Assignment 2

Zeyuan Wang

February 14, 2024

1 Problems

1.1

A numerical method is called A-stable if its region of absolute stability include the entire complex half-plane with negative real part, where absolute stability concern the stability for a fixed Δt when $t \to \infty$. For the simplification sake, we only discuss linear ODE here, modeled as

$$y'(t) = \lambda y(t), \ y(t_0) = y_0, \ \lambda \in \mathbb{C} \tag{1}$$

Note that it only makes sense to consider the stability for a numerical solution when the ODE itself is stable, so we should constraint $Re(\lambda) \le 0$ here.

1. For Backward Euler, we have

$$
\begin{aligned}
y_{n+1} &= y_n + \Delta t f(t_{n+1}, y_{n+1}) \\
&= y_n + \Delta t \lambda y_{n+1} \\
&= \frac{1}{1 - \lambda \Delta t} y_n
\end{aligned} \tag{2}
$$

By induction we get $y_{n+1} = \frac{1}{(1-\lambda \Delta t)^n} y_0$. Clearly to make the solution decay as the original problem does, we need $\frac{1}{|1-\lambda \Delta t|^n} < 1$, or equivalently $|1 - \lambda \Delta t| > 1$. Obviously for any $\lambda \in \mathbb{C}$, $Re(\lambda) < 0$, this condition is always satisfied, so Backward Euler is A-stable by definition is this case.

2. For Trapezoidal Rule, we have

$$
\begin{aligned}
y_{n+1} &= y_n + \frac{\Delta t}{2}[f(t_n, y_n) + f(t_{n+1}, y_{n+1})] \\
&= y_n + \frac{\lambda \Delta t}{2}[y_n + y_{n+1}] \\
&= \frac{|1 + \frac{\lambda \Delta t}{2}|}{|1 - \frac{\lambda \Delta t}{2}|} y_n
\end{aligned}
\tag{3}
$$

By induction we get $y_{n+1} = (\frac{|1+\frac{\lambda\Delta t}{2}|}{|1-\frac{\lambda\Delta t}{2}|})^n y_0$. Similarly we want $\frac{|1+\frac{\lambda\Delta t}{2}|}{|1-\frac{\lambda\Delta t}{2}|} < 1$.

$$
\begin{aligned}
\frac{|1 + \frac{\lambda \Delta t}{2}|}{|1 - \frac{\lambda \Delta t}{2}|} &< 1 \\
|1 + \frac{\lambda \Delta t}{2}| &< |1 - \frac{\lambda \Delta t}{2}| \\
(1 + \frac{\lambda \Delta t}{2})^2 &< (1 - \frac{\lambda \Delta t}{2})^2 \\
1 + \lambda \Delta t + \frac{(\lambda \Delta t)^2}{4} &< 1 - \lambda \Delta t + \frac{(\lambda \Delta t)^2}{4} \\
\lambda \Delta t &< 0
\end{aligned}
\tag{4}
$$

Since Δt is always greater than zero so for any $\lambda \in \mathbb{C}$, $Re(\lambda) < 0$, excepting the λ which makes $1 - \frac{\lambda \Delta t}{2}$ equals 0, we have $\lambda \Delta t < 0$. Thus the Trapezoidal Rule is also A-stable.

1.2

(a) See file **AB2TestScript_a.jl** line 9.

(b) See file **AB2TestScript_b.jl** line 9.
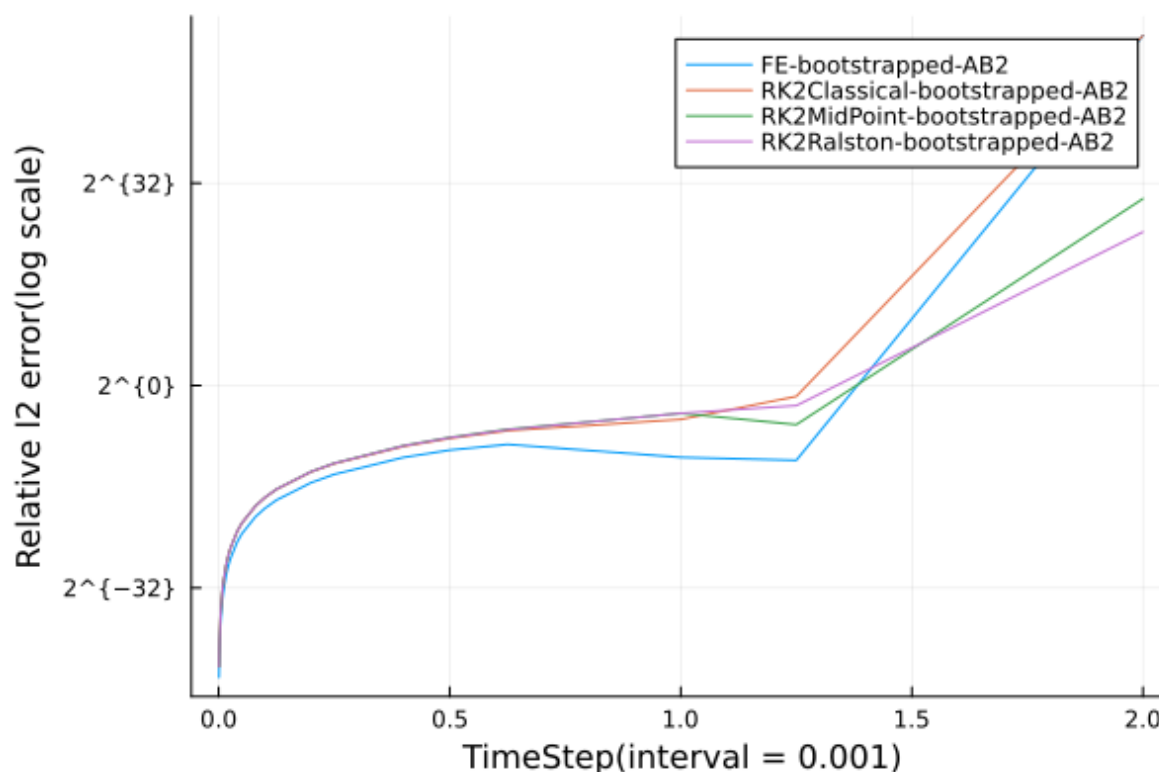
(c) See file **AB2TestScript_c.jl**

Figure 1: Accuracy measure(L2) for different bootstrap strategies

From the graph we can see that both methods convergence to the exact solution when $\Delta t \to 0$. When $\Delta t \in (0, 1]$, there is no significant difference since the error is dominated by the global truncation error of AB2, which is $O(h^2)$. Forward Euler bootstrapped performs better than all types of RK2. The reason is due to the damping property of RK2. When $\Delta t > 1.5$, AB2 stop providing meaningful result since it's not stable enough to allow larger timestep.

1.3

(a) See file **AB3TestScript.jl**.

(b) If possible, neither AB2 nor AB3 should be used to solve very stiffness problem. If choice have to be made, I would choose AB2 because its stability region is slightly larger and it's cheaper.

(c) I would choose AB3 since the stability region of AB3 covers more imaginary axis.

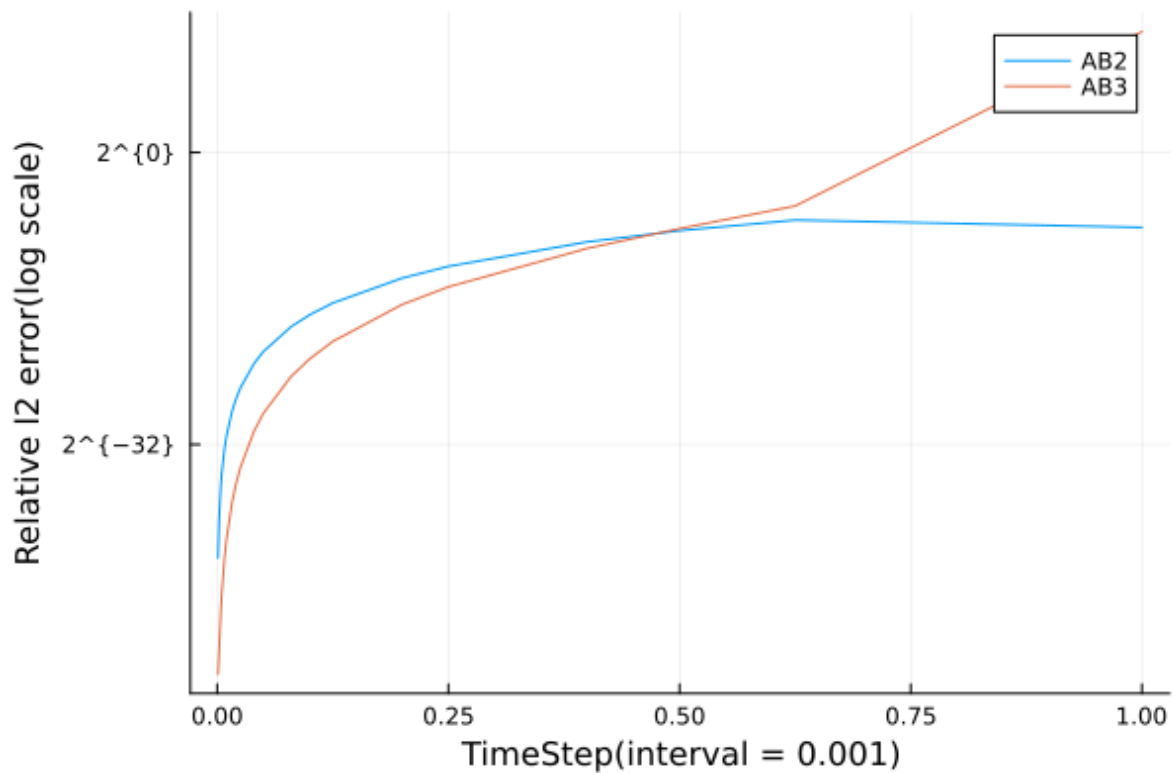(d) Since AB3 is less stable than AB2, we should focus on when $\Delta t < 0.5$.

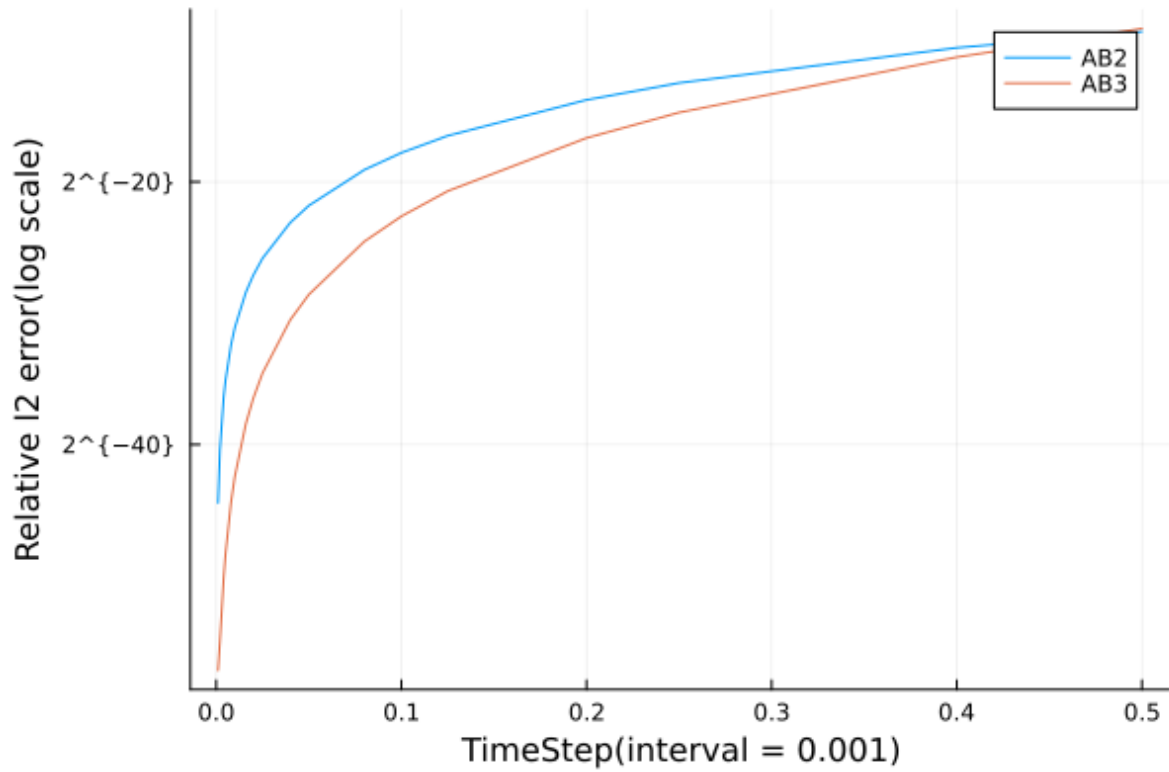Figure 2: Accuracy measure(L2) for AB2/AB3 $\Delta t \in (0, 1]$

Figure 3: Accuracy measure(L2) for AB2/AB3 $\Delta t \in (0, 0.5]$

1.4

(a) Consider the Taylor expansion of y_{n+1} with 4th order remainder:

$$y_{n+1} = y_n + \Delta t \frac{\partial}{\partial t} y_n(t) + \frac{\Delta t^2}{2} \frac{\partial^2}{\partial t^2} y_n(t) + \frac{\Delta t^3}{6} \frac{\partial^3}{\partial t^3} y_n(t) + O(\Delta t^4) \tag{5}$$

From the ODE we have $\frac{\partial}{\partial t}y = f(t, y)$. Let's obtain the expression for $\frac{\partial^2}{\partial t^2}y$ and $\frac{\partial^3}{\partial t^3}y$ respectively:

$$
\begin{aligned}
\frac{\partial^2}{\partial t^2}y &= \frac{\partial}{\partial t}f(t, y) \\
&= f_t(t, y) + f_y(t, y)\frac{\partial}{\partial t}y \\
&= f_t(t, y) + f_y(t, y)f(t, y)
\end{aligned}
\tag{6}
$$

$$
\begin{aligned}
\frac{\partial^3}{\partial t^3}y &= \frac{\partial}{\partial t}\left(\frac{\partial^2}{\partial t^2}y\right) \\
&= \frac{\partial}{\partial t}[f_t(t, y) + f_y(t, y)f(t, y)] \\
&= \frac{\partial}{\partial t}[f_t(t, y)] + \frac{\partial}{\partial t}[f_y(t, y)]f(t, y) + f_y(t, y)\frac{\partial}{\partial t}[f(t, y)] \\
&= f_{tt} + f_{ty}\frac{\partial}{\partial t}y + (f_{yt} + f_{yy}\frac{\partial}{\partial t}y)f + f_y(f_t + f_y f) \\
&= f_{tt} + 2f_{ty}f + f_{yy}ff + f_y f_t + f_y f_y f
\end{aligned}
\tag{7}
$$

Here we assume that both f_{ty} and f_{yt} are smooth so $f_{ty} = f_{yt}$. **To make the result less messy, we denote f^n as $f(t, y)$ evaluate at t_n, y_n.** Plugging back into the Taylor series for y_{n+1}, we get

$$
\begin{aligned}
y_{n+1} = y_n &+ \Delta t f^n + \frac{\Delta t^2}{2}[f_t^n + f_y^n f^n] \\
&+ \frac{\Delta t^3}{6}[f_{tt}^n + 2f_{ty}^n f^n + f_{yy}^n f^n f^n + f_y^n f_t^n + f_y^n f_y^n f^n] \\
&+ O(\Delta t^4)
\end{aligned}
\tag{8}
$$

The general form of RK3 method should look like:

$$
y_{n+1} = y_n + \Delta t(b_1 k_1 + b_2 k_2 + b_3 k_3) + O(\Delta t^4)
\tag{9}
$$

$$
k_1 = f(t_n, y_n)
\tag{10}
$$

$$k_2 = f(t_n + c_2 \Delta t, y_n + \Delta t a_{21} k_1) \tag{11}$$

$$k_3 = f(t_n + c_3 \Delta t, y_n + \Delta t a_{31} k_1 + \Delta t a_{32} k_2) \tag{12}$$

We then Taylor expand k_2 and k_3 respectively.

$$
\begin{aligned}
k_2 &= f(t_n + c_2 \Delta t, y_n + \Delta t a_{21} k_1) \\
&= f^n + c_2 \Delta t f_t^n + \Delta t a_{21} f_y^n k_1 + O(\Delta t^2) \\
&= f^n + c_2 \Delta t f_t^n + \Delta t a_{21} f_y^n f^n + O(\Delta t^2)
\end{aligned}
\tag{13}
$$

$$
\begin{aligned}
k_3 &= f(t_n + c_3 \Delta t, y_n + \Delta t a_{31} k_1 + \Delta t a_{32} k_2) \\
&= f^n + c_3 \Delta t f_t^n + (a_{31} k_1 + a_{32} k_2) \Delta t f_y^n \\
&= f^n + c_3 \Delta t f_t + [a_{31} f^n + a_{32}(f^n + c_2 \Delta t f_t^n + \Delta t a_{21} f_y^n f^n)] \Delta t f_y^n + O(\Delta t^2)
\end{aligned}
\tag{14}
$$

Put (13) (14) back to (9), we get

$$
\begin{aligned}
y_{n+1} &= y_n + \Delta t \{ b_1 f^n + b_2(f^n + c_2 \Delta t f_t^n + \Delta t a_{21} f_y^n f^n) + b3(f^n + c_3 \Delta t f_t \\
&\quad + [a_{31} f^n + a_{32}(f^n + c_2 \Delta t f_t^n + \Delta t a_{21} f_y^n f^n)] \Delta t f_y^n) \} \\
&= y_n + (b_1 + b_2 + b3) \Delta t f^n \\
&\quad + (b_2 c_2 + b_3 c_3) \Delta t^2 f_t^n + (b_2 a_{21} + b_3(a_{31} + a_{32})) \Delta t^2 f^n f_y^n \\
&\quad + b_3 a_{32} c_2 \Delta t^3 f_t^n f_y^n + b_3 a_{21} a_{32} \Delta t^3 f_y^n f_y^n f^n
\end{aligned}
\tag{15}
$$

Compare coefficients with (8), we finally get conditions:

$$b_1 + b_2 + b_3 = 1$$
$$b_2 c_2 + b_3 c_3 = \frac{1}{2}$$
$$b2a_{21} + b_3(a_{31} + a_{32}) = \frac{1}{2} \tag{16}$$
$$b_3 a_{32} c_2 = \frac{1}{6}$$
$$b_3 a_{21} a_{32} = \frac{1}{6}$$

We can take $b_1 = \frac{1}{6}, b_2 = \frac{2}{3}, b_3 = \frac{1}{6}$, $c_2 = \frac{1}{1}$, $c_3 = 1$, $a_{21} = \frac{1}{2}$, $a_{31} = -1$, $a_{32} = 2$, which as known as Kutta's third order method.

(b)

0	0	0	0
1/2	1/2	0	0
1	−1	2	0
	1/6	2/3	1/6

(c) The Butcher tableau of RK 4 3/8 is

0	0	0	0	0
1/3	1/3	0	0	0
2/3	−1/3	1	0	0
1	1	−1	1	0
	1/8	3/8	3/8	1/8

For linear ODE $\frac{dy}{dt} = \lambda y$, for fourth order RK method, we have

$$y_{n+1} = y_n + \Delta t(b_1 k_1 + b_2 k_2 + b_3 k_3 + b_4 k_4)$$

$$y_{n+1} = y_n + \Delta t \mathbf{b}^T \mathbf{k}$$

(17)

where $\mathbf{b} = [b_1, b_2, b_3, b_4]^T$, $\mathbf{k} = [k_1, k_2, k_3, k_4]^T$. Also we have

$$k_1 = \lambda y_n$$

$$k_2 = \lambda y_n + \lambda \Delta t(a_{21} k_1)$$

$$k_3 = \lambda y_n + \lambda \Delta t(a_{31} k_1 + a_{32} k_2)$$

$$k_4 = \lambda y_n + \lambda \Delta t(a_{41} k_1 + a_{42} k_2 + a43 k_3)$$

(18)

Therefore

$$\begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} = \lambda \Delta t \begin{bmatrix} 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ a_{31} & a_{32} & 0 & 0 \\ a_{41} & a_{42} & a_{43} & 0 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} + \lambda \begin{bmatrix} y_n \\ y_n \\ y_n \\ y_n \end{bmatrix}$$

$$\mathbf{k} = \lambda \Delta t \mathbf{A} \mathbf{k} + \lambda y_n \mathbf{1}$$

$$(\mathbf{I} - \lambda \Delta t \mathbf{A})\mathbf{k} = \lambda y_n \mathbf{1}$$

$$\mathbf{k} = (\mathbf{I} - \lambda \Delta t \mathbf{A})^{-1} \lambda y_n \mathbf{1}$$

(19)

So we get

$$y_{n+1} = y_n + \lambda \Delta t \mathbf{b}^T (\mathbf{I} - \lambda \Delta t \mathbf{A})^{-1} y_n \mathbf{1}$$

$$y_{n+1} = [1 + \lambda \Delta t \mathbf{b}^T (\mathbf{I} - \lambda \Delta t \mathbf{A})^{-1} \mathbf{1}] y_n$$

(20)

As usual we need $|1 + \lambda\Delta t\mathbf{b}^T(\mathbf{I} - \lambda\Delta t\mathbf{A})^{-1}\mathbf{1}| < 1$ to maintain stability. $\lambda\Delta t\mathbf{b}^T(\mathbf{I} - \lambda\Delta t\mathbf{A})^{-1}\mathbf{1} < 0$. Again we suppose $\lambda < 0$ so $\mathbf{b}^T(\mathbf{I} - \lambda\Delta t\mathbf{A})^{-1}\mathbf{1} > 0$.

Thankfully $\mathbf{I} - \lambda\Delta t\mathbf{A}$ is lower triangle matrix and is only 4x4, so we can explicitly compute the result

$$
\begin{bmatrix}
1 & 0 & 0 & 0 \\
1 - \lambda\Delta t a_{21} & 1 & 0 & 0 \\
1 - \lambda\Delta t a_{31} & 1 - \lambda\Delta t a_{32} & 1 & 0 \\
1 - \lambda\Delta t a_{41} & 1 - \lambda\Delta t a_{42} & 1 - \lambda\Delta t a_{43} & 1
\end{bmatrix}
\begin{bmatrix}
x_1 \\ x_2 \\ x_3 \\ x_4
\end{bmatrix}
=
\begin{bmatrix}
1 \\ 1 \\ 1 \\ 1
\end{bmatrix}
\tag{21}
$$

$$
\begin{bmatrix}
x_1 \\ x_2 \\ x_3 \\ x_4
\end{bmatrix}
=
\begin{bmatrix}
1 \\
\lambda\Delta t a_{21} \\
\lambda\Delta t a_{31} + x_2(\lambda\Delta t a_{32} - 1) \\
\lambda\Delta t a_{41} + x_2(\lambda\Delta t a_{42} - 1) + x_3(\lambda\Delta t a_{43} - 1)
\end{bmatrix}
\tag{22}
$$

$$
\mathbf{b}^T
\begin{bmatrix}
1 \\
\lambda\Delta t a_{21} \\
\lambda\Delta t a_{31} + x_2(\lambda\Delta t a_{32} - 1) \\
\lambda\Delta t a_{41} + x_2(\lambda\Delta t a_{42} - 1) + x_3(\lambda\Delta t a_{43} - 1)
\end{bmatrix}
> 0
\tag{23}
$$

$$
b_1 + b_2(\lambda\Delta t a_{21}) + b_3[\lambda\Delta t a_{31} + x_2(\lambda\Delta t a_{32} - 1)] + b_4[\lambda\Delta t a_{41} + x_2(\lambda\Delta t a_{42} - 1) + x_3(\lambda\Delta t a_{43} - 1)] > 0
\tag{24}
$$

Take the coefficients of RK4 3/8 method into (24), we finally gain

$$
3\mu^3 + 2\mu^2 + 43\mu + 64 > 0, \mu = \lambda\Delta t
\tag{25}
$$