High-Fidelity Simulation for Turbulent Flows

Contents

1 Solution of the advection-diffusion equation

Consider the steady-state 1D advection-diffusion equation

$$a\frac{\partial w}{\partial x} = \nu\frac{\partial^2 w}{\partial x^2} \qquad \text{with} \quad w \in [0, L], \quad a > 0, \quad \nu > 0 \tag{1}$$

with boundary conditions $w(0) = \alpha$ and $w(L) = \beta$.

1. Show that the analytical solution of equation (1) reads

$$w(x) = \alpha + (\beta - \alpha)\frac{\exp(\mathrm{Re}_x) - 1}{\exp(\mathrm{Re}_L) - 1} \tag{2}$$

with $\mathrm{Re}_x = \dfrac{ax}{\nu}$ and $\mathrm{Re}_L = \dfrac{aL}{\nu}$, respectively.
Hint: The general form of the solution is $w(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}$, where λ_1 and λ_2 are the roots of the characteristic equation.

Solution: By setting $R = \dfrac{a}{\nu}$, equation (1) can be written as

$$\frac{\partial^2 w}{\partial x^2} - R\frac{\partial w}{\partial x} = 0 \tag{3}$$

Injecting the general solution $e^{\lambda x}$ in (3), one has:

$$\lambda^2 e^{\lambda x} - R\lambda e^{\lambda x} = 0 \quad \Longrightarrow \quad \lambda(\lambda - R) = 0 \quad \Longrightarrow \quad \lambda_1 = 0 \quad \text{and} \quad \lambda_2 = R \tag{4}$$

Therefore, the solution is of the form

$$w(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} = c_1 + c_2 e^{Rx} \tag{5}$$

The constants c_1 and c_2 are found by applying the boundary conditions:

$$\begin{cases} w(x{=}0) = c_1 + c_2 = \alpha \\ w(x{=}L) = c_1 + c_2 e^{RL} = \beta \end{cases} \quad \Longrightarrow \quad \begin{cases} c_1 = \alpha - \dfrac{\beta - \alpha}{e^{RL} - 1} \\ c_2 = \dfrac{\beta - \alpha}{e^{RL} - 1} \end{cases} \tag{6}$$

which results in

$$w(x) = \alpha - \frac{\beta-\alpha}{e^{RL}-1} + \frac{\beta-\alpha}{e^{RL}-1}e^{Rx} = \alpha + (\beta-\alpha)\frac{\exp(Rx)-1}{\exp(RL)-1} \tag{7}$$

$$= \boxed{\alpha + (\beta-\alpha)\frac{\exp(\mathrm{Re}_x)-1}{\exp(\mathrm{Re}_L)-1}} \tag{8}$$

2. We consider the simple case for which α=0 and β=1. A Cartesian regular mesh is considered, such that x_j=$j\Delta x$, for $j\in[0, N]$. Express the formula of the exact solution on the discretized mesh as a function of j, N and $R_m = \dfrac{a\Delta x}{\nu}$, where R_m is the mesh Reynolds number.

Solution: Re_x and Re_L can be expressed as

$$\mathrm{Re}_x = \frac{ax}{\nu} = \frac{aj\Delta x}{\nu} = jR_m \tag{9}$$

$$\mathrm{Re}_L = \frac{aL}{\nu} = \frac{aN\Delta x}{\nu} = NR_m \tag{10}$$

and thus, considering $\alpha = 0$ and $\beta = 1$:

$$w_j = \frac{e^{\mathrm{Re}_x}-1}{e^{\mathrm{Re}_L}-1} = \boxed{\frac{e^{jR_m}-1}{e^{NR_m}-1}} \tag{11}$$

3. Equation (1) is now discretized by means of the following numerical scheme:

$$a\frac{w_j - w_{j-1}}{\Delta x} = \nu\frac{w_{j+1} - 2w_j + w_{j-1}}{\Delta x^2} \tag{12}$$

Show that the analytical solution of the discretized equation reads:

$$w_j = \frac{1-(1+R_m)^j}{1-(1+R_m)^N} \tag{13}$$

Solution: Equation (12) can be rewritten as

$$\frac{a\Delta x}{\nu}(w_j - w_{j-1}) = w_{j+1} - 2w_j + w_{j-1} \tag{14}$$

leading to

$$w_{j+1} - (R_m+2)w_j + (1+R_m)w_{j-1} = 0 \tag{15}$$

We search for a solution of the form $w_j = q^j$. Replacing in (15) and dividing by q^{j-1}, one obtains:

$$q^{j+1} - (R_m+2)q^j + (1+R_m)q^{j-1} = 0 \tag{16}$$

$$q^2 - (R_m+2)q + (1+R_m) = 0 \tag{17}$$

Solving for q:

$$\Delta = b^2 - 4ac = (R_m + 2)^2 - 4(1 + R_m) = R_m^2 \tag{18}$$

$$\implies q_{1,2} = \frac{-b \pm \sqrt{\Delta}}{2a} = \frac{R_m + 2 \pm R_m}{2} \implies \quad q_1 = 1 \quad \text{and} \quad q_2 = 1 + R_m \tag{19}$$

The general solution is of the form $w_j = c_1 q_1^j + c_2 q_2^j$, thus one obtains

$$w_j = c_1 + c_2(1 + R_m)^j \tag{20}$$

Applying the boundary conditions:

$$\begin{cases} w_0 = c_1 + c_2(1 + R_m)^0 = c_1 + c_2 = 0 \\ w_N = c_1 + c_2(1 + R_m)^N = 1 \end{cases} \implies \begin{cases} c_1 = -c_2 \\ c_2[-1 + (1 + R_m)^N] = 1 \end{cases} \tag{21}$$

$$\implies \begin{cases} c_1 = \dfrac{1}{1 - (1 + R_m)^N} \\ c_2 = -\dfrac{1}{1 - (1 + R_m)^N} \end{cases} \tag{22}$$

And replacing in equation (20) one has:

$$w_j = \frac{1}{1 - (1 + R_m)^N} - \frac{(1 + R_m)^j}{1 - (1 + R_m)^N} = \boxed{w_j = \frac{1 - (1 + R_m)^j}{1 - (1 + R_m)^N}} \tag{23}$$

4. Repeat the analysis by considering the following scheme:

$$a\frac{w_{j+1} - w_{j-1}}{2\Delta x} = \nu \frac{w_{j+1} - 2w_j + w_{j-1}}{\Delta x^2} \tag{24}$$

and show that the analytical solution reads

$$w_j = \frac{\left(\frac{2+R_m}{2-R_m}\right)^j - 1}{\left(\frac{2+R_m}{2-R_m}\right)^N - 1} \tag{25}$$

Solution: Equation (24) can be rewritten as

$$\frac{a\Delta x}{\nu}(w_{j+1} - w_{j-1}) = 2w_{j+1} - 4w_j + 2w_{j-1} \tag{26}$$

leading to

$$(R_m - 2)w_{j+1} + 4w_j - (R_m + 2)w_{j-1} = 0 \tag{27}$$

We search for a solution of the form $w_j = q^j$. Replacing in (27) and dividing by q^{j-1}, one obtains:

$$(R_m - 2)q^{j+1} + 4q^j - (R_m + 2)q^{j-1} = 0 \tag{28}$$

$$(R_m - 2)q^2 + 4q - (R_m + 2) = 0 \tag{29}$$

Solving for q:

$$\Delta = b^2 - 4ac = 16 + 4(R_m - 2)(R_m + 2) = 4R_m^2 \tag{30}$$

$$\implies q_{1,2} = \frac{-b \pm \sqrt{\Delta}}{2a} = \frac{-4 \pm 2R_m}{2(R_m - 2)} \quad \implies \quad q_1 = 1 \quad \text{and} \quad q_2 = \frac{2+R_m}{2-R_m} \tag{31}$$

The general solution is of the form $w_j = c_1 q_1^j + c_2 q_2^j$, thus one obtains

$$w_j = c_1 + c_2 \left(\frac{2+R_m}{2-R_m}\right)^j \tag{32}$$

Applying the boundary conditions:

$$\begin{cases} w_0 = c_1 + c_2 \left(\frac{2+R_m}{2-R_m}\right)^0 = c_1 + c_2 = 0 \\ w_N = c_1 + c_2 \left(\frac{2+R_m}{2-R_m}\right)^N = 1 \end{cases} \implies \begin{cases} c_1 = -c_2 \\ c_2 \left[-1 + \left(\frac{2+R_m}{2-R_m}\right)^N\right] = 1 \end{cases} \tag{33}$$

$$\implies \begin{cases} c_1 = -\dfrac{1}{\left(\frac{2+R_m}{2-R_m}\right)^N - 1} \\ c_2 = \dfrac{1}{\left(\frac{2+R_m}{2-R_m}\right)^N - 1} \end{cases} \tag{34}$$

And replacing in equation (32) one has:

$$w_j = -\frac{1}{\left(\frac{2+R_m}{2-R_m}\right)^N - 1} + \frac{\left(\frac{2+R_m}{2-R_m}\right)^j}{\left(\frac{2+R_m}{2-R_m}\right)^N - 1} = \boxed{\frac{\left(\frac{2+R_m}{2-R_m}\right)^j - 1}{\left(\frac{2+R_m}{2-R_m}\right)^N - 1}} \tag{35}$$

5. Compute the relative error of the two chosen discretizations with respect to the exact solution, at the grid point x_{N-1} for R_m=1.5 and R_m=2.5. Consider $N = 1000$.

Solution: For $N \gg 1$ the value of w_{N-1} given by the three analytical solutions is:

Exact solution:

$$w_j = \frac{e^{jR_m} - 1}{e^{NR_m} - 1} \quad \implies \quad w_{N-1} = \frac{e^{(N-1)R_m} - 1}{e^{NR_m} - 1} = \frac{1}{e^{R_m}} \quad \text{for} \quad N \gg 1 \tag{36}$$

$$R_m = 1.5: \quad w_{N-1} = \frac{1}{e^{1.5}} = 0.22313 \tag{37}$$

$$R_m = 2.5: \quad w_{N-1} = \frac{1}{e^{2.5}} = 0.082085 \tag{38}$$

Upwind scheme solution:

$$w_j = \frac{1 - (1+R_m)^j}{1 - (1+R_m)^N} \implies w_{N-1} = \frac{1 - (1+R_m)^{N-1}}{1 - (1+R_m)^N} = \frac{1}{1+R_m} \quad \text{for} \quad N \gg 1 \tag{39}$$

$$R_m = 1.5: \quad w_{N-1} = \frac{1}{1+1.5} = 0.4 \qquad \Longrightarrow \varepsilon = 79.3\% \tag{40}$$

$$R_m = 2.5: \quad w_{N-1} = \frac{1}{1+2.5} = 0.28571 \Longrightarrow \varepsilon = 248\% \tag{41}$$

Centred scheme solution:

$$w_j = \frac{\left(\frac{2+R_m}{2-R_m}\right)^j - 1}{\left(\frac{2+R_m}{2-R_m}\right)^N - 1} \Longrightarrow w_{N-1} = \frac{\left(\frac{2+R_m}{2-R_m}\right)^{N-1} - 1}{\left(\frac{2+R_m}{2-R_m}\right)^N - 1} = \frac{2-R_m}{2+R_m} \quad \text{for} \quad N \gg 1 \tag{42}$$

$$R_m = 1.5: \quad w_{N-1} = \frac{2-1.5}{2+1.5} = 0.14286 \qquad \Longrightarrow \varepsilon = -36\% \tag{43}$$

$$R_m = 2.5: \quad w_{N-1} = \frac{2-2.5}{2+2.5} = -0.11111 \Longrightarrow \varepsilon = -235\% \tag{44}$$

The use of a second-order centred scheme allows to reduce the relative error with respect to the analytical solution. However, oscillations appears when $R_m > 2$, leading to unphysical results.

6. Does the value of N change the relative errors? What should be done to remove the oscillating solution for the centred scheme?

Solution: The value of N does not change the solution, provided that the condition $N \gg 1$ holds. The only way to remove the numerical oscillations is to reduce the grid size Δx in order to obtain a value of the mesh Reynolds number $R_m < 2$.

2 Semi-implicit Lax-Wendroff scheme

We want to study the linear advection equation

$$\frac{\partial w}{\partial t} + a\frac{\partial w}{\partial x} = 0 \qquad a \in \mathbb{R} \tag{1}$$

To this aim, we consider the semi-implicit Lax-Wendroff scheme:

$$\frac{w_j^{n+1} - w_j^n}{\Delta t} + a\frac{w_{j+1}^n - w_{j-1}^n}{2\Delta x} = |a|^2 \Delta t \frac{w_{j+1}^{n+1} - 2w_j^{n+1} + w_{j-1}^{n+1}}{2\Delta x^2} \tag{2}$$

1. Perform a stability analysis of the chosen discretization.

Solution: Defining $\dot{a} = a\dfrac{\Delta t}{\Delta x}$, a Von Neumann analysis leads to

$$\begin{cases} w_j^n = e^{at}e^{ikx} \\ w_{j+1}^n = e^{at}e^{ik(x+\Delta x)} \\ w_{j-1}^n = e^{at}e^{ik(x-\Delta x)} \end{cases} \qquad \begin{cases} w_j^{n+1} = e^{a(t+\Delta t)}e^{ikx} \\ w_{j+1}^{n+1} = e^{a(t+\Delta t)}e^{ik(x+\Delta x)} \\ w_{j-1}^{n+1} = e^{a(t+\Delta t)}e^{ik(x-\Delta x)} \end{cases} \tag{3}$$

The amplification factor being defined as $G = \dfrac{w_j^{n+1}}{w_j^n} = e^{a\Delta t}$. Replacing the expressions in (2) and dividing by $w_j^n = e^{at}e^{ikx}$, one has

$$e^{a\Delta t} - 1 + \frac{\dot{a}}{2}(e^{ik\Delta x} - e^{-ik\Delta x}) = \frac{|\dot{a}|^2}{2} e^{a\Delta t}(e^{ik\Delta x} - 2 + e^{-ik\Delta x}) \tag{4}$$

Developing the relation:

$$\left[1 - \frac{|\dot{a}|^2}{2}(e^{ik\Delta x} - 2 + e^{-ik\Delta x})\right] e^{a\Delta t} = 1 - \frac{\dot{a}}{2}(e^{ik\Delta x} - e^{-ik\Delta x}) \tag{5}$$

Defining the reduced wavenumber $\beta = k\Delta x$ and recalling that $e^{\pm i\beta} = \cos\beta \pm i\sin\beta$, the amplification factor reads:

$$G = e^{a\Delta t} = \frac{1 - \frac{\dot{a}}{2}(e^{i\beta} - e^{-i\beta})}{1 - \frac{|\dot{a}|^2}{2}(e^{i\beta} - 2 + e^{-i\beta})} = \frac{1 - i\dot{a}\sin\beta}{1 + |\dot{a}|^2(1 - \cos\beta)} \tag{6}$$

Considering that $\sin\beta = 2\sin\dfrac{\beta}{2}\cos\dfrac{\beta}{2}$ and $1 - \cos\beta = 2\sin^2\dfrac{\beta}{2}$:

$$|G|^2 = \frac{1 + \dot{a}^2\sin^2\beta}{[1 + |\dot{a}|^2(1 - \cos\beta)]^2} = \frac{1 + 4\dot{a}^2\sin^2\frac{\beta}{2}\cos^2\frac{\beta}{2}}{\left(1 + 2|\dot{a}|^2\sin^2\frac{\beta}{2}\right)^2} \tag{7}$$

$$= \frac{1 + 4\dot{a}^2\sin^2\frac{\beta}{2} - 4\dot{a}^2\sin^4\frac{\beta}{2}}{1 + 4|\dot{a}|^2\sin^2\frac{\beta}{2} + 4|\dot{a}|^4\sin^4\frac{\beta}{2}} \leq 1 \quad \forall\dot{a} \tag{8}$$

The scheme is then unconditionally stable.

2. By using the standard finite difference operators:

$$\delta(\bullet)_j = (\bullet)_{j+\frac{1}{2}} - (\bullet)_{j-\frac{1}{2}} \qquad \mu(\bullet)_j = \frac{1}{2}\left[(\bullet)_{j+\frac{1}{2}} + (\bullet)_{j-\frac{1}{2}}\right] \tag{9}$$

Rewrite the scheme in Δ-form as:

$$A_{\mathrm{LW}}\Delta w_j^n = \Delta w_{\mathrm{exp},j}^{\mathrm{LW},n} \tag{10}$$

and give the expressions of the mass matrix A_{LW} and of the explicit phase $\Delta w_{\mathrm{exp},j}^{\mathrm{LW},n}$.

Solution: Using the finite difference operators, equation (2) can be recast into

$$\Delta w_j^n = -\dot{a}\delta\mu w_j^n + \frac{|\dot{a}|^2}{2}\delta^2 w_j^{n+1} \tag{11}$$

To build the mass matrix, we add and subtract the term $\dfrac{|\dot{a}|^2}{2}\delta^2 w_j^n$ to obtain:

$$\Delta w_j^n = -\dot{a}\delta\mu w_j^n + \frac{|\dot{a}|^2}{2}\delta^2 w_j^{n+1} + \frac{|\dot{a}|^2}{2}\delta^2 w_j^n - \frac{|\dot{a}|^2}{2}\delta^2 w_j^n \tag{12}$$

$$= -\dot{a}\delta\mu w_j^n + \frac{|\dot{a}|^2}{2}\delta^2 w_j^n + \frac{|\dot{a}|^2}{2}\delta^2\Delta w_j^n \tag{13}$$

Lastly, regrouping Δw_j^n one obtains:

$$\left[1 - \frac{|\dot{a}|^2}{2}\delta^2\right] \Delta w_j^n = \left[-\dot{a}\delta\mu + \frac{|\dot{a}|^2}{2}\delta^2\right] w_j^n \tag{14}$$

Thus

$$A_{\text{LW}} = \left[1 - \frac{|\dot{a}|^2}{2}\delta^2\right] \quad \text{and} \quad \Delta w_{\exp,j}^{\text{LW},n} = \left[-\dot{a}\delta\mu + \frac{|\dot{a}|^2}{2}\delta^2\right] w_j^n \tag{15}$$

3. Specify the form of the mass matrix A_{LW}. Verify if it is diagonal dominant.

Solution: A_{LW} is a tridiagonal matrix of the form

$$A_{\text{LW}} = \begin{bmatrix} \ddots & \cdots & \cdots & \cdots & 0 \\ -\frac{|\dot{a}|^2}{2} & 1+|\dot{a}|^2 & -\frac{|\dot{a}|^2}{2} & 0 & 0 \\ 0 & -\frac{|\dot{a}|^2}{2} & 1+|\dot{a}|^2 & -\frac{|\dot{a}|^2}{2} & 0 \\ 0 & 0 & -\frac{|\dot{a}|^2}{2} & 1+|\dot{a}|^2 & -\frac{|\dot{a}|^2}{2} \\ 0 & \cdots & \cdots & \cdots & \ddots \end{bmatrix} \tag{16}$$

The matrix $A_{\text{LW}} = [a_{ij}]$ is said to be diagonally dominant if $a_{ii} \geq \sum_{j \neq i} |a_{ij}| \forall i$. In this case:

$$|a_{ii}| = 1 + |\dot{a}|^2 > \left|\frac{\dot{a}}{2}\right| + \left|\frac{\dot{a}}{2}\right| = |\dot{a}|^2 \quad \forall \dot{a} \tag{17}$$

A_{LW} is then a diagonally dominant matrix.

4. Is the scheme appropriate for computing the steady-state solution of the problem? Why? For unsteady problems, would it be more suited for slow or rapid cases?

Solution: The scheme is appropriate for the study of steady-state configurations since:

1. It is unconditionally stable $\forall \dot{a}$, meaning that large CFL can be used without stability issues;
2. $G \to 0$ for $|\dot{a}| \to \infty$, meaning that numerical errors are quickly damped as the CFL gets larger;
3. The mass matrix is strictly diagonally dominant, meaning that convergence is ensured by using any efficient iterative algorithm (such as Gauss-Seidel or SOR).

For unsteady configurations, the scheme would be more suited for slow cases which could benefit from the large stability region of the discretization, provided that sufficient temporal accuracy is achieved.

3 Adams-Bashforth 3-step

We consider the semi-discrete ODE

$$\frac{\mathrm{d}w}{\mathrm{d}t} = R(w) \tag{1}$$

The three-step explicit Adams-Bashforth method can be written under the form:

$$w^{n+1} = w^n + a\Delta t R(w^n) + b\Delta t R(w^{n-1}) + c\Delta t R(w^{n-2}) \tag{2}$$

1. Derive the coefficients of the method by equating the coefficients of the Taylor series for the l.h.s. and r.h.s. of equation (2).

Solution: Since $\frac{\mathrm{d}w}{\mathrm{d}t} = R(w)$, we can write equation (2) as

$$w^{n+1} = w^n + a\Delta t w_t^n + b\Delta t w_t^{n-1} + c\Delta t w_t^{n-2} \tag{3}$$

The subscript t denoting the time derivative of w. Expanding both sides of equation (3) in Taylor series about w^n, one has

$$\begin{aligned}
w + \Delta t w_t + \frac{\Delta t^2}{2} w_{tt} + \frac{\Delta t^3}{6} w_{ttt} + \mathcal{O}(\Delta t^4) = & \qquad (4)\\
= w + a\Delta t w_t + b\Delta t \left(w_t - \Delta t w_{tt} + \frac{\Delta t^2}{2} w_{ttt} + \mathcal{O}(\Delta t^3) \right) & \qquad (5)\\
+ c\Delta t \left(w_t - 2\Delta t w_{tt} + \frac{4\Delta t^2}{2} w_{ttt} + \mathcal{O}(\Delta t^3) \right) & \qquad (6)\\
= w + (a+b+c)\Delta t w_t + (-b-2c)\Delta t^2 w_{tt} + \left[\frac{b}{2} + 2c \right] \Delta t^3 w_{ttt} + \mathcal{O}(\Delta t^4) & \qquad (7)
\end{aligned}$$

Equating the coefficients one has

$$\begin{cases} a+b+c=1 \\ -b-2c = \frac{1}{2} \\ \frac{b}{2} + 2c = \frac{1}{6} \end{cases} \implies a = \frac{23}{12}, \quad b = -\frac{16}{12}, \quad c = \frac{5}{12} \tag{8}$$

Plugging them into equation (2), one obtains the 3-step Adams-Bashforth method:

$$w^{n+1} = w^n + \frac{\Delta t}{12} \left[23R(w^n) - 16R(w^n - 1) + 5R(w^{n-2}) \right] \tag{9}$$

2. Compute the temporal order of accuracy of the discretization.

Solution: The temporal order of accuracy is therefore $\mathcal{O}(\Delta t^3)$

4 Analysis of a finite element scheme

By applying the Galerkin method with linear elements to the conservation equation

$$\frac{\partial w}{\partial t} + \frac{\partial f}{\partial x} = 0 \tag{1}$$

one obtains the following implicit formulation:

$$\frac{1}{6}\left[\frac{\mathrm{d}w_{j-1}}{\mathrm{d}t}+4\frac{\mathrm{d}w_j}{\mathrm{d}t}+\frac{\mathrm{d}w_{j+1}}{\mathrm{d}t}\right]=\frac{f_{j+1}-f_{j-1}}{2\Delta x} \tag{2}$$

Using a central time integration (leapfrog) and considering the linear equation $f = aw$, the following scheme is obtained:

$$(w_{j-1}^{n+1}-w_{j-1}^{n-1})+4(w_j^{n+1}-w_j^{n-1})+(w_{j+1}^{n+1}-w_{j+1}^{n-1})+6\dot{a}(w_{j+1}^n-w_{j-1}^n)=0 \tag{3}$$

with $\dot{a}=\dfrac{a\Delta t}{\Delta x}$.

1. Compute the amplification factor of the scheme and obtain a stability condition of the type

$$\dot{a}^2 \leq f(\beta) \tag{4}$$

where $f(\beta)$ is a function of $\beta = k\Delta x$.

Solution: From a Von Neumann analysis, one obtains:

$$Ge^{-i\beta}-\frac{e^{-i\beta}}{G}+4\left(G-\frac{1}{G}\right)+Ge^{i\beta}-\frac{e^{i\beta}}{G}+6\dot{a}(e^{i\beta}-e^{-i\beta})=0 \tag{5}$$

$$\left[G-\frac{1}{G}\right]\left[e^{-i\beta}+4+e^{i\beta}\right]+6\dot{a}(e^{i\beta}-e^{-i\beta})=0 \tag{6}$$

$$\iff \left[G-\frac{1}{G}\right](4+2\cos\beta)+12\dot{a}i\sin\beta=0 \tag{7}$$

$$\iff \left[G-\frac{1}{G}\right]+\frac{6\dot{a}i\sin\beta}{2+\cos\beta}=0 \tag{8}$$

$$\iff G^2+\frac{6\dot{a}i\sin\beta}{2+\cos\beta}G-1=0 \quad\iff\quad G^2+2i\gamma G-1=0 \tag{9}$$

where $\gamma=\dfrac{3\dot{a}\sin\beta}{2+\cos\beta}$. Solving for G:

$$\Delta=-4\gamma^2+4 \quad\Longrightarrow\quad G_{1,2}=-i\gamma\pm\sqrt{1-\gamma^2} \tag{10}$$

For the scheme to be stable, we need $|G_{1,2}|\leq 1$. Thus

$$|G|^2\leq 1 \quad\Longrightarrow\quad GG^*\leq 1 \quad\Longrightarrow\quad \gamma^2+1-\gamma^2\leq 1 \quad \forall\gamma \tag{11}$$

This is true provided that $\sqrt{1-\gamma^2}$ is real; that is, $\gamma^2\leq 1$, obtaining

$$\frac{9\dot{a}^2\sin^2\beta}{(2+\cos\beta)^2}\leq 1 \quad\Longrightarrow\quad \dot{a}^2\leq\frac{(2+\cos\beta)^2}{9\sin^2\beta} \tag{12}$$

2. Derive the most restrictive condition for satisfying equation (4). *Hint*: compute the minimum of the function $f(\beta)$.

Solution: In order to find the most restrictive condition, one could compute the minimum of the function $f(\beta) = \dfrac{(2+\cos\beta)^2}{9\sin^2\beta}$. Therefore:

$$\frac{\mathrm{d}f(\beta)}{\mathrm{d}\beta} = \frac{-2(2+\cos\beta)\sin^2\beta - 2(2+\cos\beta)^2\cos\beta}{9\sin^3\beta} \tag{13}$$

$$= -\frac{2(2+\cos\beta)}{9\sin^2\beta}\left[\sin\beta + \frac{(2+\cos\beta)\cos\beta}{\sin\beta}\right] \tag{14}$$

and then find the roots of $\dfrac{\mathrm{d}f(\beta)}{\mathrm{d}\beta} = 0$, thus

$$\cancel{-\frac{2(2+\cos\beta)}{9\sin^2\beta}}\left[\sin\beta + \frac{(2+\cos\beta)\cos\beta}{\sin\beta}\right] = 0 \tag{15}$$

$$\sin^2\beta + 2\cos\beta + \cos^2\beta = 0 \quad \Longrightarrow \quad 1 + 2\cos\beta = 0 \tag{16}$$

One obtains $\cos\beta = -\dfrac{1}{2}$ and then $\beta_{\min} = \dfrac{4\pi}{3}$ (one should verify that it is a true minimum by checking that $\dfrac{\mathrm{d}^2 f(\beta_{\min})}{\mathrm{d}\beta^2} > 0$). Replacing in the condition (12) one obtains:

$$\dot{a}^2 \le \frac{(2-\frac{1}{2})^2}{9\left(\frac{\sqrt{3}}{2}\right)^2} = \frac{\frac{9}{4}}{9\frac{3}{4}} = \frac{1}{3} \quad \Longrightarrow \quad |\dot{a}| \le \frac{1}{\sqrt{3}} \tag{17}$$

3. Compute the dispersion and diffusion errors of the scheme.

Solution: By injecting the general solution $w = e^{ct}e^{ikx}$ in equation (1), one obtains

$$w = e^{ik(x-at)} \qquad (\text{since} \quad c = -ika) \tag{18}$$

The exact amplification factor then reads

$$G_{\text{ex}} = \frac{w(t+\Delta t)}{w(t)} = \frac{e^{ik[x-a(t+\Delta t)]}}{e^{ik(x-at)}} = e^{-ika\Delta t} \tag{19}$$

which can be recast into

$$G_{\text{ex}} = |G_{\text{ex}}|e^{i\Phi_{\text{ex}}} \quad \text{with} \quad |G_{\text{ex}}| = 1 \quad \text{and} \quad \Phi_{\text{ex}} = -ka\Delta t = -\dot{a}\beta \tag{20}$$

Similarly, the numerical amplification factor in equation (10) reads

$$G_{\text{num}} = |G_{\text{num}}|e^{i\Phi_{\text{num}}} \quad \text{with} \quad |G_{\text{num}}| = 1 \quad \text{and} \quad \Phi_{\text{num}} = \tan^{-1}\left[\frac{-\gamma}{-\sqrt{1-\gamma^2}}\right] \tag{21}$$

From which $\Phi_{\text{num}} = \sin^{-1}\gamma$. The dispersion error represents the error on the phase of the solution, and is defined as

$$\varepsilon_\beta = \frac{\Phi_{\text{num}}}{\Phi_{\text{ex}}} = \frac{\sin^{-1}\gamma}{\dot{a}\beta} = \frac{\frac{3\dot{a}\sin\beta}{2+\cos\beta}}{\dot{a}\beta} = \frac{3\sin\beta}{\beta(2+\cos\beta)} \tag{22}$$

whereas the diffusion error is

$$\varepsilon_D = \frac{|G_{\text{num}}|}{|G_{\text{ex}}|} = 1 \tag{23}$$

5 Stability of the Leapfrog scheme

We apply the Leapfrog scheme (central difference in time) for the general semi-discrete ODE:

$$\frac{\partial w}{\partial t} = R(w) \quad \Longrightarrow \quad \frac{w_j^{n+1} - w_j^{n-1}}{2\Delta t} = R(w) \tag{1}$$

1. First, we consider the heat-conduction equation (thus $R(w) = \dfrac{\partial^2 w}{\partial x^2}$) where the space operator is discretized by centred second-order finite differences, giving:

$$w_j^{n+1} - w_j^{n-1} = 2\nu \frac{\Delta t}{\Delta x^2}(w_{j+1}^n - 2w_j^n + w_{j-1}^n) \tag{2}$$

Discuss the stability of the numerical discretization.

Solution: Imposing $\dot{\nu} = \dfrac{\nu \Delta t}{\Delta x^2}$, from a Von Neumann analysis one obtains:

$$G - \frac{1}{G} = 2\dot{\nu}\left(e^{i\beta} - 2 + e^{-i\beta}\right) \tag{3}$$

$$G^2 + 4\dot{\nu}(1 - \cos\beta)G - 1 = 0 \tag{4}$$

$$G^2 + 8\dot{\nu}\sin^2\frac{\beta}{2}G - 1 = 0 \tag{5}$$

Since $1 - \cos\beta = 2\sin^2\dfrac{\beta}{2}$. Solving for G:

$$\Delta = 64\dot{\nu}^2\sin^4\frac{\beta}{2} + 4 = 4\left[16\dot{\nu}\sin^4\frac{\beta}{2} + 1\right] \tag{6}$$

Then

$$G_{1,2} = -4\dot{\nu}\sin^2\frac{\beta}{2} \pm \sqrt{16\dot{\nu}\sin^4\frac{\beta}{2} + 1} = -4\dot{\nu}\sin^2\frac{\beta}{2} \pm \sqrt{\gamma} \tag{7}$$

where $\gamma = 16\dot{\nu}\sin^4\dfrac{\beta}{2} + 1$ for compactness. The roots are real since $\gamma > 0$. For the numerical method to be stable, we need $|G_{1,2}| \leq 1$. We start from G_1:

$$|G_1| = |-4\dot{\nu}\sin^2\frac{\beta}{2} - \sqrt{\gamma}| \leq 1 \quad \Longrightarrow \quad -1 \leq -4\dot{\nu}\sin^2\frac{\beta}{2} - \sqrt{\gamma} \leq 1 \tag{8}$$

Considering the lower bound, it is soon evident that

$$-1 \leq -4\dot{\nu}\sin^2\frac{\beta}{2} - \sqrt{\gamma} \tag{9}$$

$$\sqrt{\gamma} \leq 1 - 4\dot{\nu}\sin^2\frac{\beta}{2} \tag{10}$$

$$\gamma \leq 1 + 16\dot{\nu}^2\sin^4\frac{\beta}{2} - 8\dot{\nu}\sin^2\frac{\beta}{2} \tag{11}$$

$$\cancel{\gamma} \leq \cancel{\gamma} - 8\dot{\nu}\sin^2\frac{\beta}{2} \tag{12}$$

$$8\dot{\nu}\sin^2\frac{\beta}{2} \leq 0 \tag{13}$$

Since $\dot{\nu} > 0$, the scheme is always unstable.

2. Now we consider the advection equation (thus $R(w) = -a\dfrac{\partial w}{\partial x}$, with $a > 0$), where the space operator is discretized by upwind first-order differences, giving:

$$w_j^{n+1} - w_j^{n-1} = -2a\frac{\Delta t}{\Delta x}(w_j^n - w_{j-1}^n) \tag{14}$$

Discuss the stability of the numerical discretization. (*Hint:* what happens for $k\Delta x{=}\beta{=}\pi$?)

Solution: Imposing $\dot{a} = \dfrac{a\Delta t}{\Delta x}$, from a Von Neumann analysis one obtains:

$$G - \frac{1}{G} = -2\dot{a}(1 - e^{i\beta}) \tag{15}$$

$$G^2 + 2\dot{a}(1 - e^{i\beta})G - 1 = 0 \tag{16}$$

Solving for G:

$$\Delta = 4\dot{a}^2(1 - e^{i\beta})^2 + 4 \tag{17}$$

Then

$$G_{1,2} = -\dot{a}(1 - e^{i\beta}) \pm \sqrt{\dot{a}^2(1 - e^{i\beta})^2 + 1} \tag{18}$$

In this case, the development is longer; nevertheless, if we show that for a given value of β it results $|G| > 1$, we can conclude that the numerical scheme is unstable for any value of $\dot{a}$. The interval of variation of β being $[-\pi, \pi]$, we start considering $\beta = \pi$. Then $e^{i\beta} = e^{i\pi} = -1$ and the roots of equation (18) become

$$G_{1,2} = -2\dot{a} \pm \sqrt{4\dot{a}^2 + 1} \tag{19}$$

For G_1 then:

$$|G_1| = |-2\dot{a} \pm \sqrt{4\dot{a}^2 + 1}| \leq 1 \quad \Longrightarrow \quad 1 \leq -2\dot{a} \pm \sqrt{4\dot{a}^2 + 1} \leq 1 \tag{20}$$

Consider the lower bound, it is simple to show that

$$1 \leq -2\dot{a} - \sqrt{4\dot{a}^2 + 1} \tag{21}$$

$$\cancel{1 + 4\dot{a}^2} + 4\dot{a} \leq \cancel{4\dot{a}^2 + 1} \quad \Longrightarrow \quad 4\dot{a} \leq 0 \tag{22}$$

Which is impossible since $a > 0$. The method is then unstable

3. Justify the previous results by sketching the absolute stability region of the leapfrog scheme. What properties should have $R(w)$?

Solution: The region of absolute stability of the numerical scheme is given by the set of complex numbers z for which the stability polynomial $\pi(\xi; z)$ associated to the method satisfies the root condition. By considering the model equation $\dfrac{\mathrm{d}w}{\mathrm{d}t} = \lambda w$, we obtain:

$$w_j^{n+1} - w_j^{n-1} = 2\Delta t R(w) = 2\lambda\Delta t w_j^n = 2z w_j^n \tag{23}$$

Injecting the general solution ξ^n and dividing by ξ^{n-1}:

$$\xi_j^{n+1} - \xi_j^{n-1} = 2z\xi_j^n \quad \Longrightarrow \quad \xi^2 - 2z\xi - 1 = 0 \tag{24}$$

The stability polynomial reads then $\pi(\xi; z) = \xi^2 - 2z\xi - 1$. The roots $\xi_{1,2}$ are given by:

$$\Delta = 4(z^2 + 1) \qquad \Longrightarrow \qquad \xi_{1,2} = z \pm \sqrt{z^2 + 1} \tag{25}$$

It is clear that for Re(z) $\neq 0$, the root condition is never satisfied. For $z = 0 \pm i$, one has $\xi_1 = \xi_2$ hence the root condition is also not satisfied (two repeated roots of modulus 1). Therefore, the interval of stability is the segment on the imaginary axis $]$-$i, i[$. One can thus conclude that the method, when coupled with a dissipative spatial operator, will always result in an unstable discretization.

6 2D Burgers equation

We aim to solve the coupled system of Burgers' equations:

$$\frac{\partial w}{\partial t} + \frac{\partial f(w)}{\partial x} + \frac{\partial g(w)}{\partial y} = \nu \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right) \tag{1}$$

with

$$w = \begin{bmatrix} u \\ v \end{bmatrix}, \quad f = \begin{bmatrix} u^2 \\ uv \end{bmatrix}, \quad g = \begin{bmatrix} uv \\ v^2 \end{bmatrix}, \quad \nu \geq 0 \tag{2}$$

We will consider a regular Cartesian grid with sizes $\Delta x, \Delta y$ such that the coordinates for a generic grid point (i, j) are $x_{i,j} = j\Delta x$ and $y_{i,j} = j\Delta y$. The spatial derivatives are discretized by means of finite-difference centred second-order operator:

$$\left.\frac{\partial f(w)}{\partial x}\right|_{i,j} \approx \frac{\delta_x \mu_x f_{i,j}}{\Delta x} + \mathcal{O}(\Delta x^2) \qquad \left.\frac{\partial g(w)}{\partial y}\right|_{i,j} \approx \frac{\delta_y \mu_y f_{i,j}}{\Delta y} + \mathcal{O}(\Delta y^2) \tag{3}$$

$$\left.\frac{\partial^2 w}{\partial x^2}\right|_{i,j} \approx \frac{\delta_x^2 w_{i,j}}{\Delta x^2} + \mathcal{O}(\Delta x^2) \qquad \left.\frac{\partial^2 w}{\partial y^2}\right|_{i,j} \approx \frac{\delta_y^2 w_{i,j}}{\Delta y^2} + \mathcal{O}(\Delta y^2) \tag{4}$$

with the directional difference operators

$$\delta_x(\bullet)_{i,j} = (\bullet)_{i+\frac{1}{2},j} - (\bullet)_{i-\frac{1}{2},j} \qquad \mu_x(\bullet)_{i,j} = \frac{1}{2}[(\bullet)_{i+\frac{1}{2},j} + (\bullet)_{i-\frac{1}{2},j}] \tag{5}$$

$$\delta_y(\bullet)_{i,j} = (\bullet)_{i,j+\frac{1}{2}} - (\bullet)_{i,j-\frac{1}{2}} \qquad \mu_y(\bullet)_{i,j} = \frac{1}{2}[(\bullet)_{i,j+\frac{1}{2}} + (\bullet)_{i,j-\frac{1}{2}}] \tag{6}$$

First, we will use the following semi-implicit scheme for the time integration:

$$\frac{w_{i,j}^{n+1} - w_{i,j}^n}{\Delta t} = -R(w_{i,j}^n) + \nu \left[\frac{\delta_x^2 w}{\Delta x^2} + \frac{\delta_y^2 w}{\Delta y^2} \right]_{i,j}^{n+1} \tag{7}$$

where $R(w) = \dfrac{\partial f(w)}{\partial x} + \dfrac{\partial g(w)}{\partial y}$ and the index n refers to the time instant $t^n = n\Delta t$.

1. Compute the temporal order of accuracy of the scheme (7).

Solution: By expanding in Taylor series one has

$$w^{n+1} = w^n + \Delta t w_t + \frac{\Delta t^2}{2} w_{tt} + \frac{\Delta t^3}{6} w_{ttt} + \mathcal{O}(\Delta t^4) \tag{8}$$

The temporal derivative gives then

$$\frac{w_{i,j}^{n+1}-w_{i,j}^{n}}{\Delta t}=w_t+\frac{\Delta t}{2}w_{tt}+\mathcal{O}(\Delta t^2) \tag{9}$$

So the scheme is first-order-accurate in time. This is rapidly seen also by considering that the convective part is discretized by means of an explicit Euler, whereas the viscous part correspond to an implicit Euler. In both cases, one obtain a first-order-accurate scheme in time. Concerning the spatial discretization, one has:

$$w_{i+1}=w_i+\Delta x w_x+\frac{\Delta x^2}{2}w_{xx}+\frac{\Delta x^3}{6}w_{xxx}+\frac{\Delta x^4}{24}w_{4x}+\mathcal{O}(\Delta x^5) \tag{10}$$
$$w_{i-1}=w_i-\Delta x w_x+\frac{\Delta x^2}{2}w_{xx}-\frac{\Delta x^3}{6}w_{xxx}+\frac{\Delta x^4}{24}w_{4x}+\mathcal{O}(\Delta x^5) \tag{11}$$

and similarly for the y direction. Therefore, the convective part reads

$$\frac{\delta_x\mu_x f}{\Delta x}=\frac{f_{i+1}-f_{i-1}}{2\Delta x}=f_x+\Delta x^2 6 f_{xxx}+\mathcal{O}(\Delta x^3) \tag{12}$$
$$\frac{\delta_y\mu_y g}{\Delta y}=\frac{g_{j+1}-g_{j-1}}{2\Delta y}=g_y+\Delta y^2 6 g_{yyy}+\mathcal{O}(\Delta y^3) \tag{13}$$

whereas for the viscous part one has

$$\frac{\delta_x^2 w}{\Delta x^2}=\frac{w_{i-1}-2w_i+w_{i+1}}{\Delta x^2}=w_{xx}+2\frac{\Delta x^2}{24}w_{4x}+\mathcal{O}(\Delta x^6) \tag{14}$$
$$\frac{\delta_y^2 w}{\Delta y^2}=\frac{w_{j-1}-2w_j+w_{j+1}}{\Delta y^2}=w_{yy}+2\frac{\Delta y^2}{24}w_{4y}+\mathcal{O}(\Delta y^6) \tag{15}$$

Thus, both the convective and viscous part are second-order-accurate in space.

2. Write the scheme 7 in Δ-form by introducing the increment $\Delta w_{i,j}^n = w_{i,j}^{n+1}-w_{i,j}^n$.

Solution: Starting from equation (7), one obtains:

$$\frac{\Delta w_{i,j}^n}{\Delta t}=-R(w_{i,j}^n)+\nu\left[\frac{\delta_x^2 w}{\Delta x^2}+\frac{\delta_y^2 w}{\Delta y^2}\right]_{i,j}^{n+1} \tag{16}$$
$$=-R(w_{i,j}^n)+\nu\left[\frac{\delta_x^2 w_{i,j}^n}{\Delta x^2}+\frac{\delta_y^2 w_{i,j}^n}{\Delta y^2}\right]+\nu\left[\frac{\delta_x^2 \Delta w_{i,j}^n}{\Delta x^2}+\frac{\delta_y^2 \Delta w_{i,j}^n}{\Delta y^2}\right] \tag{17}$$

Regrouping and multiplying by Δt:

$$\left[\mathcal{I}-\dot{\nu}_x\delta_x^2-\dot{\nu}_y\delta_y^2\right]\Delta w_{i,j}^n=-\Delta t R(w_{i,j}^n)+\nu\Delta t\left[\frac{\delta_x^2 w_{i,j}^n}{\Delta x^2}+\frac{\delta_y^2 w_{i,j}^n}{\Delta y^2}\right] \tag{18}$$

where $\dot{\nu}_x=\dfrac{\nu\Delta t}{\Delta x^2}$ and $\dot{\nu}_y=\dfrac{\nu\Delta t}{\Delta y^2}$.

3. Show that, at each time step, the computation of the increment $\Delta w_{i,j}^n$ requires the inversion of a linear system with a scalar band matrix. Give the expression of the matrix.

Solution: In order to compute $\Delta w_{i,j}^n$, one has indeed to invert the matrix

$$A = \left[\mathcal{I} - \nu_x \delta_x^2 - \nu_y \delta_y^2\right] \tag{19}$$

which clearly represents a scalar band matrix.

4. In order to speed up the computation, we rewrite the scheme in the factorized form:

$$(\mathcal{I} - \dot{\nu}_x \delta_x^2)(\mathcal{I} - \dot{\nu}_y \delta_y^2)\Delta w_{i,j}^n = -\Delta t R(w_{i,j}^n) + \nu \Delta t \left[\frac{\delta_x^2 w}{\Delta x^2} + \frac{\delta_y^2 w}{\Delta y^2}\right]_{i,j}^n \tag{20}$$

where $\dot{\nu}_x = \dfrac{\nu \Delta t}{\Delta x^2}$ and $\dot{\nu}_y = \dfrac{\nu \Delta t}{\Delta y^2}$, which allows to compute the increment $\Delta w_{i,j}^n$ by means of the inversion of a tridiagonal scalar matrix in each direction. Show that the factorization generates an error of order Δt^2, and give the expression of the error term.

Solution: Expanding the factorized term, one obtains:

$$(\mathcal{I} - \dot{\nu}_x \delta_x^2)(\mathcal{I} - \dot{\nu}_y \delta_y^2) = \mathcal{I} - \dot{\nu}_x \delta_x^2 - \dot{\nu}_y \delta_y^2 + \dot{\nu}_x \dot{\nu}_y \delta_x^2 \delta_y^2 \tag{21}$$

The additional term is therefore

$$\dot{\nu}_x \dot{\nu}_y \delta_x^2 \delta_y^2 = \frac{\nu^2 \Delta t^2}{\Delta x^2 \Delta y^2} \delta_x^2 \delta_y^2 \tag{22}$$

which of course generates an error depending on Δt^2.

5. Now, we want to make implicit also the nonlinear operator $R(w_{i,j}^n)$. Write the fully-implicit version of the scheme by using the Δ-form.

Solution: One has

$$(\mathcal{I} - \dot{\nu}_x \delta_x^2)(\mathcal{I} - \dot{\nu}_y \delta_y^2)\Delta w_{i,j}^n = -\Delta t R(w_{i,j}^n + \Delta w_{i,j}^n) + \nu \Delta t \left[\frac{\delta_x^2 w}{\Delta x^2} + \frac{\delta_y^2 w}{\Delta y^2}\right]_{i,j}^n \tag{23}$$

Since the operator $R(w)$ is non-linear, $R(w_{i,j}^n + \Delta w_{i,j}^n)$ cannot be decomposed.

6. Recall two possible methods that one can apply to compute the solution at the time t^{n+1} for the fully-implicit scheme.

Solution: Since R is non-linear, there are two possibilities:

1. Avoid linearization, and then use an algorithm such as the Dual Time Stepping or Newton Methods;
2. Linearize R and then use any kind of iterative method, either stationary (such as Jacobi, Gauss-Seidel or SOR) or unsteady. The problem being 2D, one could also apply the Alternating-Direction-Implicit (ADI) method.

7. In order to end up with a linear system, we decide to linearize $R(w_{i,j})$ at the time t^n, *i.e.* we introduce the approximation

$$R(w_{i,j}^{n+1}) = R(w_{i,j}^n) + \left.\frac{\partial R}{\partial w}\right|_{i,j}^n \Delta w_{i,j}^n + \mathcal{O}(\Delta t^2) \tag{24}$$

To simplify the problem, we consider the inviscid case ($\nu = 0$). Give the expression of the linearized scheme as a function of the operator R and of its derivative.

Solution: Starting from the inviscid case:

$$\frac{\partial w}{\partial t} + \frac{\partial f(w)}{\partial x} + \frac{\partial g(w)}{\partial y} = 0 \tag{25}$$

The fully implicit discretized scheme reads

$$\frac{\Delta w_{i,j}^n}{\Delta t} + \frac{\delta_x \mu_x f(w^{n+1})}{\Delta x} + \frac{\delta_y \mu_y g(w^{n+1})}{\Delta y} = 0 \tag{26}$$

A linearization leads to

$$\frac{\Delta w_{i,j}^n}{\Delta t} + \frac{\delta_x \mu_x}{\Delta x}\left[f(w^n) + \frac{\partial f(w^n)}{\partial w}\Delta w_{i,j}^n\right] + \frac{\delta_y \mu_y}{\Delta y}\left[g(w^n) + \frac{\partial g(w^n)}{\partial w}\Delta w_{i,j}^n\right] = 0 \tag{27}$$

where $\frac{\partial f(w^n)}{\partial w} = J_x$ and $\frac{\partial g(w^n)}{\partial w} = J_y$ are the two Jacobian matrices of the problem. By multiplying by Δt, one obtains thus

$$\left[\mathcal{I} + \dot{J}_x \delta_x \mu_x + \dot{J}_y \delta_y \mu_y\right] \Delta w_{i,j}^n = \frac{\delta_x \mu_x f(w^n)}{\Delta x} + \frac{\delta_y \mu_y g(w^n)}{\Delta y} \tag{28}$$

where $\dot{J}_x = \frac{J_x \Delta t}{\Delta x}$ and $\dot{J}_y = \frac{J_y \Delta t}{\Delta y}$. This is exactly equivalent to:

$$\left[\mathcal{I} + \Delta t \left.\frac{\partial R}{\partial w}\right|_{i,j}^n\right] \Delta w_{i,j}^n = R(w^n) \tag{29}$$

8. Compute the Jacobian matrix of the operator R for the numerical scheme considered.

Solution:

$$\frac{\partial R(w)}{\partial w} = \frac{\delta_x \mu_x J_x}{\Delta x} + \frac{\delta_y \mu_y J_y}{\Delta y} = \frac{\delta_x \mu_x}{\Delta x}\begin{bmatrix} 2u & 0 \\ v & u \end{bmatrix} + \frac{\delta_y \mu_y}{\Delta y}\begin{bmatrix} v & u \\ 0 & 2v \end{bmatrix} \tag{30}$$

9. Specify the type of linear system to be solved (band matrix or not? How many non-zero diagonals?) and suggest a suitable algorithm for solving the problem.

Solution: Because of $\delta_x \mu_x$ and $\delta_y \mu_y$, the resulting matrix is a band matrix, with five non-zero diagonals (pentadiagonal), of type

$$\begin{bmatrix} \ddots & \ddots & 0 & \ddots & 0 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & 0 & \ddots & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \ddots & 0 & 0 \\ \ddots & 0 & \ddots & \ddots & \ddots & 0 & \ddots & 0 \\ 0 & \ddots & 0 & \ddots & \ddots & \ddots & 0 & \ddots \\ 0 & 0 & \ddots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 0 & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & \ddots & 0 & \ddots & \ddots \end{bmatrix} \tag{31}$$

In the factorized case, instead, two different tridiagonal should be solved, for instance by means of the efficient Thomas algorithm.

7 Analysis of a Dual Time Stepping technique

We aim to solve the linear advection equation:

$$w_t + aw_x = 0 \qquad \text{with} \quad a \in \mathbb{R} \tag{1}$$

To this purpose, we consider a Cartesian grid such that $x_j = j\Delta x$. The following spatial semi-discrete numerical scheme is used:

$$\frac{\partial w}{\partial t} + \frac{a}{\Delta x}\delta\left[\mu w - \frac{1}{6}\delta^2 \mu w\right]^n_j = -\nu|a|\frac{\delta^4 w^n_j}{\Delta x} \qquad \text{with} \quad \nu \in \mathbb{R} \tag{2}$$

Where the classical finite-difference operators δ and μ are used:

$$\delta(\bullet)_j = (\bullet)_{j+\frac{1}{2}} - (\bullet)_{j-\frac{1}{2}} \qquad \text{and} \qquad \mu(\bullet)_j = \frac{1}{2}[(\bullet)_{j+\frac{1}{2}} + (\bullet)_{j-\frac{1}{2}}] \tag{3}$$

1. Determine the order of accuracy of the proposed spatial discretization.

Solution: Expanding the operators one obtain

$$\delta^2(w_j) = w_{j+1} - 2w_j + w_{j-1} \qquad \delta\mu(w_j) = \frac{1}{2}(w_{j+1} - w_{j-1}) \tag{4}$$

$$\delta^2[\delta\mu(w_j)] = \frac{1}{2}\left[w_{j+2} - 2w_{j+1} + 2w_{j-1} - w_{j-2}\right] \tag{5}$$

$$\delta^2[\delta^2(w_j)] = w_{j+2} - 4w_{j+1} + 6w_j - 4w_{j-1} + w_{j-2} \tag{6}$$

The scheme (2) can be rewritten as $w_t + R(w) = 0$, where $R(w)$ is the spatial approximation:

$$R(w) = \frac{a}{\Delta x}\left[\frac{1}{2}(w_{j+1} - w_{j-1}) - \frac{1}{12}(w_{j+2} - 2w_{j+1} + 2w_{j-1} - w_{j-2})\right] \tag{7}$$

$$+ \frac{\nu|a|}{\Delta x}(w_{j+2} - 4w_{j+1} + 6w_j - 4w_{j-1} + w_{j-2}) \tag{8}$$

The spatial order of accuracy of the discretization can be obtained by computing the truncation error $\varepsilon = aw_x - R(w)$. By expanding in Taylor series one has

$$w_{j+1} = w^n_j + \Delta x w_x + \frac{\Delta x^2}{2}w_{xx} + \frac{\Delta x^3}{6}w_{3x} + \frac{\Delta x^4}{24}w_{4x} + \frac{\Delta x^5}{120}w_{5x} + \mathcal{O}(\Delta x^6) \tag{9}$$

$$w_{j-1} = w^n_j - \Delta x w_x + \frac{\Delta x^2}{2}w_{xx} - \frac{\Delta x^3}{6}w_{3x} + \frac{\Delta x^4}{24}w_{4t} - \frac{\Delta x^5}{120}w_{5t} + \mathcal{O}(\Delta x^6) \tag{10}$$

$$w_{j+2} = w^n_j + 2\Delta x w_x + \frac{4\Delta x^2}{2}w_{xx} + \frac{8\Delta x^3}{6}w_{3x} + \frac{16\Delta x^4}{24}w_{4x} + \frac{32\Delta x^5}{120}w_{5x} + \mathcal{O}(\Delta x^6) \tag{11}$$

$$w_{j-2} = w^n_j - 2\Delta x w_x + \frac{4\Delta x^2}{2}w_{xx} - \frac{8\Delta x^3}{6}w_{3x} + \frac{16\Delta x^4}{24}w_{4x} - \frac{32\Delta x^5}{120}w_{5x} + \mathcal{O}(\Delta x^6) \tag{12}$$

and therefore

$$R(w) = \frac{a}{\Delta x}\left[\Delta x w_x + \frac{\Delta x^3}{6}w_{3x} + \frac{\Delta x^5}{120}w_{5x} + \mathcal{O}(\Delta x^7) - \frac{1}{12}\left(\cancel{4\Delta x w_x} + \frac{16\Delta x^3}{6}w_{3x}\right.\right. \tag{13}$$

$$\left.\left.+\frac{64\Delta x^5}{120}w_{5x} - \cancel{4\Delta x w_x} - \frac{4\Delta x^3}{6}w_{3x} - \frac{4\Delta x^5}{120}w_{5x} + \mathcal{O}(\Delta x^7)\right)\right] \tag{14}$$

$$+\frac{\nu|a|}{\Delta x}\left[\cancel{2w_j} + \cancel{4\Delta x^2 w_{xx}} + \frac{32\Delta x^4}{24}w_{4x} + \mathcal{O}(\Delta x^6) + \cancel{6w_j} - \cancel{8w_j}\right. \tag{15}$$

$$\left.-\cancel{4\Delta x^2 w_{xx}} - \frac{8\Delta x^4}{24}w_{4x} + \mathcal{O}(\Delta x^6)\right] \tag{16}$$

Then

$$R(w) = \frac{a}{\Delta x}\left[\Delta x w_x + \cancel{\frac{\Delta x^3}{6}w_{3x}} + \frac{\Delta x^5}{120}w_{5x} - \cancel{\frac{\Delta x^3}{6}w_{3x}} - \frac{5\Delta x^5}{120}w_{5x} + \mathcal{O}(\Delta x^7)\right] \tag{17}$$

$$+\frac{\nu|a|}{\Delta x}\left[\Delta x^4 w_{4x} + \mathcal{O}(\Delta x^6)\right] \tag{18}$$

$$\implies R(w) = a w_x - \frac{\Delta x^4}{30}w_{5x} + \nu|a|\Delta x^3 w_{4x} + \mathcal{O}(\Delta^5) \tag{19}$$

The truncation error reads then

$$\varepsilon = a w_x - R(w) = \cancel{a w_x} - \cancel{a w_x} - \nu|a|\Delta x^3 w_{4x} + \mathcal{O}(\Delta x^4) \tag{20}$$

The spatial discretization is then 3$^{\text{rd}}$-order accurate.

2. We rewrite the semi-discrete scheme under the form of an ordinary differential equation:

$$\frac{\mathrm{d}w}{\mathrm{d}t} + R(w) = 0 \tag{21}$$

and solve the equation by means of the Gear (BDF2) scheme, which reads:

$$\frac{\mathrm{d}w}{\mathrm{d}t} + R(w^{n+1}) = 0 \qquad \text{with} \qquad \frac{\mathrm{d}w}{\mathrm{d}t} = \frac{3w_j^{n+1} - 4w_j^n + w_j^{n-1}}{2\Delta t} \tag{22}$$

Compute the order of accuracy of the temporal approximation chosen.

Solution: We now expand in Taylor series in time:

$$w^{n+1} = w + \Delta t w_t + \frac{\Delta t^2}{2}w_{tt} + \frac{\Delta t^3}{6}w_{ttt} + \mathcal{O}(\Delta t^4) \tag{23}$$

$$w^{n-1} = w - \Delta t w_t + \frac{\Delta t^2}{2}w_{tt} - \frac{\Delta t^3}{6}w_{ttt} + \mathcal{O}(\Delta t^4) \tag{24}$$

$$-R(w^{n+1}) = w_t^{n+1} = w_t + \Delta t w_{tt} + \frac{\Delta t^2}{2}w_{3t} + \frac{\Delta t^3}{6}w_{4t} + \mathcal{O}(\Delta t^4) \tag{25}$$

The order of accuracy of the temporal approximation is then

$$\varepsilon = \frac{1}{2\Delta t}\left[\cancel{3w} + 3\Delta t w_t + \frac{3\Delta t^2}{2}w_{tt} + \frac{3\Delta t^3}{6}w_{3t} - \cancel{4w} + \cancel{w} - \Delta t w_t + \frac{\Delta t^2}{2}w_{tt}\right. \tag{26}$$

$$\left.-\frac{\Delta t^3}{6}w_{3t} + \mathcal{O}(\Delta t^4)\right] - w_t - \Delta t w_{tt} - \frac{\Delta t^2}{2}w_{3t} - \frac{\Delta t^3}{6}w_{4t} + \mathcal{O}(\Delta t^4) \tag{27}$$

$$\implies \varepsilon = \cancel{w_t} + \cancel{\Delta t w_{tt}} + \frac{\Delta t^2}{6}w_{ttt} + \mathcal{O}(\Delta t^3) - \cancel{w_t} - \cancel{\Delta t w_{tt}} - \frac{\Delta t^2}{2}w_{ttt} + \mathcal{O}(\Delta t^3) \tag{28}$$

$$= \mathcal{O}(\Delta t^2) \tag{29}$$

The discretization is then second-order accurate in time.

3. Knowing that the Gear scheme is A-stable, compute the Fourier symbol of the spatial operator $\widehat{R}$, and comment the conditional (or unconditional) stability (or instability) of the fully-discrete scheme (22), as a function of the sign of the coefficient ν.

Solution: Remembering that

$$\widehat{\delta\mu} = i\sin\beta \tag{30}$$

$$\widehat{\delta^2} = 2(\cos\beta - 1) \tag{31}$$

$$\widehat{\delta^2(\delta\mu)} = \widehat{\delta^2}\widehat{\delta\mu} = 2i\sin\beta(\cos\beta - 1) \tag{32}$$

$$\widehat{\delta^4} = \widehat{\delta^2}\widehat{\delta^2} = 4(\cos\beta - 1)^2 \tag{33}$$

Plugging these relations in $\widehat{R(w)}$, one has

$$\widehat{R(w)} = \frac{a}{\Delta x}\left[i\sin\beta - \frac{1}{6}i\sin\beta(\cos\beta - 1)\right] + \frac{\nu|a|}{\Delta x}4(\cos\beta - 1)^2 \tag{34}$$

The real part of the Fourier symbol of $R(w)$ reads $+\frac{\nu|a|}{\Delta x}4(\cos\beta - 1)^2$. Separating the temporal and spatial operators of the fully-discrete scheme, one has:

$$\widehat{\frac{\mathrm{d}w}{\mathrm{d}t}} = -\widehat{R(w)} \tag{35}$$

For the scheme to be stable, the stability region of the spatial operator (on the r.h.s) must be contained in the one of the temporal operator (on the l.h.s.), which is A-stable in this case. Therefore, it results that the fully-discrete scheme is unconditionally stable if

$$\mathrm{Re}(-\widehat{R(w)}) < 0 \qquad \forall\beta \tag{36}$$

Since $\mathrm{Re}(-\widehat{R(w)}) = -\frac{\nu|a|}{\Delta x}4(\cos\beta - 1)^2$, one has that the scheme is unconditionally stable for $\nu > 0$, and unconditionally unstable for $\nu < 0$.

4. We now want to solve the equation (22), with the spatial operator shown in equation (2), by means of the Dual Time Stepping (DTS) method. In order to do so, we add to the equation (22) a derivative with respect to a fictitious time τ, obtaining

$$\frac{\mathrm{d}w}{\mathrm{d}\tau} + R^*(w) = 0 \qquad \text{with} \quad R^*(w) = \frac{Dw}{\Delta t} + R(w) \tag{37}$$

The ODE (37) is approximated by means of the implicit Euler scheme:

$$\frac{w^{n+1,m+1} - w^{n+1,m}}{\Delta\tau} + R^*(w^{n+1,m+1}, w^n, w^{n-1}) = 0 \tag{38}$$

where $w^{n+1,m+1}$ denotes the approximation of w at the physical time step $n+1$, at the iteration $m+1$ with respect to the fictitious time. The approximation in τ becomes then:

$$\frac{w_j^{m+1} - w_j^m}{\Delta\tau} + \lambda w_j^{m+1} + \dot{a}\left[\delta\mu w_j^{m+1} - \frac{1}{6}\delta^3\mu w_j^{m+1}\right] + \nu|\dot{a}|\delta^4 w_j^{m+1} = C(w^n, w^{n-1}) \tag{39}$$

where $C(w^n, w^{n-1})$ is a constant (in the pseudo-time) function, λ a numerical constant and $\dot{a} = a\dfrac{\Delta t}{\Delta x}$. Give the expression of λ and $C(w^n, w^{n-1})$.

Solution: Since $R^*(w) = \dfrac{\mathrm{d}w}{\mathrm{d}t} + R(w) = 0$, we can multiply it by Δt obtaining:

$$R^*(w) = \frac{3w_j^{n+1} - 4w_j^n + w_j^{n-1}}{2} + \dot{a}\left[\delta\mu w_j^{n+1} - \frac{1}{6}\delta^3\mu w_j^{n+1}\right] + \nu|\dot{a}|\delta^4 w_j^{n+1} = 0 \tag{40}$$

Injecting equation (40) in (38), one obtains

$$\frac{w^{n+1,m+1} - w^{n+1,m}}{\Delta\tau} + \frac{3w_j^{n+1,m+1} - 4w_j^n + w_j^{n-1}}{2} \tag{41}$$

$$+ \dot{a}\left[\delta\mu w_j^{n+1,m+1} - \frac{1}{6}\delta^3\mu w_j^{n+1,m+1}\right] + \nu|\dot{a}|\delta^4 w_j^{n+1,m+1} = 0 \tag{42}$$

Since we are iterating in the pseudo-time τ, the values of w_j^n and w_j^{n-1} are constant and can be grouped in a constant term:

$$-C(w^n, w^{n-1}) = \frac{-4w_j^n + w_j^{n-1}}{2} \tag{43}$$

On the contrary, w_j^{n+1} is updated at each m-th iteration. Therefore, we can drop the index n obtaining

$$\frac{w^{m+1} - w^m}{\Delta\tau} + \frac{3}{2}w_j^{m+1} + C + \dot{a}\left[\delta\mu w_j^{m+1} - \frac{1}{6}\delta^3\mu w_j^{m+1}\right] + \nu|\dot{a}|\delta^4 w_j^{m+1} = C(w^n, w^{n-1}) \tag{44}$$

Thus, a direct comparison with equation (39) shows that $\lambda = \dfrac{3}{2}$.

5. Analyse the linear stability of eq. (39) as a function of the parameter λ, assuming $C = 0$.

Solution: Assuming $C = 0$, eq. (39) reads:

$$\frac{w_j^{m+1} - w_j^m}{\Delta\tau} + \lambda w_j^{m+1} + \dot{a}\left[\delta\mu w_j^{m+1} - \frac{1}{6}\delta^3\mu w_j^{m+1}\right] + \nu|\dot{a}|\delta^4 w_j^{m+1} = 0 \tag{45}$$

A classical Von Neumann analysis leads to

$$G - 1 + \Delta\tau\lambda G + \Delta\tau\widehat{R(w)}G = 0 \quad \Longrightarrow \quad G = \frac{1}{1 + \Delta\tau\lambda + \Delta\tau\widehat{R(w)}} \tag{46}$$

The stability imposes that

$$|G|^2 = \frac{1}{\left|1 + \Delta\tau\lambda + \Delta\tau\widehat{R(w)}\right|^2} \leq 1 \tag{47}$$

Therefore, if the denominator of (47) is strictly greater than unity, the scheme is unconditionally stable. Remember that

$$\mathrm{Re}(\widehat{R(w)}) = \nu|\dot{a}|4(\cos\beta - 1)^2 \tag{48}$$

$$\mathrm{Im}(\widehat{R(w)}) = \dot{a}\sin\beta\left[1 + \frac{1}{6}(1 - \cos\beta)\right] \tag{49}$$

Then

$$\left|1 + \Delta\tau\lambda + \Delta\tau\widehat{R(w)}\right|^2 = [1 + \lambda + \Delta\tau\mathrm{Re}(\widehat{R(w)})]^2 + \Delta\tau[\mathrm{Im}(\widehat{R(w)})]^2 \tag{50}$$

We have already shown that $\mathrm{Re}(\widehat{R(w)}) > 0\ \forall\beta$, then the real part is strictly larger than zero. Note that the presence of the source term $\lambda > 0$ improves the stability of the scheme. For the imaginary part, one has:

$$[\mathrm{Im}(\widehat{R(w)})]^2 = \dot{a}^2\sin^2\beta\left[1 + \frac{1}{6}(1 - \cos\beta)\right]^2 \tag{51}$$

Which is also strictly larger than unity. Therefore, the scheme is unconditionally stable.

6. Does the chosen fictitious time approximation, in combination with the time and space approximations, allow to converge rapidly towards the steady state with respect to τ? Is it more suited for slow or rapid problems? Justify the answer.

Solution: The numerical scheme allows to converge rapidly towards the steady state with respect to τ since the discretization is unconditionally stable (no limitations on $\Delta\tau$). Therefore, it is well suited for slow problems.