

Numerical Solutions for DEs HW2

YANG, Ze (5131209043)

March 22, 2017

A Note to TA:

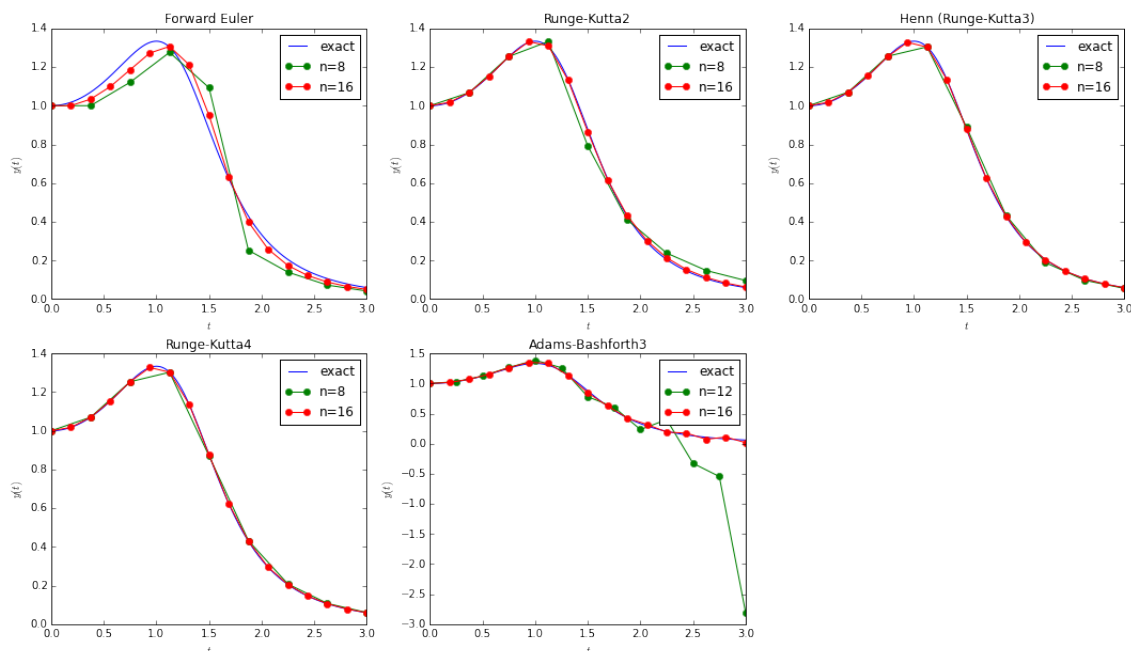
Hi, this is the senior student from Antai College who did not register for this course. I would like to do all the assignments for practice, but feel free to just skip my homework if you don't have time. Thank you again for allowing me to access the assignments and other class material! :)
- Ze

- Problem 1.** a. Implement *RK2*, *Heun's Method* and the classical *RK4*, and justify the rate of convergence numerically.
b. Choose appropriate Runge-Kutta method to initialize *Adams-Bashforth* method of order 3,

$$y_{n+3} = y_{n+2} + h[\frac{23}{12}f(t_{n+2}, y_{n+2}) - \frac{4}{3}f(t_{n+1}, y_{n+1}) + \frac{5}{12}f(t_n, y_n)]$$

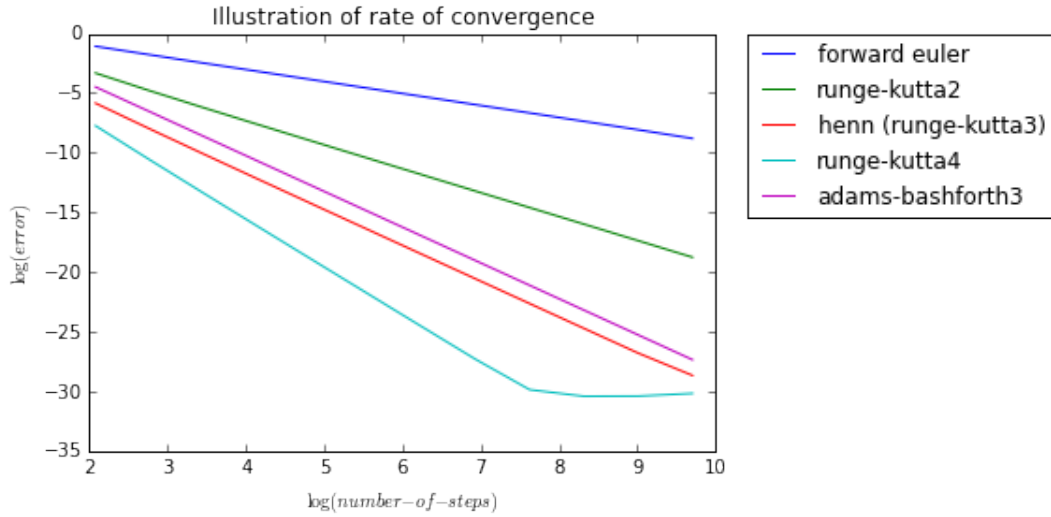
Justify the rate of convergence numerically.

Solution. We solve the test case 6: $y' = t^2 - y$ with all the five methods above.



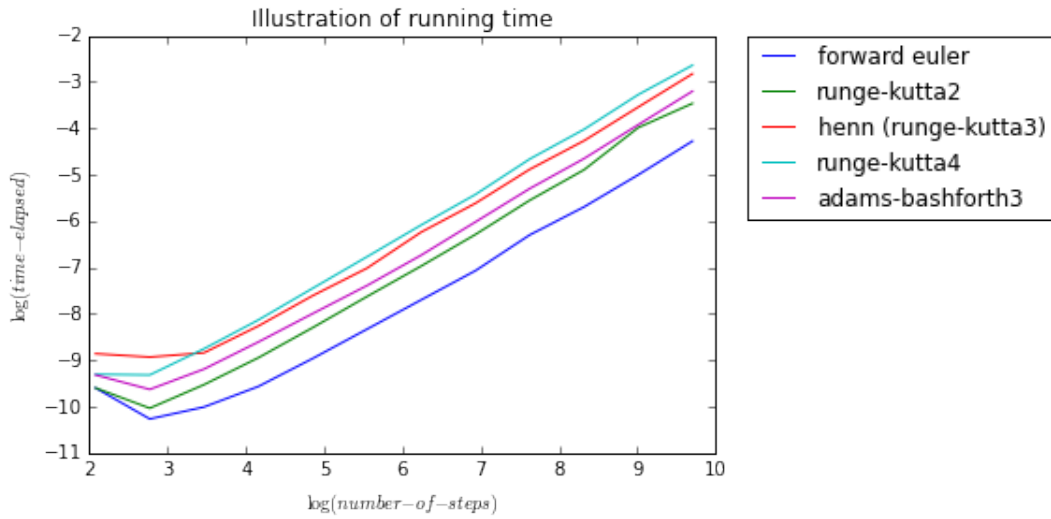
Where the *Adams-Bashforth* multistep method is initialized with a 3-order method (Heun's method).

We plot $\log(n)$ against $\log(\text{error})$ to check the rate of convergence. Theoretically, for a method of order k , we expect to get a line with slope $-k$.



From the figure above we know that it is indeed the case. The curve for 4-th order Runge-Kutta method becomes flat in the end because the round-off error will dominate the error of the algorithm for n sufficiently large. And from the figure we know that the round-off error may be $\sim 2^{-30}$.

We also investigate the running time. The slope of these curves are 1, because all the algorithms take $O(n)$ time, regardless of their orders.



Please also see the attached code for details.

Problem 2. Consider the RK method (I just use the notations in the class)

$$\frac{y_{n+1} - y_n}{h} = c_1 k_1 + c_2 k_2$$

where $k_1 = f(t_n, y_n)$, $k_2 = f(t_n + b_{21}h, y_n + b_{21}hk_1)$, and $c_1, c_2, b_{21} \in \mathbb{R}$.

a. Show that there is a choice of these parameters such that the truncation error of the method

$$\frac{T_n}{\Delta t} = \frac{y(t_{n+1}) - y(t_n)}{h} - c_1 f(t_n, y(t_n)) - c_2 f(t_n + b_{21}h, y(t_n) + b_{21}hf(t_n, y(t_n)))$$

is of order 2 as $h \rightarrow 0$.

- b. Suppose that a second order method of the above form is applied to the IVP $y' = -\lambda y$, $y(0) = 1$, where λ is *real positive* number. Show that the sequence $\{y_n\}_{n \geq 0}$ is bounded $\iff h \leq 2/\lambda$ (like the interval of absolute stability). Show further that, for such λ : $|y(t_n) - y_n| \leq \frac{1}{6}\lambda^3 h^2 t_n$, $n \geq 0$.

Proof. (a) We have (by f we mean $f(t_n, y(t_n))$, and by y we mean $y(t_n)$):

$$\begin{aligned} y' &= f \\ y'' &= f_t + f_y y' = f_t + f_y f \\ y''' &= f_{tt} + f_{ty} y' + y'' f_y + (f_y)' y' = f_{tt} + 2f_{ty} f + f^2 f_{yy} + f_y f_t + f_y^2 f \end{aligned} \quad (1)$$

Therefore use taylor expansion of $y(t_{n+1})$ at t_n :

$$\frac{y(t_{n+1}) - y(t_n)}{h} = f + \frac{h}{2}(f_t + f_y f) + \frac{h^2}{6}(f_{tt} + 2f_{ty} f + f^2 f_{yy} + f_y f_t + f_y^2 f) \quad (2)$$

And that

$$\begin{aligned} b_1 k_1 + b_2 k_2 &= b_1 f + b_2 f(t_n + c_2 h, y(t_n) + h a_{21} f(t_n, y(t_n))) \\ &= b_1 f + b_2 (f + f_t c_2 h + f_y h a_{21} f + \frac{1}{2}(f_{tt} c_2^2 h^2 + 2f_{ty} c_2 h \cdot a_{21} h f + f_{yy} h^2 a_{21}^2 f^2 + O(h^3))) \\ &= (b_1 + b_2) f + h(b_2 c_2 f_t + b_2 a_{21} f_y f) + \frac{h^2}{2}(c_2^2 f_{tt} + 2c_2 a_{21} f_{ty} f + a_{21}^2 f_{yy} f^2) + O(h^3) \end{aligned} \quad (3)$$

So we subtract $b_1 k_1 + b_2 k_2$ from $\frac{y(t_{n+1}) - y(t_n)}{h}$. To achieve a truncation error of order 2, we have to let $O(1)$ and $O(h)$ terms equal to 0:

$$\begin{aligned} O(1) \text{ terms : } & f - (b_1 + b_2) f = 0 \\ O(h) \text{ terms : } & \frac{h}{2}(f_t + f_y f) - h(b_2 c_2 f_t + b_2 a_{21} f_y f) = 0 \\ \Rightarrow & b_1 + b_2 = 1; \quad b_2 c_2 = \frac{1}{2}; \quad b_2 a_{21} = \frac{1}{2} \end{aligned} \quad (4)$$

Clearly, there exists such parameters. For example, $b_1 = b_2 = \frac{1}{2}$, $a_{21} = c_2 = 1$. \square

- (b.) Apply RK2 to $y' = -\lambda y$: $k_1 = -\lambda y_n$. $k_2 = -\lambda(y_n + a_{21} h k_1)$. So let $z := \lambda h$

$$y_{n+1} = y_n + h(-b_1 \lambda y_n - b_2 \lambda y_n + b_2 a_{21} h \lambda^2 y_n) = y_n(1 - b_1 z - b_2 z + b_2 a_{21} z^2) = y_n(1 - z + \frac{1}{2} z^2)$$

Therefore $y_n = (1 - z + \frac{1}{2} z^2)^n y_0$ is bounded $\Rightarrow |1 - z + \frac{1}{2} z^2| \leq 1$. The solution is $0 \leq z \leq 2$, i.e. $h \leq 2/\lambda$. The exact solution $y(t_n) = e^{-\lambda t} y(0) = e^{-nz}$. Hence

$$y(t_n) - y_n = e^{-nz} - (1 - z + \frac{1}{2} z^2)^n$$

Use taylor expansion to e^{-nz} , and trinomial expansion to the second term:

$$\begin{aligned} e^{-nz} &= 1 - nz + \frac{n^2 z^2}{2} - \frac{n^3 z^3}{6} + \frac{n^4 z^4}{24} + h.o.t. \\ (1 - z + \frac{1}{2} z^2)^n &= 1 + n(-z) + n(\frac{z^2}{2}) + \frac{n(n-1)}{2!}(-z)^2 + \frac{n(n-1)}{1!1!}(-z)(\frac{z^2}{2}) + \frac{n(n-1)(n-2)}{3!}(-z)^3 \\ &\quad + \frac{n(n-1)(n-2)(n-3)}{4!}(-z)^4 + \frac{n(n-1)}{2!}(\frac{z^2}{2})^2 + \frac{n(n-1)(n-2)}{2!1!}(-z)^2(\frac{z^2}{2}) + h.o.t. \\ &= 1 - nz + \frac{(n^2 - n + n)}{2} z^2 - \frac{[(n^3 - 3n^2 + 2n) + (3n^2 - 3n)]}{6} z^3 \\ &\quad + \frac{[(n^4 - 6n^3 + 11n^2 - 6n) + (3n^2 - 3n) + (6n^3 - 18n^2 + 12n)]}{24} z^4 + h.o.t \\ &= 1 - nz + \frac{n^2}{2} z^2 - \frac{n^3 - n}{6} z^3 + \frac{n^4 - 4n^2 + 3n}{24} z^4 + h.o.t. \end{aligned} \quad (5)$$

Hence

$$\begin{aligned} y(t_n) - y_n &= -\frac{n}{6} z^3 + \frac{4n^2 - 3n}{24} z^4 + h.o.t. \\ &= -\frac{1}{6} \lambda^3 h^2 t_n + \frac{4n^2 - 3n}{24} z^4 + h.o.t. \end{aligned} \quad (6)$$

For $n \in \mathbb{N}^+ \cup \{0\}$, we have $4n^2 - 3n \geq 0$. Therefore

$$|y(t_n) - y_n| = \left| -\frac{1}{6} \lambda^3 h^2 t_n + \frac{4n^2 - 3n}{24} z^4 + h.o.t. \right| \leq \left| \frac{1}{6} \lambda^3 h^2 t_n \right|$$

As $h \rightarrow 0$, $z = \lambda h \rightarrow 0$. \square

Problem 3. Find the general solution $\{y_n\}$ for the homogeneous recurrence relation

$$a_0 y_n + a_1 y_{n+1} + \dots + a_s y_{n+s} = 0$$

Proof. WLOG we assume the recurrence is of order s , i.e. $a_s \neq 0$. Then rearrange the terms as:

$$y_{n+s} + \frac{a_{s-1}}{a_s} y_{n+s-1} + \dots + \frac{a_1}{a_s} y_{n+1} + \frac{a_0}{a_s} y_n = y_{n+s} + c_{s-1} y_{n+s-1} + \dots + c_1 y_{n+1} + c_0 y_n = 0$$

with $c_k = \frac{a_k}{a_s}$, $k = 0, 1, \dots, s-1$. We start with a lemma on *companion matrix*.

Lemma. The characteristic polynomial of n -degree $P(\lambda) = \lambda^s + c_{s-1} \lambda^{s-1} + \dots + c_1 \lambda + c_0 = \det(\lambda \mathbf{I} - \mathbf{C})$, where

$$\mathbf{C} = \begin{pmatrix} -c_{s-1} & -c_{s-2} & \dots & -c_1 & -c_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

Proof of Lemma: We prove by induction. For degree $s = 1$ case, $P(\lambda) = \lambda + c_0$. And $\lambda \mathbf{I} - \mathbf{C}$ is just $\lambda - (-c_0)$.

Now assume the relation is true for $s-1$, at s , we have $\det(\lambda \mathbf{I} - \mathbf{C}) =$

$$\begin{vmatrix} c_{s-1} & c_{s-2} & \dots & c_1 & c_0 \\ -1 & \lambda & \dots & 0 & 0 \\ 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & \lambda \end{vmatrix}_s = \lambda \begin{vmatrix} c_{s-1} & c_{s-2} & \dots & c_2 & c_1 \\ -1 & \lambda & \dots & 0 & 0 \\ 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & \lambda \end{vmatrix}_{s-1} + (-1)^{s+1} c_0 \begin{vmatrix} -1 & \lambda & \dots & 0 & 0 \\ 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & \lambda \\ 0 & 0 & \dots & 0 & -1 \end{vmatrix}_{s-1} \quad (\dagger)$$

with our inductive assumption, the determinant in the first term is exactly $\det(\lambda \mathbf{I} - \mathbf{C})$ in the degree $s-1$ case. So,

$$(\dagger) = \lambda(\lambda^{s-1} + c_{s-1} \lambda^{s-2} + \dots + c_2 \lambda + c_1) + (-1)^{s+1} c_0 (-1)^{s-1} = P(\lambda) \quad \blacksquare$$

The lemma suggests that finding the roots of $P(\lambda) = 0$ is equivalent to finding the eigenvalues of \mathbf{C} .

We denote $\mathbf{y}_n = (y_{n+s}, \dots, y_{n+2}, y_{n+1})^\top$; $\mathbf{c} = (c_{s-1}, \dots, c_1, c_0)^\top$, and unit vector with 1 on position k as \mathbf{e}_k . We find that the linear recurrence can be vectorized as $y_{n+s} = -\mathbf{c}^\top \mathbf{y}_{n-1}$. Moreover $y_{n+k} = \mathbf{e}_k^\top \mathbf{y}_{n-1}$ for $k = 1, \dots, s-1$.

$$\mathbf{y}_n = \begin{pmatrix} y_{n+s} \\ y_{n+s-1} \\ y_{n+s-2} \\ \vdots \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} -\mathbf{c}^\top \mathbf{y}_n \\ \mathbf{e}_1^\top \mathbf{y}_n \\ \mathbf{e}_2^\top \mathbf{y}_n \\ \vdots \\ \mathbf{e}_{s-1}^\top \mathbf{y}_n \end{pmatrix} = \begin{pmatrix} -c_{s-1} & -c_{s-2} & \dots & -c_1 & -c_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} y_{n+s-1} \\ y_{n+s-2} \\ y_{n+s-3} \\ \vdots \\ y_n \end{pmatrix} = \mathbf{C} \mathbf{y}_{n-1} = \mathbf{C}^n \mathbf{y}_0$$

where $\mathbf{y}_0 = (y_s, y_{s-1}, \dots, y_2, y_1)^\top$ is the initial condition.

Assume \mathbf{C} has eigenvalues $\lambda_1, \dots, \lambda_p$, $p \leq s$ with algebraic multiplicity m_1, \dots, m_p , where $\sum_p m_p = s$. Hence we decompose $\mathbf{C} = \mathbf{P} \mathbf{J} \mathbf{P}^{-1}$ in its Jordan Normal Form:

$$\mathbf{C}^n \mathbf{y}_0 = \mathbf{P} \mathbf{J}^n \mathbf{P}^{-1} \mathbf{y}_0 = \mathbf{P} \begin{pmatrix} \mathbf{J}_{\lambda_1}^n & & \\ & \mathbf{J}_{\lambda_2}^n & \\ & & \ddots \\ & & & \mathbf{J}_{\lambda_p}^n \end{pmatrix} \mathbf{P}^{-1} \mathbf{y}_0 \quad (*)$$

$$\text{where } \mathbf{J}_{\lambda_i}^n = \begin{pmatrix} \lambda_i^n & \binom{n}{1} \lambda_i^{n-1} & \binom{n}{2} \lambda_i^{n-2} & \dots & \binom{n}{m_i-1} \lambda_i^{n-(m_i-1)} \\ & \lambda_i^n & \binom{n}{1} \lambda_i^{n-1} & \dots & \binom{n}{m_i-2} \lambda_i^{n-(m_i-2)} \\ & & \ddots & \ddots & \vdots \\ & & & \lambda_i^n & \binom{n}{1} \lambda_i^{n-1} \\ & & & & \lambda_i^n \end{pmatrix}_{m_i \times m_i}$$

The typical entry is $\binom{n}{r}\lambda_i^{n-r} = \frac{\binom{n}{r}}{\lambda_i^r}\lambda_i^n$, which we can regard as $p_i^{[r]}(n)\lambda_i^n$. Here $p_i^{[r]}(n)$ is a polynomial of n to the degree r . Therefore, we regard $(*)$ as applying linear operations on these entries. We know that applying linear operations on a collection of polynomials $\{p_i^{[r]}(n)\}_{r=1}^{m_i-1}$ will result in another polynomial with degree no more than $m_i - 1$, denote it as $g_i^{[m_i-1]}(n)$. So we have

$$\begin{aligned} y_{n+s} &= \sum_{i=1}^p g_i^{[m_i-1]}(n)\lambda_i^n = \sum_{i=1}^p \sum_{r=0}^{m_i-1} \alpha_{ir} n^r \lambda_i^n \\ &= (\alpha_{10} + \alpha_{11}n + \dots + \alpha_{1,m_1-1}n^{m_1-1})\lambda_1^n + \dots + (\alpha_{p0} + \alpha_{p1}n + \dots + \alpha_{p,m_p-1}n^{m_p-1})\lambda_p^n \end{aligned} \quad (7)$$

Where $\{\alpha_{ir}\}_{i=1, r=0}^{p, m_i-1}$ are constants. Since $\sum_{i=1}^p m_i = s$, there are s constants in total, which are solved by inserting initial condition $\mathbf{y}_0 = (y_s, y_{s-1}, \dots, y_1)^\top$ into the general solution. If we index the initial condition as $(y_0, y_{-1}, \dots, y_{-(s-1)})^\top$ instead, the general solution will be

$$y_n = (C_{10} + C_{11}n + \dots + C_{1,m_1-1}n^{m_1-1})\lambda_1^n + \dots + (C_{p0} + C_{p1}n + \dots + C_{p,m_p-1}n^{m_p-1})\lambda_p^n$$

Where $\{C_{ir}\}$ is another set of constants. By lemma, $\{\lambda_i\}$ are the roots of $P(\lambda) = 0$, with multiplicity $\{m_i\}$. And this general solution holds for all $n \geq 0$. □

Problem 4. Find the region of A-stability for the multistep method:

$$y_{n+2} - y_n = \frac{1}{3}h[f(t_{n+2}, y_{n+2}) + 4f(t_{n+1}, y_{n+1}) + f(t_n, y_n)]$$

Proof. The first- and second-characteristic polynomial of the method is:

$$\rho(z) = z^2 - z; \quad \sigma(z) = \frac{1}{3}z^2 + \frac{4}{3}z + \frac{1}{3}$$

So the method is absolute stable \iff the roots z of $\Pi(z, \bar{h}) = \rho(z) - \bar{h}\sigma(z)$ have modulus < 1 .

$$\Pi(z, \bar{h}) = z^2 - z - \bar{h}\left(\frac{1}{3}z^2 + \frac{4}{3}z + \frac{1}{3}\right) = \left(1 - \frac{\bar{h}}{3}\right)z^2 + \left(-1 - \frac{4\bar{h}}{3}\right)z + \left(-\frac{\bar{h}}{3}\right) = 0 \quad (8)$$

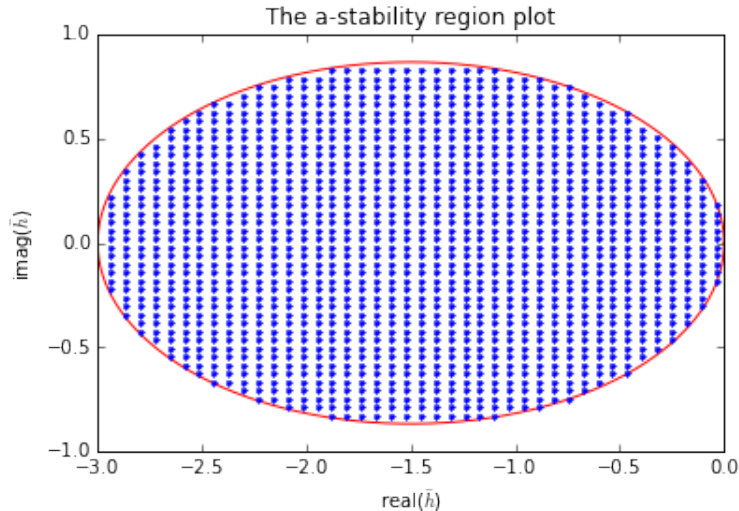
We solve the equation and obtain

$$z = \frac{\pm\sqrt{3(4\bar{h}^2 + 12\bar{h} + 3)} - 4\bar{h} - 3}{2(\bar{h} - 3)}$$

Or $\bar{h} = (z^2 - z)/(\frac{1}{3}z^2 + \frac{4}{3}z + \frac{1}{3})$. Denote the stability region as \mathcal{S} , then on $\partial\mathcal{S}$, the polynomial has root $|z| = 1$. Hence we can find the boundary by

$$\partial\mathcal{S} = \left\{ \frac{\rho(e^{i\theta})}{\sigma(e^{i\theta})}, \theta \in [0, 2\pi] \right\} = \left\{ \frac{e^{2i\theta} - e^{i\theta}}{\frac{1}{3}e^{2i\theta} + \frac{4}{3}e^{i\theta} + \frac{1}{3}}, \theta \in [0, 2\pi] \right\}$$

The region is give as the following plot (please also see the attached code).



□

Problem 5. (Iserles 3.4) Apply the classical RK4 method (with coefficient table shown as below) to scalar autonomous ode $y' = f(y)$, check that this method is indeed of order 4.

$$\begin{array}{c|cccc} c_1 & & & & \\ c_2 & a_{21} & & & \\ c_3 & a_{31} & a_{32} & & \\ c_4 & a_{41} & a_{42} & a_{43} & \end{array} = \begin{array}{c|cccc} 0 & & & & \\ 1/2 & & & & \\ 1/2 & 0 & 1/2 & & \\ 1 & 0 & 0 & 1 & \end{array}$$

$$\begin{array}{c|cccc} & b_1 & b_2 & b_3 & b_4 \end{array} \quad \begin{array}{c|cccc} & 1/6 & 1/3 & 1/3 & 1/6 \end{array}$$

Proof. We want to calculate the truncation error:

$$\frac{T_n}{h} = \frac{y(t_{n+1}) - y(t_n)}{h} - (b_1 k_1 + b_2 k_2 + b_3 k_3 + b_4 k_4)$$

We first calculate the first term using Taylor expansion. We calculate the derivatives with **Mathematica**:

$$\begin{aligned} y' &= f(y(t)) = f, \quad \frac{1}{2!}y^{(2)} = y'(t)f'(y(t)) = \frac{1}{2}f_y f \\ \frac{1}{3!}y^{(3)} &= \frac{1}{6}y'(t)^2 f''(y(t)) + \frac{1}{6}y''(t)f'(y(t)) = \frac{1}{6}f_{yy}f^2 + \frac{1}{6}f_y^2 f \\ \frac{1}{4!}y^{(4)} &= \frac{1}{24}[f^{(3)}(y(t))y'(t)^3 + 3y'(t)y''(t)f''(y(t)) + y^{(3)}(t)f'(y(t))] \\ &= \frac{1}{24}[f_{yyy}f^3 + 3f(ff_y)f_{yy} + (f^2 f_{yy} + f_y^2 f)f_y] \\ &= \frac{1}{24}f_{yyy}f^3 + \frac{1}{6}f_{yy}f_y f^2 + \frac{1}{24}f_y^3 f \end{aligned} \tag{9}$$

Then we expand the second part, the k 's are given by

$$\begin{aligned} k_1 &= f \\ k_2 &= f(y_n + \frac{1}{2}hf(y_n)) = \{f + f_y \cdot (\frac{1}{2}hf) + \frac{f_{yy}}{2} \cdot (\frac{1}{2}hf)^2 + \frac{f_{yyy}}{6} \cdot (\frac{1}{2}hf)^3 + O(h^4)\} \\ &= f + \frac{1}{2}f_y f h + \frac{1}{8}f_{yy}f^2 h^2 + \frac{1}{48}f_{yyy}f^3 h^3 + O(h^4) \end{aligned} \tag{10}$$

$$\begin{aligned} k_3 &= f(y_n + \frac{1}{2}h(f + \frac{1}{2}f_y f h + \frac{1}{8}f_{yy}f^2 h^2 + \frac{1}{48}f_{yyy}f^3 h^3 + O(h^4))) \\ &= \left\{f + f_y \cdot \frac{1}{2}h(f + \frac{1}{2}f_y f h + \frac{1}{8}f_{yy}f^2 h^2 + O(h^3)) + \frac{f_{yy}}{2} \cdot \frac{1}{4}h^2(f + \frac{1}{2}f_y f h + O(h^2))^2 + \right. \\ &\quad \left. + \frac{f_{yyy}}{6} \cdot \frac{1}{8}h^3(f + O(h))^3 + O(h^4)\right\} \\ &= f + h\left(\frac{f_y f}{2} + \frac{1}{4}f_y^2 f h + \frac{1}{16}f_{yy}f_y f^2 h^2 + O(h^3)\right) + h^2\left(\frac{f_{yy}f^2}{8} + \frac{1}{8}f_{yy}f_y f^2 h + O(h^2)\right) \\ &\quad + h^3\left(\frac{1}{48}f_{yyy}f^3 + O(h)\right) + O(h^4) \\ &= f + \frac{1}{2}f_y f h + (\frac{1}{4}f_y^2 f + \frac{1}{8}f_{yy}f^2)h^2 + (\frac{3}{16}f_{yy}f_y f^2 + \frac{1}{48}f_{yyy}f^3)h^3 + O(h^4) \end{aligned} \tag{11}$$

$$\begin{aligned} k_4 &= f(y_n + h[f + \frac{1}{2}f_y f h + (\frac{1}{4}f_y^2 f + \frac{1}{8}f_{yy}f^2)h^2 + (\frac{3}{16}f_{yy}f_y f^2 + \frac{1}{48}f_{yyy}f^3)h^3 + O(h^4)]) \\ &= \left\{f + f_y \cdot h[f + \frac{1}{2}f_y f h + (\frac{1}{4}f_y^2 f + \frac{1}{8}f_{yy}f^2)h^2 + O(h^3)] + \frac{f_{yy}}{2} \cdot h^2[f + \frac{1}{2}f_y f h + O(h^2)]^2 + \right. \\ &\quad \left. + \frac{f_{yyy}}{6} \cdot h^3[f + O(h)]^3 + O(h^4)\right\} \\ &= f + h[f_y f + \frac{1}{2}f_y^2 f h + (\frac{1}{4}f_y^3 f + \frac{1}{8}f_{yy}f_y f^2)h^2 + O(h^3)] + h^2[\frac{1}{2}f_{yy}f^2 + \frac{1}{2}f_{yy}f_y f^2 h + O(h^2)] \\ &\quad + h^3(\frac{1}{6}f_{yyy}f^3 + O(h)) + O(h^4) \\ &= f + f_y f h + (\frac{1}{2}f_y^2 f + \frac{1}{2}f_{yy}f^2)h^2 + (\frac{1}{4}f_y^3 f + \frac{5}{8}f_{yy}f_y f^2 + \frac{1}{6}f_{yyy}f^3)h^3 + O(h^4) \end{aligned} \tag{12}$$

Hence we have:

$$\begin{aligned} \frac{1}{6}k_1 &= \frac{1}{6}f \\ \frac{1}{3}k_2 &= \frac{1}{3}f + \frac{1}{6}f_y f h + \frac{1}{24}f_{yy}f^2 h^2 + \frac{1}{144}f_{yyy}f^3 h^3 + O(h^4) \\ \frac{1}{3}k_3 &= \frac{1}{3}f + \frac{1}{6}f_y f h + (\frac{1}{12}f_y^2 f + \frac{1}{24}f_{yy}f^2)h^2 + (\frac{1}{16}f_{yy}f_y f^2 + \frac{1}{144}f_{yyy}f^3)h^3 + O(h^4) \\ \frac{1}{6}k_4 &= \frac{1}{6}f + \frac{1}{6}f_y f h + (\frac{1}{12}f_y^2 f + \frac{1}{12}f_{yy}f^2)h^2 + (\frac{1}{24}f_y^3 f + \frac{5}{48}f_{yy}f_y f^2 + \frac{1}{36}f_{yyy}f^3)h^3 + O(h^4) \end{aligned} \tag{13}$$

$$\begin{aligned}
\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 &= \left(\frac{1}{6} + \frac{1}{3} + \frac{1}{3} + \frac{1}{6}\right)f + \left(\frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right)f_y f h \\
&\quad + \left[\left(\frac{1}{24} + \frac{1}{24} + \frac{1}{12}\right)f_{yy}f^2 + \left(\frac{1}{12} + \frac{1}{12}\right)f_y^2 f\right]h^2 \\
&\quad + \left[\left(\frac{1}{144} + \frac{1}{144} + \frac{1}{36}\right)f_{yyy}f^3 + \left(\frac{1}{16} + \frac{5}{48}\right)f_{yy}f_y f^2 + \left(\frac{1}{24}\right)f_y^3 f\right]h^3 + O(h^4) \\
&= f + \frac{1}{2}f_y f h + \left(\frac{1}{6}f_{yy}f^2 + \frac{1}{6}f_y^2 f\right)h^2 + \left(\frac{1}{24}f_{yyy}f^3 + \frac{1}{6}f_{yy}f_y f^2 + \frac{1}{24}f_y^3 f\right)h^3 + O(h^4) \quad (\dagger)
\end{aligned} \tag{14}$$

Compare (\dagger) with the Taylor expansion of $\frac{y(t_{n+1}) - y(t_n)}{h}$, and notice that the coefficients of $O(1)$, $O(h)$, $O(h^2)$ and $O(h^3)$ match together, hence they vanish. The truncation error $T_n/h = O(h^4)$. \square

Problem 6. (Iserles 3.5) Suppose that a ν -stage explicit RK method of order ν is applied to linear scalar ode $y' = \lambda y$, show that

$$y_n = \left(\sum_{k=0}^{\nu} \frac{1}{k!} (h\lambda)^k \right)^n y_0, \quad n = 0, 1, \dots$$

Proof. The scheme of explicit RK method is

$$y_{n+1} = y_n + h\Phi(t_n, y_n; h)$$

And the truncation error is

$$\frac{T_n}{h} = \frac{y(t_{n+1}) - y(t_n)}{h} - \Phi(t_n, y(t_n); h)$$

As usual we expand the first term via Taylor series, and since y is solution of the linear scalar ode, we have $y'' = \lambda y' = \lambda^2 y$, ..., $y^{(n)} = \lambda^n y$.

$$\begin{aligned}
\frac{y(t_{n+1}) - y(t_n)}{h} &= y'(t_n) + \frac{y''(t_n)}{2!}h + \dots + \frac{y^{(\nu)}(t_n)}{\nu!}h^{\nu-1} + O(h^\nu) \\
&= y(t_n) \left(\lambda + \frac{\lambda^2}{2!}h + \dots + \frac{\lambda^\nu}{\nu!}h^{\nu-1} \right) + O(h^\nu) \\
&= \frac{y(t_n)}{h} \sum_{k=1}^{\nu} \frac{1}{k!} (\lambda h)^k + O(h^\nu)
\end{aligned} \tag{15}$$

Since the method is of order ν , the truncation error is $O(h^\nu)$, which implies that $\Phi(t_n, y(t_n); h) = \frac{y(t_n)}{h} \sum_{k=1}^{\nu} \frac{1}{k!} (\lambda h)^k$. Therefore the RK method formula is

$$y_{n+1} = y_n + h\Phi(t_n, y_n; h) = y_n + y_n \sum_{k=1}^{\nu} \frac{1}{k!} (\lambda h)^k = y_n \sum_{k=0}^{\nu} \frac{1}{k!} (\lambda h)^k$$

i.e.

$$y_n = \left(\sum_{k=0}^{\nu} \frac{1}{k!} (\lambda h)^k \right)^n y_0$$

\square

Problem 7. (Iserles 4.1) Let $\mathbf{y}' = \mathbf{\Lambda}\mathbf{y}$, $\mathbf{y}(t_0) = \mathbf{y}_0$ be solved (with a constant step size $h > 0$) by a one-step method with a function $r(\cdot)$ that obeys the relation (4.12). Suppose that a nonsingular matrix \mathbf{V} and a diagonal matrix \mathbf{D} exists such that $\mathbf{\Lambda} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$, show that there exists vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_d \in \mathbb{R}^d$ such that

$$\mathbf{y}(t_n) = \sum_{j=1}^d e^{t_n \lambda_j} \mathbf{x}_j, \quad n = 0, 1, \dots$$

And

$$\mathbf{y}_n = \sum_{j=1}^d [r(h\lambda_j)]^n \mathbf{x}_j, \quad n = 0, 1, \dots$$

where $\lambda_1, \dots, \lambda_d$ are eigenvalues of $\mathbf{\Lambda}$. Deduce that the values of \mathbf{x}_1 and \mathbf{x}_2 , given in (4.3) and (4.4) are identical.

Proof. (a.) The exact solution of the linear ode system is $\mathbf{y}(t) = e^{\Lambda t} \mathbf{y}_0$. Since Λ is diagonalizable, we have $\Lambda = \mathbf{V} \mathbf{D} \mathbf{V}^{-1}$, where the entries of \mathbf{D} are the d linear independent eigenvalues of Λ and the columns of \mathbf{V} are the corresponding eigenvectors. Therefore

$$e^{\Lambda t} = e^{\mathbf{V} \mathbf{D} \mathbf{V}^{-1} t} = \sum_{k=0}^{\infty} \frac{t^k}{k!} (\mathbf{V} \mathbf{D} \mathbf{V}^{-1})^k = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbf{V} \mathbf{D}^k \mathbf{V}^{-1} = \mathbf{V} \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbf{D}^k \mathbf{V}^{-1} = \mathbf{V} e^{\mathbf{D} t} \mathbf{V}^{-1}$$

We denote $\mathbf{w} = \mathbf{V}^{-1} \mathbf{y}_0 = (w_1, \dots, w_d)^\top$ is a $d \times 1$ vector, then

$$\mathbf{y}(t_n) = \mathbf{V} e^{\mathbf{D} t_n} \mathbf{V}^{-1} \mathbf{y}_0 = \mathbf{V} \begin{pmatrix} e^{\lambda_1 t_n} & & & \\ & e^{\lambda_2 t_n} & & \\ & & \ddots & \\ & & & e^{\lambda_d t_n} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{pmatrix} = \mathbf{V} \begin{pmatrix} w_1 e^{\lambda_1 t_n} \\ w_2 e^{\lambda_2 t_n} \\ \vdots \\ w_d e^{\lambda_d t_n} \end{pmatrix} = \sum_{j=1}^d w_j e^{\lambda_j t_n} \mathbf{v}_j$$

where $\mathbf{w} = \mathbf{V}^{-1} \mathbf{y}_0$ and \mathbf{v}_j is the j -th eigenvector. If we denote $\mathbf{x}_j := w_j \mathbf{v}_j$, then we come to the desired results. $\mathbf{y}(t_n) = \sum_{j=1}^d e^{\lambda_j t_n} \mathbf{x}_j$. \square

(b.) We apply the ν -stage RK method on this linear ode system, the formulation of IRK is given by (c.f. Iserles eq.3.9)

$$\begin{cases} \xi_j = \mathbf{y}_n + h \sum_{i=1}^{\nu} a_{ji} \mathbf{f}(t_n + c_i h, \xi_i) & j = 1, 2, \dots, \nu \\ \mathbf{y}_{n+1} = \mathbf{y}_n + h \sum_{j=1}^{\nu} b_j \mathbf{f}(t_n + c_j h, \xi_j) \end{cases} \quad (16)$$

And apply $\mathbf{f}(\cdot, \mathbf{y}) = \Lambda \mathbf{y}$, the scheme becomes

$$\begin{cases} \xi_j = \mathbf{y}_n + h \Lambda \sum_{i=1}^{\nu} a_{ji} \xi_i & j = 1, 2, \dots, \nu \\ \mathbf{y}_{n+1} = \mathbf{y}_n + h \Lambda \sum_{j=1}^{\nu} b_j \xi_j \end{cases} \quad (17)$$

Notice that:

- $\sum_{j=1}^{\nu} b_j \xi_j$ is a linear combination of $\{\xi_j\}_1^\nu$ s, with weights being $\mathbf{b} = (b_1, \dots, b_\nu)^\top$. So we write it as $\Xi \mathbf{b}$, where Ξ is $d \times \nu$ matrix, whose columns are the vectors $\{\xi_j\}_1^\nu$ s. Then the RK iteration formula can be written as

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h \Lambda \Xi \mathbf{b} \quad (\dagger)$$

- $\sum_{i=1}^{\nu} a_{ji} \xi_i$ can be considered in the same way, i.e. $\sum_{i=1}^{\nu} a_{ji} \xi_i = \Xi \mathbf{a}_j$, $\mathbf{a}_j = (a_{j1}, a_{j2}, \dots, a_{j\nu})^\top$. Stack $\xi_j = \mathbf{y}_n + h \Lambda \Xi \mathbf{a}_j$ together on columns:

$$\Xi = (\xi_1, \dots, \xi_\nu) = (\mathbf{y}_n + h \Lambda \Xi \mathbf{a}_1, \dots, \mathbf{y}_n + h \Lambda \Xi \mathbf{a}_\nu) = \mathbf{y}_n \mathbf{1}^\top + h \Lambda \Xi \mathbf{A} \quad (18)$$

Where $\mathbf{1}^\top = (1, 1, \dots, 1)$ is an $1 \times \nu$ iota vector. So we have $\mathbf{y}_n \mathbf{1}^\top = \Xi - h \Lambda \Xi \mathbf{A} = (\mathbf{I}_d - h \Lambda) \Xi (\mathbf{I}_\nu - \mathbf{A})$, where $\mathbf{I}_d, \mathbf{I}_\nu$ are $d \times d$ and $\nu \times \nu$ identity matrices respectively. Assume $\mathbf{I}_d - h \Lambda$ and $\mathbf{I}_\nu - \mathbf{A}$ are nonsingular, we have

$$\Xi = (\mathbf{I}_d - h \Lambda)^{-1} \mathbf{y}_n \mathbf{1}^\top (\mathbf{I}_\nu - \mathbf{A})^{-1} \quad (\ddagger)$$

Insert (\ddagger) into (\dagger) , using diagonalization $\Lambda = \mathbf{V} \mathbf{D} \mathbf{V}^{-1}$:

$$\begin{aligned} \mathbf{y}_{n+1} &= \mathbf{y}_n + h \Lambda (\mathbf{I}_d - h \Lambda)^{-1} \mathbf{y}_n \mathbf{1}^\top (\mathbf{I}_\nu - \mathbf{A})^{-1} \mathbf{b} \\ &= \mathbf{y}_n + h \mathbf{V} \mathbf{D} \mathbf{V}^{-1} (\mathbf{V} (\mathbf{I}_d - h \mathbf{D}) \mathbf{V}^{-1})^{-1} \mathbf{y}_n \mathbf{1}^\top (\mathbf{I}_\nu - \mathbf{A})^{-1} \mathbf{b} \\ &= \mathbf{V} \mathbf{I}_d \mathbf{V}^{-1} \mathbf{y}_n \cdot \mathbf{1} + h \mathbf{V} \mathbf{D} (\mathbf{I}_d - h \mathbf{D})^{-1} \mathbf{V}^{-1} \mathbf{y}_n \mathbf{1}^\top (\mathbf{I}_\nu - \mathbf{A})^{-1} \mathbf{b} \\ &= \mathbf{V} (\mathbf{I}_d + h \mathbf{D} (\mathbf{I}_d - h \mathbf{D})^{-1}) \mathbf{V}^{-1} \mathbf{y}_n (\mathbf{1} + \mathbf{1}^\top (\mathbf{I}_\nu - \mathbf{A})^{-1} \mathbf{b}) \end{aligned} \quad (19)$$

Notice that:

- $(\mathbf{1} + \mathbf{1}^\top (\mathbf{I}_\nu - \mathbf{A})^{-1} \mathbf{b})$ is just a scalar, totally determined by the coefficients of the RK method. We denote it as ζ and bring it forward, i.e. consider $\mathbf{y}_{n+1} = \mathbf{V} \zeta (\mathbf{I}_d + h \mathbf{D} (\mathbf{I}_d - h \mathbf{D})^{-1}) \mathbf{V}^{-1} \mathbf{y}_n$
- $\zeta (\mathbf{I}_d + h \mathbf{D} (\mathbf{I}_d - h \mathbf{D})^{-1})$ is a diagonal matrix. And denote $\mathbf{D}_{\bar{h}} = h \mathbf{D} = \text{diag}\{h\lambda_1, \dots, h\lambda_d\}$, this matrix is $\zeta (\mathbf{I}_d + \mathbf{D}_{\bar{h}} (\mathbf{I}_d - \mathbf{D}_{\bar{h}})^{-1})$. By the lemma on the book, it must be of the form:

$$\zeta (\mathbf{I}_d + \mathbf{D}_{\bar{h}} (\mathbf{I}_d - \mathbf{D}_{\bar{h}})^{-1}) = \begin{pmatrix} r(h\lambda_1) & & \\ & \ddots & \\ & & r(h\lambda_d) \end{pmatrix} \quad (20)$$

$r(\cdot)$ is the function of the type defined in (Iserles) section 4.3.

Therefore, by the discussion above, and using the same notation in our derivation of the exact solution:

$$\begin{aligned} \mathbf{y}_n &= \mathbf{V} \begin{pmatrix} r(h\lambda_1) & & \\ & \ddots & \\ & & r(h\lambda_d) \end{pmatrix} \mathbf{V}^{-1} \mathbf{y}_{n-1} = \mathbf{V} \begin{pmatrix} r(h\lambda_1) & & \\ & \ddots & \\ & & r(h\lambda_d) \end{pmatrix}^n \mathbf{V}^{-1} \mathbf{y}_0 \\ &= \mathbf{V} \begin{pmatrix} r^n(h\lambda_1) & & \\ & \ddots & \\ & & r^n(h\lambda_d) \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_d \end{pmatrix} = \mathbf{V} \begin{pmatrix} w_1 [r(h\lambda_1)]^n \\ \vdots \\ w_d [r(h\lambda_d)]^n \end{pmatrix} = \sum_{j=1}^d [r(h\lambda_j)]^n w_j \mathbf{v}_j \end{aligned} \quad (21)$$

Which equals to $\sum_{j=1}^d [r(h\lambda_j)]^n \mathbf{x}_j$ with same $\mathbf{x}_j = w_j \mathbf{v}_j$, finished the proof. \square

Problem 8. (Iserles 4.2) Consider the solution of $\mathbf{y}' = \mathbf{\Lambda} \mathbf{y}$ where

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad \lambda \in \mathbb{C}^-$$

a. Show that

$$\mathbf{\Lambda}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}, \quad n = 0, 1, \dots$$

b. Let g be an arbitrary function that is analytic about the origin. The 2×2 matrix $g(\mathbf{\Lambda})$ can be defined by substituting powers of $\mathbf{\Lambda}$ into the Taylor expansion of g , show that

$$g(t\mathbf{\Lambda}) = \begin{pmatrix} g(t\lambda) & tg'(t\lambda) \\ 0 & g(t\lambda) \end{pmatrix}$$

c. By letting $g(z) = e^z$, show that $\lim_{t \rightarrow \infty} \mathbf{y}(t) = \mathbf{0}$.

d. Suppose that $\mathbf{y}' = \mathbf{\Lambda} \mathbf{y}$ is solved with RK method, using a constant step size $h > 0$. Let r be the function from lemma 4.1. Letting $g = r$, obtain the explicit form of $[r(h\mathbf{\Lambda})]^n$, $n = 0, 1, \dots$

e. Show that if $h\lambda \in \mathcal{D}$, where \mathcal{D} is the linear stability domain of the RK method, then $\lim_{n \rightarrow \infty} \mathbf{y}_n = \mathbf{0}$.

Proof. (a.) We can decompose $\mathbf{\Lambda} = \lambda \mathbf{I} + \mathbf{B}$, with $\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. It is easy to see that $\mathbf{B}^n = \mathbf{0}$ for $n \geq 2$. Hence

$$\mathbf{\Lambda}^n = (\lambda \mathbf{I})^n + \binom{n}{1} (\lambda \mathbf{I})^{n-1} \mathbf{B} = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} \quad \square$$

(b.) By the given definition of $g(t\mathbf{\Lambda})$ (with $\mathbf{\Lambda}^0 := \mathbf{I}$, $\mathbf{\Lambda}^n$ being the expression we obtained in (a)):

$$\begin{aligned} g(t\mathbf{\Lambda}) &= \sum_{n=0}^{\infty} \frac{g^{(n)}(0)}{n!} (t\mathbf{\Lambda})^n = \sum_{n=0}^{\infty} \frac{g^{(n)}(0)}{n!} \begin{pmatrix} (t\lambda)^n & nt(t\lambda)^{n-1} \\ 0 & (t\lambda)^n \end{pmatrix} \\ &= \begin{pmatrix} \sum_{n=0}^{\infty} \frac{g^{(n)}(0)}{n!} (t\lambda)^n & t \sum_{n=0}^{\infty} \frac{g^{(n)}(0)}{(n-1)!} (t\lambda)^{n-1} \\ 0 & \sum_{n=0}^{\infty} \frac{g^{(n)}(0)}{n!} (t\lambda)^n \end{pmatrix} \\ &= \begin{pmatrix} g(t\lambda) & tg'(t\lambda) \\ 0 & g(t\lambda) \end{pmatrix} \quad \square \end{aligned} \quad (22)$$

(c.) The solution of ode is $\mathbf{y}(t) = e^{\mathbf{\Lambda}t} \mathbf{y}_0$. Let $g(\mathbf{\Lambda}t) = e^{\mathbf{\Lambda}t}$, by the result of (b.) we get:

$$\mathbf{y}(t) = e^{\mathbf{\Lambda}t} \mathbf{y}_0 = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} \mathbf{y}_0 \quad (23)$$

Since $\lambda \in \mathbb{C}^-$, it is clear that $\lim_{t \rightarrow \infty} e^{\lambda t} = 0$, and is of higher order of decay than the order of growth of t , i.e. $\lim_{t \rightarrow \infty} te^{\lambda t} = 0$ too. Hence in the expression above, $\lim_{t \rightarrow \infty} e^{\mathbf{\Lambda}t} = \mathbf{O}$, so $\lim_{t \rightarrow \infty} e^{\mathbf{\Lambda}t} \mathbf{y}_0 = \mathbf{0}$.

(d.) By the result of b., we have

$$[r(h\mathbf{\Lambda})]^n = [g(h\mathbf{\Lambda})]^n = \begin{pmatrix} g(h\lambda) & hg'(h\lambda) \\ 0 & g(h\lambda) \end{pmatrix}^n \quad (*) \quad (24)$$

And note that this is of the same form as $\mathbf{\Lambda}$ itself, we decompose

$$(*) = \left[g(h\lambda)\mathbf{I} + \begin{pmatrix} 0 & hg'(h\lambda) \\ 0 & 0 \end{pmatrix} \right]^n = \begin{pmatrix} [g(h\lambda)]^n & n[g(h\lambda)]^{n-1}hg'(h\lambda) \\ 0 & [g(h\lambda)]^n \end{pmatrix} = \begin{pmatrix} [r(h\lambda)]^n & n[r(h\lambda)]^{n-1}hr'(h\lambda) \\ 0 & [r(h\lambda)]^n \end{pmatrix} \quad \square \quad (25)$$

(e.) If $h\lambda \in \mathcal{D}$ (the stability region of RK method), by lemma 4.2 of (Iserles), $\mathcal{D} = \{z : |r(z)| < 1\}$. So we have $|r(h\lambda)| < 1$. Hence

$$\lim_{n \rightarrow \infty} [r(h\lambda)]^n = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} n[r(h\lambda)]^{n-1} = 0$$

So, $\lim_{n \rightarrow \infty} \mathbf{y}_n = \lim_{n \rightarrow \infty} [r(h\mathbf{\Lambda})]^n \mathbf{y}_0 = \mathbf{0}$, finished the proof. \square
