Numerical Methods for ODEs
Initial Value Problems

Solutions to Odd-Numbered Exercises

David F. Griffiths and Desmond J. Higham

Springer-Verlag, London 2010

Solutions to all exercises are available to approved instructors
by contacting the publishers.

Springer Undergraduate Mathematics Series

ISBN: 978-0-85729-147-9, e-ISBN: 978-0-85729-148-6

Contents

Exercises 1

1. The equation
$$x'(t) = \sin(t) - x(t) \tag{1}$$
has integrating factor $g(t) = \mathrm{e}^{-t}$ and, since $f(t) = \sin(t)$, the variation of constants formula gives
$$\begin{aligned} x(t) &= Ag(t) + g(t)\int_0^t \frac{f(s)}{g(s)}\mathrm{d}s \\ &= A\mathrm{e}^{-t} + \mathrm{e}^{-t}\int_0^t \mathrm{e}^s \sin s \mathrm{d}s. \end{aligned}$$
Integrating by parts twice we find
$$\int_0^t \mathrm{e}^s \sin s \mathrm{d}s = \tfrac{1}{2}\mathrm{e}^t(\sin t - \cos t) + \tfrac{1}{2}$$
so $x(t) = A\mathrm{e}^{-t} + \frac{1}{2}(\sin t - \cos t) + \frac{1}{2}\mathrm{e}^{-t}$.

---

3. Dividing both sides of the differential equation by $x(t)(1 - x(t))$ leads to
$$\int \frac{1}{x(1-x)}\mathrm{d}x = t + A,$$
where A is an arbitrary constant of integration. Using partial fractions on the integrand on the left hand side
$$\int \left(\frac{1}{x} + \frac{1}{1-x}\right)\mathrm{d}x = t + A$$
which gives
$$\log\left|\frac{x}{1-x}\right| = t + A$$
leading to
$$\left|\frac{x}{1-x}\right| = \mathrm{e}^{t+A} = \mathrm{e}^A\mathrm{e}^t$$
and, therefore,
$$\frac{x}{1-x} = \mathrm{e}^{t+A} = C\mathrm{e}^t,$$
where $C = \pm\exp A$ is a new arbitrary constant. Applying the initial condition $x = -1/5$ at $t = 10$, we find $C = -\mathrm{e}^{-10}/6$ then, solving for x gives
$$x = \frac{C\mathrm{e}^t}{1 + C\mathrm{e}^t} = \frac{1}{1 + \mathrm{e}^{-t}/C} = \frac{1}{1 - 6\mathrm{e}^{10-t}}.$$
The denominator approaches zero as $\mathrm{e}^{10-t} \to 1/6$, i.e., when $t = 10 + \log 6$.

---

5. (a) Let $u(t) = \theta(t)$ and $v(t) = \theta'(t)$ so $v'(t) = \theta''(t) = -\theta(t)$ and

$$\begin{aligned} u'(t) &= v(t) \\ v'(t) &= -u(t) \end{aligned} \quad \text{and} \quad \begin{aligned} u(0) &= \pi/10 \\ v(0) &= 0 \end{aligned}.$$

This is a 2-d system $\boldsymbol{x} = \mathbf{f}(t, \boldsymbol{x})$, $\boldsymbol{x}(0) = \boldsymbol{\eta}$ with

$$\boldsymbol{x} = \begin{bmatrix} u \\ v \end{bmatrix}, \qquad \boldsymbol{f}(t, \boldsymbol{x}) = \begin{bmatrix} v \\ -u \end{bmatrix}, \qquad \boldsymbol{\eta} = \begin{bmatrix} \pi/10 \\ 0 \end{bmatrix}.$$

(b) With $u = x(t)$ and $v = x'(t)$ we obtain $\boldsymbol{x}' = \boldsymbol{f}(t, \boldsymbol{x})$, $\boldsymbol{x}(0) = \boldsymbol{\eta}$ with

$$\boldsymbol{f}(t, \boldsymbol{x}) = \begin{bmatrix} v \\ v + 2u + 1 + 2t \end{bmatrix} \quad \text{and} \quad \boldsymbol{\eta} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Note that $\boldsymbol{f}(t, \boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$, where

$$\boldsymbol{x} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 0 \\ 1 + 2t \end{bmatrix}.$$

(c) With $u = x(t)$, $v = x'(t)$ and $w = x''(t)$ we find $u' = v, v' = w$ and $\boldsymbol{x}' = \boldsymbol{f}(t, \boldsymbol{x})$, $\boldsymbol{x}(0) = \boldsymbol{\eta}$ with

$$\boldsymbol{f}(t, \boldsymbol{x}) = \begin{bmatrix} v \\ w \\ 2w + v - 2u + 1 - 2t \end{bmatrix} = A\boldsymbol{x} + \boldsymbol{b},$$

where

$$\boldsymbol{x} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 0 \\ 0 \\ 1 - 2t \end{bmatrix} \text{ and } \boldsymbol{\eta} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

---

7. With $u = t$, $v = x(t)$ and $\boldsymbol{x} = [u, v]^T$, we have

$$\boldsymbol{x}'(t) = \begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} 1 \\ \sin(u) - 2v/(1 + u^4) \end{bmatrix}$$

with initial condition $\boldsymbol{x}(0) = [0, x(0)]^T$. Thus $\boldsymbol{f}(u, v) = [1, \sin(u) - 2v/(1 + u^4)]^T$.

---

9. With $u = x(t)$ and $v = x'(t)$ we obtain $\boldsymbol{x}' = \boldsymbol{f}(t, \boldsymbol{x})$, $\boldsymbol{x}(0) = \boldsymbol{\eta}$ with

$$\boldsymbol{f}(t,\boldsymbol{x}) = \begin{bmatrix} v \\ f(t) + av(t) + bu(t) \end{bmatrix} = A\boldsymbol{x} + \boldsymbol{b}, \text{ where } A = \begin{bmatrix} 0 & 1 \\ b & a \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 0 \\ f(t) \end{bmatrix}.$$

and $\boldsymbol{\eta} = [\xi, \eta]^T$. The characteristic polynomial of A is given by

$$|A - \lambda I| = \begin{vmatrix} -\lambda & 1 \\ b & a - \lambda \end{vmatrix} = \lambda^2 - a\lambda - b.$$

If we suppose that this quadratic function of λ has distinct roots (eigenvalues) λ_1 and λ_2 then the original ODE has complementary function $x(t) = A\mathrm{e}^{\lambda_1 t} + B\mathrm{e}^{\lambda_2 t}$ in which A and B are arbitrary constants.

Notice that the characteristic polynomial of A may be deduced by substituting $x(t) = \mathrm{e}^{\lambda t}$ into the homogeneous form of the ODE.

---

11. Adding together the ODEs for S, C and P we find $S'(t) + C'(t) + P'(t) = 0$, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}(S(t) + C(t) + P(t)) = 0$$

so, on integrating with respect to t, we find that $S(t) + C(t) + P(t) = \text{constant}$.

---

13. Differentiating $u(t) = \frac{1675}{21}\mathrm{e}^{-8t} + \frac{320}{21}\mathrm{e}^{-t/8} + 5$ gives

$$u'(t) = -8\tfrac{1675}{21}\mathrm{e}^{-8t} - \tfrac{40}{21}\mathrm{e}^{-t/8}$$

so

$$u'(t) + 8u(t) = (-\tfrac{40}{21} + 8\tfrac{320}{21})\mathrm{e}^{-t/8} + 40 = 120\mathrm{e}^{-t/8} + 40.$$

Also $u(0) = 100$ so the IVP (1.15) is satisfied by this function $u(t)$.

Similarly,

$$u'(t) + \tfrac{1}{8}u(t) = (-8\tfrac{1675}{21} + \tfrac{1675}{8\times 21})\mathrm{e}^{-8t} + \tfrac{5}{8} = -\tfrac{5025}{8}\mathrm{e}^{-8t} + \tfrac{5}{8}$$

and (1.16) is also satisfied.

---

15. With $x(t) = Xu(\tau)$ and the chain rule

$$x'(t) = \frac{\mathrm{d}}{\mathrm{d}t}x(t) = \frac{\mathrm{d}}{\mathrm{d}t}Xu(\tau) = \left(\frac{\mathrm{d}}{\mathrm{d}\tau}Xu(\tau)\right)\frac{\mathrm{d}\tau}{\mathrm{d}t} = aX\frac{\mathrm{d}}{\mathrm{d}\tau}u(\tau)$$

and $ax(t)(1 - x(t)/X) = aXu(\tau)(1 - u(\tau))$ and so the ODE becomes

$$\frac{\mathrm{d}}{\mathrm{d}\tau}u(\tau) = u(\tau)(1 - u(\tau)).$$

Exercises 2

1. Euler's method is, in this case,

$$x_{n+1} = x_n + hx'_n$$

with $x'_n = t_n^2 - x_n^2$, $t_0 = 0$ and $x_0 = 1$. Thus $x'_0 = -1$ and, when $h = 0.2$,

$n = 0:$ $t_1 = 0.2$
$x_1 = x_0 + hx'_0 = 0.8$
$x'_1 = t_1^2 - x_1^2 = -0.6$

$n = 1:$ $t_2 = 0.4$
$x_2 = x_1 + hx'_1 = 0.8 - 0.2 \times 0.6 = 0.68$

With $h = 0.1$ we find

n	t_n	x_n	x'_n
0	0.0	1.000	−1.000
1	0.1	0.900	−0.800
2	0.2	0.820	−0.632
3	0.3	0.757	−0.483
4	0.4	0.708	−

---

3. $x'(t) = 1 + t - x(t)$, $t > 0$ with $x(0) = 0$.

Euler's method : $x_{n+1} = x_n + h(1 + t_n - x_n)$, $n = 0, 1, 2, ...$ with $t_n = nh$ and $x_0 = 0$. We find

$$x_1 = h, \qquad x_2 = x_1 + h\,(1 + h - x_1) = 2h, \qquad x_3 = x_2 + h\,(1 + 2h - x_2) = 3h$$

which suggests that $x_n = nh = t_n$: a result that is easily proven since, substituting this into $x_{n+1} = x_n + h\,(1 + t_n - x_n)$ gives $x_{n+1} = x_n + h = (n + 1)\,h = t_{n+1}$ and the proof follows by induction.

This suggests that the exact solution of the IVP is $x(t) = t$ (which clearly satisfies both the ODE and the initial condition).

The second derivative of the exact solution is zero so the LTE given by (2.6) is also zero. Theorem 2.4 is relevant with $\lambda = -1$ and $g(t) = 1 + t$ so, with $T_j = 0$ in (2.16), the global error e_n is zero at time t_n.

---

5. With $u = x(t)$ and $v = x'(t)$ we have

$$\begin{bmatrix} u \\ v \end{bmatrix}' = \begin{bmatrix} v \\ t^2 - u(4 + v) \end{bmatrix}, \qquad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

With $h = 0.1$, $u'_0 = 1$ and $v'_0 = 0$ (from the ODEs)

$$n = 0: \quad \begin{aligned} t_1 &= 0.1, \\ u_1 &= u_0 + 0.1u'_0 = 0.1, \\ v_1 &= v_0 + 0.1v'_0 = 1, \\ u'_1 &= \ \ 1.0, \quad v'_1 = -0.49, \end{aligned} \qquad\qquad n = 1: \quad \begin{aligned} t_2 &= 0.2, \\ u_2 &= 0.2, \\ v_2 &= 0.951. \end{aligned}$$

---

7. (a) With $u = x(t)$ and $v = x'(t)$ we have

$$\begin{bmatrix} u \\ v \end{bmatrix}' = \begin{bmatrix} v \\ t^2 - 2u - 3v \end{bmatrix}, \qquad \begin{bmatrix} u(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Euler's method for this system is

$$\begin{bmatrix} u_{n+1} \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} u_n \\ v_n \end{bmatrix} + h \begin{bmatrix} v_n \\ t_n^2 - 2u_n - 3v_n \end{bmatrix}, \qquad \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

with $t_n = nh$. Then $u_{n+1} \approx x(t_{n+1})$ and $v_{n+1} \approx x'(t_{n+1})$

(b) Differentiating the first equation and substituting for y' from the second:

$$x'' = y' - 2x' = (t^2 - y) - 2x'$$

but, from the first ODE $y = x' + 2x$, so $x'' = t^2 - 3x' + 2x$, which is a re-arrangement of equation (2.18).
$y(0) = x'(0) + 2x(0) = 2$ from the given initial conditions.

(c) Euler's method for this system is

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} x_n \\ y_n \end{bmatrix} + h \begin{bmatrix} y_n - 2x_n \\ t_n^2 - 2y_n \end{bmatrix}, \qquad \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

(d) For the method in part (a):

$$\begin{aligned} u_1 &= u_0 + hv_0 = 1, \\ v_1 &= v_0 + h(t_0^2 - 3v_0 - 2u_0) = -2h, \\ u_2 &= u_1 + hv_1 = 1 - 2h^2. \end{aligned}$$

For the method in part (c):

$$\begin{aligned} x_1 &= x_0 + h(y_0 - 2x_0) = 1, \\ y_1 &= y_0 - 2h = 2 - 2h, \\ x_2 &= 1 + h(y_1 - 2x_1) = 1 - 2h^2 \end{aligned}$$

and so $u_2 = x_2$ and both are approximations of $x(t_2)$.

---

9. Summing both sides of

$$\frac{e_j}{(1+\lambda h)^j} - \frac{e_{j-1}}{(1+\lambda h)^{j-1}} = \frac{T_j}{(1+\lambda h)^j}.$$

from $j = 1$ to $j = n$ gives the telescoping sum

$$\left(\frac{e_1}{(1+\lambda h)} - \frac{e_0}{(1+\lambda h)^0}\right) + \left(\frac{e_2}{(1+\lambda h)^2} - \frac{e_1}{(1+\lambda h)^1}\right) + \cdots + \left(\frac{e_n}{(1+\lambda h)^n} - \frac{e_{n-1}}{(1+\lambda h)^{n-1}}\right) = \sum_{j=1}^{n} \frac{T_j}{(1+\lambda h)^j}$$

in which there are multiple cancellations (which are easier to see if the terms on the left are written in decreasing order, from $j = n$ to $j = 1$), leaving

$$-\frac{e_0}{(1+\lambda h)^0} + \frac{e_n}{(1+\lambda h)^n} = \sum_{j=1}^{n} \frac{T_j}{(1+\lambda h)^j}$$

from which (2.16) follows since $e_0 = 0$.

Exercises 3

1. $x'(t) = 2x(1-x)$ where $x \equiv x(t)$ and so

$$x''(t) = \frac{\mathrm{d}}{\mathrm{d}t} 2x(1-x) = 2(1-2x)x'(t) = 4x(1-2x)(1-x).$$

Hence, the TS(2) method is

$$\begin{aligned} x_{n+1} &= x_n + hx_n' + \tfrac{1}{2}h^2 x_n'' \\ x_{n+1}' &= 2x_{n+1}(1-x_n) \\ x_{n+1}'' &= 2x_{n+1}'(1-2x_n) \end{aligned}$$

n	t_n	x_n	x_n'	x_n''	$x(t_n)$
0	10.0000	0.2000	0.3200	0.3840	0.2000
1	10.5000	0.4080	0.4831	0.1778	0.4046
2	11.0000	0.6718	0.4410	−0.3030	0.6488

The rightmost column gives the exact solution of the IVP. The GE with TS(2) at $t = 11$ is $x(11) - x_2 = -0.022$ while that for Euler's method with $h = 0.2$ is $0.6488 - 0.6295 = 0.0193$ (see Table 2.2). The GEs are therefore roughly comparable—the fact that only 2 steps were required here compared to 5 steps in Example 2.2 is offset to a great extent by the need to compute x_n''.

---

3. We apply Euler's method to the IVP's of Exercise 1.1. For systems, Euler's Method (TS(1)) reads

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + h\boldsymbol{f}(t_n, \boldsymbol{x}_n), \quad n = 0, 1, ...$$

From Exercise 1.5a), $u = \theta(t)$ and $v = \theta'(t)$ and we have the 2–d system $\boldsymbol{x} = \boldsymbol{f}(t, \boldsymbol{x})$, $\boldsymbol{x}(0) = \boldsymbol{\eta}$ with

$$\boldsymbol{x} = \begin{bmatrix} u \\ v \end{bmatrix}, \qquad \boldsymbol{f}(t, \boldsymbol{x}) = \begin{bmatrix} v \\ -u \end{bmatrix}, \qquad \boldsymbol{\eta} = \begin{bmatrix} \pi/10 \\ 0 \end{bmatrix}.$$

So Euler's Method is, in terms of the components:

$$u_{n+1} = u_n + hv_n, \qquad v_{n+1} = v_n - hu_n$$

with initial values $u_0 = \pi/10 = .31416$ and $v_0 = 0$. We use $h = 0.1$.

n	t_n	u_n	v_n	
0	0	0.3142	0.0000	(starting values)
1	0.1	$u_1 = u_0 + .1v_0 = 0.3142$	$v_1 = v_0 - .1u_0 = -0.0314$	
2	0.2	$u_2 = u_1 + .1v_1 = 0.3110$	$v_2 = v_1 - .1u_1 = -0.0628$	

In order to define TS(2) we require the second derivatives of the exact solution. Since $u' = v$ and $v' = -u$ we find $u'' = v' = -u$ and $v'' = -u' = -v$ so

$$u_{n+1} = u_n + hv_n - \tfrac{1}{2}h^2 u_n, \qquad v_{n+1} = v_n - hu_n - \tfrac{1}{2}h^2 v_n \tag{2}$$

n	t_n	u_n	v_n	
0	0	0.3142	0.0000	(starting values)
1	0.1	$u_1 = 0.3126$	$v_1 = -0.0314$	
2	0.2	$u_2 = 0.3079$	$v_2 = -0.0625$	

Comparing the results of Euler's method and TS(2) at $t = 0.2$ in one table:

	Euler	TS(2)
$\boldsymbol{u}_2 =$	$\begin{bmatrix} 0.3110 \\ -0.0628 \end{bmatrix}$	$\begin{bmatrix} 0.3079 \\ -0.0625 \end{bmatrix}$

---

5. (a) The exact solution of the IVP is $x(t) = \mathrm{e}^{\lambda t}$ so $\lambda < 0$ is necessary to ensure $x(t) \to 0$ as $t \to \infty$.

(b) For Euler's method $x_{n+1} = (1 + h\lambda)x_n$ from which it follows that

$$x_n = (1 + h\lambda)^n$$

since $x_0 = 1$. Hence $|x_n| \to 0$ as $n \to \infty$ (which corresponds to $t_n \to \infty$ if h is a fixed number) if, and only if $|1 + h\lambda| < 1$. This leads to

$$-1 < 1 + h\lambda < 1 \Leftrightarrow -2 < h\lambda < 0.$$

With $\lambda = -1$ this implies $0 < h < 2$, while for $\lambda = -100$ we require the much smaller value $0 < h < 0.02$.
For TS(2) $x_{n+1} = (1 + h\lambda + \frac{1}{2}h^2\lambda^2)x_n$ so

$$x_n = (1 + h\lambda + \tfrac{1}{2}h^2\lambda^2)^n$$

since $x_0 = 1$. Hence $|x_n| \to 0$ as $n \to \infty$ if, and only if

$$-1 < 1 + h\lambda + \tfrac{1}{2}h^2\lambda^2 < 1.$$

Completing the square, we find

$$1 + h\lambda + \tfrac{1}{2}h^2\lambda^2 = \tfrac{1}{2}(1 + h\lambda)^2 + \tfrac{1}{2}$$

so the inequalities require $-3 < (1 + h\lambda)^2 < 1$. The left inequality is always satisfied while the right inequality is equivalent to $h\lambda(2 + h\lambda) < 0$ which requires $-2 < h\lambda < 0$—the same conditions as for TS(1).

---

7. For TS(1) we have to calculate, for each n: t_n (1 flop) x'_n (3 flops—see the previous exercise) and $x_{n+1} = x_n + hx'_n$ (2 flops) giving a grand total of 6 flops.

   For TS(2) we have the additional cost of computing x''_n (3 flops—$(1-2t)$ was evaluated for x' and does not need to be recalculated) and adding $\frac{1}{2}h^2x''_n$ (only 2 flops since the number $\frac{1}{2}h^2$ need only be calculated once, at the start of the exercise, and not on each step). Thus TS(2) costs 5 flops more than TS(1).

   By a similar argument, TS(3) costs 5 flops more than TS(2).

	TS(1)	TS(2)	TS(3)
Cost	6	11	16

   For the data in Table 3.1: assuming that the GE for TS(1) is proportional to h: at $h = .15$ the GE is $-.1148$, so the constant of proportionality is $-.1148/.15 \approx 0.77$. Thus, GE$\approx .77h$ and this will achieve the target GE of 0.01 with $h = .01/.77 \approx 0.013$. Thus, to integrate to $t = 1.2$ will require $1.2/.013 \approx 93$ steps. Each step costs 6 flops, so the cost of this calculation is about $6 \times 93 = 558$ flops.

   The final column in Table 3.1 suggests that the GE for TS(2) $\approx 0.138h^2$. This will achieve the target GE of 0.01 with $h = \sqrt{0.01/0.138} \approx 0.27$. Thus, to integrate to $t = 1.2$ will require $1.2/.27 \approx 5$ steps. Each step costs 11 flops, so the cost of this calculation is about 55 flops.

   Thus TS(1) requires about 10 times as much computational effort as TS(2) in order to achieve the required accuracy.

---

9. By the chain rule

$$x''(t) = \frac{\mathrm{d}}{\mathrm{d}t}x'(t) = \frac{\mathrm{d}}{\mathrm{d}t}f(t,x(t)) = f_t(t,x) + f_x(x,t)x'(t) = f_t(t,x) + f(t,x)f_x(t,x).$$

---

11. Differentiating the ODE $x'(t) = \lambda x(t) + g(t)$ we find, when we evaluate the results at $t = t_n$,

$$\begin{aligned} x'(t_n) &= \lambda x(t_n) + g(t_n) \\ x''(t_n) &= \lambda x'(t_n) + g'(t_n) \\ x'''(t_n) &= \lambda x''(t_n) + g''(t_n), \text{ etc.} \end{aligned}$$

   These derivatives are approximated by

$$\begin{aligned} x'_n &= \lambda x_n + g(t_n) \\ x''_n &= \lambda x'_n + g'(t_n) \\ x'''_n &= \lambda x''_n + g''(t_n), \text{ etc} \end{aligned}$$

and, subtracting the corresponding expressions, we have

$$\begin{aligned} x'(t_n) - x'_n &= \lambda(x(t_n) - x_n) \\ x''(t_n) - x''_n &= \lambda(x'(t_n) - x'_n) \\ x'''(t_n) - x'''_n &= \lambda(x''(t_n) - x''_n), \text{ etc} \end{aligned}$$

from which we deduce that

$$x^{(j)}(t_n) - x_n^{(j)} = \lambda^j (x(t_n) - x_n). \tag{3}$$

From (3.4) at $t = t_n$we have

$$x(t_{n+1}) = x(t_n) + hx'(t_n) + \tfrac{1}{2!}h^2 x''(t_n) + \cdots + \tfrac{1}{p!}h^p x^{(t_n)}(t) + T_{n+1}, \tag{4}$$

where $T_{n+1} = \mathscr{O}(h^{p+1})$, while the TS(p) method is (see equation (3.5))

$$x_{n+1} = x_n + hx'_n + \tfrac{1}{2}h^2 x''_n + \cdots + \tfrac{1}{p!}h^p x_n^{(p)}.$$

Subtracting this from equation (4) and using (3) we obtain

$$e_{n+1} = r(\lambda h)e_n + T_{n+1}, \qquad n = 0, 1, 2, \ldots$$

where $e_n = x(t_n) - x_n$ and $r(s) = 1 + s + \frac{1}{2!}s^2 + \cdots + \frac{1}{p!}s^p$.
Now, following the derivation of (2.16), we find

$$e_n = \sum_{j=1}^{n} r(\lambda h)^{n-j} T_j$$

and, since $r(s) \le \mathrm{e}^s$ (for $s \ge 0$) by Exercise 3.10,

$$|r(\lambda h)^{n-j}| = |r(\lambda h)|^{n-j} \le r(|\lambda| h)^{n-j} \le \mathrm{e}^{(n-j)|\lambda|h} \le \mathrm{e}^{|\lambda| t_{\mathrm{f}}}.$$

Consequently,

$$|e_n| \le \left| \sum_{j=1}^{n} r(\lambda h)^{n-j} T_j \right| \le \mathrm{e}^{|\lambda| t_{\mathrm{f}}} \sum_{j=1}^{n} |T_j|$$

and, since $|T_j| \le Ch^{p+1}$,

$$|e_n| \le \mathrm{e}^{|\lambda| t_{\mathrm{f}}}\, nCh^{p+1} \le Ch^p t_{\mathrm{f}} \mathrm{e}^{|\lambda| t_{\mathrm{f}}}.$$

Hence, the method is convergent if $p > 0$ since $|e_n| \to 0$ as $h \to 0$ for any time $t_n \in [0, t_{\mathrm{f}}]$. Moreover, the method converges at a pth order rate.

Exercises 4

1. Explicit: Euler, 2–step Adams–Bashforth, Dahlquist. The other four methods are implicit.

---

3. The backward Euler method applied to $x'(t) = 1 + x^2(t)$ leads to

$$x_{n+1} = x_n + h(1 + x_{n+1}^2) \Rightarrow \; hx_{n+1}^2 - x_{n+1} + (h + x_n) = 0.$$

Employing the formula for the roots of a quadratic, we find,

$$x_{n+1} = \frac{1 \pm \sqrt{1 - 4hx_n - 4h^2}}{2h}.$$

With the + sign:

$$x_{n+1} = \frac{1 + \sqrt{1 - 4hx_n - 4h^2}}{2h} \to \frac{1}{h} \to \infty$$

as $h \to 0$.

With the − sign:

$$\begin{aligned} x_{n+1} &= \frac{1 - \sqrt{1 - 4hx_n - 4h^2}}{2h} \frac{1 + \sqrt{1 - 4hx_n - 4h^2}}{1 + \sqrt{1 - 4hx_n - 4h^2}} \\ &= \frac{4h(h + x_n)}{2h(1 + \sqrt{1 - 4hx_n - 4h^2})} \\ &= \frac{2(h + x_n)}{1 + \sqrt{1 - 4hx_n - 4h^2}} \to x_n \end{aligned}$$

as $h \to 0$. With $x_0 = 0$ we obtain

$$x_1 = \frac{2h}{1 + \sqrt{1 - 4h^2}}.$$

---

5. For backward Euler:

$$\mathscr{L}_h z(t) = z\,(t + h) - z(t) - hz'\,(t + h)\,.$$

Substituting

$$\begin{aligned} z\,(t+h) &= z(t) + hz'(t) + \tfrac{1}{2}h^2 z''(t) + \mathscr{O}(h^3) \\ z'\,(t+h) &= z'(t) + hz''(t) + O(h^2) \end{aligned}$$

we find

$$\begin{aligned}
\mathscr{L}_h z(t) &= \left[z(t) + hz'(t) + \tfrac{1}{2}h^2 z''(t) + \mathscr{O}(h^3)\right] - z(t) - h\left[z'(t) + hz''(t) + O(h^2)\right] \\
&= -\tfrac{1}{2}h^2 z''(t) + \mathscr{O}(h^3)
\end{aligned}$$

So the error constant for backward Euler is $C_2 = -\frac{1}{2}$ and the order is $p = 1$ since $\mathscr{L}_h z(t) = O(h^2) = O(h^{p+1})$.

---

7. (a) $x_{n+2} - ax_{n+1} - 2x_n = hbf_n$ so $\rho(r) = r^2 - ar - 2$ and $\rho(1) = 0 \Leftrightarrow a = -1$. $\sigma(r) = b$ and $\rho'(r) = 2r - a$ so $\rho'(1) = \sigma(1)$ if $2 - a = b$, i.e., $b = 3$, leading to the method $x_{n+2} + x_{n+1} - 2x_n = 3hf_n$.

   (b) $x_{n+2} + x_{n+1} + ax_n = h(f_{n+2} + bf_n)$ so $\rho(r) = r^2 + r + a$ and $\rho(1) = 0 \Leftrightarrow a = -2$. $\sigma(r) = r^2 + b$ and $\rho'(r) = 2r + 1$ so $\rho'(1) = \sigma(1)$ if $3 = 1 + b$, i.e., $b = 2$, leading to the method $x_{n+2} + x_{n+1} - 2x_n = h(f_{n+2} + 2f_n)$.

---

9. For the method $x_{n+1} = x_n + 2hf_n$ the linear difference operator is

$$\begin{aligned}
\mathscr{L}_h z(t) &= z(t+h) - z(t) + 2hz'(t) \\
&= hz'(t) + O\left(h^2\right)
\end{aligned}$$

so $\mathscr{L}_h z(t) = O(h)$ but, for consistency we must have $\mathscr{L}_h z(t) = O(h^{p+1})$ with $p > 0$. Thus the method is not consistent.

Alternatively, based on Definition 4.6, the first and second characteristic polynomials are

$$\rho(r) = r - 1, \quad \sigma(r) = 2$$

from which $\rho(1) = 0$ but $\rho'(1) = 1 \neq \sigma(1) = 2$. Hence, the method is not consistent.

The IVP $x'(t) = 1$, $t \in (0, 1]$, $x(0) = 0$, has the exact solution is $x(t) = t$. The method applied to this IVP gives

$$x_{n+1} = x_n + 2h$$

so $x_1 = 2h$, $x_2 = 4h$, $x_3 = 6h, \ldots$ $x_n = 2nh = 2t_n$ while $x(t_n) = t_n$. The numerical solution x_n is therefore always twice as large as the exact solution and so convergence cannot take place. The Global Error at $t = 1$ is $x(t_n) - x_n = -1$, which does not tend to zero with h.

---

11. The associated linear difference operator is

$$\mathscr{L}_h z(t) = z(t + 2h) + \alpha_1 z(t + h) - az(t) - \beta_2 hz'(t + 2h)$$

and Taylor expansion gives

$$\begin{aligned}
\mathscr{L}_h z(t) &= z(t) + 2hz'(t) + 2h^2 z''(t) + \tfrac{4}{3}h^3 z'''(t) + \mathscr{O}(h^4)\\
&\quad + \alpha_1 \left(z(t) + hz'(t) + \tfrac{1}{2}h^2 z''(t) + \tfrac{1}{6}h^3 z'''(t) + \mathscr{O}(h^4) \right)\\
&\quad - az(t)\\
&\quad - \beta_2 h \left(z' + 2hz''(t) + 2h^2 z'''(t) + \mathscr{O}(h^3) \right)\\
&= (1 + \alpha_1 - a)z(t) + (2 + \alpha_1 - \beta_2)hz'(t) + \left(2 + \tfrac{1}{2}\alpha_1 - 2\beta_2\right) h^2 z''(t)\\
&\qquad + \left(\tfrac{4}{3} + \tfrac{1}{6}\alpha_1 - 2\beta_2\right) h^3 z'''(t) + \mathscr{O}(h^4)
\end{aligned}$$

so that we can choose

$$\alpha_1 = a - 1, \qquad \beta_2 = 2 + \alpha_1 = 1 + a$$

to give consistency (order 1) yet retain a as a free parameter. The error constant is then

$$C_2 = 2 + \tfrac{1}{2}\alpha_1 - 2\beta_2 = -\tfrac{1}{2}\left(1 + 3a\right).$$

The order increases to 2 when $a = -\frac{1}{3}$ for which we have the error constant $C_3 = \frac{4}{3} + \frac{1}{6}\alpha_1 - 2\beta_2 = -\frac{2}{9}$ (this is the BDF(2) method of the previous exercise).

---

13. The linear difference operator associated with Simpson's rule

$$x_{n+2} - x_n = \tfrac{1}{3}h(f_{n+2} + 4f_{n+1} + f_n)$$

is $\mathscr{L}_h z(t) = z(t+2h) - z(t) - \frac{1}{3}h\left[z'(t+2h) + 4z'(t+h) + z'(t)\right]$. When the Taylor series

$$\begin{aligned}
z(t+2h) &= z(t+h) + hz'(t+h) + \tfrac{1}{2!}h^2 z''(t+h) + \ldots\\
z(t) &= z(t+h) - hz'(t+h) + \tfrac{1}{2!}h^2 z''(t+h) - \ldots
\end{aligned}$$

are subtracted, terms with even powers of h cancel leaving

$$z(t+2h) - z(t) = 2hz'(t+h) + \tfrac{2}{3!}h^3 z'''(t+h) + \tfrac{2}{5!}h^5 z^{(5)}(t+h) + \mathscr{O}(h^7).$$

When added, the odd powers of h in the corresponding series for z' cancel leaving

$$z'(t+2h) + z'(t) = 2z'(t) + \tfrac{2}{2!}h^2 z'''(t+h) + \tfrac{2}{4!}h^4 z^{(5)}(t+h) + \mathscr{O}(h^6).$$

Hence

$$\begin{aligned}
\mathscr{L}_h z(t) &= \left(2hz'(t+h) + \tfrac{2}{3!}h^3 z'''(t+h) + \tfrac{2}{5!}h^5 z^{(5)}(t+h) + \mathscr{O}(h^7)\right)\\
&\quad - \tfrac{1}{3}h\left(6z'(t) + \tfrac{2}{2!}h^2 z'''(t+h) + \tfrac{2}{4!}h^4 z^{(5)}(t+h) + \mathscr{O}(h^6)\right)\\
&= 2h^5\left(\tfrac{1}{5!} - \tfrac{1}{3\times 4!}\right) h^5 z^{(5)}(t+h) + \mathscr{O}(h^7).
\end{aligned}$$

Thus, since $z^{(5)}(t+h) = z^{(5)}(t) + \mathscr{O}(h)$,

$$\mathscr{L}_h z(t) = -h^5 z^{(5)}(t+h)/90 + \mathscr{O}(h^7) = -h^5 z^{(5)}(t)/90 + \mathscr{O}(h^6)$$

and the method has order $p = 4$ with error constant $C_5 = -1/90$.

The advantage of using the non–standard expansions is that they allow us to exploit the symmetry of the coefficients which results in significant cancellation of terms.

---

15. With

$$\mathscr{L}_h z(t) = z(t+2h) + \alpha_1 z(t+h) + \alpha_0 z(t) - h(\beta_2 z'(t+2h) + \beta_1 z'(t+h) + \beta_0 z'(t))$$

and $z(t) = 1$,

$$\mathscr{L}_h 1 = 1 + \alpha_1 + \alpha_0 = \rho(1)$$

and, with $z(t) = t$,

$$\begin{aligned}\mathscr{L}_h t &= (t+2h) + \alpha_1(t+h) + \alpha_0 t - h(\beta_2 + \beta_1 + \beta_0)\\ &= t(1 + \alpha_1 + \alpha_0) + h\big(2 + \alpha_1 - (\beta_2 + \beta_1 + \beta_0)\big)\\ &= t\rho(1) + h(\rho'(1) - \sigma(1))\end{aligned}$$

so $\mathscr{L}_h 1 = \mathscr{L}_h t = 0$ if, and only if, $\rho(1) = 0$ and $\rho'(1) = \sigma(1)$ which, according to Theorem 4.7, are conditions equivalent to the consistency conditions (4.15).

---

17. The limit

$$\lim_{h\to 0} \frac{x_{n+2} + \alpha_1 x_{n+1} + \alpha_0 x_n}{h}$$

is of the form $\frac{0}{0}$ since, for a convergent method, $x_{n+2} \to x(t^*+2h)$, $x_{n+1} \to x(t^*+h)$ and $x_n \to x(t^*)$ so the numerator converges to

$$x_{n+2} + \alpha_1 x_{n+1} + \alpha_0 x_n \to (1 + \alpha_1 + \alpha_0)x(t^*) = 0$$

(using the 1st of the consistency conditions (4.15): $1 + \alpha_1 + \alpha_0 = 0$).

It is necessary, therefore, to invoke l'Ĥopital's rule which states that

$$\lim_{h\to 0} \frac{f(h)}{g(h)} = \frac{f'(0)}{g'(0)}$$

if $f(0) = g(0) = 0$ and $g'(0) \neq 0$. With $f(h) = x(t^*+2h) + \alpha_1 x(t^*+h) + \alpha_0 x(t^*)$ and $g(h) = h$, we find $f'(h) = 2x'(t^*+2h) + \alpha_1 x'(t^*+h)$ so

$$\lim_{h\to 0} \frac{x_{n+2} + \alpha_1 x_{n+1} + \alpha_0 x_n}{h} = (2 + \alpha_1)x'(t^*).$$

The proof given in the text is then complete.

---

19. Suppose that $p(t) = A + B(t - t_n) + C(t - t_n)^2$ (It is more convenient in this case to use $1, (t - t_n), (t - t_n)^2$ as a basis for quadratic polynomials). The conditions

$$p(t_n) = x_n, \quad p'(t_n) = f_n, \quad p'(t_{n+1}) = f_{n+1}.$$

then lead to the algebraic equations

$$\begin{aligned} A \qquad\qquad &= x_n \\ B \qquad &= f_n \\ B \ + 2hC &= f_{n+1} \end{aligned}$$

which solve to give $A = x_n$, $B = f_n$, $hC = \frac{1}{2}(f_{n+1} - f_n)$ so that

$$p(t) = x_n + (t - t_{n+1})f_n + \tfrac{1}{2}(t - t_{n+1})^2(f_{n+1} - f_n)/h.$$

The prediction $x_{n+1} = p(t_{n+1})$ then leads to $x_{n+1} = x_n + \frac{1}{2}(f_{n+1} + f_n)$ which is the Trapezoidal rule.

Exercises 5

1. a) $\rho(r) = r^2 - 4r + 3 = (r-1)(r-3)$ so it has roots $r = 1$ and $r = 3$. These do not satisfy the root condition and so the method is not zero–stable.

   b) $\rho(r) = 3r^2 - 4r + 1 = (3r-1)(r-1)$ so it has roots $r = 1$ and $r = \frac{1}{3}$ which satisfy the root condition and so the method is zero–stable.

   In b) we must also have consistency: $\rho(1) = 0$ (which is satisfied) and also $\rho'(1) = \sigma(1)$ with $\sigma(r) = a$ and $\rho'(1) = 2$ so that $a = 2$ is required for convergence (this leads to the BDF(2) method).

---

3. $x_{n+2} + 2ax_{n+1} - (2a-1)x_n = h\left[(a+2)f_{n+1} + af_n\right]$ has linear difference operator

$$\mathscr{L}_h z(t) = z(t+2h) + 2az(t+h) - (2a-1)z(t) - h\left[(a+2)z'(t+2h) + az'(t)\right].$$

We first check the order

$$\begin{aligned}
\mathscr{L}_h z(t) = \quad & z(t) + 2hz'(t) + 2h^2z''(t) + \tfrac{4}{3}h^3z'''(t) + \mathscr{O}(h^4) \\
& + 2a\left(z(t) + hz'(t) + \tfrac{1}{2}h^2z''(t) + \tfrac{1}{6}h^3z'''(t) + \mathscr{O}(h^4)\right) \\
& - (2a+1)z(t) \\
& - h(a+2)\left(z'(t) + hz''(t) + \tfrac{1}{2}h^2z'''(t) + \mathscr{O}(h^3)\right) \\
& - ha \quad z'(t) \\
& = -\tfrac{1}{6}(a-2)h^3z'''(t) + \mathscr{O}(h^4)
\end{aligned}$$

so the method is, in general, of order $p = 2$ with error constant

$$C_3 = -\tfrac{1}{6}(a-2).$$

When $a = 2$ the method is at least 3rd order. However, the first characteristic polynomial

$$\rho(r) = r^2 + 2ar - (2a+1) = (r-1)(r-2a-1)$$

has a root $r = 2a + 1 = 5$ in this case and so the resulting method is not zero–stable. The first characteristic polynomial is the same as that for the method in Example 4.11.

The method will be zero–stable provided $-1 \le 2a + 1 < 1$ (strict inequality at the right endpoint to avoid a double root $r = 1$). Thus $-1 \le a < 0$ and then $|C_3| = |a-2|/6$ there is no smallest error constant since $|C_3| \to 1/3$ as $a \to 0$ but the resulting method will have a double root $r = 1$ thus violating the root condition.

---

5. The most general 1–step LMM has the form

$$x_{n+1} + \alpha_0 x_n = h(\beta_1 f_{n+1} + \beta_0 f_n)$$

which has first characteristic polynomial $\rho(r) = r + \alpha_0$. Consistency requires $\rho(1) = 1 + \alpha_0 = 0$. Hence $\alpha_0 = -1$ and all consistent 1–step LMMs have the same first characteristic polynomial $\rho(r) = r - 1$ which clearly satifies the root condition.

---

7. $x_{n+3} + x_{n+2} - x_{n+1} - x_n = 4hf_n$

   (a) $\rho(r) = r^3 + r^2 - r - 1 = (r-1)(r+1)^2$ and $\sigma(r) = 4$.
   $\rho(1) = 0$ and $\rho'(1) = 4 = \sigma(1)$ so the method is consistent.

   (b) The general solution as given in the text is $x_n = A+(B+Cn)(-1)^n$. Invoking the starting conditions $x_0 = 1$, $x_1 = 1-h$, $x_2 = 1-2h$ we find

   $$\begin{aligned} A + B \qquad &= 1 \\ A - B - \ \ C &= 1 - h \\ A + B + 2C &= 1 - 2h \end{aligned}$$

   whose solution is $A = 1-h$, $B = h$, $C = -h$ so

   $$x_n = 1 - h + (-1)^n(h - t_n)$$

   since $t_n = nh$.

   (c) $\rho(r)$ has a double root at $r = -1$ and violates the root condition—it cannot therefore be zero–stable.

---

9. $x_{n+3} - x_{n+2} + x_{n+1} - x_n = \frac{1}{2}h(f_{n+3} + f_{n+2} + f_{n+1} + f_n)$

   $\rho(r) = r^3 - r^2 + r - 1 = (r-1)(r^2+1)$ and $\sigma(r) = \frac{1}{2}(r^3 + r^2 + r + 1)$.

   $\rho(1) = 0$ and $\rho'(1) = 2 = \sigma(1)$ so the method is consistent.

   $\rho(r) = (r-1)(r^2+1)$ has roots $r = 1, \pm\mathrm{i}$ and these satisfy the root condition. The method is therefore convergent.

---

11. For $x_{n+2} - x_n = h\,(\beta_1 f_{n+1} + \beta_0 f_n)$,

    $\rho(r) = r^2 - 1 = (r-1)(r+1)$ and $\sigma(r) = \beta_1 r + \beta_0$.

    $\rho(r)$ has roots $r = \pm 1$ so satisfies the root condition. It is therefore zero–stable.

    $\rho(1) = 0$ for all members of the family and $\rho'(1) = \sigma(1)$ if $2 = \beta_1 + \beta_0$. Setting $\beta_1 = b$, we have consistency if $\beta_0 = 2 - b$. Thus, the one–parameter family of methods

    $$x_{n+2} - x_n = h\,(bf_{n+1} + (2-b)f_n)$$

is convergent for all values of the parameter b.

$$\begin{aligned}\mathscr{L}_h z(t) = {} & \quad z(t) + 2hz'(t) + 2h^2 z''(t) + \tfrac{4}{3}h^3 z'''(t) + \mathscr{O}(h^4) \\ & - z(t) \\ & - \theta\big(\qquad hz'(t) + 2h^2 z''(t) + 2h^3 z'''(t) + \mathscr{O}(h^4)\big) \\ & - \quad (2-\theta)hz'(t) \\ = {} & (2-b)h^2 z''(t) - \tfrac{1}{6}(3b-8)h^3 z'''(t) + \mathscr{O}(h^4)\end{aligned}$$

so the method has order $p = 1$, in general, with error constant $C_2 = 2 - b$.

The method will have order $p = 3$ when $b = 2$ and the error constant is then $C_3 = 1/3$.

---

13. For $x_{n+2} + (\theta - 2)x_{n+1} + (1-\theta)x_n = \frac{1}{4}h\left((6+\theta)f_{n+2} + 3(\theta-2)f_n\right)$,

$$\begin{aligned}\mathscr{L}_h z(t) = {} & \qquad z(t) + 2hz'(t) + 2h^2 z''(t) + \mathscr{O}(h^3) \\ & (\theta-2)\big(z(t) + hz'(t) + \tfrac{1}{2}h^2 z''(t)\big) + \mathscr{O}(h^3) \\ & + (1-\theta)z(t) \\ & - \tfrac{1}{4}(6+\theta)\big(\quad hz'(t) + 2h^2 z''(t) + \mathscr{O}(h^3)\big) \\ & - \tfrac{3}{4}(\theta-2)\big(\quad hz'(t)\big) \\ = {} & -2h^2 z''(t) + \mathscr{O}(h^3)\end{aligned}$$

so the method has order $p = 1$ and error constant $C_2 = -2$, both independent of θ.

The method has been shown to be consistent, so it will be convergent if, and only if, it is zero–stable. For this we have to check the root condition: the first characteristic polynomial is

$$\rho(r) = r^2 + (\theta-2)r + (1-\theta) = (r-1)(r-1+\theta)$$

and has roots $r = 1$ and $r = 1 - \theta$. These satisfy the root condition if, and only if, $0 < \theta \leq 2$.

---

15. $\rho(r) = \frac{1}{c}\sum_{j=1}^{k}\frac{1}{j}r^{k-j}(1-r)^j$ and $\sigma_3 = \beta_k r^k$. Consistency requires $\sigma(1) = \rho'(1)$. Since

$$\rho'(r) = \frac{1}{c}\sum_{j=1}^{k}\frac{1}{j}\Big((k-j)r^{k-j-1}(1-r)^j - jr^{k-j}(1-r)^{j-1}\Big)$$

so, when evaluating $\rho'(1)$, all terms vanish except for $jr^{k-j}(1-r)^{j-1}$ with $j = 1$. Consequently $\beta_k = \sigma(1) = \rho'(1) = -1/c$.

Check:
with $k = 1$, $c = 1$ and $\beta_1 = \sigma(1) = \rho'(1) = -1$ (backward Euler).
with $k = 2$, $c = 1 + 1/2 = 3/2$ and $\beta_2 = \sigma(1) = \rho'(1) = -2/3$ (BDF(2)—(4.25)).
with $k = 3$, $c = 11/6$ and $\beta_3 = \sigma(1) = \rho'(1) = -6/11$ (BDF(3)—see Exercise 5.14).

Exercises 6

1. With $\widehat{h} = 2X + 2\mathrm{i}Y$ in Example 6.9,

$$r_1 = \frac{1+\frac{1}{2}\widehat{h}}{1-\frac{1}{2}\widehat{h}} = \frac{(1+X)+\mathrm{i}Y}{(1-X)-\mathrm{i}Y} \Rightarrow |r_1|^2 = \frac{(1+X)^2+Y^2}{(1-X)^2+Y^2}$$

so

$$|r_1|^2 - 1 = \frac{4X}{(1-X)^2+Y^2}.$$

Clearly $|r_1| < 1$ if, and only if, $X < 0$, i.e., $\Re(\widehat{h}) < 0$.

It follows that the interval of absolute stability (relevant when $Y = 0$) is $(-\infty, 0)$.

---

3. $x_{n+1} - x_n = h(\theta f_{n+1} + (1-\theta) f_n)$ applied to $x'(t) = \lambda x(t)$ leads to the stability polynomial $p(r) = (1-\theta\widehat{h})r - (1-\theta)\widehat{h})$ so the strict root condition leads to

$$|r_1| = \left|\frac{1+(1-\theta)\widehat{h})}{1-\theta\widehat{h}}\right| < 1.$$

The boundary of the region of absolute stability is given by $|r| = 1$ so, with $\widehat{h} = \widehat{x} + \mathrm{i}\widehat{y}$,

$$\begin{aligned}
\frac{(1+\widehat{x}(1-\theta))^2 + (1-\theta)^2\widehat{y}^2}{(1-\widehat{x}\theta)^2 + \theta^2\widehat{y}^2} &= 1\\
2\widehat{x} + \widehat{x}^2(1-2\theta) + \widehat{y}^2(1-2\theta) &= 0\\
\left(\frac{2}{(1-2\theta)}\widehat{x} + \widehat{x}^2\right) + \widehat{y}^2 &= 0\\
\left(\frac{1}{(1-2\theta)^2} + \frac{2}{(1-2\theta)}\widehat{x} + \widehat{x}^2\right) + \widehat{y}^2 &= \frac{1}{(1-2\theta)^2}\\
\left(\frac{1}{1-2\theta} + \widehat{x}\right)^2 + \widehat{y}^2 &= \frac{1}{(1-2\theta)^2}
\end{aligned}$$

which is a circle of radius $1/|1-2\theta|$ centred at $\widehat{x} = -1/(1-2\theta)$, $\widehat{y} = 0$.

When $\theta < \frac{1}{2}$ the boundary of this circle lies in the left half plane. At its centre, $\widehat{h} = -1/(1-2\theta)$ and

$$|r_1| = \left|\frac{\theta}{1-\theta}\right| < 1$$

so points inside the circle correspond to points where the method is absolutely stable.

When $\theta > \frac{1}{2}$ a similar calculation shows that the boundary lies in the right half plane and at its centre, $|r_1| > 1$ so points inside the circle correspond to points where the method is not absolutely stable. The region of absolute stability is therefore the exterior of the circle.

When $\theta = \frac{1}{2}$ the boundary of the region of absolute stability coincides with the imaginary axis, as discussed in Exercise 6.1.

---

5. The stability polynomial of the LMM $x_{n+2} - x_n = \frac{1}{2}h\,(f_{n+1} + 3f_n)$ is $p(r) = r^2 - 1 - \frac{1}{2}\widehat{h}(r+3)$.

   We only need the interval of absolute stability which means that we can assume $\widehat{h}$ to be real and therefore Lemma 6.10 may be used with

   $$a = -\tfrac{1}{2}\widehat{h}, \qquad b = -1 - \tfrac{3}{2}\widehat{h}.$$

   The conditions $b < 1$ and $p(\pm 1) > 0$ lead to

   $$\begin{array}{lllll} b < 1: & -1 - \frac{3}{2}\widehat{h} < 1 & \Longleftrightarrow & \widehat{h} > -\frac{4}{3} \\ 1 + a + b > 0: & -2\widehat{h} > 0 & \Longleftrightarrow & \widehat{h} < 0 \\ 1 - a + b > 0: & -\widehat{h} > 0 & \Longleftrightarrow & \widehat{h} < 0. \end{array}$$

   The interval of absolute stability is therefore $(-\frac{4}{3}, 0)$.

---

7. For $x_{n+2} - 4\theta x_{n+1} - (1-4\theta)x_n = h\big((1-\theta)f_{n+2} + (1-3\theta)f_n\big)$

   (a) the associated linear difference operator is

   $$\mathscr{L}_h z(t) = z(t+2h) - 4\theta z(t+h) - (1-4\theta)z(t) - h\big((1-\theta)z'(t+2h) + (1-3\theta)z'(t)\big).$$

   Expanding in powers of h we find

   $$\mathscr{L}_h z(t) = \tfrac{2}{3}h^3 z'''(t)(2\theta - 1) + \mathscr{O}(h^4)$$

   so the method has order $p = 2$ with error constant $C_3 = 2(2\theta-1)/3$.
   The highest order ($p = 3$) is achieved when $\theta = \frac{1}{2}$ and the resulting error constant is $C_4 = -1/12$.

   (b) The method will be convergent if, and only if, it is both consistent and zero–stable. The first of these conditions has been shown to hold in part (a) (order $p \geq 1$). For zero stability the roots of the 1st characteristic polynomial

   $$\rho(r) = r^2 - 4\theta r - (1-4\theta) = (r-1)(r+1-4\theta)$$

   must satisfy the root condition—this requires $-1 \leq -1 + 4\theta < 1$, i.e., $0 \leq \theta < \frac{1}{2}$.
   The method of highest order is not convergent since the root condition is violated at $\theta = \frac{1}{2}$ ($\rho(r)$ has a double root $r = 1$).

(c) For A_0–stability we must show that the interval of absolute stability includes the negative real axis: $\Re(\widehat{h}) < 0$. Applying the method to the ODE $x'(t) = \lambda x(t)$ when $\theta = 1/4$ leads to the stability polynomial

$$p(r) = \big(1 - \tfrac{3}{4}\widehat{h}\big)r^2 - r - \tfrac{1}{4}\widehat{h}.$$

We are interested only in cases where $\widehat{h}$ is real and negative, in which case the coefficient of r^2 is always > 0. To apply Lemma 6.10 we have to divide by this coefficient, then $b = -\widehat{h}/(4 - 3\widehat{h})$ and

$$\begin{aligned} b < 1 &\Longleftrightarrow -\widehat{h} < 4 - 3\widehat{h}, \qquad \Longleftrightarrow \widehat{h} < 2 \\ p(1) > 0 &\Longleftrightarrow -\widehat{h} > 0 \\ p(-1) > 0 &\Longleftrightarrow 2 - \widehat{h} > 0 \end{aligned}$$

all of which are satisfied for $\widehat{h} < 0$ and so the method is A_0–stable.

---

9. The stability polynomial of AB(2) is (see Exercise 6.)$p(r) = r^2 - r - \frac{1}{2}\widehat{h}(3r - 1)$. Writing $r = \mathrm{e}^{\mathrm{i}s}$, then the boundary of the region of absolute stability is given by

$$\mathrm{e}^{2\mathrm{i}s} - \mathrm{e}^{\mathrm{i}s} - \tfrac{1}{2}\widehat{h}(3\mathrm{e}^{\mathrm{i}s} - 1) = 0.$$

Solving for $\widehat{h}$ and writing in real/imaginary from

$$\begin{aligned} \widehat{h} &= 2\mathrm{e}^{\mathrm{i}s}\frac{\mathrm{e}^{\mathrm{i}s} - 1}{3\mathrm{e}^{\mathrm{i}s} - 1} \\ &= 2\mathrm{e}^{\mathrm{i}s}\frac{(\mathrm{e}^{\mathrm{i}s} - 1)}{(3\mathrm{e}^{\mathrm{i}s} - 1)}\frac{(3\mathrm{e}^{-\mathrm{i}s} - 1)}{(3\mathrm{e}^{-\mathrm{i}s} - 1)} \\ &= -\frac{\cos 2s - 4\cos s + 3 + \mathrm{i}(\sin 2s - 4\sin s)}{5 - 3\cos s} \\ &= -2\frac{(1 - \cos s)^2 + \mathrm{i}\sin s(\cos s - 2)}{5 - 3\cos s}. \end{aligned}$$

We observe from this that $\Re(\widehat{h}) \le 0$ for all s (so the boundary lies entirely in the left half complex plane) and the boundary crosses the real axis when $\Im(\widehat{h}) = 0$. This occurs when $\sin s = 0$, i.e., when $s = 0$ (where $\widehat{h} = 0$) and again when $s = \pi$, where $\widehat{h} = -1$. The region of absolute stability must therefore be the one on the left of Figure 6.12.

For completeness we find that a similar computation for AM(2) leads to

$$\widehat{h} = -6\frac{(1 - \cos s)^2 + \mathrm{i}\sin s(\cos s - 7)}{25 + 16\cos s - 5\cos^2 s}$$

revealing that $\Re(\widehat{h}) \le 0$ for all s and the only intersections of the boundary with the real axis occur when $\sin s = 0$, i.e., when $s = 0$ (where $\widehat{h} = 0$) and again when $s = \pi$, where $\widehat{h} = -6$. This confirms that region of absolute stability of AM(2) must therefore be the one on the right of Figure 6.12.

---

11. The stability polynomial of the LMM $x_{n+2} - x_{n+1} = \frac{1}{4}h(f_{n+2} + 2f_{n+1} + f_n)$ is $p(r) = r^2 - r - \frac{1}{4}\widehat{h}(r^2 + 2r + 1)$, i.e.,

$$p(r) = \left(1 - \tfrac{1}{4}\widehat{h}\right) r^2 - \left(1 + \tfrac{1}{2}\widehat{h}\right) r - \tfrac{1}{4}\widehat{h} = 0.$$

We only need the interval of absolute stability which means that we can assume $\widehat{h}$ to be real and therefore Lemma 6.10 may be used with

$$a = -\frac{1 + \frac{1}{2}\widehat{h}}{1 - \frac{1}{4}\widehat{h}}, \quad b = -\frac{1}{4}\frac{\widehat{h}}{1 - \frac{1}{4}\widehat{h}}.$$

The conditions $b < 1$ is best tackled by computing $b - 1$:

$$b - 1 = -\tfrac{1}{4}\tfrac{\widehat{h}}{1 - \frac{1}{4}\widehat{h}} - 1 = \frac{4}{\widehat{h} - 4}$$

so $b < 1$ if, and only if, $\widehat{h} < 4$.

Since the coefficient of r^2 in $p(r)$ is always positive for $\widehat{h} < 0$, the two remaining conditions of Lemma 6.10 can be written as $p(\pm 1) > 0$.

$$p(1) = 1 - \frac{1 + \frac{1}{2}\widehat{h}}{1 - \frac{1}{4}\widehat{h}} - \tfrac{1}{4}\tfrac{\widehat{h}}{1 - \frac{1}{4}\widehat{h}} = \frac{4\widehat{h}}{\widehat{h} - 4}$$

so $p(1) > 0$ if $\widehat{h} < 0$.

$$p(-1) = 1 + \frac{1 + \frac{1}{2}\widehat{h}}{1 - \frac{1}{4}\widehat{h}} - \tfrac{1}{4}\tfrac{\widehat{h}}{1 - \frac{1}{4}\widehat{h}} = -\frac{8}{\widehat{h} - 4}$$

so $p(-1) > 0$ if, and only if, $\widehat{h} < 4$.

The conditions of Lemma 6.10 have been shown to be satisfied for all $\widehat{h} < 0$ so we can conclude that the interval of absolute stability is $(-\infty, 0)$, proving that the method is A–stable.

---

13. Applying the starting conditions $x_0 = 1$, $x_1 = \mathrm{e}^{\widehat{h}}$ to the general solution $x_n = Ar_+^n + Br_-^n$ we obtain

$$\begin{aligned} 1 &= A + B \\ \mathrm{e}^{\widehat{h}} &= Ar_+ + Br_-. \end{aligned}$$

Writing $A = \frac{1}{2}(1+a)$ and $B = \frac{1}{2}(1-a)$ then the first of these equations is identically satisfied while the second leads to

$$(r_+ + r_-) + a(r_+ - r_-) = 2e^{\widehat{h}}$$

and, on using the expressions for r_+ and r_- given in Example 6.12, we find

$$r_+ + r_- = 2\widehat{h} \quad \text{and} \quad r_+ - r_- = 2\sqrt{1 + \widehat{h}^2}$$

and so

$$a = \frac{e^{\widehat{h}} - \widehat{h}}{\sqrt{1 + \widehat{h}^2}}.$$

Using the Binomial expansion, we find $\sqrt{1 + \widehat{h}^2} = 1 + \frac{1}{2}\widehat{h}^2 + \mathcal{O}(h^4)$ so, since $e^{\widehat{h}} - \widehat{h} = 1 + \frac{1}{2}\widehat{h}^2 + \frac{1}{6}\widehat{h}^3 + \mathcal{O}(h^4)$,

$$\begin{aligned} a - 1 &= \frac{1 + \frac{1}{2}\widehat{h}^2 + \frac{1}{6}\widehat{h}^3 + \mathcal{O}(h^4)}{1 + \frac{1}{2}\widehat{h}^2 + \mathcal{O}(h^4)} - 1 \\ &= \frac{\frac{1}{6}\widehat{h}^3 + \mathcal{O}(h^4)}{1 + \mathcal{O}(h^2)} = \frac{1}{6}\widehat{h}^3 + \mathcal{O}(h^4), \end{aligned}$$

as required. Hence, $A = \frac{1}{2}(1 + a) = 1 + \frac{1}{12}\widehat{h}^3 + \mathcal{O}(h^4) = 1 + \mathcal{O}(h^3)$ and $B = \frac{1}{2}(1 - a) = -\frac{1}{12}\widehat{h}^3 + \mathcal{O}(h^4) = \mathcal{O}(h^3)$.

---

15. The LMM $x_{n+2} + (\theta - 2)x_{n+1} + (1 - \theta)x_n = \frac{1}{4}h\left((6 + \theta)f_{n+2} + 3(\theta - 2)f_n\right)$ was shown in Exercise 5.13 to be convergent for $\theta \in (0, 2]$ and we shall assume that θ is restricted to this interval.

    The method has stability polynomial

    $$p(r) = \left(1 - \tfrac{1}{4}\widehat{h}(\theta + 6)\right)r^2 + (\theta - 2)r + 1 - \theta - \tfrac{3}{4}(\theta - 2)$$

    to which we apply Lemma 6.10 in order to show that its roots satisfy $|r| < 1$ for all $\widehat{h} < 0$ and all $\theta \in (0, 2]$. These conditions lead to

    (a) For $b < 1$,

    $$\begin{aligned} b - 1 &= \frac{1 - \theta - \frac{3}{4}(\theta - 2)\widehat{h}}{1 - \frac{1}{4}(\theta + 6)\widehat{h}} - 1 \\ &= \frac{-\theta + \frac{1}{2}(6 - \theta)\widehat{h}}{1 - \frac{1}{4}(\theta + 6)\widehat{h}} \end{aligned}$$

    and the right hand side is easily shown to be negative.

(b) The coefficient of r^2 is positive for all relevant values of $\widehat{h}$ and θ so we may use the conditions $p(\pm 1) > 0$.

$$p(1) = -\theta\widehat{h} \quad \text{and} \quad p(-1) = (4 - 2\theta) - \theta\widehat{h}$$

and both are seen to be positive for all relevant values of $\widehat{h}$ and θ.

All convergent members of the family are absolutely stable for all $\widehat{h} < 0$ and are, therefore, A_0–stable.

---

17. (a) $\rho(r) = r^2 - 2ar + (2a - 1)$, $\sigma(r) = ar^2 + (2 - 3a)r$.

(b) $\rho(1) = 0$ and $\rho'(1) = \sigma(1) = 2 - 2a$, so the method is consistent for all a.

(c) $\rho(r) = r^2 - 2ar + 2a - 1 = (r - 1)(r - 2a + 1)$ so the root condition is satisfied for $-1 \leq 2a - 1 < 1$, i.e., $0 \leq a < 1$.

(d) It is convergent if it is consistent and zero–stable, i.e., for $0 \leq a < 1$.

(e) $\mathscr{L}_h z(t) = \left(-\frac{5}{6}a + \frac{1}{3}\right) z''' h^3 + \left(\frac{1}{3} - \frac{11}{12}a\right) z^{(4)} h^4 + O(h^5)$
So the order is, in general $p = 2$ with error constant $C_3 = -\frac{5}{6}a + \frac{1}{3}$.
When $a = \frac{2}{5}$, $C_3 = 0$ and $\mathscr{L}_h z(t) = -\frac{1}{30} z^{(4)} h^4 + O(h^5)$ and the order is $p = 3$ with error constant $C_4 = -\frac{1}{30}$. This method is convergent since a is in the zero-stable range.

(f) For A_0–stability, apply the method to $x' = \lambda x$ with λ real and $\lambda < 0$. This gives the stability polynomial

$$p(r) = (1 - a\widehat{h})r^2 - (2a + (2 - 3a)\widehat{h})r + 2a - 1$$

and we have to use Lemma 6.10 to check the conditions under which the roots satisfy the strict root condition: this will ensure A_0–stability—all solutions $x_n \to 0$ as $n \to \infty$. Since the coefficient of r^2 in $p(r)$ is always positive for $\widehat{h} < 0$ and $a \geq 0$ these conditions give:

i. $p(1) > 0$

$$p(1) = -a\widehat{h} - (2 - 3a)\,\widehat{h} = 2\widehat{h}\,(a - 1)$$

and $a < 1$ for zero–stability, so $p(1) > 0$ for all $\widehat{h} < 0$.

ii. $p(-1) > 0$

$$p(-1) = -a\widehat{h} + 4a + (2 - 3a)\,\widehat{h} = 4a + 2(1 - 2a)\widehat{h}$$

so, to have $p(-1) > 0$ for all $\widehat{h} < 0$, we must have $a \geq \frac{1}{2}$.

iii. $b < 1$

$$b - 1 = \frac{2a - 1}{1 - a\widehat{h}} - 1 = \frac{a\widehat{h} - 2(1 - a)}{1 - a\widehat{h}}$$

so $b - 1 < 0$ since $1 - a > 0$ by zero–stability, $a\widehat{h} \leq 0$ since $a \geq 0$ & $\widehat{h} < 0$ and the denominator is > 0.

All conditions for the strict root condition are satisfied provided $\frac{1}{2} \le a < 1$. This condition identifies all A_0–stable convergent members of the family.

(g) The BDF(2) method has $a = \frac{2}{3}$ so it is convergent ($0 \le a < 1$) and A_0–stable ($\frac{1}{2} \le a < 1$); it has order $p = 2$ with error constant $C_3 = \left[-\frac{5}{6}a + \frac{1}{3}\right]_{a=2/3} = -\frac{2}{9}$.

(h) When $a = 0$ the method is explicit but not A_0–stable—in accord with part 1 of Dahlquist's Second Barrier Theorem (Theorem 6.15).

To have 3rd order, $a = \frac{2}{5}$, but this is outside the range of A_0–stable methods. So the maximum order of A_0–stable methods is $p = 2$—in accord of part 2 of the theorem.

For part 3 we note that the modulus of the scaled error constant is

$$\left|\frac{C_3}{\sigma(1)}\right| = \tfrac{1}{12}\frac{5a-2}{1-a}$$

which is a strictly increasing function of a—it achieves its smallest value consistent with A_0–stability when $a = \frac{1}{2}$, in which case the method becomes the Trapezoidal rule.

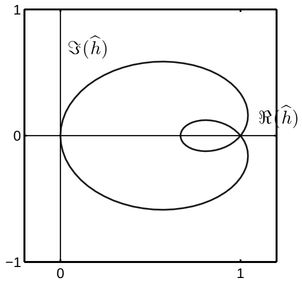

Figure 1: Boundary of the region of absolute stability for Exercise 6.18.

---

19. Since all points (a, b) inside the triangle in Figure 6.8 correspond to roots that satisfy the strict root condition ($|r| < 1$), they must also correspond to roots that satisfy the root condition ($|r| \le 1$).

    For coefficients (a, b) on the edges of the triangle at least one root of the polynomial has modulus equal to 1. However, it can only have a double root of unit modulus if there are coincident roots ($b = a^2/4$) and these occur only at $(\pm 2, 1)$.

    Hence the root condition is satisfied if $|a| - 1 \le b \le 1$, except for the case when $b = 1$ and $a = \pm 2$ (which lead to double roots $r = \pm 1$).

---

21. Applying the composite Euler method to $x'(t) = \lambda x(t)$ leads to

$$x_{2m+1} = (1 + h_0\lambda)x_{2m} \quad \text{and} \quad x_{2m+2} = (1 + h_1\lambda)x_{2m+1},$$

so $x_{2m+2} = (1 + h_1\lambda)(1 + h_0\lambda)x_{2m}$. Thus, with $h_0 = (1-\gamma)h$ and $h_1 = (1+\gamma)h$,

$$x_{2m+2} = (1 + (1+\gamma)h\lambda)(1 + (1-\gamma)h\lambda)x_{2m} = \big((1+\widehat{h})^2 - \gamma^2\widehat{h}^2)\big)x_{2m},$$

where $\widehat{h} = h\lambda$. Thus $x_{2m+2}/x_{2m} = R(\widehat{h})$, where $R(\widehat{h}) = (1+\widehat{h})^2 - \gamma^2\widehat{h}^2$.
To determine the conditions under which $R(\widehat{h}) < 1$, we consider

$$R(\widehat{h}) - 1 = (1+\widehat{h})^2 - \gamma^2\widehat{h}^2 - 1 = \widehat{h}\big(2 + (1-\gamma^2)\widehat{h}\big).$$

Thus, since $\widehat{h} < 0$, we have $R(\widehat{h}) < 1$ if, and only if (recall that $0 \le \gamma < 1$)

$$-\frac{2}{1-\gamma^2} < \widehat{h} < 0. \tag{5}$$

To determine the conditions under which $R(\widehat{h}) > -1$, we find the minimum of $R(\widehat{h})$ for $\widehat{h} < 0$ and then require this minimum to exceed -1. R has stationary points where $R'(\widehat{h}) = 2\big((1+\widehat{h}) - \gamma^2\widehat{h}\big) = 0$, that is, at

$$\widehat{h} = -1/(1-\gamma^2).$$

Since $R''(\widehat{h}) = 2(1-\gamma^2) > 0$ the stationary point is a minimum. The value of R at this point is

$$R_{\min} = -\gamma^2/(1-\gamma^2)$$

and

$$R_{\min} + 1 = -\frac{\gamma^2}{1-\gamma^2} + 1 = -\frac{1-2\gamma^2}{1-\gamma^2}$$

and so $R_{\min} > -1$ for $\gamma^2 < \frac{1}{2}$. When $\gamma^2 = \frac{1}{2}$ the inequalities (5) become $-4 < \widehat{h} < 0$ which is the interval of absolute stability[1]. The region of absolute stability is shown in Figure 2.

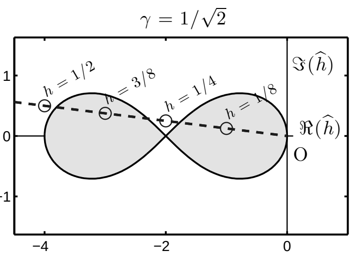

Figure 2: The region of absolute stability for the composite Euler method for Exercise 16.21 when $\gamma = 1/\sqrt{2}$.

[1]Strictly speaking, we have to weaken the conditions for absolute stability to $|R(\widehat{h})| \le 1$ in order to allow the value $\gamma^2 = \frac{1}{2}$. In practice, we need only choose a value of γ that is marginally smaller than $1/\sqrt{2}$.

Exercises 7

1. From (7.8)

$$\begin{bmatrix} u \\ v \end{bmatrix} = \tfrac{11}{20}\begin{bmatrix} 10 \\ 1 \end{bmatrix} \mathrm{e}^{-t} - \tfrac{9}{20}\begin{bmatrix} 10 \\ -1 \end{bmatrix} \mathrm{e}^{-21t}.$$

so

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = -\tfrac{11}{20}\begin{bmatrix} 10 \\ 1 \end{bmatrix} \mathrm{e}^{-t} + 21\tfrac{9}{20}\begin{bmatrix} 10 \\ -1 \end{bmatrix} \mathrm{e}^{-21t}.$$

and

$$\begin{bmatrix} u' \\ v' \end{bmatrix} + 11\begin{bmatrix} u \\ v \end{bmatrix} = \tfrac{11}{2}\begin{bmatrix} 10 \\ 1 \end{bmatrix} \mathrm{e}^{-t} + 9\begin{bmatrix} 10 \\ -1 \end{bmatrix} \mathrm{e}^{-21t} = \begin{bmatrix} 100v \\ u \end{bmatrix}$$

as required.

---

3. The system $u'(t) = v(t)$, $v'(t) = -200u(t) - 20v(t)$ has coefficient matrix

$$A = \begin{bmatrix} 0 & 1 \\ -200 & -20 \end{bmatrix}.$$

Its eigenvalues are the roots of the equation

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -200 & -20-\lambda \end{vmatrix} = \lambda^2 + 20\lambda + 200 = (\lambda+10)^2 + 100$$

which gives $\lambda = -10 \pm 10\mathrm{i}$. For absolute stability of Euler's method we require

$$|\widehat{h} + 1| < 1,$$

where $\widehat{h} = h\lambda$, for every eigenvalue λ of A. To determine the allowable range of stepsizes h, consider

$$\begin{aligned} |\widehat{h}+1|^2 - 1 &= |h(-10 \pm 10\mathrm{i})h + 1|^2 - 1 \\ &= |(1-10h) \pm 10\mathrm{i}h|^2 - 1 \\ &= (1-10h)^2 + 100h^2 - 1 \\ &= -20h + 200h^2 = -20h(1-10h). \end{aligned}$$

Hence $|\widehat{h}+1| < 1$ and we have absolute stability if, and only if, $0 < h < 1/10$.

---

5. The coefficient matrix is

$$A = \begin{bmatrix} -8 & 8 \\ 0 & -1/8 \end{bmatrix}$$

whose eigenvalues are[2] $\lambda = -8$ and $-1/8$. These are real so the condition for absolute stability of Euler's method is

$$-1 < 1 + h\lambda < 1$$

[2]The eigenvalues of an upper triangular matrix are simply its diagonal entries.

for each eigenvalue of A. With $\lambda = -8$ we find $h < 1/4$ and for $\lambda = -1/8$ we find $h < 16$. Both have to be satisfied so we must have $h < 1/4$.

7.

$$\begin{aligned} u_{n+1} &= \frac{1}{4+h^2}\big((4-h^2)u_n - 4hv_n\big), \\ v_{n+1} &= \frac{1}{4+h^2}\big(4hu_n + (4-h^2)hv_n\big), \end{aligned} \tag{6}$$

so, squaring both sides of the two equations and adding gives

$$\begin{aligned} u_{n+1}^2 + v_{n+1}^2 &= \frac{1}{(4+h^2)^2}\Big(\big((4-h^2)u_n - 4hv_n\big)^2 + \big(4hu_n + (4-h^2)hv_n\big)^2\Big) \\ &= \frac{1}{(4+h^2)^2}\Big(\big((4-h^2)^2 + 16h^2\big)u_n^2 + \big((4-h^2)^2 + 16h^2\big)v_n^2\Big) \\ &= \frac{1}{(4+h^2)^2}\big((4+h^2)^2u_n^2 + (4+h^2)^2v_n^2\big) = u_n^2 + v_n^2. \end{aligned}$$

This implies that $u_n^2 + v_n^2 = u_0^2 + v_0^2$.

With $u_n = R\cos(\theta_n)$ and $v_n = R\sin(\theta_n)$, clearly $u_n^2 + v_n^2 = R^2$, where $R^2 = u_0^2 + v_0^2$. Substituting into the right of (6), we find

$$\begin{aligned} \cos(\theta_{n+1}) &= \frac{1}{4+h^2}\big((4-h^2)\cos(\theta_n) - 4h\sin(\theta_n)\big), \\ \sin(\theta_{n+1}) &= \frac{1}{4+h^2}\big(4h\cos(\theta_n) + (4-h^2)h\sin(\theta_n)\big), \end{aligned}$$

so that, dividing the 2nd by the 1st,

$$\tan(\theta_{n+1}) = \frac{(4-h^2)\tan(\theta_n) + 4h}{(4-h^2) - 4h\tan(\theta_n)}.$$

We use the standard identity

$$\tan(\theta_{n+1} - \theta_n) = \frac{\tan(\theta_{n+1}) - \tan(\theta_n)}{1 + \tan(\theta_{n+1})\tan(\theta_n)}$$

and calculate the numerator and denominator separately.

$$\tan(\theta_{n+1}) - \tan(\theta_n) = \frac{(4-h^2)\tan(\theta_n) + 4h}{(4-h^2) - 4h\tan(\theta_n)} - \tan(\theta_n) = \frac{4h\sec^2(\theta_n)}{(4-h^2) - 4h\tan(\theta_n)},$$

where we have used $1 + \tan^2(\theta_n) = \sec^2(\theta_n)$. Also,

$$1 + \tan(\theta_{n+1})\tan(\theta_n) = 1 + \tan(\theta_n)\frac{\tan(\theta_{n+1}) - \tan(\theta_n)}{1 + \tan(\theta_{n+1})\tan(\theta_n)} = \frac{(4-h^2)\sec^2(\theta_n)}{(4-h^2) - 4h\tan(\theta_n)}$$

and so

$$\tan(\theta_{n+1} - \theta_n) = \frac{4h}{4 - h^2} = \frac{h}{1 - h^2/4}.$$

Since

$$\tan 2\varphi = \frac{2\tan\varphi}{1 - \tan^2\varphi}$$

we see that $\tan\frac{1}{2}(\theta_{n+1} - \theta_n) = \frac{1}{2}h$. Using the expansion $\tan^{-1}(x) = x - \frac{1}{3}x^3 + \mathscr{O}(x^5)$ we have

$$\begin{aligned}\theta_{n+1} - \theta_n &= 2\tan^{-1}\tfrac{1}{2}h \\ &= 2\left(\tfrac{1}{2}h - \tfrac{1}{3}\left(\tfrac{1}{2}h\right)^3 + \mathscr{O}(h^5)\right) \\ &= h - \tfrac{1}{12}h^3 + \mathscr{O}(h^5)\end{aligned}$$

as required.

---

9. Applying the mid–point rule $x_{n+2} - x_n = 2hf_{n+1}$ to the ODE $x'(t) = \mathrm{i}x(t)$ leads to

$$x_{n+2} - 2\mathrm{i}hx_{n+1} - x_n = 0,$$

a difference equation with auxiliary equation (stability polynomial with $\lambda = \mathrm{i}$)

$$p(r) = r^2 - 2\mathrm{i}hr - 1$$

whose roots are

$$r_\pm = \mathrm{i}h \pm \sqrt{1 - h^2}.$$

When $h \leq 1$ the argument of the square root is positive and so $|r_\pm| = 1$.

However, when $h > 1$, the roots become $r_\pm = \mathrm{i}h \pm \mathrm{i}\sqrt{h^2 - 1}$ the modulus of the larger root is $|r_+|^2 = h + \sqrt{h^2 - 1} > 1$ so that solutions become unstable.

Exercises 8

1. Let $u = x_{n+1}$ then $x_{n+1} = x_n + 2hx_{n+1}(1 - x_{n+1})$ becomes

$$2hu^2 + (1-2h)u - x_n = 0$$

whose solutions are

$$u_\pm = \frac{1}{4h}\left(-(1-2h) \pm \sqrt{(1-2h)^2 + 8hx_n}\right).$$

Then

$$u_- = \frac{1}{4h}\left(-(1-2h) - \sqrt{(1-2h)^2 + 8hx_n}\right) \to -\frac{1}{2h} \to -\infty$$

as $h \to 0_+$. In contrast, by a process known as rationalization, we find

$$\begin{aligned} u_+ &= \frac{1}{4h}\left(\sqrt{(1-2h)^2 + 8hx_n} - (1-2h)\right)\frac{\sqrt{(1-2h)^2 + 8hx_n} + (1-2h)}{\sqrt{(1-2h)^2 + 8hx_n} + (1-2h)} \\ &= \frac{\left((1-2h)^2 + 8hx_n\right) - (1-2h)^2}{\sqrt{(1-2h)^2 + 8hx_n} + (1-2h)} \\ &= \frac{2x_n}{\sqrt{(1-2h)^2 + 8hx_n} + (1-2h)} \to x_n \end{aligned}$$

as $h \to 0$. Since x_n and x_{n+1} are meant to be approximations to $x(t_n)$ and $x(t_{n+1})$, respectively, and $x(t_{n+1}) \to x(t_n)$ as $h \to 0$, it is appropriate to choose $x_{n+1} = u_+$.

---

3. Including additional rows in Table 8.1 to accommodate $E^{[\ell]}(= u^{[\ell]} - x_1)$ and $E^{[\ell+1]}/E^{[\ell]}$ leads to the results shown in the following table.

ℓ	0	1	2	3	4
$u^{[\ell]}$	0.2	0.232	0.2356	0.2360	0.2360
$u^{[\ell+1]} - u^{[\ell]}$	0.032	0.0036	0.0004	0.0000	
$E^{[\ell]}$	$-3.61_{10}{-2}$	$-4.07_{10}{-3}$	$-4.33_{10}{-4}$	$-4.57_{10}{-5}$	$-4.83_{10}{-6}$
$E^{[\ell+1]}/E^{[\ell]}$	0.113	0.106	0.106	0.106	

($3.61_{10}{-2} \equiv 3.61 \times 10^{-2}$, etc.) The results indicate that $E^{[\ell]} \to 0$ and that $E^{[\ell+1]}/E^{[\ell]} \to 0.106$ (approximately). The Jacobian at x_2 is

$$B = 2(1 - 2x_2) \approx 1.056$$

so the calculations confirm that $E^{[\ell+1]}/E^{[\ell]} \to hB$ as shown in equation (8.6). Because $hB \approx 0.1$, $E^{[\ell+1]} \approx 0.1 \times E^{[\ell]}$ gets smaller by a factor of about 0.1 per iteration, that is, $u^{[\ell]}$ gains about one decimal place per iteration.

---

5. Using the ODEs $x'(t) = -2y(t)^3$ and $y'(t) = 2x(t) - y(t)^3$ and the chain rule,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\big(x(t)^2 + \tfrac{1}{2}y(t)^4\big) &= 2x(t)x'(t) + 2y(t)^3y'(t) \\ &= 2x(t)(-2y(t)^3) + 2y(t)^3(2x(t) - y(t)^3) = -2y(t)^6.\end{aligned}$$

---

7. For $x'(t) = \lambda x(t)$, a typical step in the forward/backward Euler PECE method is

$$\begin{aligned}
&\text{P: } x_{n+1}^{[0]} &&= x_n + hf_n &&= (1+\widehat{h})x_n, \\
&\text{E: } f_{n+1}^{[0]} &&= f(t_{n+1}, x_{n+1}^{[0]}) &&= \lambda x_{n+1}^{[0]}, \\
&\text{C: } x_{n+1} &&= x_n + hf_{n+1}^{[0]} &&= x_n + \widehat{h}x_{n+1}^{[0]}, \\
&\text{E: } f_{n+1} &&= f(t_{n+1}, x_{n+1}) &&= \lambda x_{n+1}
\end{aligned}$$

and so $x_{n+1} = x_n + \widehat{h}(1+\widehat{h})x_n = (1+\widehat{h}+\widehat{h}^2)x_n$.

For absolute stability, we require $|x_{n+1}/x_n| < 1$, i.e.,

$$-1 < 1 + \widehat{h} + \widehat{h}^2 < 1.$$

The right hand inequality requires $\widehat{h}(1+\widehat{h}) < 0$, i.e., $\widehat{h} \in (-1, 0)$. The left hand inequality requires $2 + \widehat{h} + \widehat{h}^2 > 0$ which is true for all real $\widehat{h}$ since $2 + \widehat{h} + \widehat{h}^2 = (\widehat{h} + \frac{1}{2})^2 + \frac{7}{4} > 0$.

The interval of absolute stability is, therefore, $-1 < \widehat{h} < 0$.

---

9. The backward Euler method applied to $x'(t) = f(x(t))$ leads to $x_{n+1} = x_n + f_{n+1}$. Suppose that this equation has two solutions u and v, so

$$u = x_n + hf(u) \quad \text{and} \quad v = x_n + hf(v).$$

Subtracting these gives $u - v = h\big(f(u) - f(v)\big)$. Multiplying both sides by $(u - v)$ and assuming that f satisfies a one–sided Lipschitz condition, we find

$$(u-v)^2 = h(u-v)\big(f(u) - f(v)\big) \le h\gamma(u-v)^2$$

from which we deduce that $(1 - h\gamma)(u-v)^2 \le 0$. The left hand side of this inequality is non–negative when $h\gamma < 1$ and consequently both sides must vanish, from which we deduce that $u = v$. The equation must therefore have a unique solution.

---

11. With $f(u) = u - 0.2 - 0.2u(1-u)$, $f'(u) = 1 - 0.2(1-2u)$, the Newton–Raphson method for equation (8.8) is

$$u^{[\ell+1]} = u^{[\ell]} - \frac{f(u^{[\ell]})}{f'(u^{[\ell]})}$$

with $u_0 = 0.2$. The results are shown below

ℓ	0	1	2
$u^{[\ell]}$	0.2000	0.2364	0.2361
$u^{[\ell+1]} - u^{[\ell]}$	0.0364	−0.0003	

so $|u^{[2]} - u^{[1]}| < 0.001$ and two iterations are sufficient to meet the convergence criterion.

---

13. We extend Table 8.2 by adding two further rows which show that $(p^{[\ell]})^2/p^{[\ell+1]} \approx 4.75$ and $(q^{[\ell]})^2/q^{[\ell+1]} \approx 3.36$.

ℓ	0	1	2	3
$\boldsymbol{u}^{[\ell]}$	1.00	0.774647887	0.773901924	0.773901807
	1.00	1.042253521	1.041731347	1.041731265
$\widehat{\boldsymbol{E}}^{[\ell]}$	$2.2535_{10}{-1}$	$7.4596_{10}{-4}$	$1.1711_{10}{-7}$	$2.9131_{10}{-15}$
	$-4.2254_{10}{-2}$	$5.2217_{10}{-4}$	$8.1975_{10}{-8}$	$1.9628_{10}{-15}$
$\boldsymbol{E}^{[\ell]}$	$2.2610_{10}{-1}$	$7.4608_{10}{-4}$	$1.1711_{10}{-7}$	$2.8866_{10}{-15}$
	$-4.1731_{10}{-2}$	$5.2226_{10}{-4}$	$8.1975_{10}{-8}$	$1.9984_{10}{-15}$
$(p^{[\ell]})^2/p^{[\ell+1]}$	68.5186	4.7532	4.7510	
$(q^{[\ell]})^2/q^{[\ell+1]}$	3.3346	3.3272	3.3627	

---

15. When applied to $x'(t) = \lambda x(t)$, the PECE method of Exercise 8.12

$$\begin{aligned}
&\text{P: } x^{[0]}_{n+2} &&= x_{n+1} + \tfrac{1}{2}h(3f_{n+1} - f_n) &&= (1 + \tfrac{3}{2}\widehat{h})x_{n+1} - \tfrac{1}{2}\widehat{h}x_n,\\
&\text{E: } f^{[0]}_{n+2} &&= f(t_{n+2}, x^{[0]}_{n+2}) &&= \lambda x^{[0]}_{n+2},\\
&\text{C: } x_{n+2} &&= x_{n+1} + \tfrac{1}{2}h(f_{n+1} + f^{[0]}_{n+2}) &&= x_{n+1} + \tfrac{1}{2}\widehat{h}(x_{n+1} + x^{[0]}_{n+2}),\\
&\text{E: } f_{n+2} &&= f(t_{n+2}, x_{n+2}) &&= \lambda x_{n+2}.
\end{aligned}$$

Combining these gives the difference equation

$$x_{n+2} = (1 + \widehat{h} + \tfrac{3}{4}\widehat{h}^2)x_{n+1} - \tfrac{1}{4}\widehat{h}^2 x_n$$

whose auxiliary equation is

$$p(r) = r^2 - (1 + \widehat{h} + \tfrac{3}{4}\widehat{h}^2)r - \tfrac{1}{4}\widehat{h}^2.$$

To find the interval of absolute stability we may use Lemma 6.10 with

$$a = -(1 + \widehat{h} + \tfrac{3}{4}\widehat{h}^2), \qquad b = -\tfrac{1}{4}\widehat{h}^2.$$

The conditions $b < 1$ and $p(\pm 1) > 0$ lead to

$$\begin{array}{lll} b < 1: & \tfrac{1}{4}\widehat{h}^2 < 1 & \Longleftrightarrow \quad -2 < \widehat{h} < 2 \\ 1 + a + b > 0: & -\tfrac{1}{2}\widehat{h}(2 + \widehat{h}) > 0 & \Longleftrightarrow \quad -2 < \widehat{h} < 0 \\ 1 - a + b > 0: & 2 + \widehat{h} + \widehat{h}^2 > 0 & \text{which is true for all real } \widehat{h}. \end{array}$$

The interval of absolute stability is therefore $(-2, 0)$.

17. We extend the table shown in the solution to Exercise 3 to include a row to show the values of $|E^{[\ell+1]} - E^{[\ell]}|/(1 - r)$, where $r = hB \approx 0.106$.

ℓ	0	1	2	3	4
$u^{[\ell]}$	0.2	0.232	0.2356	0.2360	0.2360
$u^{[\ell+1]} - u^{[\ell]}$	0.032	0.0036	0.0004	0.0000	
$E^{[\ell]}$	$-3.61_{10}{-2}$	$-4.07_{10}{-3}$	$-4.33_{10}{-4}$	$-4.57_{10}{-5}$	$-4.83_{10}{-6}$
$E^{[\ell+1]}/E^{[\ell]}$	0.113	0.106	0.106	0.106	
$\dfrac{\lvert E^{[\ell+1]} - E^{[\ell]}\rvert}{1 - r}$	$3.58_{10}{-2}$	$4.068_{10}{-3}$	$4.33_{10}{-4}$	$4.57_{10}{-5}$	

($3.61_{10}{-2} \equiv 3.61 \times 10^{-2}$, etc.) Close agreement is seen between $|E^{[\ell]}|$ and the values in the last row.

Exercises 9

1. With $\theta = \frac{1}{2}$ in the general 2nd order RK(2) method (Table 9.5) gives the improved Euler method:

$$\begin{aligned} k_1 &= f\left(t_n, x_n\right) \\ k_2 &= f\left(t_n + ah, x_n + ahk_1\right) = f\left(t_n + h, x_n + hk_1\right) \\ x_{n+1} &= x_n + h\left(b_1 k_1 + b_2 k_2\right) = x_n + \tfrac{1}{2}h\left(k_1 + k_2\right) \end{aligned}$$

so, for $x'(t) = (1-2t)x(t)$, $h = 0.2$,

$$\begin{aligned} k_1 &= (1-2t_n)x_n \\ k_2 &= \left(1 - 2(t_n + h)\right)\left(x_n + hk_1\right) \\ x_{n+1} &= x_n + \tfrac{1}{2}h\left(k_1 + k_2\right) = x_n + 0.1\left(k_1 + k_2\right). \end{aligned}$$

$$\begin{array}{llll} n=0: & x_0 = 1, \quad t_0 = 0, & k_1 = 1, \quad k_2 = .72, & x_1 = 1.172 \\ n=1: & x_1 = 1.172,\ t_1 = 0.2, & k_1 = (1 - 2\times 0.2)1.172 = .7032 & \\ & & k_2 = \left(1 - 2(0.2+0.2)\right)\left(1.172 + 0.2\times .7032\right) = .2625 & \\ & & x_2 = x_1 + 0.1(k_1 + k_2) = 1.2686 & \end{array}$$

The value $x_2 = 1.2686$ is close to that obtained by the modified Euler method (1.2757) in Example 9.1.

The exact solution is $x(t) = \exp((\frac{1}{4} - (t - \frac{1}{2})^2)$ and so $x(t_2) = x(0.4) = 1.2712$. The global error at $t = 0.4$ is therefore $e_2 = 1.2712 - 1.2686 = 0.0026$.

---

3. The general s–stage RK method applied to the IVP $\boldsymbol{u}'(t) = \boldsymbol{f}(\boldsymbol{u}(t))$, where

$$\boldsymbol{u}(t) = \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}, \qquad \boldsymbol{f}\left(\begin{bmatrix} u \\ v \end{bmatrix}\right) = \begin{bmatrix} 1 \\ f(u,v) \end{bmatrix},$$

and $u(0) = 0$, $v(0) = \eta$ leads to

$$\boldsymbol{k}_i = \boldsymbol{f}(\boldsymbol{u}_n + h\sum_{j=1}^{s} a_{i,j}\boldsymbol{k}_j), \qquad \boldsymbol{u}_{n+1} = \boldsymbol{u}_n + h\sum_{i=1}^{s} b_i \boldsymbol{k}_i$$

with $\boldsymbol{u}_0 = [0,\eta]^T$. Then, supposing that the components of $\boldsymbol{k}_i$ are denoted by ℓ_i and k_i, we find

$$\boldsymbol{k}_i = \begin{bmatrix} \ell_i \\ k_i \end{bmatrix} = \begin{bmatrix} 1 \\ f(t_n + h\sum_{j=1}^{s} a_{i,j}\ell_j,\ x_n + h\sum_{j=1}^{s} a_{i,j}k_j) \end{bmatrix},$$

i.e., $\ell_i = 1$ and so

$$k_i = f(t_n + h\sum_{j=1}^{s} a_{i,j}, x_n + h\sum_{j=1}^{s} a_{i,j}k_j).$$

This will be in agreement with (9.6) if equation (9.7) holds, i.e., $c_i = \sum_{j=1}^{s} a_{i,j}$. Then

$$\boldsymbol{u}_{n+1} = \begin{bmatrix} t_{n+1} \\ x_{n+1} \end{bmatrix} = \begin{bmatrix} t_n \\ x_n \end{bmatrix} + h \sum_{i=1}^{s} b_i \begin{bmatrix} \ell_i \\ k_i \end{bmatrix},$$

and so

$$t_{n+1} = t_n + h \sum_{i=1}^{s} b_i, \quad x_{n+1} = x_n + h \sum_{i=1}^{s} b_i k_i.$$

The update x_{n+1} agrees with (9.5) and, in order to have $t_{n+1} - t_n = h$, we clearly require $\sum_{i=1}^{s} b_i = 1$.

---

5. Using the Taylor expansion for a function of two variables,

$$k_2 = f(t_n + h, x_n + hk_1) = f(t_n, x_n) + \mathscr{O}(h)$$

we find

$$\begin{aligned} x_{n+1} &= x_n + \tfrac{1}{2}h(k_1 + k_2) \\ &= x_n + \tfrac{1}{2}h(f(t_n, x_n) + \left[f(t_n, x_n) + \mathscr{O}(h)\right]) \\ &= x_n + hf(t_n, x_n) + \mathscr{O}(h^2). \end{aligned}$$

For the exact solution of the ODE $x'(t) = f(t, x(t))$,

$$\begin{aligned} x(t_n + h) &= x(t_n) + hx'(t_n) + \mathscr{O}(h^2) \\ &= x(t_n) + hf(t_n, x(t_n)) + \mathscr{O}(h^2). \end{aligned}$$

Under the localizing assumption $x_n = x(t_n)$, the difference is easily shown to be

$$x(t_{n+1}) - x_{n+1} = \mathscr{O}(h^2)$$

so, in view of Definition 9.3, the method is of order $p = 1$—it is therefore consistent with the given ODE.

---

7. In all methods $t_0 = 0$, $x_0 = 1$, $k_1 = f(0, x_0) = 1$, where $f(t, x) = (1 - 2t)x$.

k_2	k_3	k_4	x_1
		Improved Euler	
0.880000			$x_0 + hk_2$=1.094000
		Modified Euler	
0.945000			$x_0 + \frac{1}{2}h(k_1 + k_2)$=1.094500
		Heun	
0.964444	0.922390		$x_0 + \frac{1}{4}h(k_1 + 3k_3)$=1.094179
		Kutta 3rd order	
0.945000	0.871200		$x_0 + \frac{1}{6}h(k_1 + 4k_2 + k_3)$=1.094187
		4th order	
0.945000	0.942525	0.875402	$x_0 + \frac{1}{6}h(k_1 + 2k_2 + 2k_3 + k_4)$=1.094174

The value given by the 4th order method agrees with the exact solution $x(0.1)$ to 6 decimal places.

---

9. Applying the general 2nd order, 2–stage RK method from Table 9.5 to the ODE $x'(t) = \lambda x(t)$ we obtain $k_1 = \lambda x_n$,

$$\begin{aligned} k_2 &= \lambda(x_n + ahk_1) = (1 + a\widehat{h})x_n, \qquad (\widehat{h} = \lambda h) \\ x_{n+1} &= x_n + h\big((1-\theta)k_1 + \theta k_2\big) \\ &= (1 + \widehat{h} + a\theta\widehat{h}^2)x_n = (1 + \widehat{h} + \tfrac{1}{2}\widehat{h}^2)x_n \end{aligned}$$

since $a\theta = \frac{1}{2}$ for a 2nd order method. The exact solution has the property $x(t_{n+1}) = \mathrm{e}^{\widehat{h}}x(t_n)$, so its Taylor expansion is

$$x(t_{n+1}) = (1 + \widehat{h} + \tfrac{1}{2}\widehat{h}^2 + \tfrac{1}{6}\widehat{h}^3)x(t_n) + \mathscr{O}(h^4)$$

and therefore the LTE is given by

$$T_{n+1} = x(t_{n+1}) - x_{n+1} = \tfrac{1}{6}\widehat{h}^3 x(t_n) + \mathscr{O}(h^4),$$

where we have used the localizing assumption $x_n = x(t_n)$. Thus, the method can be of order at most two.

---

11. Using the given tableau we find

$$\begin{aligned} k_1 &= f(t_n, x_n) \\ k_2 &= f(t_n + h, x_n + \tfrac{1}{2}h(k_1 + k_2) \\ x_{n+1} &= x_n + \tfrac{1}{2}h(k_1 + k_2). \end{aligned}$$

In view of the last equation and $t_{n+1} = t_n + h$, k_2 can be re–written as

$$k_2 = f(t_{n+1}, x_{n+1})$$

and therefore, $x_{n+1} = x_n + \frac{1}{2}h(f(t_n, x_n) + f(t_{n+1}, x_{n+1})$, which is the Trapezoidal rule.

---

13. The required conditions are obtained by setting $b_3 = 0$ which gives the 2–stage version of a 3–stage RK method with tableau (with $c_2 = a_{2,1}$)

2-stage version: $\begin{array}{c|ccc} 0 & 0 & & \\ a & a & 0 & \\ - & - & - & - \\ \hline & \tilde{b}_1 & \tilde{b}_2 & 0 \end{array}$, 3-stage version: $\begin{array}{c|ccc} 0 & 0 & & \\ c_2 & a_{2,1} & 0 & \\ c_3 & a_{3,1} & a_{3,2} & 0 \\ \hline & b_1 & b_2 & b_3 \end{array}$

Suppose that the parameters of the 3–stage method satisfy the order conditions of Table 9.6. Then, choosing $a = a_{2,1}$ and $\tilde{b}_2 = \frac{1}{2}a$, $\tilde{b}_1 = 1 - \tilde{b}_2$ in the 2–stage method, the parameters of the 2–stage method satisfy the order conditions of Table 9.4.

---

15. To show that Heun's method is of order three, we follow a similar pattern to Exercises 9.5 and 9.6 except that the Taylor expansions for the numerical and exact solutions need to be computed with remainder $\mathscr{O}(h^4)$. This means that the k's need to be expanded up to h^2 terms. So

$$\begin{aligned}
k_2 &= f(t_n + \tfrac{1}{3}h, x_n + \tfrac{1}{3}hk_1) \\
&= f + \tfrac{1}{3}h(f_t + k_1 f_x) + \tfrac{1}{2}(\tfrac{1}{3}h)^2\big(f_{tt} + 2k_1 f_{xt} + k_1^2 f_{xx}\big) + \mathscr{O}(h^3) \\
&= f + \tfrac{1}{3}h(f_t + f f_x) + \tfrac{1}{18}h^2\big(f_{tt} + 2f f_{xt} + f^2 f_{xx}\big) + \mathscr{O}(h^3),
\end{aligned}$$

where f and its partial derivatives are all evaluated at (t_n, x_n). Similarly,

$$\begin{aligned}
k_3 &= f(t_n + \tfrac{2}{3}h, x_n + \tfrac{2}{3}hk_2) \\
&= f + \tfrac{2}{3}h(f_t + k_1 f_x) + \tfrac{1}{2}(\tfrac{2}{3}h)^2\big(f_{tt} + 2k_2 f_{xt} + k_2^2 f_{xx}\big) + \mathscr{O}(h^3) \\
&= f + \tfrac{2}{3}h(f_t + f f_x) + \tfrac{2}{9}h^2\big(f_{tt} + 2k_2 f_{xt} + k_2^2 f_{xx}\big) + \mathscr{O}(h^3),
\end{aligned}$$

and we delay substituting for k_2 in order to simplify the calculations. Then

$$\begin{aligned}
x_{n+1} &= x_n + \tfrac{1}{4}h(k_1 + 3k_3) \\
&= x_n + hf + \tfrac{1}{2}h^2(f_t + k_2 f_x) + \tfrac{1}{6}h^3\big(f_{tt} + 2k_2 f_{xt} + k_2^2 f_{xx}\big) + \mathscr{O}(h^4).
\end{aligned}$$

For the exact solution of the ODE $x'(t) = f(t, x(t))$, using the result of the previous exercise,

$$\begin{aligned}
x(t_n + h) &= x(t_n) + hx'(t_n) + \tfrac{1}{2}h^2 x''(t_n) + \tfrac{1}{6}h^3 x'''(t_n)\mathscr{O}(h^4) \\
&= x(t_n) + hf + \tfrac{1}{2}h^2\big(f_t + f f_x) \\
&\qquad + \tfrac{1}{6}h^3\big(f_{tt} + 2f f_{xt} + f^2 f_{xx} + f(f_t + f f_x)\big) + \mathscr{O}(h^4).
\end{aligned}$$

So, under the localizing assumption $x_n = x(t_n)$ we find

$$\begin{aligned}
T_{n+1} &= x(t_{n+1}) - x_{n+1} \\
&= \tfrac{1}{2}h^2\big(f - k_2)f_x\big) + \tfrac{1}{6}h^3\big(2(f - k_2)f_{xt} + (f^2 - k_2^2)f_{xx} + f(f_t + f f_x)\big)
\end{aligned}$$

and substituting $k_2 = f + \frac{1}{3}h(f_t + f f_x) + \mathscr{O}(h^2)$, this gives $T_{n+1} = \mathscr{O}(h^4)$ which, in view of Definition 9.3, verifies that the method of consistent of order three.

Exercises 10

1. When Heun's method is applied to $x'(t) = \lambda x(t)$, we find

$$\begin{aligned}
k_1 &= f(t_n, x_n) = \lambda x_n \\
k_2 &= f(t_n + \tfrac{1}{3}h, x_n + \tfrac{1}{3}hk_1) = \lambda(x_n + \tfrac{1}{3}hk_1) \\
&= \lambda(1 + \tfrac{1}{3}\widehat{h})x_n \\
k_3 &= f(t_n + \tfrac{2}{3}h, x_n + \tfrac{2}{3}hk_2) = \lambda(x_n + \tfrac{2}{3}hk_2) \\
&= \lambda(1 + \tfrac{2}{3}\widehat{h} + \tfrac{2}{9}\widehat{h}^2)x_n \\
x_{n+1} &= x_n + \tfrac{1}{4}h(k_1 + 3k_3) \\
&= x_n + \tfrac{1}{4}\widehat{h}\Big(1 + 3(1 + \tfrac{2}{3}\widehat{h} + \tfrac{2}{9}\widehat{h}^2)\Big)x_n = (1 + \widehat{h} + \tfrac{1}{2}\widehat{h}^2 + \tfrac{1}{6}\widehat{h}^3)x_n.
\end{aligned}$$

Similarly, for Kutta's 3rd order rule,

$$\begin{aligned}
k_1 &= f(t_n, x_n) = \lambda x_n \\
k_2 &= f(t_n + \tfrac{1}{2}h, x_n + \tfrac{1}{2}hk_1) = \lambda(x_n + \tfrac{1}{2}hk_1) \\
&= \lambda(1 + \tfrac{1}{2}\widehat{h})x_n \\
k_3 &= f(t_n + h, x_n - hk_1 + 2hk_2) = \lambda(x_n - hk_1 + 2hk_2) \\
&= \lambda(1 + \widehat{h} + \widehat{h}^2)x_n \\
x_{n+1} &= x_n + \tfrac{1}{6}h(k_1 + 4k_2 + k_3) \\
&= x_n + \tfrac{1}{6}\widehat{h}\Big(1 + 4(1 + \tfrac{1}{2}\widehat{h}) + (1 + \widehat{h} + \widehat{h}^2)\Big)x_n = (1 + \widehat{h} + \tfrac{1}{2}\widehat{h}^2 + \tfrac{1}{6}\widehat{h}^3)x_n.
\end{aligned}$$

The stability function $R(\widehat{h}) = x_{n+1}/x_n = 1 + \widehat{h} + \frac{1}{2}\widehat{h}^2 + \frac{1}{6}\widehat{h}^3$ is the same for both methods.

---

3. When the 4–stage method from Table 9.8 is applied to $x'(t) = \lambda x(t)$, we find

$$\begin{aligned}
k_1 &= f(t_n, x_n) = \lambda x_n \\
k_2 &= f(t_n + \tfrac{1}{2}h, x_n + \tfrac{1}{2}hk_1) = \lambda(x_n + \tfrac{1}{2}hk_1) \\
&= \lambda(1 + \tfrac{1}{2}\widehat{h})x_n \\
k_3 &= f(t_n + \tfrac{1}{2}h, x_n + \tfrac{1}{2}hk_2) = \lambda(x_n + \tfrac{1}{2}hk_2) \\
&= \lambda(1 + \tfrac{1}{2}\widehat{h} + \tfrac{1}{4}\widehat{h}^2)x_n \\
k_4 &= f(t_n + h, x_n + hk_3) = \lambda(x_n + hk_3) \\
&= \lambda(1 + \widehat{h} + \tfrac{1}{2}\widehat{h}^2 + \tfrac{1}{4}\widehat{h}^3)x_n \\
x_{n+1} &= x_n + \tfrac{1}{6}h(k_1 + 2k_2 + 2k_3 + k_4) \\
&= (1 + \widehat{h} + \tfrac{1}{2}\widehat{h}^2 + \tfrac{1}{3!}\widehat{h}^3 + \tfrac{1}{4!}\widehat{h}^4)x_n.
\end{aligned}$$

Thus, $R(\widehat{h}) = e^{\widehat{h}} + \mathscr{O}(h^5)$, appropriate for a 4th order method. Expanding the right hand side, reveals that we can write $R(\widehat{h}) = \frac{1}{4} + \frac{1}{3}(\widehat{h} + \frac{3}{2})^2 + \frac{1}{24}\widehat{h}^2(\widehat{h} + 2)^2$.

This is the sum of positive terms and consequently $R(\widehat{h}) > 0$ for all real $\widehat{h}$. This means that the equation $R(\widehat{h}) = -1$ can have no real roots.

The equation $R(\widehat{h}) = 1$ leads to

$$\widehat{h}(1 + \tfrac{1}{2}\widehat{h} + \tfrac{1}{6}\widehat{h}^2 + \tfrac{1}{12}\widehat{h}^3) = 0$$

which has a real root $\widehat{h} = 0$ together with the roots of the cubic $F(\widehat{h}) = 0$, where

$$F(\widehat{h}) \equiv 1 + \tfrac{1}{2}\widehat{h} + \tfrac{1}{6}\widehat{h}^2 + \tfrac{1}{24}\widehat{h}^3.$$

The derivative

$$F'(\widehat{h}) = \tfrac{1}{2} + \tfrac{1}{3}\widehat{h} + \tfrac{1}{8}\widehat{h}^2 = \tfrac{1}{8}(\widehat{h} + \tfrac{4}{3})^2 + \tfrac{5}{18} > 0$$

is always strictly positive so we may conclude that $F(\widehat{h})$ is a monotonic increasing function with $F(\widehat{h}) \to \pm\infty$ as $\widehat{h} \to \pm\infty$. The function $F(\widehat{h})$ therefore has precisely one real root.

Since $F(-2) = \frac{1}{3} > 0$ and $F(-3) = -\frac{1}{8} < 0$, this root lies between -2 and -3.

The interval of absolute stability requires $-1 < R(\widehat{h}) < 1$ and, since $|R(-1)| = \frac{1}{2} - \frac{1}{6} + \frac{1}{24} < 1$, we conclude that the required interval is $(h^*, 0)$, where h^* lies between -2 and -3.

---

5. The Newton–Raphson method to solve the equation $F(\widehat{h}) \equiv R(\widehat{h}) - (-1)^s = 0$ is defined by (see Section 8.4)

$$h^{[\ell+1]} = h^{[\ell]} - \frac{F(h^{[\ell]})}{F'(h^{[\ell]})} = h^{[\ell]} - \frac{R(h^{[\ell]}) - (-1)^s}{R'(h^{[\ell]})}.$$

So, when $s = 3$, we have

$$h^{[\ell+1]} = h^{[\ell]} - \frac{2 + h^{[\ell]} + \frac{1}{2}(h^{[\ell]})^2 + \frac{1}{6}(h^{[\ell]})^3}{1 + h^{[\ell]} + \frac{1}{2}(h^{[\ell]})^2}$$

and, when $s = 4$,

$$h^{[\ell+1]} = h^{[\ell]} - \frac{h^{[\ell]} + \frac{1}{2}(h^{[\ell]})^2 + \frac{1}{6}(h^{[\ell]})^3 + \frac{1}{24}(h^{[\ell]})^4}{1 + h^{[\ell]} + \frac{1}{2}(h^{[\ell]})^2 + \frac{1}{6}(h^{[\ell]})^3}.$$

Starting each iteration with $h^{[0]} = -2.5$, we obtain the values shown below.

	$h^{[0]}$	$h^{[1]}$	$h^{[2]}$	$h^{[3]}$	$h^{[4]}$
$s = 3$	-2.5	-2.5128	-2.5127	-2.5127	-2.5127
$s = 4$	-2.5	-2.8590	-2.7889	-2.7853	-2.7853

These lead to the estimates $h^* \approx -2.5127$ for $s = 3$ and $h^* \approx -2.7853$ for $s = 4$, in accordance with Table 10.1.

7. The stage values (9.6) for the general s–stage RK method applied to $x'(t) = \lambda x(t)$ lead to

$$k_i = \lambda\big(x_n + h\sum_{j=1}^{s} a_{i,j}k_j\big), \quad i = 1:s$$

which, in matrix–vector from become

$$\boldsymbol{k} = \lambda x_n \boldsymbol{e} + \widehat{h}\mathscr{A}\boldsymbol{k}$$

and so

$$\boldsymbol{k} = \lambda(I - \widehat{h}\mathscr{A})^{-1}\boldsymbol{e}x_n.$$

Then, using (9.5),

$$x_{n+1} = x_n + h\boldsymbol{b}^T\boldsymbol{k} = \Big(1 + \widehat{h}\boldsymbol{b}^T(I - \widehat{h}\mathscr{A})^{-1}\boldsymbol{e}\Big)x_n.$$

so the stability function is $R(\widehat{h}) = 1 + \widehat{h}\boldsymbol{b}^T(I - \widehat{h}\mathscr{A})^{-1}\boldsymbol{e}$.

9. With $\widehat{h} = p + \mathrm{i}q$,

$$\begin{aligned}
4\left|1 + \widehat{h} + \tfrac{1}{2}\widehat{h}^2\right|^2 &= 4\left(1 + p + \tfrac{1}{2}(p^2 - q^2)\right)^2 + 4q^2(1 + p^2) \\
&= \left(2 + 2p + p^2 - q^2)\right)^2 + 4q^2(1 + p^2) \\
&= \left(1 + (1 + p)^2 - q^2)\right)^2 + 4q^2(1 + p^2) \\
&= 1 + (1 + p)^4 + q^4 + 2(1 + p)^2 - 2q^2 + 2(1 + p)^2q^2 \\
&= \left(1 + (1 + p)^2 + q^2)\right)^2 - 4q^2.
\end{aligned}$$

Hence the boundary $|1 + \widehat{h} + \frac{1}{2}\widehat{h}^2|^2 = 1$ then leads to

$$\left(1 + (1 + p)^2 + q^2)\right)^2 = 4(1 + q^2). \qquad (*)$$

Taking the square root of both sides (only the positive root is possible) gives

$$\begin{aligned}
1 + (1 + p)^2 + q^2 &= 2\sqrt{1 + q^2} \\
(1 + p)^2 + (1 + q^2) - 2\sqrt{1 + q^2} + 1 &= 1 \\
(1 + p)^2 + \Big(\sqrt{1 + q^2} - 1\Big)^2 &= 1,
\end{aligned}$$

as required. This equation may be parameterized by writing

$$1 + p = \cos\phi, \qquad \sqrt{1 + q^2} - 1 = \sin\phi$$

so that $p = -1 + \cos\phi$ and $q = \pm\sqrt{(2+\sin\phi)\sin\phi}$ and the square root is real for $0 \le \phi \le \pi$. The locus of these points defines the boundary of the region of absolute stability of second order RK(2) methods shown in the top right of Figure 10.1.

[Expanding both sides of equation $(*)$ and setting $p = ah$, $q = bh$ leads to equation (10.3).]

---

11. The matrix

$$A = \begin{bmatrix} -5 & 2 \\ 2 & -5 \end{bmatrix},$$

has characteristic polynomial $(\lambda + 5)^2 + 4$ and, therefore, eigenvalues $\lambda = -5 \pm 2\mathrm{i}$. With $a = -5$, $b = \pm 2$ in (10.3) we obtain

$$P(h) \equiv 841h^3 - 580h^2 + 200h - 40 = 0.$$

This polynomial has a single real root since $P'(h) > 0$ for all h. Calculating $P(0.393) \approx 0.067 > 0$ and $P(0.392) = -0.067 < 0$ shows that the root lies between 0.392 and 0.393 and the method is absolutely stable for $0 < h < 0.392$.

For the matrix

$$A = \begin{bmatrix} -50 & -20 \\ 20 & -50 \end{bmatrix}$$

the corresponding range of stable stepsizes is $0 < h < 0.0392$.

---

13. By (9.4) the method has stability function

$$R(\widehat{h}) = \frac{1 + \frac{1}{2}\widehat{h} + \frac{1}{12}\widehat{h}^2}{1 - \frac{1}{2}\widehat{h} + \frac{1}{12}\widehat{h}^2}$$

and we have to prove that $|R(\widehat{h})| < 1$ when $\Re(\widehat{h}) < 0$. Let $\bar{z}$ denote the complex conjugate of the complex number z. We recall that $z\bar{z} = |z|^2$ and $z + \bar{z} = 2\Re z$. Let $D = 1 - \frac{1}{2}\widehat{h} + \frac{1}{12}\widehat{h}^2$ denote the denominator of $R(\widehat{h})$, then

$$R(\widehat{h}) = 1 + \frac{\widehat{h}}{D}$$

and so

$$|R(\widehat{h})|^2 - 1 = \left(1 + \frac{\widehat{h}}{D}\right)\left(1 + \frac{\bar{\widehat{h}}}{\bar{D}}\right) - 1 = 2\Re\left(\frac{\widehat{h}}{D}\right) + \frac{|\widehat{h}|^2}{|D|^2}.$$

Multiplying both numerator and denominator of the first term on the right hand side by $\overline{D}$ gives

$$\begin{aligned}|R(\widehat{h})|^2 - 1 &= 2\Re\left(\frac{\widehat{h}\overline{D}}{|D|^2}\right) + \frac{|\widehat{h}|^2}{|D|^2}\\ &= \Re\left(\frac{2\widehat{h}\overline{D} + |\widehat{h}|^2}{|D|^2}\right) = \frac{2 + \frac{1}{6}|\widehat{h}|^2}{|D|^2}\Re(\widehat{h})\end{aligned}$$

since

$$2\widehat{h}\overline{D} = 2\widehat{h} - |\widehat{h}|^2 + \tfrac{1}{6}\bar{\widehat{h}}|\widehat{h}|^2$$

and $\Re(\widehat{h}) = \Re(\bar{\widehat{h}})$. Thus $|R(\widehat{h})|^2 - 1 < 0$ when $\Re(\widehat{h}) < 0$.

---

15. When the given method is applied to $x'(t) = \lambda x(t)$, we find

$$\begin{aligned}&k_1 = f(t_n, x_n) = \lambda x_n\\ &k_2 = f(t_n + \tfrac{1}{2}h, x_n + \tfrac{5}{24}hk_1 + \tfrac{1}{3}hk_2 - \tfrac{1}{24}hk_3) = \lambda(x_n + \tfrac{5}{24}hk_1 + \tfrac{1}{3}hk_2 - \tfrac{1}{24}hk_3)\\ &k_3 = f(t_n + h, x_n + hk_2) = \lambda(x_n + hk_2).\end{aligned}$$

These give three linear equations which may be solved to give

$$k_1 = \lambda x_n, \quad k_2 = \lambda\frac{1 + \frac{1}{6}\widehat{h}}{1 - \frac{1}{3}\widehat{h} + \frac{1}{24}\widehat{h}^2}x_n, \quad k_3 = \lambda\frac{1 + \frac{2}{3}\widehat{h} + \frac{5}{24}\widehat{h}^2}{1 - \frac{1}{3}\widehat{h} + \frac{1}{24}\widehat{h}^2}x_n.$$

Substituting these into

$$x_{n+1} = x_n + \tfrac{1}{6}h(k_1 + 4k_2 + k_3)$$

we find

$$x_{n+1} = \frac{1 + \frac{2}{3}\widehat{h} + \frac{5}{24}\widehat{h}^2 + \frac{1}{24}\widehat{h}^3}{1 - \frac{1}{3}\widehat{h} + \frac{1}{24}\widehat{h}^2}x_n.$$

The stability function is

$$R(\widehat{h}) = \frac{1 + \frac{2}{3}\widehat{h} + \frac{5}{24}\widehat{h}^2 + \frac{1}{24}\widehat{h}^3}{1 - \frac{1}{3}\widehat{h} + \frac{1}{24}\widehat{h}^2},$$

a rational function whose numerator has a higher degree than its denominator. It therefore follows that $|R(\widehat{h})| \to \infty$ as $\widehat{h} \to -\infty$ so the method cannot be A_0 stable since this would require $|R(\widehat{h})| < 1$ for all $\Re(\widehat{h}) < 0$. Consequently, it cannot be A–stable.

We note that the Maclaurin expansion of $R(\widehat{h})$ gives

$$R(\widehat{h}) = 1 + \widehat{h} + \tfrac{1}{2}\widehat{h}^2 + \tfrac{1}{6}\widehat{h}^3 + \tfrac{5}{144}\widehat{h}^4 + \mathscr{O}(h^5) = \mathrm{e}^{\widehat{h}} + \mathscr{O}(h^4)$$

and so the method is of order at most $p = 3$.

Exercises 11

1. As in Example 11.1, $x_1 = 0$, $t_1 = h_0$ and the GE at the end of the first step $x(t_1) - x_1$ is equal to tol when $h_0^2 = \mathsf{tol}$, that is, $h_0 = \mathsf{tol}^{1/2}$.

   For the second step,
$$x_2 = x_1 + 2h_1 t_1,$$
   where $t_1 = \mathsf{tol}^{1/2}$, so $x_2 = 2\mathsf{tol}^{1/2}h_1$ while $x(t_2) = (\mathsf{tol}^{1/2} + h_1)^2$. The GE is
$$x(t_1) - x_1 = \mathsf{tol} + h_1^2$$
   and can equal tol only if $h_1 = 0$. It is not possible to obtain a GE of tol after two steps unless the GE after one step is less than this amount.

---

3. In this case $x''(t) = \lambda x'(t) = \lambda^2 x(t)$ and so the TS(1) algorithm (11.10) and (11.12) becomes
$$x_{n+1} = (1 + h_n\lambda)x_n, \qquad h_{\text{new}} = \left|\frac{2\,\mathsf{tol}}{\lambda^2 x_n}\right|^{1/2}, \qquad x_0 = 1.$$

---

5. From (11.13) and (11.11) in Example 11.2 we find
$$x_{n+1} = \big(1 + h_n(1-2t_n) + \tfrac{1}{2}h_n^2((1-2t_n)^2 - 2)\big)x_n, \qquad x_0 = 1$$
   and $|x_{n+1}/x_n| < 1$ if, and only if,
$$-1 < 1 + h_n(1-2t_n) + \tfrac{1}{2}h_n^2((1-2t_n)^2 - 2) < 1.$$
   The left hand inequality can be shown to hold for all real h_n when $(2t_n - 1)^2 > 8/3$. In this case $(1-2t_n)^2 - 2 > 0$ and the right hand inequality leads to
$$h_n < \frac{2(2t_n - 1)}{(2t_n - 1)^2 - 2}.$$
   When $t_n = 3$ this gives the bound $h_n < 0.435$ in agreement with Figure 11.2 (Right). More generally, this inequality is violated for points (t_n, h_n) in the shaded region in Figure 3 (Right).

---

7. Using (11.5) and (11.7) with $p = 3$, we obtain
$$x_{n+1} = x_n + h_n x_n' + \tfrac{1}{2}h_n^2 x_n'' + \tfrac{1}{6}h_n^3 x_n''', \qquad t_{n+1} = t_n + h_n,$$
$$h_{\text{new}} = \left|\frac{4!\,\mathsf{tol}}{x_n^{(4)}}\right|^{1/4}.$$

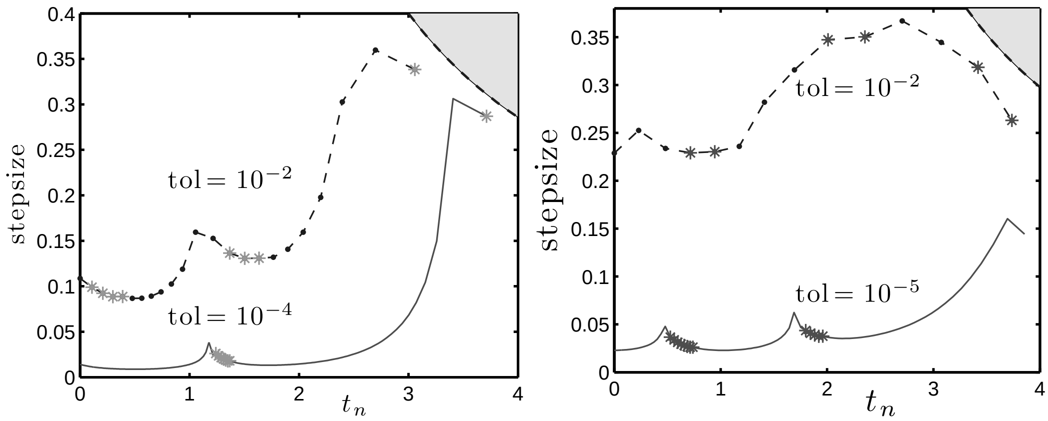

Figure 3: We reproduce parts of Figures 11.1 and 11.2 from the text showing the behaviour of h_n with t_n. The shaded regions indicate the bounds on h_n derived in Exercises 11.4 and 11.5.

Although expressions for x'_n and x''_n are available from Example 11.2 and for x'''_n from the previous exercise, it is more efficient to follow the procedure described in Exercise 3.6:

$$\begin{aligned} x'_{n+1} &= (1-2t_{n+1})x_{n+1} \\ x''_{n+1} &= (1-2t_{n+1})x'_{n+1} - 2x_{n+1} \\ x'''_{n+1} &= (1-2t_{n+1})x''_{n+1} - 4x'_{n+1} \end{aligned}$$

and differentiating $x'''(t) = (1-2t)x''(t) - 4x'(t)$ leads to

$$x_n^{(4)} = (1-2t_{n+1})x'''_{n+1} - 6x''_{n+1}$$

which enables the formula for h_{new} to be evaluated. This completes the specification of the algorithm.

---

9. The backward Euler method gives

$$x_{n+1} = x_n + h_n x'_{n+1}, \qquad t_{n+1} = t_n + h_n,$$

for which the LTE at the end of the current step is (see Example 4.9 with $\theta = 1$)

$$T_{n+1} = -\tfrac{1}{2!}h_n^2 x''(t_n) + \mathscr{O}(h_n^2).$$

Using the same argument as in Example 11.4, this can be approximated by the negative of the expression given in (11.16).

If forward and backward Euler methods are used as a predictor–corrector pair as described in Section 8.3, then (8.11) provides an estimate of the LTE of the corrector (backward Euler). From Example 4.9 we find that the error constants are $C_2^* = \frac{1}{2}$ and $C_2 = -\frac{1}{2}$ so (8.11) gives, with $k = p = 1$,

$$\widehat{T}_{n+1} = -\tfrac{1}{2}(x_{n+1} - x_{n+1}^{[0]})$$

which, because $x_{n+1}^{[0]} = x_n + h_n x_n'$ and $x_{n+1} = x_n + h_n x_{n+1}'$ gives

$$\widehat{T}_{n+1} = -\tfrac{1}{2}h_n(x_{n+1}' - x_n').$$

This is the negative of the expression given in (11.16).

---

11. For the Trapezoidal rule we follow Example 11.5. Because the estimate $\widehat{T}_n$ of the LTE involves three time levels, we have to initiate the process by computing two steps before we can calculate a new stepsize. So with $h_0 = h_1 = \mathsf{tol} = 0.01$, $t_1 = h_0 = 0.01$, $t_2 = h_0 + h_1 = 0.02$.

$$x_1 = x_0 + \tfrac{1}{2}h_0(x_0' + x_1'), \qquad x_1 = \frac{1 + \frac{1}{2}h_0}{1 - \frac{1}{2}h_0(1 - 2h_0)} = 1.0099$$

$$x_2 = x_1 + \tfrac{1}{2}h_1(x_1' + x_2'), \qquad x_2 = \frac{1 + \frac{1}{2}h_1(1 - 2t_1)}{1 - \frac{1}{2}h_1(1 - 2t_2}x_1 = 1.0198.$$

Since $x'(t) = (1 - 2t)x(t)$, we can compute $x_0' = 1$, $x_1' = 0.9897$, $x_2' = 0.9790$ so (11.20) gives $\widehat{T}_2 = 4.156 \times 10^{-7}$. This is certainly smaller than tol so the step is accepted. Finally, (11.19) gives $h_{\text{new}} = 0.2887$. (Notice from the dashed curve in Figure 11.5 (Right) that subsequent time steps are of roughly the same size as this. This suggests that it is a reasonable strategy to choose very small values for initial time steps.)

---

13. We follow the structure in Exercise 8.18. Under the localizing assumption, the LTEs of AB(2) and Trapezoidal rule are given by, respectively,

$$\begin{aligned} T_{n+1}^* &= x(t_{n+1}) - x_{n+1}^{[0]} = \tfrac{1}{12}\left(2 + 3\tfrac{h_{n-1}}{h_n}\right)h_n^3 x'''(t_n) + \mathscr{O}(h^4) \\ T_{n+1} &= x(t_{n+1}) - x_{n+1} = -\tfrac{1}{12}h_n^3 x'''(t_n) + \mathscr{O}(h^4), \end{aligned}$$

where $x_{n+1}^{[0]}$ is the result of using AB(2) as a predictor and the LTE of the Trapezoidal rule is derived in Example 4.9.

From $T_{n+1}^* - T_{n+1}$ we find

$$\begin{aligned} x_{n+1} - x_{n+1}^{[0]} &= \tfrac{1}{4}h_n^2(h_n + h_{n-1})x'''(t_n) + \mathscr{O}(h^4), \\ h_n^2 x'''(t_n) &= 4\frac{x_{n+1} - x_{n+1}^{[0]}}{h_n + h_{n-1}} + \mathscr{O}(h^2). \end{aligned}$$

The estimate $\widehat{T}_{n+1}$ of the LTE is then based on the leading term in T_{n+1}:

$$\widehat{T}_{n+1} = -\tfrac{1}{12}h_n^3 x'''(t_n) = -\tfrac{1}{3}\frac{h_n}{h_n + h_{n-1}}(x_{n+1} - x_{n+1}^{[0]}) + \mathscr{O}(h^4).$$

This reduces to the expression in the solution to Exercise 8.12 when $h_{n-1} = h_n = h$. From equation (11.23) (see previous solution)

$$x_{n+1}^{[0]} = x_n + h_n x_n' + \tfrac{1}{2}h_n^2 \frac{x_n' - x_{n-1}'}{h_{n-1}}$$

and $x_{n+1} = x_n + \frac{1}{2}h_n(x_n' + x_{n+1}')$ so

$$\begin{aligned} x_{n+1} - x_{n+1}^{[0]} &= \tfrac{1}{2}h_n(x_n' + x_{n+1}') - h_n x_n' - \tfrac{1}{2}h_n^2 \frac{x_n' - x_{n-1}'}{h_{n-1}} \\ &= \tfrac{1}{2}h_n^2\left(\frac{x_{n+1}' - x_n'}{h_n} - \frac{x_n' - x_{n-1}'}{h_{n-1}}\right) \end{aligned}$$

which, when substituted into the earlier expression for $\widehat{T}_{n+1}$ leads to (11.20).

---

15. The modified Euler method is given by Table 9.5 with $\theta = 1$:

$$\begin{aligned} k_1 &= f(t_n, x_n) \\ k_2 &= f(t_n + \tfrac{1}{2}h, x_n + \tfrac{1}{2}hk_1) \\ x_{n+1}^{\langle 2\rangle} &= x_n + hk_2, \end{aligned}$$

while Kutta's 3rd order rule is

$$\begin{aligned} k_1 &= f(t_n, x_n) \\ k_2 &= f(t_n + \tfrac{1}{2}h, x_n + \tfrac{1}{2}hk_1) \\ k_3 &= f(t_n + h, x_n - hk_1 + 2hk_2) \\ x_{n+1}^{\langle 3\rangle} &= x_n + \tfrac{1}{6}h(k_1 + 4k_2 + k_3). \end{aligned}$$

The definitions of k_1 and k_2 are clearly the same in both methods. An estimate $\widehat{T}_{n+1}$ of the LTEof the modified Euler method is then

$$\widehat{T}_{n+1} = x_{n+1}^{\langle 3\rangle} - x_{n+1}^{\langle 2\rangle} = \tfrac{1}{6}h(k_1 - 2k_2 + k_3).$$

With $x_0 = 1$, $t_0 = 0$, $h_0 = \mathsf{tol} = 0.01$, when applied to (11.9),

Step 1.

$$\begin{aligned} k_1 &= (1 - 2t_0)x_0 && = 1.0000 \\ k_2 &= (1 - 2(t_0 + \tfrac{1}{2}h_0))(x_0 + \tfrac{1}{2}h_0 k_1) && = 0.9949 \\ k_3 &= (1 - 2(t_0 + h_0))(x_0 - h_0(k_1 - 2k_2)) && = 0.9897 \\ x_1 &= x_0 + hk_2 && = 1.0099 \\ \widehat{T}_1 &= \tfrac{1}{6}h(k_1 - 2k_2 + k_3) && = 3.3163 \times 10^{-7}. \end{aligned}$$

Since $|\widehat{T}_1| < \mathsf{tol}$ the step is accepted so $t_1 = h_0 = 0.01$ and (see Example 11.7)

$$h_{\text{new}} = h_0 \left|\frac{\mathsf{tol}}{\widehat{T}_1}\right|^{1/3} = 0.3112.$$

Step 2.

$$\begin{aligned}
k_1 &= (1-2t_1)x_1 && = 0.9898 \\
k_2 &= (1-2(t_1+\tfrac{1}{2}h_1))(x_1+\tfrac{1}{2}h_1k_1) && = 0.7784 \\
k_3 &= (1-2(t_1+h_1))(x_1-h_1(k_1-2k_2)) && = 0.4242 \\
x_2 &= x_1+hk_2 && = 1.2522 \\
\widehat{T}_2 &= \tfrac{1}{6}h(k_1-2k_2+k_3) && = 7.4140\times 10^{-3}.
\end{aligned}$$

Since $|\widehat{T}_2| < \mathsf{tol}$ the step is accepted and $t_2 = t_1 + h_1 = 0.3213$,

$$h_{\text{new}} = h_1 \left|\frac{\mathsf{tol}}{\widehat{T}_2}\right|^{1/3} = 0.3439.$$

Exercises 12

1. The fixed points of the system

$$
\begin{aligned}
x'(t) &= y(t) - y^2(t) \\
y'(t) &= x(t) - x^2(t)
\end{aligned}
$$

occur when $x'(t) = y'(t) = 0$. These equations give the four points $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$.

$$
\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} y(t) - y^2(t) \\ x(t) - x^2(t) \end{bmatrix}, \qquad \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{x}) = \begin{bmatrix} 0 & 1-2y \\ 1-2x & 0 \end{bmatrix}.
$$

Fixed Point	Jacobian	Eigenvalues	Stability
$(0,0)$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	± 1	Unstable
$(1,0)$	$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$	$\pm \mathrm{i}$	Undecided
$(0,1)$	$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$	$\pm \mathrm{i}$	Undecided
$(1,1)$	$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$	± 1	Unstable

---

3. Suppose that A has an eigenvalue λ_A with corresponding eigenvector $\boldsymbol{v}$, then

$$A\boldsymbol{v} = \lambda_A \boldsymbol{v}$$

and so

$$B\boldsymbol{v} = (I + hA)\boldsymbol{v} = \boldsymbol{v} + h(A\boldsymbol{v}) = \boldsymbol{v} + h(\lambda_A \boldsymbol{v}) = (1 + h\lambda_A)\boldsymbol{v}.$$

Hence $\boldsymbol{v}$ is also an eigenvector of B corresponding to an eigenvalue $\lambda_B = 1 + h\lambda_A$.

The converse is also true: Suppose that B has an eigenvalue λ_B with corresponding eigenvector $\boldsymbol{v}$, then

$$\lambda_B \boldsymbol{v} = B\boldsymbol{v}$$

and so

$$\lambda_B \boldsymbol{v} = B\boldsymbol{v} = (I + hA)\boldsymbol{v} = \boldsymbol{v} + hA\boldsymbol{v}.$$

Rearranging, we find

$$A\boldsymbol{v} = \frac{1}{h}(\lambda_B - 1)\boldsymbol{v}.$$

Hence $\boldsymbol{v}$ is also an eigenvector of A corresponding to an eigenvalue

$$\lambda_A = \frac{1}{h}(\lambda_B - 1) \quad \Rightarrow \quad \lambda_B = 1 + h\lambda_A.$$

---

5. When $f(x) = 2x(1-x)$

$$x + \tfrac{1}{2}hf(x) = x + hx(1-x) = x(1+h-hx)$$

and so

$$\begin{aligned} x + \tfrac{1}{2}hf(x) = 0 \quad &\Rightarrow \quad x = 0, x = 1 + 1/h \\ x + \tfrac{1}{2}hf(x) = 1 \quad &\Rightarrow \quad hx(1-x) = 1 - x, \ \Rightarrow x = 1, 1/h. \end{aligned}$$

Hence (12.11) has the four fixed points

$$x_1^* = 0, \quad x_2^* = 1, \quad x_3^* = 1/h \ \text{ and } \ x_4^* = 1 + 1/h.$$

The iteration is given by $x_{n+1} = F(x_n)$, where

$$F(x) = x + hf\big(x + \tfrac{1}{2}hf(x)\big).$$

Hence, by the chain rule,

$$\begin{aligned} F'(x) &= 1 + h\frac{\mathrm{d}}{\mathrm{d}x} f\big(x + \tfrac{1}{2}hf(x)\big) \\ &= 1 + hf'\big(x + \tfrac{1}{2}hf(x)\big)\frac{\mathrm{d}}{\mathrm{d}x}\big(x + \tfrac{1}{2}hf(x)\big) = 1 + hf'\big(x + \tfrac{1}{2}hf(x)\big)(1 + \tfrac{1}{2}hf'(x)). \end{aligned}$$

---

7. Applying AB(2) to $x'(t) = f(x(t))$ gives

$$x_{n+2} = x_{n+1} + \tfrac{1}{2}h(3f(x_{n+1}) - f(x_n)).$$

With $y_n = x_{n+1}$, $z_n = x_n$ we have $z_{n+1} = y_n$ and

$$y_{n+1} = y_n + \tfrac{1}{2}h(3f(y_n) - f(z_n))$$

so

$$\boldsymbol{x}_{n+1} = \begin{bmatrix} y_{n+1} \\ z_{n+1} \end{bmatrix} = \begin{bmatrix} y_n + \tfrac{1}{2}h(3f(y_n) - f(z_n)) \\ y_n \end{bmatrix}.$$

Hence the two–step method AB(2) applied to a scalar ODE may be written as a one–step vector system $\boldsymbol{x}_{n+1} = \boldsymbol{F}(\boldsymbol{x}_n)$, where

$$\boldsymbol{F}(\boldsymbol{x}_n) = \begin{bmatrix} y_n + \tfrac{1}{2}h(3f(y_n) - f(z_n)) \\ y_n \end{bmatrix}, \qquad \boldsymbol{x}_n = \begin{bmatrix} y_n \\ z_n \end{bmatrix}.$$

The fixed points of this system are defined by

$$\boldsymbol{x} = \boldsymbol{F}(\boldsymbol{x}), \quad \Rightarrow \begin{bmatrix} y + \frac{1}{2}h(3f(y) - f(z)) \\ y \end{bmatrix} = \begin{bmatrix} y \\ z \end{bmatrix}$$

so $y = z$ and $f(y) = 0$. Hence, when $f(x) = 2x(1-x)$, there are two fixed points $(y, z) = (0, 0)$ and $(1, 1)$. When $\boldsymbol{F}(\boldsymbol{x}) = [F(y,z), G(y,z)]^T$, then its Jacobian is

$$\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}(\boldsymbol{x}) = \begin{bmatrix} F_y(y,z) & F_z(y,z) \\ G_y(y,z) & G_z(y,z) \end{bmatrix} = \begin{bmatrix} 1 + \frac{3}{2}hf'(y) & -\frac{1}{2}hf'(z) \\ 1 & 0 \end{bmatrix}.$$

When $f(x) = 2x(1-x)$,

$$\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}(\boldsymbol{x}) = \begin{bmatrix} 1 + 3h(1-2y) & -h(1-2z) \\ 1 & 0 \end{bmatrix}$$

and we find

$\boldsymbol{x}_1^* = (0,0)$: The Jacobian $\left[\begin{smallmatrix} 1+3h & -h \\ 1 & 0 \end{smallmatrix}\right]$ has characteristic polynomial

$$p(\lambda) = \lambda^2 - (1+3h)\lambda + h$$

and, by Lemma 6.10 (Jury Conditions), this has roots inside the unit circle (the strict root condition) if $p(0) < 1$ and $p(\pm 1) > 0$. Since $p(1) = -2h < 0$, this fixed point is unstable.

$\boldsymbol{x}_2^* = (1,1)$: The Jacobian $\left[\begin{smallmatrix} 1-3h & h \\ 1 & 0 \end{smallmatrix}\right]$ has characteristic polynomial $p(\lambda) = \lambda^2 - (1-3h)\lambda - h$ for which $p(0) = h < 1$, $p(1) = 2h > 0$ and $p(-1) = 2 - 4h > 0$ for $h < \frac{1}{2}$. Hence, the fixed point is stable for $0 < h < \frac{1}{2}$ in agreement with the bifurcation diagram shown in Figure 12.3.

---

9. The Trapezoidal rule applied to $x'(t) = x(t)(X - x(t))$ leads to

$$x_{n+1} = x_n + \tfrac{1}{2}h\Big[x_{n+1}(X - x_{n+1}) + x_n(X - x_n)\Big].$$

(a) Rearranging and "completing the squares" gives

$$\begin{aligned} x_{n+1} &= x_n + \tfrac{1}{2}h\Big[x_{n+1}(X - x_{n+1}) + x_n(X - x_n)\Big] \\ \tfrac{1}{2}hx_{n+1}^2 + \tfrac{1}{2}hx_n^2 &+ (1 - \tfrac{1}{2}hX)x_{n+1} - (1 + \tfrac{1}{2}hX)x_n = 0 \\ x_{n+1}^2 + x_n^2 &+ (\frac{2}{h} - X)x_{n+1} - (\frac{2}{h} + X)x_n = 0 \qquad (*) \\ \left(x_{n+1} + \frac{2}{h} - \frac{X}{2}\right)^2 + \left(x_n - \frac{2}{h} - \frac{X}{2}\right)^2 &= \left(\frac{2}{h} - \frac{X}{2}\right)^2 + \left(\frac{2}{h} + \frac{X}{2}\right)^2 \\ &= \frac{2}{h^2} + \frac{X^2}{2} \end{aligned}$$

which is the equation of a circle centred at $x_n = \frac{2}{h} + \frac{X}{2}$, $x_{n+1} = -\frac{2}{h} + \frac{X}{2}$ and passing through the origin.

(b) Fixed points are defined by by the property $x_{n+1} = x_n = x^*$ which, in the $x_n x_{n+1}$ phase plane, corresponds to the intersection of the line $x_{n+1} = x_n$ with the circle of part (a). There are two intersections given by $\boldsymbol{x}_1^* = (0,0)$ and $\boldsymbol{x}_2^* = (X,X)$.

(c) Differentiating both sides of the equation labelled $(*)$ with respect to x_n gives

$$\frac{\mathrm{d}x_{n+1}}{\mathrm{d}x_n} = \frac{\frac{2}{h} + X - 2x_n}{\frac{2}{h} - X + 2x_{n+1}}$$

so

$$\left|\frac{\mathrm{d}x_{n+1}}{\mathrm{d}x_n}\right|_{\boldsymbol{x}=\boldsymbol{x}_1^*} = \left|\frac{2+hX}{2-hX}\right| > 1$$
$$\left|\frac{\mathrm{d}x_{n+1}}{\mathrm{d}x_n}\right|_{\boldsymbol{x}=\boldsymbol{x}_2^*} = \left|\frac{2-hX}{2+hX}\right| < 1$$

and the fixed point $\boldsymbol{x}_1^* = (0,0)$ is unstable while $\boldsymbol{x}_2^* = (X,X)$ is stable. Clearly $\mathrm{d}x_{n+1}/\mathrm{d}x_n > 0$ at both fixed points when $hX < 2$.

(d) At each step x_{n+1} has to be found by solving a quadratic equation. Using the quadratic formula, the roots of the equation labelled $(*)$ are

$$x_{n+1} = \frac{X}{2} - \frac{1}{h} \pm \left(R^2 - \left(x_n - \frac{1}{h} - \frac{X}{2}\right)^2\right)^{1/2}.$$

We use the positive square root so that $x_{n+1} \to x_n$ as $h \to 0$ (see, for example, Exercises 4.2–4.4 and 8.1). This can then be used to calculate x_1 in the two cases.

Cobweb diagrams for $hX = 5$ and $hX = 1$ are shown in Figure 4.

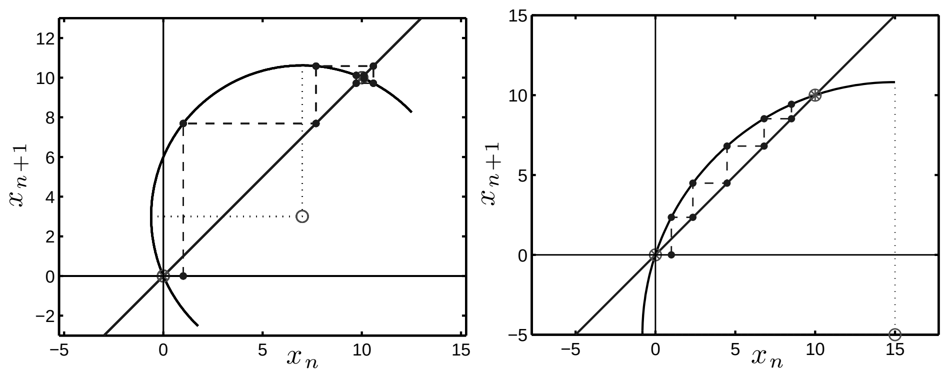

Figure 4: Cobweb diagrams for Exercise 12.9 with $X = 10$, $hX = 5$ (Left) and $hX = 1$ (Right).

Exercises 13

1. The argument used to progress from (2.15) to (2.16) may be used to deduce that (13.7) leads to
$$\hat{e}_n = \sum_{j=1}^{n}(1+h\lambda)^{n-j}\widehat{T}_j.$$
It was shown in Example 13.1 that $|\widehat{T}_j| \le Ch^3$ so, following the proof of Theorem 2.4,
$$\begin{aligned}|\hat{e}_n| &\le (Ch^3)\sum_{j=1}^{n}|1+h\lambda|^{n-j}\\ &\le (Ch^3)\mathrm{e}^{|\lambda|t_\mathrm{f}}n \le Ch^2 t_\mathrm{f}\mathrm{e}^{|\lambda|t_\mathrm{f}}\end{aligned}$$
(since $nh = t_n \le t_\mathrm{f}$) and consequently the global error for the modified equation is second order.

---

3. Suppose that $y(t)$ is a solution of the alternative modified equation $y'(t) = \mu y(t)$, where $\mu = \lambda(1-\frac{1}{2}\lambda h + \frac{1}{3}\lambda^2h^2)$. Using equation $(*)$ from the solution to the previous exercise we find, using $\mu^2 = \lambda^2(1-\lambda h) + \mathscr{O}(h^2)$,
$$\begin{aligned}\widehat{T}_{n+1} &= (1+\mu h + \tfrac{1}{2}\mu^2h^2)y(t_n) + \mathscr{O}(h^3) - (1+\lambda h)y(t_n)\\ &= (1+\lambda h(1-\tfrac{1}{2}\lambda h+\tfrac{1}{3}\lambda^2h^2) + \tfrac{1}{2}\lambda^2h^2(1-\lambda h) - (1+\lambda h))y(t_n) + \mathscr{O}(h^3)\\ &= -\tfrac{1}{6}\lambda^3h^3 + \mathscr{O}(h^3)\end{aligned}$$
and Euler's method is consistent of order $p = 2$ with the modified equation.

---

5. The backward Euler method $x_{n+1} = x_n + hf_{n+1}$ implies that $x_{n+1} = x_n + \delta_n$, where $\delta_n = hf_{n+1}$ (so that $\delta_n = \mathscr{O}(h)$). However,
$$\begin{aligned}f_{n+1} &= f(x_{n+1}) = f(x_n+\delta_n)\\ &= f(x_n) + \delta_n f'(x_n) + \mathscr{O}(h^2)\end{aligned}$$
which, together with $\delta_n = hf_{n+1}$, gives
$$\begin{aligned}(1-hf'(x_n))\delta_n &= hf_n + \mathscr{O}(h^3)\\ \delta_n &= (1-hf'(x_n))^{-1}hf_n + \mathscr{O}(h^3)\\ &= (1+hf'(x_n))hf_n + \mathscr{O}(h^3),\end{aligned}$$
where we have used the binomial expansion $(1-z)^{-1} = 1+z+\mathscr{O}(z^2)$. Therefore
$$x_{n+1} = x_n + (1+hf'(x_n))hf_n + \mathscr{O}(h^3).$$

Using (13.4) with the localizing assumption $x_n = y(t_n)$, we find

$$\begin{aligned}
\widehat{T}_{n+1} &= y(t_{n+1}) - x_{n+1} \\
&= \big(y(t_n) + hf(y(t_n)) + h^2\big(g(y(t_n)) + \tfrac{1}{2}f'(y(t_n))f(y(t_n))\big) + \mathscr{O}(h^3)\big) \\
&\qquad\qquad -\big(y(t_n) + (1 + hf'(y(t_n)))hf(y(t_n)) + \mathscr{O}(h^3)\big) \\
&= h^2\big(g(y(t_n)) - \tfrac{1}{2}f'(y(t_n))f(y(t_n))\big) + \mathscr{O}(h^3)
\end{aligned}$$

and we shall have consistency of order two ($\widehat{T}_{n+1} = \mathscr{O}(h^3)$) by choosing $g(y) - \frac{1}{2}f'(y)f(y)$. The modified equation (13.2) then becomes

$$y'(t) = f(y(t)) + \tfrac{1}{2}hf'(y(t)).$$

---

7. Suppose that $\boldsymbol{v}$ is an eigenvector of A with corresponding eigenvalue $\lambda : A\boldsymbol{v} = \lambda\boldsymbol{v}$. Then $\boldsymbol{v}$ is also an eigenvector of $\widehat{A}$ with corresponding eigenvalue $\widehat{\lambda} = \lambda(1 - \frac{1}{2}h\lambda)$ (see Exercise 10.12 and (13.6)). The respective solutions of the IVPs $\boldsymbol{x}'(t) = A\boldsymbol{x}(t)$, $\boldsymbol{x}(0) = \boldsymbol{v}$ and $\boldsymbol{y}'(t) = \widehat{A}\boldsymbol{y}(t)$, $\boldsymbol{y}(0) = \boldsymbol{v}$ are

$$\boldsymbol{x}(t) = \mathrm{e}^{\lambda t}\boldsymbol{v}, \qquad \boldsymbol{y}(t) = \mathrm{e}^{\widehat{\lambda} t}\boldsymbol{v}.$$

We now consider three possibilities.

(a) A is positive definite: Then $\lambda > 0$ and so $\widehat{\lambda} < \lambda$ and the solution of Euler's method, because it is closer to the solution of the modified IVPthan it is to the original ivp, grows more slowly than the exact solution.

(b) A is negative definite: Then $0 > \lambda > \widehat{\lambda}$ and the solution of Euler's method decays more rapidly than the exact solution.

(c) A is skew–symmetric: $A^T = -A$. Then λ is imaginary. Suppose that $\lambda = \mathrm{i}\mu$, where $\mu \in \Re$, then $\widehat{\lambda} = \frac{1}{2}h\mu^2 + \mathrm{i}\mu$. Thus, while each component of the exact solution is constant in time: $|x_j(t)| = |v_j|$, all components of the modified equation grow in time: $|y_j(t)| = \mathrm{e}^{h\mu^2/2}|v_j|$.

---

9. Following the process described in Example 13.3, We suppose that the modified equation is a system of two ODEs with dependent variables $x(t)$ and $y(t)$. The LTE of the given method is, therefore,

$$\widehat{\boldsymbol{T}}_{n+1} = \begin{bmatrix} x(t+h) - x(t) + hy(t) \\ y(t+h) - y(t) - hx(t+h) \end{bmatrix}, \quad t = nh,$$

which differs from (13.13) in that $x(t)$ in the second component is replaced by $x(t+h)$. Taylor expansion gives

$$\widehat{\boldsymbol{T}}_{n+1} = h \begin{bmatrix} x'(t) + \frac{1}{2}hx''(t) + y(t) \\ y'(t) + \frac{1}{2}hy''(t) - x(t) - hx'(t) \end{bmatrix} + \mathscr{O}(h^3)$$

We now suppose that the modified equations take the form

$$\begin{aligned} x'(t) &= -y(t) + ha(x,y), \\ y'(t) &= \quad x(t) + hb(x,y), \end{aligned}$$

where the functions $a(x,y)$ and $b(x,y)$ are to be determined. Differentiating these with respect to t gives

$$x''(t) = -y'(t) + \mathscr{O}(h) = -x(t) + \mathscr{O}(h), \qquad y''(t) = x'(t) + \mathscr{O}(h) = -y(t) + \mathscr{O}(h)$$

and substituting the results into the above expression for $\widehat{\boldsymbol{T}}_{n+1}$ leads to

$$\widehat{\boldsymbol{T}}_{n+1} = h^2 \begin{bmatrix} a(x,y) + \frac{1}{2}x''(t) \\ b(x,y) + \frac{1}{2}y''(t) + y(t) \end{bmatrix} + \mathscr{O}(h^3) = h^2 \begin{bmatrix} a(x,y) - \frac{1}{2}x(t) \\ b(x,y) + \frac{1}{2}y(t) \end{bmatrix} + \mathscr{O}(h^3).$$

Therefore $\widehat{\boldsymbol{T}}_{n+1} = \mathscr{O}(h^3)$ on choosing $a(x,y) = \frac{1}{2}x$ and $b(x,y) = -\frac{1}{2}y$. Our modified system of equations is, therefore,

$$\begin{aligned} x'(t) &= -y(t) + \tfrac{1}{2}hx(t), \\ y'(t) &= \quad x(t) - \tfrac{1}{2}hy(t) \end{aligned}$$

in agreement with (13.19). Using these we find

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(x^2(t) - hx(t)y(t) + y^2(t)\right) &= 2x(t)x'(t) + 2y(t)y'(t) - h\big(x'(t)y(t) + x(t)y'(t)\big) \\ &= 2(x(t) - hy(t))x'(t) + (2y(t) - hx(t))y'(t)) = 0 \end{aligned}$$

and so $x^2(t) - hx(t)y(t) + y^2(t) =$ constant, as required.

---

11. From (13.30)

$$u''(t) = \frac{\mathrm{d}}{\mathrm{d}t}u'(t) = -v'(t) = -f(u(t))$$

and so u satisfies $u''(t) + f(u(t)) = 0$.

Conversely, suppose that $u''(t) + f(u(t)) = 0$. If v is defined by $v(t) = -u(t)$, then

$$f(u(t)) = -u''(t) = -\frac{\mathrm{d}}{\mathrm{d}t}u'(t) = v'(t)$$

and so $u(t)$, $v(t)$ satisfy the given system.

This proves that the second order ODE and the first order system are equivalent.

Then, since $F'(u) = f(u)$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\big(2F(u(t)) + v^2(t)\big) &= 2F'(u(t))u'(t) + 2v(t)v'(t)\\ &= 2f(u(t))\Big(-v(t)\Big) + 2v(t)\Big(f(u(t))\Big) = 0\end{aligned}$$

as required.

The next stage of the solution follows Exercise 13.9. We suppose that the modified equation is a system of two ODEs with dependent variables $x(t)$ and $y(t)$. The LTE of the given method is, therefore,

$$\widehat{\boldsymbol{T}}_{n+1} = \begin{bmatrix} x(t+h) - x(t) + hy(t)\\ y(t+h) - y(t) - hf(x(t+h)) \end{bmatrix}, \quad t = nh.$$

Taylor expansion of the first component around t and the second[3] around $t+h$ gives

$$\begin{aligned}\widehat{\boldsymbol{T}}_{n+1} &= \begin{bmatrix} \Big(x(t) + hx'(t) + \tfrac{1}{2}h^2x''(t)\Big) - x(t) + hy(t)\\ y(t+h) - \Big(y(t+h) - hy'(t+h) + \tfrac{1}{2}h^2y''(t+h)\Big) - hf(x(t+h)) \end{bmatrix} + \mathscr{O}(h^3)\\ &= h\begin{bmatrix} x'(t) + \tfrac{1}{2}hx''(t) + y(t)\\ y'(t+h) - \tfrac{1}{2}hy''(t+h) - f(x(t+h)) \end{bmatrix} + \mathscr{O}(h^3)\end{aligned}$$

We now suppose that the modified equations take the form

$$\begin{aligned} x'(t) &= \quad -y(t) + ha(x,y),\\ y'(t) &= \quad f(x(t)) + hb(x,y), \end{aligned}$$

where the functions $a(x,y)$ and $b(x,y)$ are to be determined. Differentiating these with respect to t gives

$$\begin{aligned} x''(t) &= -y'(t) + \mathscr{O}(h) = -x(t) + \mathscr{O}(h),\\ y''(t) &= \quad f'(x(t))x'(t) + \mathscr{O}(h) = -y(t)f'(x(t)) + \mathscr{O}(h) \end{aligned}$$

and substituting the results into the above expression for $\widehat{\boldsymbol{T}}_{n+1}$ leads to

$$\begin{aligned}\widehat{\boldsymbol{T}}_{n+1} &= h^2\begin{bmatrix} a(x,y) + \tfrac{1}{2}x''(t)\\ b(x,y) + \tfrac{1}{2}y''(t+h) \end{bmatrix} + \mathscr{O}(h^3)\\ &= h^2\begin{bmatrix} a(x,y) - \tfrac{1}{2}x(t)\\ b(x,y) + \tfrac{1}{2}y(t+h)f'(x(t+h)) \end{bmatrix} + \mathscr{O}(h^3)\\ &= h^2\begin{bmatrix} a(x,y) - \tfrac{1}{2}x(t)\\ b(x,y) + \tfrac{1}{2}y(t)f'(x(t)) \end{bmatrix} + \mathscr{O}(h^3)\end{aligned}$$

[3]This avoids having to expand $f(x(t+h))$ and the calculations are consequently much simpler.

where $t+h$ has been replaced by t without affecting the order. Therefore $\widehat{\boldsymbol{T}}_{n+1} = \mathscr{O}(h^3)$ on choosing $a(x,y) = \frac{1}{2}x$ and $b(x,y) = -\frac{1}{2}yf'(x)$. Our modified system of equations is, therefore,

$$\begin{aligned} x'(t) &= -y(t) + \tfrac{1}{2}hx(t), \\ y'(t) &= \quad f(x(t)) - \tfrac{1}{2}hy(t)f'(x(t)) \end{aligned}$$

as required. Then

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\big(2F(x(t)) - hf(x(t))y(t) + y^2(t)\big) \\ &\quad = 2F'(x(t))x'(t) - h\Big(f'(x(t))x'(t)y(t) + f(x(t))y'(t)\Big) + 2y(t)y'(t) \\ &\quad = \big(2f(x) - hf'(x)y\big)x' + \big(2y - hf(x)\big)y' \\ &\quad = 2\big(f(x) - \tfrac{1}{2}hf'(x)y\big)\big(-y + \tfrac{1}{2}hx\big) + 2\big(y - \tfrac{1}{2}hf(x)\big)\big(f(x) - \tfrac{1}{2}hf'(x)y\big) = 0 \end{aligned}$$

and so $2F(x(t)) - hf(x(t))y(t) + y^2(t) =$ constant.

When $f(u) = u$ it follows that $F(u) = \frac{1}{2}u^2$ and the results given here simplify to those in Example 13.4.

Exercises 14

1. Differentiating the given expressions we find that

$$u'(t) = -\left(-k_1A + k_2B\right) e^{-(k_1+k_2)t},$$
$$v'(t) = -\left(k_1A - k_2B\right) e^{-(k_1+k_2)t}.$$

Substituting the expressions for u and v into the RHS of the ODE system we find

$$-k_1u + k_2v = -\left(-k_1A + k_2B\right) e^{-(k_1+k_2)t},$$
$$k_1u + k_2v = -\left(k_1A - k_2B\right) e^{-(k_1+k_2)t}.$$

Also, the expressions have $u(0) = A$ and $v(0) = B$. Hence, the ODE system and initial conditions are satisfied.

Adding the two expressions gives

$$u + v = A + B + \left(\frac{k_1A - k_2B + k_2B - k_1A}{k_1 + k_2}\right) e^{-(k_1+k_2)t} = A + B.$$

As $t \to \infty$

$$u(t) \to \frac{k_2}{k_1 + k_2}(A + B), \quad \text{and} \quad v(t) \to \frac{k_1}{k_1 + k_2}(A + B).$$

This shows that, at steady state, the concentration of each species is directly proportional to its production rate constant. This makes sense, e.g., if $k_2 > k_1$ then we would expect to have more molecules of X_1 than X_2 at steady state.

---

3. Since

$$A = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} k_1, -k_2 \end{bmatrix} = \begin{bmatrix} -k_1 & k_2 \\ k_1 & -k_2 \end{bmatrix}$$

this is the required coefficient matrix. The result $A^j = \left(-k_1 - k_2\right)^{j-1} A$ ($j = 1, 2, 3, \ldots$) may be established by induction. It is trivially true when $j = 1$ and, when $j = 2$

$$\begin{aligned} A^2 &= \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} k_1, -k_2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} k_1, -k_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \left(\begin{bmatrix} k_1, -k_2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right) \begin{bmatrix} k_1, -k_2 \end{bmatrix} \\ &= (-k_1 - k_2)A \end{aligned}$$

and the result is true. We suppose now that it holds for $j = 1, 2, \ldots, \ell$, that is, $A^\ell = \left(-k_1 - k_2\right)^{\ell-1} A$. Multiplying both sides by A gives

$$A^{\ell+1} = \left(-k_1 - k_2\right)^{\ell-1} A^2 = \left(-k_1 - k_2\right)^{\ell-1} (-k_1 - k_2) A = \left(-k_1 - k_2\right)^{\ell} A$$

(where we have used the fact that the result holds for $j = 2$). The result now follows by induction.

Since $\boldsymbol{u}^{(j)}(t) = A^j \boldsymbol{u}(t)$, the TS(p) method

$$\begin{bmatrix} u_{n+1} \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} u_n \\ v_n \end{bmatrix} + h \begin{bmatrix} u_n' \\ v_n' \end{bmatrix} + \cdots + \tfrac{1}{p!} h^p \begin{bmatrix} u_n^{(p)} \\ v_n^{(p)} \end{bmatrix}$$

becomes

$$\begin{bmatrix} u_{n+1} \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} u_n \\ v_n \end{bmatrix} + h \left(1 - \tfrac{1}{2} h(k_1 + k_2) + \cdots + \tfrac{1}{p!} h^{p-1} (-k_1 - k_2)^{p-1}\right) A \begin{bmatrix} u_n \\ v_n \end{bmatrix}.$$

Since $[1,1]^T A = [0,0]$ it follows that $u_{n+1} + v_{n+1} = u_n + v_n$ and we have conservation of the linear invariant.

---

5. For

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

we have

$$\boldsymbol{x}^T C \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{f}(\boldsymbol{x}) = (a_1 + a_2 + a_3) x_1 x_2 x_3 = 0.$$

For

$$C = \begin{bmatrix} 1/I_1 & 0 & 0 \\ 0 & 1/I_2 & 0 \\ 0 & 0 & 1/I_3 \end{bmatrix}$$

we have

$$\boldsymbol{x}^T C \boldsymbol{f}(\boldsymbol{x}) = (a_1/I_1 + a_2/I_2 + a_3/I_3) x_1 x_2 x_3 = \frac{I_2 - I_3 + I_3 - I_1 + I_1 - I_2}{I_1 I_2 I_3} = 0.$$

---

7. From the general definition of an RK method in Section 9.2, for this two-step, two-stage case we have $\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \frac{1}{2} h \boldsymbol{k}_1 + \frac{1}{2} h \boldsymbol{k}_2$, where

$$\begin{aligned} \boldsymbol{k}_1 &= \boldsymbol{f}(\boldsymbol{x}_n + h a_{11} \boldsymbol{k}_1 + h a_{12} \boldsymbol{k}_2), \\ \boldsymbol{k}_2 &= \boldsymbol{f}(\boldsymbol{x}_n + h a_{21} \boldsymbol{k}_1 + h a_{22} \boldsymbol{k}_2). \end{aligned}$$

So, for these particular coefficients,

$$\begin{aligned} \boldsymbol{k}_1 &= \boldsymbol{f}\left(\boldsymbol{x}_n + a_{11}(2\boldsymbol{x}_{n+1} - 2\boldsymbol{x}_n - h\boldsymbol{k}_2) + h a_{12} \boldsymbol{k}_2\right) \\ &= \boldsymbol{f}\left(\boldsymbol{x}_n + \tfrac{1}{2}\boldsymbol{x}_{n+1} - \tfrac{1}{2}\boldsymbol{x}_n - \tfrac{1}{4} h \boldsymbol{k}_2 + h(\tfrac{1}{4} - \tfrac{\sqrt{3}}{6})\boldsymbol{k}_2\right) \\ &= \boldsymbol{f}\left(\boldsymbol{x}_{\text{mid}} - h\tfrac{\sqrt{3}}{6}\boldsymbol{k}_2\right). \end{aligned}$$

Similarly,

$$\boldsymbol{k}_2 = \boldsymbol{f}\left(\boldsymbol{x}_{\text{mid}} + h\tfrac{\sqrt{3}}{6}\boldsymbol{k}_1\right).$$

We wish to show that $\boldsymbol{x}_{n+1}^T C \boldsymbol{x}_{n+1} = \boldsymbol{x}_n^T C \boldsymbol{x}_n$. We will show the equivalent condition $(\boldsymbol{x}_n + \boldsymbol{x}_{n+1})^T C(\boldsymbol{x}_n - \boldsymbol{x}_{n+1}) = 0$.

We have

$$(\boldsymbol{x}_n + \boldsymbol{x}_{n+1})^T C(\boldsymbol{x}_n - \boldsymbol{x}_{n+1}) = (\boldsymbol{x}_n + \boldsymbol{x}_{n+1})^T C h(\boldsymbol{k}_1 + \boldsymbol{k}_2).$$

Now, since $\boldsymbol{z}^T C \boldsymbol{f}(\boldsymbol{z}) = 0$ for all $\boldsymbol{z}$, we have

$$\boldsymbol{x}_{\text{mid}}^T C \boldsymbol{k}_2 = \left(\boldsymbol{x}_{\text{mid}} + h\tfrac{\sqrt{3}}{6}\boldsymbol{k}_1 - h\tfrac{\sqrt{3}}{6}\boldsymbol{k}_1\right)^T C\boldsymbol{k}_2 = -h\tfrac{\sqrt{3}}{6}\boldsymbol{k}_1^T C \boldsymbol{k}_2.$$

Similarly,

$$\boldsymbol{x}_{\text{mid}}^T C \boldsymbol{k}_1 = \left(\boldsymbol{x}_{\text{mid}} - h\tfrac{\sqrt{3}}{6}\boldsymbol{k}_2 + h\tfrac{\sqrt{3}}{6}\boldsymbol{k}_2\right)^T C\boldsymbol{k}_1 = h\tfrac{\sqrt{3}}{6}\boldsymbol{k}_2^T C \boldsymbol{k}_1.$$

Adding, we find that $(\boldsymbol{x}_n + \boldsymbol{x}_{n+1})^T C h(\boldsymbol{k}_1 + \boldsymbol{k}_2) = 0$, which establishes the result.

---

9. This follows with

$$C = \tfrac{1}{2}\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

---

11. We have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{F}(x(t)) = x(t)^3 x'(t) = -x(t)^6 = -(f(x(t))^2.$$

So, along any solution, $\mathscr{F}(x(t))$ strictly decreases until it becomes stationary at $x(t) = 0$.

For the given modified equation, we have

$$\begin{aligned} y''(t) &= \left(-3y(t)^2 + \tfrac{3}{2}h \times 5y(t)^4\right) y'(t) \\ &= \left(-3y(t)^2 + O(h)\right)\left(-y(t)^3 + O(h)\right) \\ &= 3y(t)^5 + O(h). \end{aligned}$$

So

$$\begin{aligned} y(t+h) &= y(t) + hy'(t) + \tfrac{1}{2}h^2 y''(t) + O(h^2) \\ &= y(t) + h\left(-y(t)^3 + \tfrac{3}{2}hy(t)^5\right) + \tfrac{1}{2}h^2\left(3y(t)^5 + O(h)\right) \\ &= y(t) - hy(t)^3 + 3h^2 y(t)^5 + O(h^3). \end{aligned}$$

For the backward Euler method, $x_{n+1} = x_n - hx_{n+1}^3$, let us find a_n and b_n in the expansion $x_{n+1} = x_n + ha_n + h^2 b_n + O(h^3)$. We have

$$\begin{aligned} x_n + ha_n + h^2 b_n + O(h^3) &= x_n - h\left(x_n + ha_n + h^2 b_n + O(h^3)\right)^3 \\ &= x_n - h\left(x_n^3 + 3x_n^2 ha_n + O(h^2)\right) \\ &= x_n - hx_n^3 - 3h^2 x_n^2 a_n + O(h^3). \end{aligned}$$

Equating powers of h gives $a_n = -x_n^3$ and $b_n = -3x_n^2 a_n = 3x_n^5$. So

$$x_{n+1} = x_n - hx_n^3 + 3h^2 x_n^5 + O(h^3),$$

confirming that the given modified equation is valid.

Now $y'(t) > 0$ when $3hy(t)^5/2 > y(t)^3$, that is

$$y(t)^2 > \frac{2}{3h}.$$

So for

$$y(0) > \sqrt{\frac{2}{3h}}$$

we have $y'(0) > 0$, and we see that $y'(t)$ increases with t. So $y(t)$ increases monotonically without bound.

Exercises 15

1. Referring to Figure 5, let $\boldsymbol{u}$ and $\boldsymbol{v}$ be represented by the vectors $\overrightarrow{\mathrm{OA}}$ and $\overrightarrow{\mathrm{OC}}$, respectively. Then $J\boldsymbol{v}$ is represented by $\overrightarrow{\mathrm{OQ}}$ and makes an angle $\frac{1}{2}\pi - \theta$ with $\overrightarrow{\mathrm{OA}}$. Using the formula for scalar product

$$\overrightarrow{\mathrm{OA}} \cdot \overrightarrow{\mathrm{OQ}} = \|\overrightarrow{\mathrm{OA}}\| \, \|\overrightarrow{\mathrm{OQ}}\| \cos \mathrm{QOA},$$

so

$$\boldsymbol{u}^T J\boldsymbol{v} = \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| \cos(\tfrac{1}{2}\pi - \theta) = \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| \sin \theta.$$

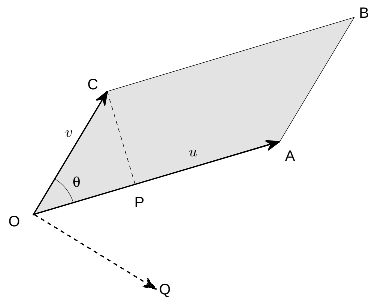

Figure 5: The area of the parallelogram OABC is, by the "base length times vertical height", OA×CP, where P is the foot of the perpendicular from C onto OA. The base length is $\|\boldsymbol{u}\|$ and the height $\|\boldsymbol{v}\| \sin\theta$ and their product is equal to the expression $\boldsymbol{u}^T J\boldsymbol{v}$ given above.

---

3. We have

$$\begin{aligned} \mathrm{area_o}(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}) &= (x_1 + z_1)y_2 - (x_2 + z_2)y_1 \\ &= x_1 y_2 - x_2 y_1 + z_1 y_2 - z_2 y_1 \\ &= \mathrm{area_o}(\boldsymbol{x}, \boldsymbol{y}) + \mathrm{area_o}(\boldsymbol{z}, \boldsymbol{y}). \end{aligned}$$

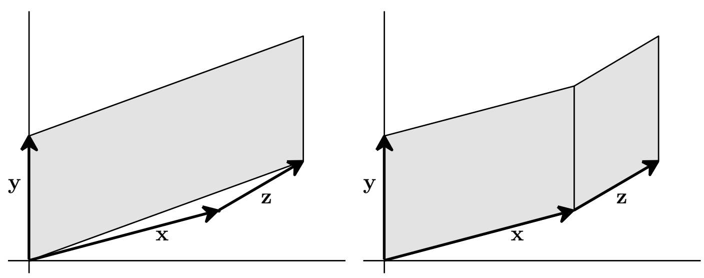

Figure 6: The two shaded regions have the same area.

---

5. We have

$$\begin{aligned}
A^T J A &= \left[\begin{array}{cc} a_{11} & a_{21} \\ a_{12} & a_{22} \end{array}\right]\left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right]\left[\begin{array}{cc} a_{11} & a_{11} \\ a_{21} & a_{22} \end{array}\right] \\
&= \left[\begin{array}{cc} -a_{21}a_{11} + a_{11}a_{21} & -a_{21}a_{12} + a_{11}a_{22} \\ -a_{22}a_{11} + a_{12}a_{21} & -a_{22}a_{12} + a_{12}a_{22} \end{array}\right] \\
&= \left[\begin{array}{cc} 0 & \det(A) \\ -\det(A) & 0 \end{array}\right].
\end{aligned}$$

---

7. We have

$$\begin{aligned}
\frac{\partial p_{n+1}}{\partial p_n} &= 1 - h\cos q_{n+1}\frac{\partial q_{n+1}}{\partial p_n}, \\
\frac{\partial p_{n+1}}{\partial q_n} &= -h\cos q_{n+1}\frac{\partial q_{n+1}}{\partial q_n}, \\
\frac{\partial q_{n+1}}{\partial p_n} &= \quad h\frac{\partial p_{n+1}}{\partial p_n}, \\
\frac{\partial q_{n+1}}{\partial q_n} &= 1 + h\frac{\partial p_{n+1}}{\partial q_n}.
\end{aligned}$$

This agrees with the matrix version given.

Taking determinants in the matrix version, using $\det(AB) = \det(A)\det(B)$, we find that

$$\left(1 + h^2\cos q_{n+1}\right)\det\left(\left[\begin{array}{cc} \frac{\partial p_{n+1}}{\partial p_n} & \frac{\partial p_{n+1}}{\partial q_n} \\ \frac{\partial q_{n+1}}{\partial p_n} & \frac{\partial q_{n+1}}{\partial q_n} \end{array}\right]\right) = 1,$$

which gives the required expression.

---

9. We first compute partial derivatives to obtain

$$\begin{aligned}
\frac{\partial p_{n+1}}{\partial p_n} &= 1 - hH_{qp}(p_n, q_{n+1}) - hH_{qq}(p_n, q_{n+1})\frac{\partial q_{n+1}}{\partial p_n}, \\
\frac{\partial p_{n+1}}{\partial q_n} &= -hH_{qq}(p_n, q_{n+1})\frac{\partial q_{n+1}}{\partial q_n}, \\
\frac{\partial q_{n+1}}{\partial p_n} &= hH_{pp}(p_n, q_{n+1}) + hH_{pq}(p_n, q_{n+1})\frac{\partial q_{n+1}}{\partial p_n}, \\
\frac{\partial q_{n+1}}{\partial q_n} &= 1 + hH_{pq}(p_n, q_{n+1})\frac{\partial q_{n+1}}{\partial q_n}.
\end{aligned}$$

Collecting these together, we have

$$\left[\begin{array}{cc} 1 & hH_{qq}(p_n, q_{n+1}) \\ 0 & 1 - hH_{pq}(p_n, q_{n+1}) \end{array}\right]\left[\begin{array}{cc} \frac{\partial p_{n+1}}{\partial p_n} & \frac{\partial p_{n+1}}{\partial q_n} \\ \frac{\partial q_{n+1}}{\partial p_n} & \frac{\partial q_{n+1}}{\partial q_n} \end{array}\right] = \left[\begin{array}{cc} 1 - hH_{qp}(p_n, q_{n+1}) & 0 \\ hH_{pp}(p_n, q_{n+1}) & 1 \end{array}\right].$$

Taking determinants, we find that

$$(1 - hH_{pq}(p_n, q_{n+1}))\det\left(\begin{bmatrix} \frac{\partial p_{n+1}}{\partial p_n} & \frac{\partial p_{n+1}}{\partial q_n} \\ \frac{\partial q_{n+1}}{\partial p_n} & \frac{\partial q_{n+1}}{\partial q_n} \end{bmatrix}\right) = 1 - hH_{qp}(p_n, q_{n+1}).$$

So, since $H_{pq} \equiv H_{qp}$, the map has determinant equal to one, as required.

In the separable case where $H(p,q) = T(p) + V(q)$, this adjoint method takes the form

$$\begin{aligned} p_{n+1} &= p_n - hV'(q_{n+1}), \\ q_{n+1} &= q_n + hT'(p_n). \end{aligned}$$

This is an explicit method—given (p_n, q_n), we may first compute q_{n+1} and then compute p_{n+1}.

---

11. In this case, symplectic Euler becomes

$$\begin{aligned} p_{n+1} &= p_n - hq_n, \\ q_{n+1} &= q_n + hp_{n+1}, \end{aligned}$$

which matches Example 13.4 if we take $u_n \equiv p_n$ and $v_n = q_n$. The ellipse in Example 13.4 then becomes $p^2 + q^2 - hpq$, or $2(T(p) + V(q) - \frac{1}{2}hT'(p)V'(q))$, as required.

---

13. On a separable problem with $H(p,q) = T(p) + V(q)$, the adjoint method takes the form

$$\begin{aligned} p_{n+1} &= p_n - hV'(q_{n+1}), \\ q_{n+1} &= q_n + hT'(p_n). \end{aligned}$$

We need to expand in the expression for p_{n+1}:

$$\begin{aligned} p_{n+1} &= p_n - hV'\left(q_n + hT'(p_n)\right) \\ &= p_n - h\left(V'(q_n) + V''(q_n)hT'(p_n) + \mathscr{O}(h^2)\right) \\ &= p_n - hV'(q_n) - h^2V''(q_n)T'(p_n) + \mathscr{O}(h^3). \end{aligned}$$

Matching the expansions for the true solution in Section 15.4, we require

$$h^2\left(A(p,q) - \tfrac{1}{2}V''(q)T'(p)\right) = -h^2V''(q)T'(p),$$

so

$$A(p,q) = -\tfrac{1}{2}V''(q)T'(p),$$

and

$$h^2 \left(B(p,q) - \tfrac{1}{2} T''(p) V'(q) \right) = 0,$$

so

$$B(p,q) = \tfrac{1}{2} T''(p) V'(q).$$

Hence, the required modified equation is

$$\begin{aligned} u'(t) &= -V'(u) - \tfrac{1}{2} h T'(u) V''(v), \\ v'(t) &= \quad T'(u) + \tfrac{1}{2} h T''(u) V'(v). \end{aligned}$$

This is of Hamiltonian form with

$$H(p,q) = T(p) + V(q) + \tfrac{1}{2} T'(p) V'(q).$$

Exercises 16

1. We have

$$
\begin{aligned}
\mathbb{E}[X] &= \frac{1}{\sqrt{2\sigma^2\pi}} \int_{-\infty}^{\infty} y \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right) \mathrm{d}y \\
&= \frac{-\sigma^2}{\sqrt{2\sigma^2\pi}} \int_{-\infty}^{\infty} \frac{-2(y-\mu)}{2\sigma^2} \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right) \mathrm{d}y \\
&\quad + \frac{1}{\sqrt{2\sigma^2\pi}} \int_{-\infty}^{\infty} \mu \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right) \mathrm{d}y \\
&= \frac{-\sigma^2}{\sqrt{2\sigma^2\pi}} \left[\exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right)\right]_{-\infty}^{\infty} + \mu \times 1 \\
&= \mu.
\end{aligned}
$$

---

3. Using integration by parts,

$$
\begin{aligned}
\mathbb{E}[X^2] &= \frac{1}{\sqrt{2\sigma^2\pi}} \int_{-\infty}^{\infty} y^2 \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right) \mathrm{d}y \\
&= \frac{-\sigma^2}{\sqrt{2\sigma^2\pi}} \int_{-\infty}^{\infty} y \frac{-2(y-\mu)}{2\sigma^2} \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right) \mathrm{d}y \\
&\quad + \frac{1}{\sqrt{2\sigma^2\pi}} \int_{-\infty}^{\infty} \mu y \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right) \mathrm{d}y \\
&= \frac{\sigma^2}{\sqrt{2\sigma^2\pi}} \int_{-\infty}^{\infty} \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right) \mathrm{d}y - \left[y \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right)\right]_{-\infty}^{\infty} + \mu \mathbb{E}[X] \\
&= \sigma^2 + \mu^2.
\end{aligned}
$$

Then

$$\mathsf{var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2.$$

---

5. We have about two significant digits on the bottom row, with $M = 10^7$ samples. Because the interval width scales like $1/\sqrt{M}$, we need about a factor of 10^6 as many samples to get three more digits; that is, six more rows.

---

7. For $nh = t_f$,

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[x_n] &= \lim_{n\to\infty} (1+ha)^n x_0 \\
&= \lim_{n\to\infty} \left(1 + \frac{at_f}{n}\right)^n x_0 \\
&= \mathrm{e}^{at_f} x_0.
\end{aligned}$$

Also, since x_0 is deterministic we have $\mathbb{E}[x_0^2] = x_0^2$, so

$$\mathbb{E}[x_n^2] = \left(1 + 2ha + hb^2 + h^2a^2\right)^n x_0^2.$$

Then, using $\log(1+x) = 1 + x + O(x^2)$ as $x \to 0$, we may write

$$\begin{aligned}
\log \mathbb{E}[x_n^2] &= n \log\left(1 + 2ha + hb^2 + h^2a^2\right) + \log x_0^2 \\
&= n\left(2ha + hb^2 + \mathscr{O}(h^2)\right) + \log x_0^2 \\
&= t_f(2a + b^2) + \log x_0^2 + \mathscr{O}(h),
\end{aligned}$$

where we also used $nh^2 = ht_f = \mathscr{O}(h)$. So

$$\lim_{n\to\infty} \log \mathbb{E}[x_n^2] - t_f(2a+b^2) - \log x_0^2 = 0.$$

Hence

$$\lim_{n\to\infty} \exp\left(\log \mathbb{E}[x_n^2] - t_f(2a+b^2) - \log x_0^2\right) = 1,$$

that is,

$$\lim_{n\to\infty} \mathbb{E}[x_n^2] e^{-t_f(2a+b^2)}/x_0^2 = 1,$$

which implies that

$$\lim_{n\to\infty} \mathbb{E}[x_n^2] = e^{t_f(2a+b^2)} x_0^2.$$

Finally,

$$\mathsf{var}[x_n] = \mathbb{E}[x_n^2] - (\mathbb{E}[x_n])^2 = \mathrm{e}^{t_f(2a+b^2)} x_0^2 - \mathrm{e}^{2at_f} x_0^2 = \mathrm{e}^{t_f 2a} x_0^2 \left(\mathrm{e}^{b^2 t_f} - 1\right).$$

---

9. In this case we have $x_{k+1} = x_k - hax_k + a\mu h + \sqrt{h} b \sqrt{x_k} Z_k$, so

$$\mathbb{E}[x_{k+1}] = \mathbb{E}[x_k](1 - ha) + a\mu h.$$

Let $y_k = \mathbb{E}[x_k]$, $r = (1 - ha)$ and $s = a\mu h$, so

$$y_{k+1} = ry_k + s.$$

Then, as in the previous exercise,

$$y_n = r^n y_0 + s\frac{r^n - 1}{r - 1}.$$

So

$$\begin{aligned} y_n &= r^n \left(y_0 + \frac{s}{r-1} \right) - \frac{s}{r-1} \\ &= r^n \left(y_0 + \frac{a\mu h}{-ha} \right) - \frac{a\mu h}{-ha} \\ &\to \mathrm{e}^{-at_f} \left(y_0 - \mu \right) + \mu. \end{aligned}$$

Exercises B

1.

$$\begin{aligned} g(x) &= xe^{1-x^2} & g(1) &= 1 \\ g'(x) &= e^{1-x^2} - 2x^2e^{1-x^2} & g'(1) &= 1 \\ g''(x) &= -6xe^{1-x^2} + 4x^3e^{1-x^2} & g''(1) &= -4 \\ g'''(x) &= -6e^{1-x^2} + 24x^2e^{1-x^2} - 8x^2e^{1-x^2} & g'''(1) &= 10 \end{aligned}$$

so, using (B.5) with $a = 1$,

$$g(x) \approx 1 - (x-1) - (x-1)^2 + \tfrac{5}{3}(x-1)^3.$$

Exercises D

1. (a) $2x_{n+1} - x_n = 3$.
   CF: $x_n = A(\frac{1}{2})^n$, PS: try $x_n = C$ and substitute into the left hand side of the $\triangle$E:
   $$2x_{n+1} - x_n = 2C - C = C = 3$$
   so the PS is $x_n = 3$ and the GS(general solution) is $x_n = A(\frac{1}{2})^n + 3$. This will satisfy $x_0 = 5$ if $x_0 = A(\frac{1}{2})^0 + 3 = 5$, that is $A = 2$. The required solution is $x_n = 2(\frac{1}{2})^n + 3$ and $x_n \to 3$ as $n \to \infty$.

   (b) $x_{n+1} - 2x_n = 3$.
   CF: $x_n = A2^n$, PS: try $x_n = C$ and substitute into the left hand side of the $\triangle$E:
   $$x_{n+1} - 2x_n = C - 2C = -C = 3$$
   so the PS is $x_n = -3$ and the GS(general solution) is $x_n = A2^n - 3$. This will satisfy $x_0 = 5$ if $x_0 = A2^0 - 3 = 5$, that is $A = 8$. The required solution is $x_n = 8 \times 2^n - 3 = 2^{n+3} - 3$ and $x_n \to \infty$ as $n \to \infty$.

   (c) $x_{n+1} - x_n = 3$.
   CF: $x_n = A1^n = A$, PS: we could try $x_n = C$ but a constant term is already in the CFso we increase the degree of the PS to $x_n = Cn$. We substitute into the left hand side of the $\triangle$E:
   $$x_{n+1} - x_n = C(n+1) - Cn = C = 3$$
   so the PS is $x_n = 3n$ and the GS(general solution) is $x_n = A + 3n$. This will satisfy $x_0 = 5$ if $x_0 = A = 5$. The required solution is $x_n = 5 + 3n$ and $x_n \to \infty$ as $n \to \infty$.

---

3. CFs for these problems have been given in the solution to Exercise D.1 so we need only find PSs.

   (a) $2x_{n+1} = x_n + 3 \times 2^n$. Try $x_n = C\,2^n$:
   $$2x_{n+1} - x_n = 2C2^{n+1} - C\,2^n = 3C\,2^n = 3 \times 2^n$$
   when $C = 1$. Hence, the GS is $x_n = A(\frac{1}{2})^n + 2^n$.

   (b) $x_{n+1} = 2x_n + 3 \times 2^n$ has CF$x_n = A\,2^n$ which already contains a term constant$\times 2^n$ so, for a PSwe try $x_n = C\,n2^n$:
   $$x_{n+1} - 2x_n = C\,2^n(2(n+1) - 2n) = 2C\,2^n = 3 \times 2^n$$
   when $C = 3/2$. The GSis $x_n = A\,2^n + \frac{3}{2}2^n$.

(c) Try $x_n = C\,2^n$:

$$x_{n+1} - x_n = C\,2^{n+1} - C\,2^n = C\,2^n = 3 \times 2^n$$

when $C = 3$. Hence, the GS is $x_n = A + 3 \times 2^n$.

---

5. The AE is

$$p(r) = r^5 - 8r^4 + 25r^3 - 38r^2 + 28r - 8.$$

A straightforward calculation gives $p(1) = p'(1) = 0$ while $p''(1) = -2 \neq 0$. $p(r)$ therefore has a double root at $r = 1$.

$p(2) = p'(2) = p''(1) = 0$ while $p'''(1) = 6 \neq 0$. $p(r)$ therefore has a triple root at $r = 2$.

Since $p(r)$ is of degree five and the coefficient of r^5 is one, it follows that $p(r) = (r-2)^3(r-1)^2$. The $\triangle$E therefore has GS

$$x_n = A + Bn + (B + Cn + Dn^2)2^n.$$

Springer

http://www.springer.com/978-0-85729-147-9

Numerical Methods for Ordinary Differential Equations
Initial Value Problems
Griffiths, D.; Higham, D.J.
2010, XIV, 271 p. 69 illus., Softcover
ISBN: 978-0-85729-147-9