrt{2}. \quad (16.3.6b)$$

4. Next

$$(g_\theta)_2 = f_1(t_0 + h, \theta_{\text{tmp}}, v_{\text{tmp}}) = v_{\text{tmp}} = \sqrt{2}, \quad (16.3.7a)$$

$$(g_v)_2 = f_2(t_0 + h, \theta_{\text{tmp}}, v_{\text{tmp}}) = -k \sin(\theta_{\text{tmp}}) = -k \sin(\sqrt{2}h). \quad (16.3.7b)$$

5. Then

$$\theta_1 = \theta_0 + \frac{h}{2}[(g_\theta)_1 + (g_\theta)_2] = \sqrt{2}h, \quad (16.3.8a)$$

$$v_1 = v_0 + \frac{h}{2}[(g_v)_1 + (g_v)_2] = \sqrt{2} - \frac{hk}{2} \sin(\sqrt{2}h). \quad (16.3.8b)$$

- Second step $i = 1$.

1. Yuck.

16.3.5 Runge–Kutta RK4

- In this scheme, we have

$$\mathbf{g}_1 = \mathbf{f}(t_i, \mathbf{y}_i), \quad (16.3.1a)$$

$$\mathbf{g}_2 = \mathbf{f}(t_i + \frac{1}{2}h, \mathbf{y}_i + \frac{1}{2}h\mathbf{g}_1), \quad (16.3.1b)$$

$$\mathbf{g}_3 = \mathbf{f}(t_i + \frac{1}{2}h, \mathbf{y}_i + \frac{1}{2}h\mathbf{g}_2), \quad (16.3.1c)$$

$$\mathbf{g}_4 = \mathbf{f}(t_i + h, \mathbf{y}_i + h\mathbf{g}_3), \quad (16.3.1d)$$

$$\mathbf{y}_{i+1} = \mathbf{y}_i + \frac{h}{6}(\mathbf{g}_1 + 2\mathbf{g}_2 + 2\mathbf{g}_3 + \mathbf{g}_4). \quad (16.3.1e)$$

- First step $i = 0$.

1. Then $t_0 = 0$, $\theta_0 = 0$ and $v_0 = \sqrt{2}$.

2. Hence

$$(g_\theta)_1 = f_1(t_0, \theta_0, v_0) = v_0 = \sqrt{2}, \quad (16.3.2a)$$

$$(g_v)_1 = f_2(t_0, \theta_0, v_0) = -k \sin \theta_0 = 0. \quad (16.3.2b)$$

3. Then

$$\theta_{\text{tmp}} = \theta_0 + \frac{1}{2}h (g_\theta)_1 = h/\sqrt{2}, \quad (16.3.3a)$$

$$v_{\text{tmp}} = v_0 + \frac{1}{2}h (g_v)_1 = \sqrt{2}. \quad (16.3.3b)$$

4. Next

$$(g_\theta)_2 = f_1(t_0 + \frac{1}{2}h, \theta_{\text{tmp}}, v_{\text{tmp}}) = v_{\text{tmp}} = \sqrt{2}, \quad (16.3.4a)$$

$$(g_v)_2 = f_2(t_0 + \frac{1}{2}h, \theta_{\text{tmp}}, v_{\text{tmp}}) = -k \sin(\theta_{\text{tmp}}) = -k \sin(h/\sqrt{2}). \quad (16.3.4b)$$

5. Then define

$$\tilde{\theta}_{\text{tmp}} = \theta_0 + \frac{1}{2}h (g_\theta)_2 = h/\sqrt{2}, \quad (16.3.5a)$$

$$\tilde{v}_{\text{tmp}} = v_0 + \frac{1}{2}h (g_v)_2 = \sqrt{2} - \frac{hk}{2} \sin(h/\sqrt{2}). \quad (16.3.5b)$$

6. Next

$$(g_\theta)_3 = f_1(t_0 + \frac{1}{2}h, \tilde{\theta}_{\text{tmp}}, \tilde{v}_{\text{tmp}}) = \tilde{v}_{\text{tmp}} = \sqrt{2} - \frac{hk}{2} \sin(h/\sqrt{2}), \quad (16.3.6a)$$

$$(g_v)_3 = f_2(t_0 + \frac{1}{2}h, \tilde{\theta}_{\text{tmp}}, \tilde{v}_{\text{tmp}}) = -k \sin(\tilde{\theta}_{\text{tmp}}) = -k \sin(h/\sqrt{2}). \quad (16.3.6b)$$

7. Then define

$$\hat{\theta}_{\text{tmp}} = \theta_0 + h (g_\theta)_3 = \sqrt{2}h - \frac{h^2k}{2} \sin(h/\sqrt{2}), \quad (16.3.7a)$$

$$\hat{v}_{\text{tmp}} = v_0 + h (g_v)_3 = \sqrt{2} - hk \sin(h/\sqrt{2}). \quad (16.3.7b)$$

8. Next

$$(g_\theta)_4 = f_1(t_0 + h, \hat{\theta}_{\text{tmp}}, \hat{v}_{\text{tmp}}) = \hat{v}_{\text{tmp}} = \sqrt{2} - hk \sin(h/\sqrt{2}), \quad (16.3.8a)$$

$$(g_v)_4 = f_2(t_0 + h, \hat{\theta}_{\text{tmp}}, \hat{v}_{\text{tmp}}) = -k \sin(\hat{\theta}_{\text{tmp}}) = -k \sin(\sqrt{2}h - \frac{h^2k}{2} \sin(h/\sqrt{2})). \quad (16.3.8b)$$

9. Then

$$\theta_1 = \theta_0 + \frac{h}{6} [(g_\theta)_1 + 2(g_\theta)_2 + 2(g_\theta)_3 + (g_\theta)_4], \quad (16.3.9a)$$

$$v_1 = v_0 + \frac{h}{6} [(g_v)_1 + 2(g_v)_2 + 2(g_v)_3 + (g_v)_4]. \quad (16.3.9b)$$

- Second step $i = 1$.

1. Oh dear.