Tuesday, January 11, 2022

MAE290B HW1 (Opal Issan)

Problem 1

A function f(x) is known at points x_i, i = 1,2,...,n.

a) Use points x_i, x_{i+1}, and x_{i+2} to obtain the best approximation to $f'(x)$ at x_i. Give the truncation error (including coefficient).

The objective is to find the optimal coefficients $\{a_0, a_1, a_2\}$ where we minimize the truncation error (TE), .i.e.

$f'(x_j) + a_0 f(x_j) + a_1 f(x_{j+1}) + a_2 f(x_{j+2}) = TE$

This can be systematically solved using the Taylor Table

	$f(x_j)$	$f'(x_j)$	$f''(x_j)$	$f'''(x_j)$	$f''''(x_j)$	$f'''''(x_j)$
$f'(x_j)$	0	1	0	0	0	0
$a_0 f(x_j)$	a_0	0	0	0	0	0
$a_1 f(x_{j+1})$	a_1	$a_1 h$	$a_1 \frac{h^2}{2}$	$a_1 \frac{h^3}{6}$	$a_1 \frac{h^4}{24}$	$a_1 \frac{h^5}{120}$
$a_2 f(x_{j+2})$	a_2	$a_2(2h)$	$a_2(2h^2)$	$a_2 \frac{4h^3}{3}$	$a_2 \frac{2h^4}{3}$	$a_2 \frac{4h^5}{15}$
Total	0	0	0	??	??	??

Which results in the following system of linear equations

$$a_0 + a_1 + a_2 = 0$$

$$1 + a_1 h + a_2(2h) = 0$$

$$a_1 \frac{h^2}{2} + a_2(2h^2) = 0$$

The solution of this system is

$$a_0 = \frac{3}{2h} \qquad a_1 = \frac{-4}{2h} \qquad a_2 = \frac{1}{2h}$$

With that, the fourth column in the Taylor table sums up to $\frac{1}{3}h^2 f'''(x)$. Hence, $TE \sim O(h^2)$. The finite difference approximation is of 2nd-order accurate resulting in $f'(x_j) = \dfrac{-3f(x_j) + 4f(x_{j+1}) - f(x_{j+2})}{2h} + O(h^2)$.

b) Consider the function e^{ikx}. Compare the accuracy of the FDA in (a) and the 3-point central FDA for $f'(x)$ using modified wavenumber analysis. Plot the real part of the non-dimensional modified wavenumber $k'h$ as a function of kh for these two FDA schemes.

The 3-point central FDA solves for the following coefficients a_0, a_1, a_2, such that

$f'(x_j) + a_0 f(x_j) + a_1 f(x_{j-1}) + a_2 f(x_{j+1}) = TE$

Similar to the analysis performed in (a), we use the Taylor Table to obtain the coefficients

	$f(x_j)$	$f'(x_j)$	$f''(x_j)$	$f'''(x_j)$	$f''''(x_j)$	$f'''''(x_j)$
$f'(x_j)$	0	1	0	0	0	0
$a_0 f(x_j)$	a_0	0	0	0	0	0
$a_1 f(x_{j-1})$	a_1	$-a_1 h$	$a_1 \frac{h^2}{2}$	$-a_1 \frac{h^3}{6}$	$a_1 \frac{h^4}{24}$	$-a_1 \frac{h^5}{120}$
$a_2 f(x_{j+1})$	a_2	$a_2 h$	$a_2 \frac{h^2}{2}$	$a_2 \frac{h^3}{6}$	$a_2 \frac{h^4}{24}$	$a_2 \frac{h^5}{120}$
Total	0	0	0	??	??	??

Which results in the following system of linear equations

$$a_0 + a_1 + a_2 = 0$$

$$1 - a_1 h + a_2 h = 0$$

$$a_1 \frac{h^2}{2} + a_2 \frac{h^2}{2} = 0$$

The solution of this system is

$$a_0 = 0 \qquad a_1 = \frac{1}{2h} \qquad a_2 = \frac{-1}{2h}$$

Which leads to $f'(x_j) = \dfrac{f(x_{j+1}) - f(x_{j-1})}{2h} - \dfrac{h^2}{6} f'''(x_j)$. The truncation error is $O(h^2)$. The modified wavenumber analysis compares the exact derivative with the numerical derivative based on the wavenumber frequency k and the mesh grid spacing h. The central FDA to f' where $f(x) = e^{ikx}$ is

$$\left.\frac{df}{dx}\right|_{x=x_j} = \frac{f(x_{j+1}) - f(x_{j-1})}{2h} = \frac{e^{ikx_{j+1}} - e^{ikx_{j-1}}}{2h} = \frac{e^{ik(x_j+h)} - e^{ik(x_j-h)}}{2h} = e^{ikx_j}\frac{e^{ikh} - e^{-ikh}}{2h} = e^{ikx_j}\frac{2i\sin(kh)}{2h} = i\frac{\sin(kh)}{h}e^{ikx_j}$$

The exact derivative is $f'_{exact}(x) = ike^{ikx}$. Therefore $k' = \dfrac{\sin(kh)}{h}$ for the central finite difference scheme. The (forward) FDA described in (a) to f' where $f(x) = e^{ikx}$ is

$$\left.\frac{df}{dx}\right|_{x=x_j} = \frac{-3f(x_j) + 4f(x_{j+1}) - f(x_{j+2})}{2h} = \frac{-3e^{ikx_j} + 4e^{ik(x_j+h)} - e^{ik(x_j+2h)}}{2h} = e^{ikx_j}\frac{-3 + 4e^{ikh} - e^{2ikh}}{2h}$$

Hence,

$$hk' = \frac{-3 + 4e^{ikh} - e^{2ikh}}{2i} = \frac{-3 + 4\cos(kh) - \cos(2kh) + i(4\sin(kh) - \sin(2kh))}{2i} = \frac{[3 - 4\cos(kh) + \cos(2kh)]i + (4\sin(kh) - \sin(2kh))}{2}$$

The real part of $Re(hk') = \dfrac{4\sin(kh) - \sin(2kh)}{2} = \dfrac{4\sin(kh) - 2\cos(kh)\sin(kh)}{2} = \sin(kh)(2 - \cos(kh))$

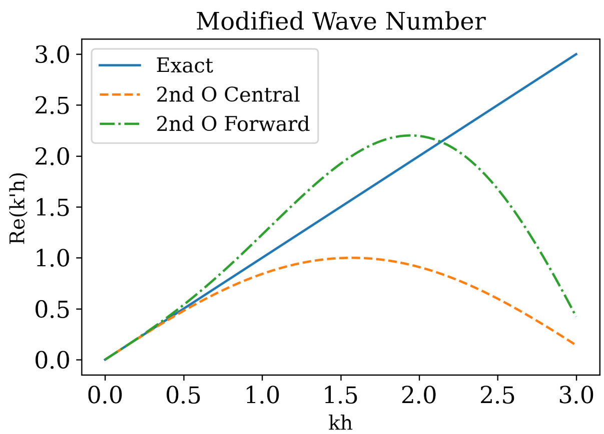

Fig. 1: The real part of the non-dimensional modified wave number $k'h$ as a function of kh for the FDA in (a) and the central FDA. h is the grid spacing.

Based on Figure 1, The forward 2nd order FDA derived in (a) outperforms the 2nd order central FDA in approximating the modified wave number. However, for large kh both methods would not be suitable in achieving high accuracy as they diverge from the exact $kh = k'h$.

It is important to mention that although the FDA in (a) is a better approximation to the modified wave number than the central 2nd-order FDA, the central 2nd-order FDA is more accurate based on its TE coefficient ($\frac{1}{3} vs. \frac{-1}{6}$). In fact, it is a factor of 2 more accurate.

c) Perform a modified wavenumber analysis for the central 3-point formula for $f''(x)$. Plot the non-dimensional modified wavenumber $k'h$ as a function of kh in the same plot as (b). Comment on the accuracy of the 2nd order central FDAs to $f'(x)$ and $f''(x)$ at high wave numbers.

The 3-point central FDA solves for the following coefficients a_0, a_1, a_2, such that
$f''(x_j) + a_0 f(x_j) + a_1 f(x_{j-1}) + a_2 f(x_{j+1}) = TE$

Similar to the analysis performed in (a), we use the Taylor Table to obtain the coefficients

	$f(x_j)$	$f'(x_j)$	$f''(x_j)$	$f'''(x_j)$	$f''''(x_j)$	$f'''''(x_j)$
$f''(x_j)$	0	0	1	0	0	0
$a_0 f(x_j)$	a_0	0	0	0	0	0
$a_1 f(x_{j-1})$	a_1	$-a_1 h$	$a_1 \frac{h^2}{2}$	$-a_1 \frac{h^3}{6}$	$a_1 \frac{h^4}{24}$	$-a_1 \frac{h^5}{120}$
$a_2 f(x_{j+1})$	a_2	$a_2 h$	$a_2 \frac{h^2}{2}$	$a_2 \frac{h^3}{6}$	$a_2 \frac{h^4}{24}$	$a_2 \frac{h^5}{120}$
Total	0	0	0	??	??	??

Which results in the following system of linear equations

$$a_0 + a_1 + a_2 = 0$$

$$-a_1 h + a_2 h = 0$$

$$1 + a_1 \frac{h^2}{2} + a_2 \frac{h^2}{2} = 0$$

The solution of this system is

$$a_0 = \frac{2}{h^2} \qquad a_1 = \frac{-1}{h^2} \qquad a_2 = \frac{-1}{h^2}$$

Which leads to $f''(x_j) = \dfrac{f(x_{j+1}) - 2f(x_j) + f(x_{j-1})}{h^2} - \dfrac{h^2}{12} f''''(x_j)$. The truncation error is $O(h^2)$. The central finite difference approximation to f' where $f(x) = e^{ikx}$ is

$$\left.\frac{d^2 f}{dx^2}\right|_{x=x_j} = \frac{f(x_{j+1}) - 2f(x_j) + f(x_{j-1})}{h^2} = \frac{e^{ik(x_j+h)} - 2e^{ikx_j} + e^{ik(x_j-h)}}{h^2} = e^{ikx_j}\frac{e^{ikh} - 2 + e^{-ikh}}{h^2} = e^{ikx_j}\frac{2\cos(kh) - 2}{h^2}$$

The exact derivative is $f''_{exact}(x) = -k^2 e^{ikx}$. Therefore $k'h = \sqrt{2(1-\cos(kh))}$ for the central finite difference scheme.

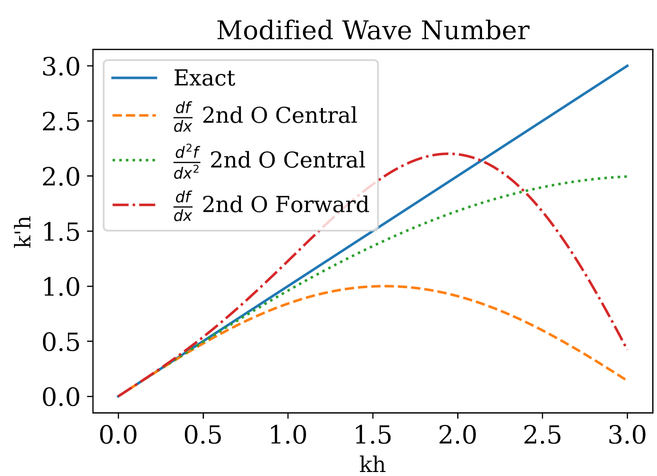

Fig 2. Non-dimensional modified wave number $k'h$ as a function of kh. The second derivative approximation using the central FDA outperformed approximating the first derivative using the central FDA. The second derivative approximation using the central FDA is in better agreement with the exact wavenumber.

As illustrate in Fig 2., the three stencil central FDA is better suited for approximating the second derivate in comparison to the first derivative. This statement holds because of the TE coefficient ($\frac{-1}{6}$ vs. $\frac{-1}{12}$) and the modified wave number analysis.

d) Use Taylor series to show that the Pade formula given by (2.19) in the text book is fourth-order accurate, and give the truncation error.

The objective is to find the optimal coefficients $\{a_0, a_1, a_2, a_3, a_4\}$ where we minimize the truncation error (TE), .i.e.

$f''(x_j) + a_0 f(x_j) + a_1 f(x_{j-1}) + a_2 f(x_{j+1}) + a_3 f''(x_{j-1}) + a_4 f''(x_{j+1}) = TE$

This can be systematically solved using the Taylor Table

	$f(x_j)$	$f'(x_j)$	$f''(x_j)$	$f'''(x_j)$	$f''''(x_j)$	$f^{(5)}(x_j)$	$f^{(6)}(x_j)$
$f''(x_j)$	0	0	1	0	0	0	0
$a_0 f(x_j)$	a_0	0	0	0	0	0	0
$a_1 f(x_{j-1})$	a_1	$-a_1 h$	$a_1 \frac{h^2}{2}$	$-a_1 \frac{h^3}{6}$	$a_1 \frac{h^4}{24}$	$-a_1 \frac{h^5}{120}$	$a_1 \frac{h^6}{720}$
$a_2 f(x_{j+1})$	a_2	$a_2 h$	$a_2 \frac{h^2}{2}$	$a_2 \frac{h^3}{6}$	$a_2 \frac{h^4}{24}$	$a_2 \frac{h^5}{120}$	$a_2 \frac{h^6}{720}$
$a_3 f''(x_{j-1})$	0	0	a_3	$-a_3 h$	$a_3 \frac{h^2}{2}$	$-a_3 \frac{h^3}{6}$	$a_4 \frac{h^4}{24}$
$a_4 f''(x_{j+1})$	0	0	a_4	$a_4 h$	$a_4 \frac{h^2}{2}$	$a_4 \frac{h^3}{6}$	$a_4 \frac{h^4}{24}$
Total	0	0	0	0	0	??	??

Then, by solving the system of five linear equations:

$$
\begin{aligned}
a_0 + a_1 + a_2 &= 0 \\
-a_1 h + a_2 h &= 0 \\
1 + a_1 \frac{h^2}{2} + a_2 \frac{h^2}{2} + a_3 + a_4 &= 0 \\
-a_1 \frac{h^3}{6} + a_2 \frac{h^3}{6} - a_3 h + a_4 h &= 0 \\
a_1 \frac{h^4}{24} + a_2 \frac{h^4}{24} + a_3 \frac{h^2}{2} + a_4 \frac{h^2}{2} &= 0
\end{aligned}
$$

The solution of this system is

$$a_0 = \frac{12}{5h^2} \qquad a_1 = a_2 = \frac{-12}{10h^2} \qquad a_3 = a_4 = \frac{1}{10}$$

Resulting in

$$f''(x_j) + \frac{24}{10h^2} f(x_j) - \frac{12}{10h^2} f(x_{j-1}) - \frac{12}{10h^2} f(x_{j+1}) + \frac{1}{10} f''(x_{j-1}) + \frac{1}{10} f''(x_{j+1}) = TE$$

The truncation error can be found by summing the last column in the Taylor table which is $\frac{5}{1000} h^4 f^{(6)}(x_j)$ (the fifth column sum is zero hence we sum the last column). Hence, the truncation error is $O(h^4)$. The pade approximation to the second derivative is 4th-order accurate.

Problem 2

The following IVP is to be numerically integrated to obtain y at t = 500 s:

$$\dot{y} = \left(\frac{-1}{\tau} + i\omega\right) y; y(0) = 1.$$

Use the θ method for solving $\dot{y} = f(y)$ as follows:

$$y_{n+1} = y_n + h[\theta f_{n+1} + (1-\theta)f_n]$$

a) Using the model problem $\dot{y} = \lambda y$ perform a stability analysis for $\theta = [0, \frac{1}{2}, 1]$ choice of θ affect the stability? Sketch the stability regions for the three cases. Recall that we are assuming that λ_R < 0 i.e the exact solution is bounded.

From the θ approximation we get that

$$\begin{aligned} y_{n+1} &= y_n + h[\theta f_{n+1} + (1-\theta)f_n] \\ y_{n+1} &= y_n + h\theta\lambda y_{n+1} + h(1-\theta)\lambda y_n \\ y_{n+1}(1-h\theta\lambda) &= y_n(1+h\lambda - h\theta\lambda) \\ \sigma = \frac{y_{n+1}}{y_n} &= \frac{1+h\lambda-h\theta\lambda}{1-h\theta\lambda} = 1 + \frac{h\lambda}{1-h\theta\lambda} \end{aligned}$$

The IVP is stable only if $\lambda = \lambda_R + i\lambda_I, \lambda_R < 0$ and the numerical method is stable only if the amplification factor, σ, is with less than or equal to 1.

Case #1 : $\theta = 0$.

$$|\sigma|^2 = [1 + h(\lambda_R + i\lambda_I)]^2 = (1+\lambda_R h)^2 + (\lambda_I h)^2 = (\lambda_R h - (-1))^2 + (\lambda_I h)^2$$

This is the explicit Euler's (EE) method. Therefore, the region of convergence is a circle in the complex plane ($\lambda_R h$, $\lambda_I h$) centered at (-1, 0) with a radius 1.

Case #2: $\theta = \frac{1}{2}$.

$$\sigma = 1 + \frac{2\lambda h}{2-\lambda h} = \frac{2-\lambda h + 2\lambda h}{2-\lambda h} = \frac{2+\lambda h}{2-\lambda h} = \frac{(2+\lambda_R h) + (\lambda_I h)i}{(2-\lambda_R h) - (\lambda_I h)i} \cdot \frac{(2-\lambda_R h) + (\lambda_I h)i}{(2-\lambda_R h) + (\lambda_I h)i}$$

For convenience, let's label $\lambda_R h = x$ & $\lambda_I h = y$.

$$\sigma = \frac{(2+x)+yi}{(2-x)-yi} \cdot \frac{(2-x)+yi}{(2-x)+yi} = \frac{(4-x^2) + (2+x)yi + (2-x)yi - y^2}{(2-x)^2+y^2} = \frac{(4-x^2-y^2)+4yi}{(2-x)^2+y^2}$$

Then, the modulus square of the amplification factor σ is $|\sigma|^2 = [\frac{4-x^2-y^2}{(2-x)^2+y^2}]^2 + [\frac{4y}{(2-x)^2+y^2}]^2 \leq 1$

$$\frac{x}{(x-2)^2+y^2} \leq 0$$
$$x \leq 0$$

Hence, it is unconditionally stable for $\theta = \frac{1}{2}$. This is called the trapezoid method. The region of stability is when $\lambda_R h \leq 0$, similar to the analytic stability requirement.

Case #3: $\theta = 1$.

$$\sigma = 1 + \frac{\lambda h}{1-\lambda h} = \frac{1}{1-\lambda h} = \frac{1}{(1-\lambda_R h)-(\lambda_I h)i} \cdot \frac{(1-\lambda_R h)+(\lambda_I h)i}{(1-\lambda_R h)+(\lambda_I h)i} = \frac{(1-\lambda_R h)-(\lambda_I h)i}{(1-\lambda_R h)^2+(\lambda_I h)^2}$$

Therefore, the region of convergence is defined as

$$(1-x)^2+y^2 \leq ((1-x)^2+y^2)^2$$
$$-x^4+4x^3-2x^2y^2-5x^2+4xy^2+2x-y^4-y^2 \leq 0$$

This holds for the entire plane. This is in fact the implicit Euler (IE) formula. It is unconditionally stable. The region of stability is the entire second and third quadrant, essentially $\lambda_R h \leq 0$.

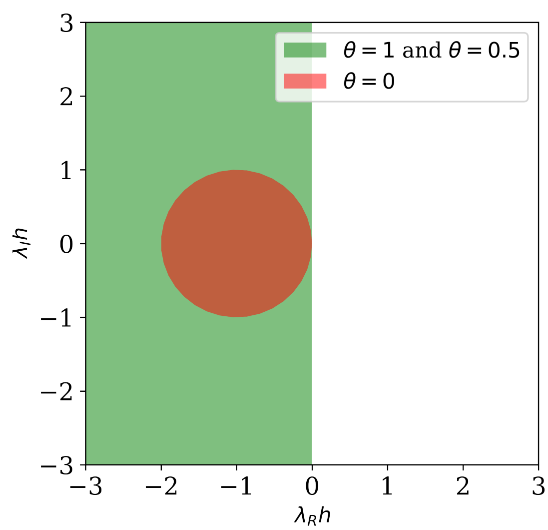

Fig 3. Region of stability for the three cases of $\theta = [0, 0.5, 1]$.

b) Can θ method be implemented with θ larger than 1? Derive the amplification factor for θ method for θ ≫ 1. Comment on its suitability for solving Eq.(1).

If $\theta > 1$ then the weighted mean between f_{n+1} and f_n no longer holds. More tersely, the weight of f_n is negative. For $\theta \gg 1$, the amplification number becomes

$$\sigma = \lim_{\theta \to \infty} 1 + \frac{h\lambda}{1-h\theta\lambda} = 1$$

With that, as mentioned earlier, the theta method is only suitable for values of theta between 0 and 1, since the purpose of the method is to take a weighted average of the two points.

c) Write a computer program that implements the θ method. Use it to obtain the solution for the initial times 0 ≤ t ≤ 50 and the later times 450 ≤ t ≤ 500 with a user-prescribed θ, ω and h. Let ω = 0.2 s^{-1}, τ = 400 s and consider two cases: θ = 1/4 and θ = 4/5. Vary the time step h = 0.01, 0.08, 1 s and discuss the influence on the solution. Compare the two methods to the exact solution by plotting the amplitude and phase.

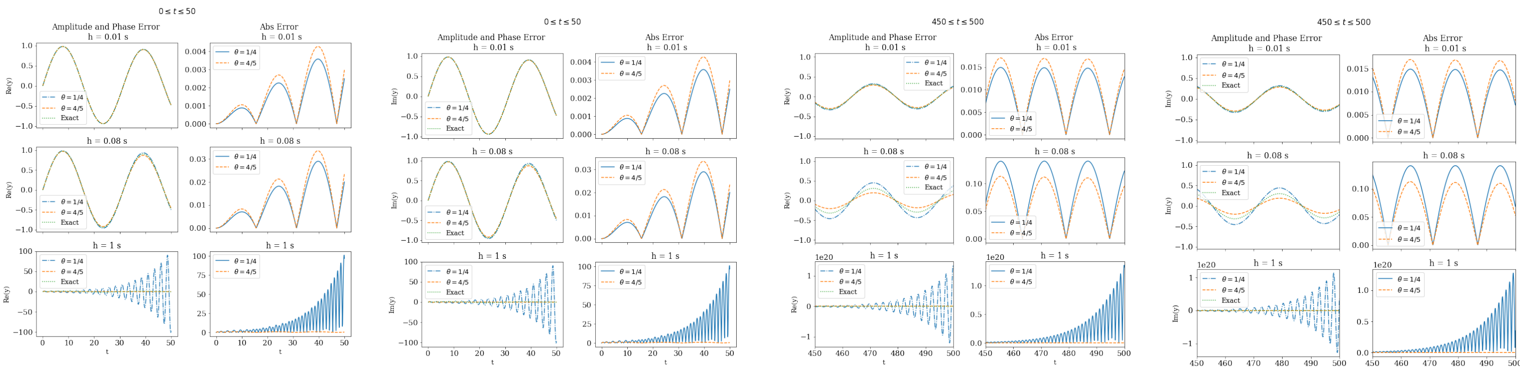

Fig 4. The absolute error between the numerical approximation and the exact solution (real and imaginary parts) using $\theta = \frac{1}{4}, \frac{4}{5}$ and $h = 0.01, 0.08, 1s$.

As seen in Fig 4., as we decrease the step size h, we are able to approximate the exact solution better. In particular, using $h = 0.08s$ vs. $h = 0.01s$ results in an absolute error that is approximately a factor of 10 larger, respectively. When we set $h = 1s$ the algorithm becomes <u>unstable</u> and the solution diverges (in the case when $\theta = 1/4$) or converges to 0 (in the case when $\theta = 4/5$). In the computer program, I also compared the difference between using $\theta = 4/5$ and $\theta = 1/4$. In the case where h = 0.01s, it is apparent that $\theta = 1/4$ outperforms $\theta = 4/5$ in the real and imaginary components by evaluating the absolute error between the two. Whereas, in the case where h = 0.08s, $\theta = 1/4$ outperforms $\theta = 4/5$ in short term prediction $0 \leq t \leq 50$ and in the later times $450 \leq t \leq 500$ the opposite holds.

d) Assume the characteristic decay time τ → ∞ so that Eq. (1) simplifies to $y' = i\omega y$; y(0)=1. Analytically obtain the amplitude and phase error of the method with θ = 3/4 and time step h. Simplify the error expressions assuming that ωh << 1.

$$\sigma = \frac{y_{n+1}}{y_n} = \frac{1 + h\lambda - h\theta\lambda}{1 - h\theta\lambda} = \frac{1 + ih\omega - 3/4ih\omega}{1 - i3/4h\omega} = \frac{1 + 1/4ih\omega}{1 - i3/4h\omega}\frac{1 + i3/4h\omega}{1 + i3/4h\omega} = \frac{16ih\omega}{16 + 9h^2\omega^2} + \frac{16 - 3h^2\omega^2}{16 + 9h^2\omega^2}$$

$$|\sigma| = \frac{\sqrt{1 + 1/16h^2\omega^2}}{\sqrt{1 + 9/16h^2\omega^2}}$$

The amplitude error

$$|\sigma| - 1 = \frac{\sqrt{1 + 1/16h^2\omega^2}}{\sqrt{1 + 9/16h^2\omega^2}} - 1$$

For $\omega h \ll 1$, the amplitude error is

$$|\sigma| - 1 = -(h^2\omega^2)/4 + (7h^4\omega^4)/64 + O(h^5)$$

The phase error

$$\omega h - \theta = \omega h - \tan^{-1}(\frac{Im(\sigma)}{Re(\sigma)}) = \omega h - \tan^{-1}(\frac{16h\omega}{16 - 3h^2\omega^2})$$

For $\omega h \ll 1$, the phase error is

$$PE = \omega h - \theta = \omega h - \frac{16h\omega}{16 - 3h^2\omega^2} - 1/3(\frac{16h\omega}{16 - 3h^2\omega^2})^3 + \ldots = \frac{-3h^3w^3}{16 - 3h^2\omega^2} + \ldots$$

Problem 3

3. A third-order Runge-Kutta scheme (RK3) is used to integrate the model linear problem: $\dot{y} = \lambda y$;$y(0) = 1$.
a) Obtain the stability restriction on the time step h for λ ∈ C. Plot the solution as a stability diagram. What is the restriction on h when λ ∈ R?

By RK3, the relation between y_{n+1} and y_n is

$$y_{n+1} = y_n[1 + \lambda h + \frac{\lambda^2 h^2}{2} + \frac{\lambda^3 h^3}{6}]$$

The above confirms that the method is third-order accurate. For stability, we must have $|\sigma| \leq 1$, where

$$\sigma = 1 + \lambda h + \frac{\lambda^2 h^2}{2} + \frac{\lambda^3 h^3}{6}$$

For complex $\lambda \in \mathbb{C}$ with $\lambda = \lambda_I i$,

$$\sigma = 1 - \frac{\lambda_I^2 h^2}{2} + i(\lambda_I h - \frac{\lambda_I^3 h^3}{6})$$

$$|\sigma| = \sqrt{(1 - \frac{\lambda_I^2 h^2}{2})^2 + (\lambda_I h - \frac{\lambda_I^3 h^3}{6})^2} = \sqrt{1 - 3/36 h^4 \lambda_I^4 + 1/36 h^6 \lambda_I^6}$$

The method is conditionally stable in the regions where $\sqrt{1 - 3/36h^4\lambda_I^4 + 1/36h^6\lambda_I^6} \leq 1$ for purely imaginary λ.
In particular, $|h^2\lambda_I^2| \leq 3$;

$$|h| \leq \sqrt{\frac{3}{\lambda_I^2}}$$

For real $\lambda \in \mathbb{R}$, $\sigma = 1 + \lambda_R h + \frac{\lambda_R^2 h^2}{2} + \frac{\lambda_R^3 h^3}{6}$. Hence, the stability restriction is $|1 + \lambda_R h + \frac{\lambda_R^2 h^2}{2} + \frac{\lambda_R^3 h^3}{6}| \leq 1$, leading to $|\lambda_R h| \leq 2.513$. Hence, for the algorithm to be stable $|h| < \frac{2.153}{|\lambda_R|}$ must hold.

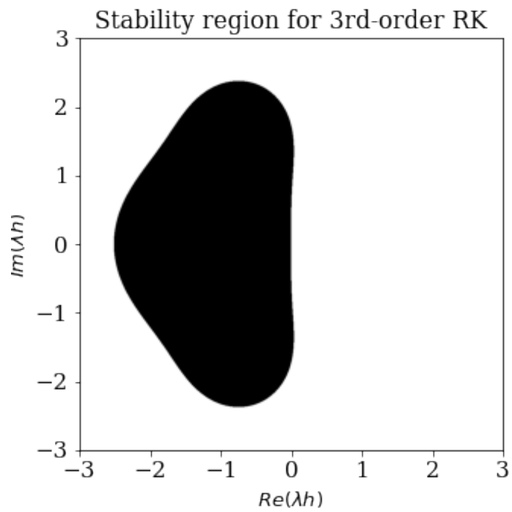

Fig 5. The stability region on the complex plane of λh of the 3rd-order RK method to solve the IVP $\dot{y} = \lambda y$.

b) Implement the RK3 scheme to solve Equation (1). Verify the order of truncation error in your RK3 solution by comparing it to the exact solution.

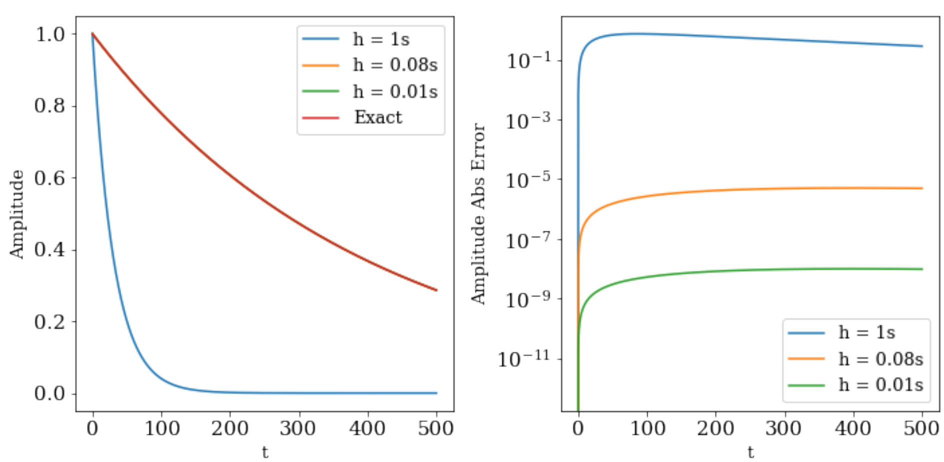

Fig 6. The RK3 scheme to solve Equation (1) in the homework using different discretization intervals.

The RK3 method is 3rd-order accurate. We can verify this by the numerical results presented in Fig 6. For all results we set $\lambda = -1/\tau + \omega i$, where $\omega = 0.2$ and $\tau = 400$. The magnitude of λ is approximately 0.2. In the case where

$h = 0.01s$, we expect the error to plateau at $(h\lambda)^3/6$ which is 8e-09 as seen in Fig 5. Similarly, for $h = 0.08s$ and $h = 1s$ the amplitude error reaches a plateau at 4.096e-06 and 0.008, respectively.