Exam

Exercise 1 *To approximate the solution to the ODE*

$$\begin{cases} y'(t) = f\big(t, y(t)\big), & (1a)\\ y(0) = 2. & (1b)\end{cases}$$

we propose to use the Adams-Moulton scheme

$$y_{n+2} - y_{n+1} = \Delta t\left(\frac{5}{12}f(t^{n+2}, y_{n+2}) + \frac{2}{3}f(t^{n+1}, y_{n+1}) - \frac{1}{12}f(t^n, y_n)\right). \tag{2}$$

1. (a) *Scheme* (2) *is a 2-step scheme (see Definition 2.9 in the lecture notes) characterized by*

$$\alpha_0 = 0,\ \alpha_1 = -1,\ \alpha_2 = 1,\ \beta_0 = \frac{-1}{12},\ \beta_1 = \frac{2}{3} \text{ and } \beta_2 = \frac{5}{12}.$$

According to Prop. 2.6, this scheme is consistant iff

$$\sum_{k=0}^{2}\alpha_k = 0 \text{ and } \sum_{k=0}^{2}k\alpha_k = \sum_{k=0}^{2}\beta_k.$$

We easily check that

$$\sum_{k=0}^{2}\alpha_k = 0 - 1 + 1 = 0 \text{ and } \sum_{k=0}^{2}k\alpha_k = 0\times 0 + 1\times(-1) + 2\times 1 = 1 = \frac{-1}{12} + \frac{2}{3} + \frac{5}{12} = \sum_{k=0}^{2}\beta_k.$$

(b) *The Dahlquist statement (Prop. 2.7) ensures that the scheme is stable if the polynomial* $\rho(x) = \sum_{0\le k\le 2}\alpha_k x^k$ *has its roots in* $[-1,1]$ *and roots of modulus* 1 *are simple. Here, we have* $\rho(x) = x^2 - x = x(x-1)$ *whose roots are* 0 *and* 1*. Hence the stability.*

(c) *Stability and consistency imply the convergence of the scheme due to the Lax–Richtmyer theorem (Th. 2.2). The fact that a scheme is convergent means that if we denote by* (y_n) *the numerical solution and by* $\hat{y}$ *the exact solution, then* $\hat{y}(t^n) - y_n$ *goes to* 0 *as* Δt *goes to* 0*: the more* Δt *decreases, the more accurate the numerical solution (close to the exact solution).*

(d) *We apply Prop. 2.8 and assess each* $i \ge 2$ *until the relation in Prop. 2.8 is not satisfied:*
- $i = 2$: $\sum_{0\le k\le 2} k^2\alpha_k = 0\times 0 + 1\times(-1) + 4\times 1 = 3 = 2\big(0\times\frac{-1}{12} + 1\times\frac{2}{3} + 2\times\frac{5}{12}\big) = 2\sum_{0\le k\le 2} k\beta_k$. *The scheme is at least of order 2.*
- $i = 3$: $\sum_{0\le k\le 2} k^3\alpha_k = 0\times 0 + 1\times(-1) + 8\times 1 = 7 = 3\big(0\times\frac{-1}{12} + 1\times\frac{2}{3} + 4\times\frac{5}{12}\big) = 3\sum_{0\le k\le 2} k^2\beta_k$. *The scheme is at least of order 3.*
- $i = 4$: $\sum_{0\le k\le 2} k^4\alpha_k = 0\times 0 + 1\times(-1) + 16\times 1 = 15 \ne 16 = 4\big(0\times\frac{-1}{12} + 1\times\frac{2}{3} + 8\times\frac{5}{12}\big) = 4\sum_{0\le k\le 2} k^3\beta_k$. ***The scheme is exactly of order 3.***

(e) *Judging from the definition* $t^n = (n-1)\Delta t$*, we have* $t^1 = 0$*,* $t^2 = \Delta t$ *and* $t^N = 1$*, which is coherent with the domain of study* $[0,1]$.

(f) *As Scheme* (2) *is a 2-step scheme, we need* ***two initializing values*** *(to compute* y_3*, we need to know* y_2 *and* y_1*).*

(g) *As* y_1 *must be an approximation of* $\hat{y}(t^1) = \hat{y}(0) = 2$ *according to* (1b)*, we choose* $\boxed{y_1 = 2}$*. We now tackle the computation of* $y_2 \approx \hat{y}(t^2) = \hat{y}(\Delta t)$ *which we do not know. As Scheme* (2) *is of order 3,* *we must choose a value which is accurate at order 3, that is to say* $y_2 = \hat{y}(\Delta t) + \mathcal{O}(\Delta t^3)$*. To do so, we perform a Taylor expansion (Prop. 1.17):*

$$\hat{y}(\Delta t) = \hat{y}(0) + \Delta t \hat{y}'(0) + \frac{\Delta t^2}{2}\hat{y}''(0) + \mathcal{O}(\Delta t^3).$$

$\hat{y}'(0)$ *is deduce from ODE* (1a) *and from the initial condition* (1b)*:* $\hat{y}'(0) = f\big(0, \hat{y}(0)\big) = f(0,2)$*. Likewise, to compute* $\hat{y}''(0)$*, we differentiate ODE* (1a) *once to obtain by the chain rule*

$$\hat{y}''(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left[f\big(t,\hat{y}(t)\big)\right] = \frac{\partial f}{\partial t}\big(t,\hat{y}(t)\big) + \hat{y}'(t)\frac{\partial f}{\partial y}\big(t,\hat{y}(t)\big) = \frac{\partial f}{\partial t}\big(t,\hat{y}(t)\big) + f\big(t,\hat{y}(t)\big)\frac{\partial f}{\partial y}\big(t,\hat{y}(t)\big).$$

Hence we set $\boxed{y_2 = 2 + \Delta t f(0,2) + \big(\partial_t f(0,2) + f(0,2)\partial_y f(0,2)\big)\frac{\Delta t^2}{2}}$.

(h) *In Scheme* (2)*,* y_{n+2} *is present in both sides of the equality. Thus, given* y_n *and* y_{n+1}*, it may not be possible to compute* y_{n+2}*. The scheme is* ***implicit****.*

2. (a) *In this particular case, ODE* (1) *reads*

$$\begin{cases} y'(t) = y(t), & \text{(3a)} \\ y(0) = 2. & \text{(3b)} \end{cases}$$

It is thus a 1st-order linear ODE with constant coefficients (see (2.2) in the lecture notes). The exact solution $\hat{y}$ *is given by* $\boxed{\hat{y}(t) = 2e^t}$.

(b) *Scheme* (2) *applied to ODE* (3a) *is*

$$y_{n+2} - y_{n+1} = \Delta t\left(\frac{5}{12}y_{n+2} + \frac{2}{3}y_{n+1} - \frac{1}{12}y_n\right). \tag{4}$$

(c) *As stated in Q. 1(h), the Adams-Moulton scheme is implicit. It may not be possible to compute* y_n*. But in the linear case of Q. 2,* (4) *leads to*

$$\left(1 - \frac{5\Delta t}{12}\right)y_{n+2} - \left(1 + \frac{2\Delta t}{3}\right)y_{n+1} + \frac{\Delta t}{12} = 0. \tag{5}$$

Provided that $1 - \frac{5\Delta t}{12} \neq 0$[1] *and* $\frac{\Delta t}{12} \neq 0$*, relation* (5) *defines a 2nd-order linear induction with constant coefficients (p. 10 in the lecture notes). It is characterized by*

$$\alpha = 2,\ \beta = 2\left(1 + \Delta t + \frac{\Delta t^2}{2}\right),\ \zeta = 1 - \frac{5\Delta t}{12},\ \eta = -1 - \frac{2\Delta t}{3}\ \textit{and}\ \theta = \frac{\Delta t}{12}.$$

To compute y_n*, we study the characteric equation* $\zeta r^2 + \eta r + \theta = 0$*. Given that*

$$\Delta = \eta^2 - 4\zeta\theta = 1 + \Delta t + \frac{7}{12}\Delta t^2 > 0,$$

we deduce that $y_n = \kappa r_1^n + \lambda r_2^n$ *where* κ *and* λ *can be expressed from* α *and* β*.*

(d) *The* 3rd-order Runge-Kutta *scheme (see p. 27 in the lecture notes) applied to ODE* (3a) *reads*

$$z_{n+1} = z_n + \frac{\Delta t}{6}(k_1 + 4k_2 + k_3)$$

[1] Which is always the case since $1 - \frac{5\Delta t}{12} \neq 0 \Longleftrightarrow N \neq \frac{5}{12} + 1 \notin \mathbb{N}$.

with

$$\begin{aligned} k_1 &= f(t^n, z_n) = z_n, \\ k_2 &= f\left(t^n + \frac{\Delta t}{2}, z_n + \frac{\Delta t}{2} k_1\right) = z_n \times \left(1 + \frac{\Delta t}{2}\right), \\ k_3 &= f\left(t^{n+1}, z_n + \Delta t(2k_2 - k_1)\right) = z_n \times (1 + \Delta t + \Delta t^2). \end{aligned}$$

Hence

$$z_{n+1} = z_n \times \left(1 + \Delta t + \frac{\Delta t^2}{2} + \frac{\Delta t^3}{6}\right).$$

(e) *The previous relation shows that (z_n) is a geometric sequence (p. 10 in the lecture notes). We deduce that*

$$z_n = z_1 \left(1 + \Delta t + \frac{\Delta t^2}{2} + \frac{\Delta t^3}{6}\right)^{n-1} = 2\left(1 + \Delta t + \frac{\Delta t^2}{2} + \frac{\Delta t^3}{6}\right)^{n-1}.$$

(f) *Both* Adams-Moulton *and* Runge-Kutta 3 *(RK3) are 3rd-order accurate. However, in the general case, the Adams-Moulton scheme is implicit (unlike RK3) which is more expensive from a computational point of view. Thus RK3 seems preferable.*

Exercise 2 1. *The Cholesky decomposition (Prop. 1.9): If $A \in \mathscr{M}_n(\mathbb{R})$ is symmetric and positive-definite (Def. 1.8 and Prop. 1.12), then there exists $B \in \mathscr{M}_n(\mathbb{R})$ upper triangular such that $A = B^T B$. Moreover, there exists a unique B such that $B_{ii} > 0$ for all $i \in \{1, \dots, n\}$.*

2. *To build the algorithm, we notice that*

$$A_{ij} = \sum_{k=1}^{n} (B^T)_{ik} B_{kj} = \sum_{k=1}^{n} B_{ki} B_{kj} = \sum_{k=1}^{\min(i,j)} B_{ki} B_{kj}$$

since B is upper triangular (which implies $B_{ij} = 0$ if $i < j$). Hence we can compute components of B row by row. The MATLAB *function reads*

```
function B=chol_yp(A)
% Compute the Cholesky decomposition
% of a symmetric positive-definite matrix A

n=length(A(1,:)); % number of columns
m=length(A(:,1)); % number of rows

% Tests: does A satisfy the Cholesky hypotheses?
if m~=n
    disp('A␣is␣not␣a␣square␣matrix');
    return
end

if A'~=A
    disp('A␣is␣not␣symmetric');
    return
end

B=zeros(n,n);
```

```
% Algorithm
for i=1:n
    sum_B=sum(B(1:i-1,i).*B(1:i-1,i));
    if A(i,i)-sum_B<0
        disp('A is not positive-definite');
        return
    else
        B(i,i)=sqrt(A(i,i)-sum_B);
    end
    for j=i+1:n
        sum_B=sum(B(1:i-1,i).*B(1:i-1,j));
        B(i,j)=(A(i,j)-sum_B)/B(i,i);
    end
end
```

3. (a) *A is obviously symmetric. To prove that A is positive-definite, we determine its eigenvalues (Prop. 1.12). Due to the Sarrhus law for* 3×3 *determinants, we have*

$$\det(A - X\mathscr{I}_3) = \begin{vmatrix} 2-X & -1 & 0 \\ -1 & 2-X & -1 \\ 0 & -1 & 2-X \end{vmatrix} = (2-X)^3 - (2-X) - (2-X)$$

$$= (2-X)[(2-X)^2 - 2] = (2-X)(X^2 - 4X + 2) = (2-X)(X - 2 - \sqrt{2})(X - 2 + \sqrt{2}).$$

The eigenvalues are $2-\sqrt{2}$, 2 *and* $2+\sqrt{2}$ *which are all positive. We can thus apply the Cholesky decomposition to A.*

(b) *According to the algorithm from Q. 2, we deduce that*

$$B = \begin{pmatrix} \sqrt{2} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & \sqrt{\frac{3}{2}} & -\sqrt{\frac{2}{3}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{pmatrix}$$

4. *The Cholesky decomposition helps solve linear systems when the underlying matrix is symmetric and positive-definite. The corresponding algorithm is less expensive than standard algorithms for linear systems.*

5. *Given* $b \in \mathbb{R}^n$ *and* $A \in \mathscr{M}_n(\mathbb{R})$ *symmetric and positive-definite, we aim at solving* $Ax = b$. *To do so, we use the Cholesky decomposition of A and notice that the linear system is equivalent to* $B^T Bx = b$ *which can be decomposed into two triangular systems:* $B^T y = b$ *and then* $Bx = y$. *Hence:*

Algorithm 1 Resolution of a symmetric linear system

1: Data: $A \in \mathscr{M}_n(\mathbb{R})$ symmetric and positive-definite, $b \in \mathbb{R}^n$
2: Compute B such that $A = B^T B$ ▷ Use Algorithm from Q. 2
3: Solve the lower triangular linear system $B^T y = b$ ▷ Alg. 1, p. 16 in the lecture notes
4: Solve the upper triangular linear system $Bx = y$ ▷ Alg. 2, p. 16 in the lecture notes