1D Numerical Methods With Finite Volumes

Guillaume Riflet
MARETEC IST

1 The advection-diffusion equation

The original concept, applied to a property within a control volume V, from which is derived the integral advection-diffusion equation, states as

$$\begin{aligned} \{\text{Rate of change in time}\} &= \{\text{Ingoing} - \text{Outgoing fluxes}\} \\ &+ \{\text{Created} - \text{Destroyed}\}. \end{aligned} \tag{1}$$

Annotated in a correct mathematical encapsulation, equation 1 yields

$$\begin{aligned} \frac{d\int_{V_t} P dV}{dt} = &- \oint_{\partial V_t} P\left(\mathbf{v}_r \cdot \mathbf{n}\right) dS \\ &- \oint_{\partial V_t} -K\ \left(\nabla P \cdot \mathbf{n}\right) dS \\ &+ \int_{V_t} \left(\mathrm{Sc} - \mathrm{Sk}\right) dV, \end{aligned} \tag{2}$$

where P is the transported property concentration, K is the diffusivity coefficient, ∂V_t is the surface of the moving control volume V_t, $\mathbf{v}_r$ is the flow velocity relative to the moving surface of the control volume V_t, $\mathbf{n}$ is the outward normal to the control surface and Sc and Sk are the source and sink terms, respectively. The first term on the right hand side (RHS) of equation 2 states ingoing and outgoing fluxes due to advection, the second term in the RHS states ingoing and outgoing fluxes due to diffusion, and the last term on the RHS accounts for source and sink terms. Note that Fick's law, $-K\,\nabla P$, is applied to mathematically describe diffusion [1].

By resorting to the Reynolds transport theorem [2], illustrated in figure 1, stating that the derivative in time of a property in a given moving control volume V_A is equal to the derivative in time of the same property in another given moving control volume V_B, coincident in one time instant with V_A, summed to the flow of the property through the control volumes given by

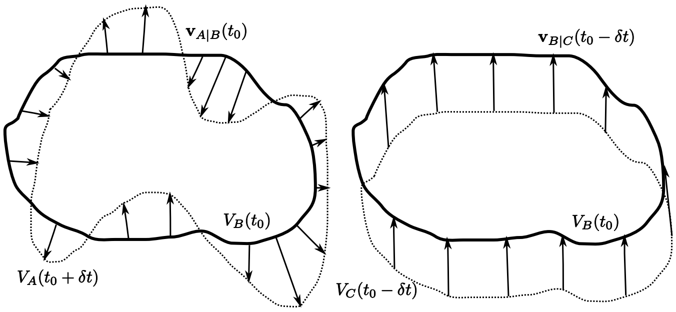

Fig. 1: Illustration of control volumes V_A, V_B and V_C near time instants $t = t_0 + \delta t$, $t = t_0$ and $t = t_0 - \delta t$, respectively, and their relative deformation rate of change, represented by the vector fields $\mathbf{v}_{A|B}$ and $\mathbf{v}_{B|C}$ plotted on the surfaces of volumes V_B and V_C, respectively. The control volumes have different motions and are coincident at time instant $t = t_0$.

their relative velocity $\mathbf{v}_{A|B}$, i.e.

$$\frac{d}{dt}\int_{V_A} P\,dV = \frac{d}{dt}\int_{V_B} P\,dV + \oint_{\partial V_B} P\,\mathbf{v}_{A|B}\cdot\mathbf{n}\,dS, \tag{3}$$

and applying it to equation 2,

$$\begin{aligned} \frac{d\int_V PdV}{dt} + \oint_{\partial V} P\left(\mathbf{v}_{V_t|V}\cdot\mathbf{n}\right)dS = &-\oint_{\partial V_t} P\left(\mathbf{v}_r\cdot\mathbf{n}\right)dS \\ &-\oint_{\partial V_t} -K\,\left(\nabla P\cdot\mathbf{n}\right)dS \\ &+\int_{V_t}\left(\mathrm{Sc}-\mathrm{Sk}\right)dV, \end{aligned} \tag{4}$$

where V is a control volume held fixed in time relative the laboratory reference frame and $\mathbf{v}_{V_t|V}$ is the relative velocity of the moving control volume V_t relative to the laboratory reference frame or to V, which is the same.

The flow velocity relative the laboratory reference frame $\mathbf{v}$ is then given by

$$\mathbf{v} = \mathbf{v}_r + \mathbf{v}_{V_t|V},$$

and minding that $\partial V = \partial V_t$ and $V = V_t$ at one time instant, equation 4 rewrites

$$\frac{d\int_V PdV}{dt} = -\oint_{\partial V} P\,(\mathbf{v}\cdot\mathbf{n})\,dS - \oint_{\partial V} -K\;(\nabla P\cdot\mathbf{n})\,dS + \int_V (\mathrm{Sc}-\mathrm{Sk})\,dV, \tag{5}$$

By resorting to the divergence theorem, stating that the divergence of a vector field inside any finite gaussian volume is equal to its flux through the boundary of the volume, i.e.

$$\int_V \nabla.\mathbf{E}\,dV = \int_{\partial V} \mathbf{E}\cdot\mathbf{n}\,dS, \tag{6}$$

being V the gaussian volume, ∂V its boundary, $\mathbf{E}$ the vector field, and $\mathbf{E}\cdot\mathbf{n}$ its flux through the boundary. $\mathbf{n}$ is defined as the external normal unit vector to the boundary. Hence, by applying equation(6) to equation(2), the finite volume formulation of equation (7) is obtained:

$$\frac{d\int_V PdV}{dt} = -\int_V \nabla\cdot(\mathbf{v}\,P)\,dV - \int_V \nabla\cdot(-\,K\,\nabla P)\,dV + \int_V (\mathrm{Sc}-\mathrm{Sk})\,dV. \tag{7}$$

By resorting to the Leibniz integration rule,

$$\frac{d}{dt}\int_V f(\mathbf{x},\mathbf{t})dV = \int_V \frac{d\,f(\mathbf{x},\mathbf{t})}{dt}dV,$$

which is true as long as the integral volume V (and consequently $\mathbf{x}$) is held fixed in time, in one hand. By joining all the integrals in V into a single integral on the other, equation 7 yields

$$\int_V \left\{\frac{\partial P}{\partial t} + \nabla\cdot(\mathbf{v}\,P) + \nabla\cdot(-\,K\,\nabla P) - \mathrm{Sc} + \mathrm{Sk}\right\} dV = 0.$$

Since the above integral holds zero for all gaussian volume V, then its integrand must be zero, thus yielding the differential equation of advection-diffusion:

$$\frac{\partial P}{\partial t} + \nabla\cdot(\mathbf{v}\,P) + \nabla\cdot(-\,K\,\nabla P) - \mathrm{Sc} + \mathrm{Sk} = 0. \tag{8}$$

A few comments: first, the control volume is held fixed in time. Second, the time derivative applied to a function varying in time only in its explicit component (such as $P(\mathbf{x}, t)$), may be annotated as a partial derivative, $\frac{\partial}{\partial t}$, without loss of generality. Third, the total derivative and the partial derivative are related by the Leibniz chain rule,

$$\begin{aligned} \left.\frac{d\,P(\mathrm{x}(t),t)}{dt}\right|_{t_0} = & \left.\frac{\partial P(\mathrm{x}(t_0),t)}{\partial t}\right|_{t_0} \\ & + \left.\frac{d\,x(t)}{dt}\right|_{t_0} \frac{\partial P(\mathrm{x}(t_0),t_0)}{\partial x} \\ & + \left.\frac{d\,y(t)}{dt}\right|_{t_0} \frac{\partial P(\mathrm{x}(t_0),t_0)}{\partial y} \\ & + \left.\frac{d\,z(t)}{dt}\right|_{t_0} \frac{\partial P(\mathrm{x}(t_0),t_0)}{\partial z}. \end{aligned}$$

If $\frac{d\mathrm{x}}{dt} = 0$, then the partial derivative equals the total derivative, which is the case of P within the fixed volume V.

By noting that

$$\nabla \cdot (\mathbf{v}\,P) = \mathbf{v} \cdot \nabla P + P\,\nabla \cdot \mathbf{v},$$

equation 8 returns

$$\frac{\partial P}{\partial t} + \mathbf{v} \cdot \nabla P + P\,\nabla \cdot \mathbf{v} + \nabla \cdot (-\,K\,\nabla P) - \mathrm{Sc} + \mathrm{Sk} = 0.$$

By noting that the material derivative is defined by

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla,$$

equation 8 finally yields

$$\frac{DP}{Dt} + P\,\nabla \cdot \mathbf{v} + \nabla \cdot (-\,K\,\nabla P) - \mathrm{Sc} + \mathrm{Sk} = 0.$$

In incompressible fluids we have $\nabla \cdot \mathbf{v} = 0$, hence in such case

$$\frac{DP}{Dt} + \nabla \cdot (-\,K\,\nabla P) - \mathrm{Sc} + \mathrm{Sk} = 0. \tag{9}$$

Thus, to sum up, equation (5) is the integral equation of advection and diffusion (with source and sink terms) in flux form, equation (7) is the

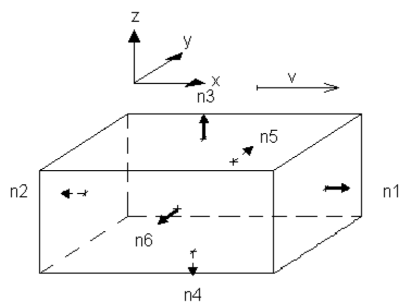

Fig. 2: Finite cartesian volume, under uniform flow $\mathbf{v}$, defined by its normal vectors n_1, n_2, n_3, n_4, n_5 and n_6 and respective faces A_1, A_2, A_3, A_4, A_5 and A_6.

integral equation of advection-diffusion in volume form, equation (8) is the differential equation of advection-diffusion (where the advective field is the velocity of the fluid particles relative to a fixed reference frame) and equation (9) is the differential equation of advection-diffusion for an incompressible fluid. Mind, however, that equation (2) is the flux form counterpart of equation (5) for moving control volumes. In that case, the kinematics of the moving control volume must also be given in order to solve the equation. For example, consider the control volumes of a undimensional channel where the free surface of the control volume moves in time according to the water elevation...

1.1 Integrating in a cartesian finite volume

Let us integrate each term of equation (2) for each face of the cartesian finite-volume of figure 2:

$$\begin{array}{lll} \mathbf{v} = (u,\, 0,\, 0) & \nabla P = \left(\frac{\partial P}{\partial x},\, 0,\, 0\right) & V = \Delta x \Delta y \Delta z \\ \mathbf{n1} = (1,\, 0,\, 0) & \mathbf{n2} = (-1,\, 0,\, 0) & \mathbf{n3} = (0,\, 0,\, 1) \\ \mathbf{n4} = (0,\, 0,\, -1) & \mathbf{n5} = (0,\, 1,\, 0) & \mathbf{n6} = (0,\, -1,\, 0) \\ A_1 = A_2 = \Delta y \Delta z & A_3 = A_4 = \Delta x \Delta y & A_5 = A_6 = \Delta x \Delta z \end{array} \tag{10}$$

Only fluxes through faces defined by $\mathbf{n1}$ and $\mathbf{n2}$ are non-null:

$$\begin{array}{ll} \mathbf{v} \cdot \mathbf{n1} = u & \mathbf{v} \cdot \mathbf{n2} = -u \\ \nabla P \cdot \mathbf{n1} = \frac{\partial P}{\partial x} & \nabla P \cdot \mathbf{n2} = -\frac{\partial P}{\partial x}. \end{array}$$

Thus we get, if we define $\overline{P} \equiv \frac{\int_V P dV}{V}$,

$$\begin{aligned} \frac{\partial \int_V P dV}{\partial t} &= \frac{\partial \left(\overline{P}\, V \right)}{\partial t} \\ &= \frac{\partial \overline{P}}{\partial t} V, \end{aligned}$$

furthermore, if we define $\tilde{P}_i \equiv \frac{\int_{A_i} P\, dS}{A_i}$, we obtain

$$\begin{aligned} \int_A P\ (\mathbf{v} \cdot \mathbf{n})\, dS &= \sum_{i=1}^{6} \int_{A_i} P\ (\mathbf{v} \cdot \mathbf{ni})\, dS \\ &= u \int_{A_1} P\, dS - u \int_{A_2} P\, dS \\ &= u\, \tilde{P}_1\, A_1 - u\, \tilde{P}_2\, A_2, \end{aligned}$$

finally, if we approximate $\frac{\int_{A_i} \frac{\partial P}{\partial x}\, dS}{A_i} \approx \frac{\partial \overline{P}}{\partial x}$, we have

$$\begin{aligned} \int_A -K\ (\nabla P \cdot \mathbf{n})\, dS &= \sum_{i=1}^{6} \int_{A_i} -K\ (\nabla P \cdot \mathbf{ni})\, dS \\ &= -K \int_{A_1} \frac{\partial P}{\partial x} dS + K \int_{A_2} \frac{\partial P}{\partial x} dS \\ &= -K\, \frac{\partial \overline{P}}{\partial x}_{(A_1)} A_1 + K\, \frac{\partial \overline{P}}{\partial x}_{(A_2)} A_2. \end{aligned}$$

Gathering up the terms above back into equation (2) yields

$$\frac{\partial \overline{P}}{\partial t} V = -u\, \tilde{P}_1\, A_1 + u\, \tilde{P}_2\, A_2 + K\, \frac{\partial \overline{P}}{\partial x}_{(A_1)} A_1 - K\, \frac{\partial \overline{P}}{\partial x}_{(A_2)} A_2 + \text{Sc'} - \text{Sk'}$$

$$\frac{\partial \overline{P}}{\partial t} = -u\, \frac{A_1}{V}\, \tilde{P}_1 + u\, \frac{A_2}{V}\, \tilde{P}_2 + K\, \frac{A_1}{V}\, \frac{\partial \overline{P}}{\partial x}_{(A_1)} - K\, \frac{A_2}{V}\, \frac{\partial \overline{P}}{\partial x}_{(A_2)} + \frac{\text{Sc'}}{V} - \frac{\text{Sk'}}{V}. \quad (11)$$

Now the whole of the game is to find numerical schemes that evaluate the time and space derivatives, in equation (11), of the volume averaged $\overline{P}$, $\frac{\partial \overline{P}}{\partial t}$ and $\frac{\partial \overline{P}}{\partial x}$ properties; as well as to evaluate the surface averaged property $\tilde{P}$.

1.2 Discretizing in a cartesian 1D finite volume

The 1D finite volume is described in figure 2 and the 1D uniform mesh is described in figure 4. Let it be defined that $\overline{P} \equiv P_i$ and that $\tilde{P}_1 \equiv P_{i+1/2}$, $\tilde{P}_2 \equiv P_{i-1/2}$.

1.2.1 Time derivatives

- Time forward or explicit method

$$\frac{\partial P}{\partial t}(t) \approx \frac{P(t+\Delta t) - P(t)}{\Delta t} \tag{12}$$

- Time backward or implicit method

$$\frac{\partial P}{\partial t}(t) \approx \frac{P(t) - P(t-\Delta t)}{\Delta t} \tag{13}$$

- Midpoint method

$$\frac{\partial P}{\partial t}(t) \approx \frac{P(t+\Delta t) - P(t-\Delta t)}{2\,\Delta t}. \tag{14}$$

As a note, it can be shown ([3], [4]) that the forward and backward method have first-order precision and that the midpoint method has second-order precision, on an uniform grid.

1.2.2 Space derivatives

- Diffusion typically uses the midpoint method for the spatial derivative $\frac{\partial P}{\partial x}$ (see equation (14)).

$$\begin{cases} \frac{\partial P}{\partial x}_{(i+1/2)} & = & \frac{P_{i+1}-P_i}{\Delta x} \\ \frac{\partial P}{\partial x}_{(i-1/2)} & = & \frac{P_i-P_{i-1}}{\Delta x} \end{cases},$$

$$\begin{aligned} & -K\,\frac{A_1}{V}\,\frac{\partial P}{\partial x}_{(i+1/2)} + K\,\frac{A_2}{V}\,\frac{\partial P}{\partial x}_{(i-1/2)} \\ & = -\frac{K}{\Delta x}\left(\frac{P_{i+1}-P_i}{\Delta x} - \frac{P_i - P_{i-1}}{\Delta x}\right) \\ & = -\frac{K}{\Delta x^2}\left(P_{i+1} - 2\,P_i + P_{i-1}\right), \end{aligned} \tag{15}$$

- Advection typically uses the upwind method (backward method if $u < 0$, and forward method if $u \geq 0$),

$$\begin{array}{rcl} P_{i+1/2} & = & \begin{cases} P_i & , \text{ if } \quad u \geq 0 \\ P_{i+1} & , \text{ if } \quad u < 0 \end{cases} \\ P_{i-1/2} & = & \begin{cases} P_{i-1} & , \text{ if } \quad u \geq 0 \\ P_i & , \text{ if } \quad u < 0 \end{cases} \end{array} , \tag{16}$$

$$\begin{array}{rcl} & & u\,\frac{A_1}{V}\,P_{i+1/2} - u\,\frac{A_2}{V}\,P_{i-1/2} \\ = & 0 \quad P_{i+1} \quad + \quad \frac{u}{\Delta x} \quad P_i \quad + \quad -\frac{u}{\Delta x} \quad P_{i-1}, & \text{if} \quad u \geq 0 \\ = & \frac{u}{\Delta x} \quad P_{i+1} \quad + \quad -\frac{u}{\Delta x} \quad P_i \quad + \quad 0 \quad P_{i-1}, & \text{if} \quad u < 0 \end{array}$$

$$= -\frac{\left\| \frac{u}{\Delta x} \right\| - \frac{u}{\Delta x}}{2}\,P_{i+1} + \left\| \frac{u}{\Delta x} \right\|\,P_i - \frac{\left\| \frac{u}{\Delta x} \right\| + \frac{u}{\Delta x}}{2}\,P_{i-1}, \tag{17}$$

- or the central i method (i.e. midpoint method),

$$\begin{array}{rcl} P_{i+1/2} & = & \frac{P_{i+1}+P_i}{2} \\ P_{i-1/2} & = & \frac{P_i+P_{i-1}}{2} \end{array} , \tag{18}$$

$$u\,\frac{A_1}{V}\,P_{i+1/2} - u\,\frac{A_2}{V}\,P_{i-1/2} = \frac{u}{2\,\Delta x}\,P_{i+1} + 0\,P_i - \frac{u}{2\,\Delta x}P_{i-1}. \tag{19}$$

2 Some notes on numerical methods

This section aims at commenting on stability (robustness), positivity, mass conservation, boundary conditions and initial conditions for the several types of schemes reviewed in this work. We'll only look at the fluxes through the boundaries of the control volumes and consider null sources and sinks in order to properly discuss mass conservation. But, in the first hand a quick revision on finite-difference methods are suggested so as clearly derive the approximation order of the forward, backward and mid-point discretization of the derivative.

2.1 Finite-difference solvers

In order to solve numerically equations like equation (8) it is necessary to discretize the derivative operators. After discretizing the derivative operators, it is also useful to have an estimate of the error of the solution. When one knows its analytical solution, then it suffices to compute the RMS (root mean square error) in order to know exactly what is the error. However, often, numerical solvers are used because there are no known analytical solution to the PDE (partial differential equation), thus there is no alternative but to estimate the error. A traditional approach consists in using constant discrete time steps dt and constant discrete spatial steps dx combined with the Taylor serie expansion definition of analytical functions [4]. The Taylor serie expansion of an analytical real function forward in its variable coordinate is

$$f\,(t+dt) = \sum_{n=0}^{\infty} f^{(n)}(t)\frac{(dt)^n}{n!} \tag{20}$$

and backward in its variable coordinate is

$$f\,(t-dt) = \sum_{n=0}^{\infty} f^{(n)}(t)\frac{(-dt)^n}{n!}. \tag{21}$$

2.1.1 Explicit method

The Taylor series in equation (20) expanded up to the second order (with third order error),

$$f\,(t+dt) = f(t) + f^{'}(t)\,dt + f^{''}(t)\frac{(dt)^2}{2} + o\left((dt)^3\right). \tag{22}$$

A first derivative approximation may be obtained from equation 22,

$$f^{'}(t) = \frac{f\,(t+dt) - f\,(t)}{dt} + o\,(dt)\,. \tag{23}$$

The first derivative approximation in equation (23) yields first order error. This numerical method is often referred to as the explicit method or the forward in time method. Note that

$$\frac{o((dt)^3)}{dt} = o((dt)^2),$$

$$o(dt) = o(-dt) = -o(dt),$$

are general properties of truncature error operator.

2.1.2 Implicit method

Likewise, the Taylor series in equation (21) expanded up to the second order (with third order error) yields,

$$f\,(t-dt) = f(t) + f^{'}(t)\,(-dt) + f^{''}(t)\frac{(-dt)^2}{2} + o\left((-dt)^3\right). \tag{24}$$

A first derivative approximation may be obtained from equation 24,

$$f^{'}(t) = \frac{f(t) - f(t - dt)}{dt} + o(dt). \tag{25}$$

This numerical method is referred to as the implicit method or the backward in time method and is also first order error.

2.1.3 Mid-point method

By subtracting equation (24) from equation (22), the mid-point method is obtained:

$$f(t + dt) - f(t - dt) = f^{'}(t)\, 2dt + o\left((dt)^3\right), \tag{26}$$

$$f^{'}(t) = \frac{f(t + dt) - f(t - dt)}{2dt} + o\left((dt)^2\right). \tag{27}$$

Equation (27) is referred to as the mid-point method and has a second order error.

2.1.4 Second derivative discretization

By summing equation (24) from equation (22), a discretization of the second derivative is obtained:

$$f(t + dt) + f(t - dt) = 2\, f(t) + f^{''}(t)\,(dt)^2 + o\left((dt)^4\right). \tag{28}$$

Note that the third order terms in the forward and backward Taylor series expansion cancel out. In fact, every odd order term are cancelled when summing the forward and backward Taylor series expansion, whereas its the even terms that are cancelled when subtracting the backward Taylor series from the forward Taylor series. Thus, equation (28) yields

$$f^{''}(t) = \frac{f(t + dt) - 2\, f(t) + f(t - dt)}{(dt)^2} + o\left((dt)^2\right). \tag{29}$$

2.2 Explicit schemes

Unidimensional meshes yield, by discretizing equation (2) forward in time,

$$P_i(t + \Delta t) = A_i\, P_{i-1}(t) + (1 - B_i) P_i(t) + C_i\, P_{i+1}(t). \tag{30}$$

The adimensional coefficients A_i and C_i are associated with the ingoing fluxes from neighbouring cells, while coefficient B_i is associated with the outgoing fluxes to neighbouring cells. Physically, they represent the percentage of mass coming from, and going to, neighbouring cells and, as such, they should be bounded between 0% and 100%.

2.2.1 positivity

Indeed, if the coefficients affected to P were negative, that would violate the principle of positive mass, or positivity, and the numerical solution would return negative oscillations of the property concentration. Thus, to ensure that positivity is not violated, it is required that:

$$\forall i \begin{cases} A_i & \geq & 0, \\ B_i & \geq & 0, \\ C_i & \geq & 0. \end{cases} \tag{31}$$

2.2.2 stability and robustness

On the other hand, if the absolute value of the coefficients were above unity, that would transport more mass than there is (credit is risky) and would turn the scheme unstable. When instabilities occur, they are immediately spotted, due to the disproportionate growth of the property concentration, and the computer usually quickly returns an overflow error. Thus, to ensure the stability of the method, it is required that:

$$\forall i \begin{cases} \|A_i\| & \leq & 1, \\ \|B_i\| & \leq & 1, \\ \|C_i\| & \leq & 1. \end{cases} \tag{32}$$

Some numerical schemes can only ensure this condition for a limited range of parameters. A numerical scheme robustness is measured by the range of values that the parameters may take that ensure the stability condition in equation 32.

2.2.3 mass conservation

Furthermore, all outward fluxes in one control volume should become inward fluxes, in the same exact proportion, in the neighbouring volumes. If the outward/inward symmetric fluxes aren't exactly the same, then the numerical scheme gains or looses mass over time, and is considered "not conservative". Thus, to ensure that the scheme is conservative, it is required that:

$$\forall i \quad A_{i+1} - B_i + C_{i-1} = 0, \tag{33}$$

as long as the B coefficients don't contain any sources nor sink terms. Equation 33 states the following:

What P_{i+1} gains from P_i (A_{i+1}) and what P_{i-1} gains from P_i (C_{i-1}) is equal to what P_i looses (B_i).

2.2.4 numerical diffusion

Once the continuous equation of advection-diffusion are discretized with finite-difference solvers, errors are introduced. A particular expression of these errors is that the finite-difference solver will always show a certain degree of spurious diffusion along the advective direction, which is called numerical diffusion. Numerical diffusion is maximum when the diffusive term is zeroed and only advection occurs. The mecanism goes that any particular tracer that doesn't undergo diffusive processes, but is mainly advected by the tracer, should follow a steady path and the particule of tracer should never grow in size nor dilute. However, whenever a finite-difference solver for the advection equation is applied, the particulate tracer present in a control volume is always split, at each time increment, between the remaining part in the control volume and the outgoing part to the neighbouring control volumes. This mechanism, which is illustrated in figure 3, creates numerical diffusion and is the main reason modelers often try to implement higher-order schemes that exhibit less numerical diffusion.

2.3 Implicit schemes

We can adapt the time forward equation into a time backward equation by simply considering that all outgoing and ingoing fluxes occur at $t+\Delta t$, thus yielding

$$-A_i\,P_{i-1}(t+\Delta t)+(1+B_i)P_i(t+\Delta t)-C_i\,P_{i+1}(t+\Delta t)=P_i(t). \quad (34)$$

This system of N equations, where N is the number of cells in the domain, is harder to solve because we have $N+2$ unknowns. Boundary conditions give the other two missing equations. The time forward scheme was simpler because it was a one equation to one unknown problem. The time backward scheme describes a tridiagonal matrix system of the type

$$M\,\underline{x}=T,$$

where M is the tridiagonal matrix, $\underline{x}$ is the unknown and T is the independent term. Such systems can be solved resorting to the Gauss-Seidel elimination method or, more efficiently, to the Thomas algorithm. The gain is that the implicit method, applied to equation 2, is unconditionally stable, although it can still violate positivity should the conditions of equation 31 not be met.

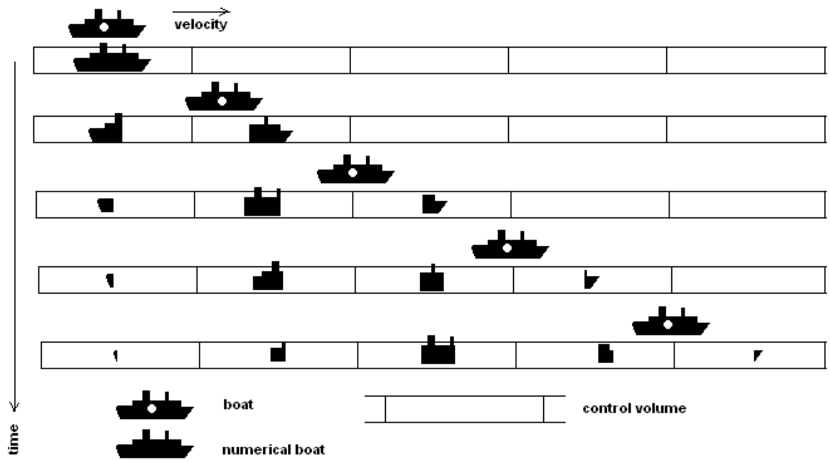

Fig. 3: Advected boat by a uniform current along time (from top to bottom). The dotted boat represents a realistic advection. The all-black boat represents advection by an upwind time-forward scheme in a unidimensional grid. The all-black boat falls in pieces, after each time iteration, as a consequence of numerical diffusion. Thus, finite-difference solvers aren't adequate to model pure advection. However, they work adequately to model advection and diffusion.

2.3.1 Thomas algorithm

The Thomas algorithm is used to solve tridiagonal systems described by

$$a_i x_{i-1} + b_i x_i + c_i x_{i+1} = d_i \quad \forall i \in \{1,\, 2 \ldots n\}\,,$$

where x_0 and x_{n+1} is the reference solution for the boundary conditions. In matricial form, we can establish that

$$\begin{bmatrix} 1 & & & & & & 0 \\ a_1 & b_1 & c_1 & & & & \\ & a_2 & b_2 & c_2 & & & \\ & & a_3 & b_3 & c_3 & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & a_n & b_n & c_n \\ 0 & & & & & & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \\ x_{n+1} \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \cdot \\ \cdot \\ d_n \\ d_{n+1} \end{bmatrix};$$

which suitably simplifies to

$$\begin{bmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & a_3 & b_3 & \cdot & \\ & & \cdot & \cdot & c_{n-1} \\ 0 & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 - a_1 d_0 \\ d_2 \\ \cdot \\ \cdot \\ d_n - c_n d_{n+1} \end{bmatrix} = \begin{bmatrix} d'_1 \\ d'_2 \\ \cdot \\ \cdot \\ d'_n \end{bmatrix},$$

after noting that $d_0 = x_0$ and $d_{n+1} = x_{n+1}$.

Thus, any non-trivial boundary conditions must go into the independent term in d'_1 and d'_n. The algorithm follows two sweeps, one forward and one backwards. First, the forward sweep:

$$Q_i = \begin{cases} \dfrac{c_1}{b_1}, & i = 1, \\ \dfrac{c_i}{b_i - Q_{i-1} a_i}, & i = 2,\, 3, \ldots,\, n-1\,; \end{cases}$$

$$R_i = \begin{cases} \dfrac{d'_1}{b_1}, & i = 1, \\ \dfrac{d'_i - R_{i-1} a_i}{b_i - Q_{i-1} a_i}, & i = 2,\, 3, \ldots,\, n; \end{cases}$$

which yields the following system:

$$\begin{bmatrix} 1 & Q_1 & & & 0 \\ & 1 & Q_2 & & \\ & & \cdot & \cdot & \\ & & & 1 & Q_{n-1} \\ 0 & & & & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ \cdot \\ \cdot \\ R_n \end{bmatrix}.$$

Then, the backward sweep is applied,

$$x_i = \begin{cases} R_i\,, & i = n \\ R_i - Q_i\, x_{i+1}\,, & i = n-1,\, n-2, \ldots, 1\,. \end{cases}$$

2.4 Boundary conditions

When solving the discretized equations (30) or (34) in for a finite domain with $i \in \{1, 2, \cdots, N\}$, necessity urges to impose a boundary condition for P_0 and P_{N+1}, as well as to establish the advective fluxes in coefficients A_1 and C_N. Imposed boundary conditions are also known as Dirichlet boundary conditions and can be described as

$$P(t) = f(t),$$

at the boundaries, where f is an imposed solution for P. A particular case is the null-value condition,

$$P(0) = 0 = P(N+1).$$

Gradient dependent boundary conditions are also known as Neumann boundary conditions and are best described by

$$\nabla P(t) = g(t),$$

at the boundaries, where g is an imposed solution for ∇P. A particular case is the null-gradient condition,

$$\begin{array}{lcl} P(0) & = & P(1) \\ P(N+1) & = & P(N). \end{array}$$

In a closed system (with walls at the boundary), null-gradient is adequate to simulate correctly diffusion, and nullyfying advection in coefficients A_1 and C_N is adequate to simulate correctly advection. In an open system however, the correct value of P must be known at the boundaries in order to correctly simulate advection-diffusion. In a uniform steady flow, admitting that the initial value of P is null outside the domain where the flow is incoming, the null-value condition at the open boundaries allows to correctly simulate advection, but diffusion will always be overestimated at the boundaries.

A more interesting way to look at the open-boundaries is to build a balance of the fluxes that cross the finite volume's faces. Each flux gathers a physical meaning, such as inwards advection or outwards diffusion. Depending on the modeler's boundary conditions, the modeler must choose to keep or to nullify outward fluxes, or to keep or nullify or impose inward fluxes.

2.5 Initial conditions

In order to completely define a problem, besides equation (2), we also need to define its boundary conditions and its initial conditions. In the time backward or time forward scheme, one initial condition suffices.

$$\begin{aligned} f'(t) &\approx \frac{f\,(t+\Delta t) - f\,(t)}{\Delta t} \\ \Leftrightarrow f\,(t+\Delta t) &\approx f\,(t) + \Delta t\, f'(t). \end{aligned}$$

$$\begin{aligned} f'(t+\Delta t) &\approx \frac{f\,(t+\Delta t) - f\,(t)}{\Delta t} \\ \Leftrightarrow f'(t+\Delta t)\,\Delta t - f\,(t+\Delta t) &\approx f\,(t)\,. \end{aligned}$$

For both previous methods, in order to compute $f\,(m\,\Delta t)$ for any m , $f((m-1)\ \Delta t)$ is required. But with the midpoint method, two initial conditions are required as,

$$\begin{aligned} f'\,(t+\Delta t) &\approx \frac{f\,(t+2\,\Delta t) - f\,(t)}{2\,\Delta t} \\ \Leftrightarrow f\,(t+2\,\Delta t) &\approx 2\,\Delta t\, f'\,(t+\Delta t) + f\,(t)\,, \end{aligned}$$

thus, in order to compute $f\,(m\,\Delta t)$ for any m, $f\,((m-2)\ \Delta t)$ and $f\,((m-1)\ \Delta t)$ are both required. Another issue with the midpoint method is that consecutive time steps may come decoupled if the two initial conditions are unsimilar. To re-enable coupling, use of filters such as Robert-Asselin filters may be required [5], at the expense of some precision in the solution,

$$f^*(t) = f(t) + \alpha\,(f\,(t+\Delta t) - 2\,f\,(t) + f\,(t-\Delta t))\,,$$

where α typically ranges the 1%.

3 The problem

We want to model the transport of a sourceless and sinkless tracer in a square-shaped duct and subjected to a uniform one-dimensional flow. The advection-diffusion model of the tracer is mathematically described by equation (2). Also, we want to model the system using a finite volume approach. The tracer is injected at initial time at some position inside the channel. The boundary condition in this problem are two-fold: i) the duct is open at both ends and inward and outward fluxes of the tracer are allowed; ii) the

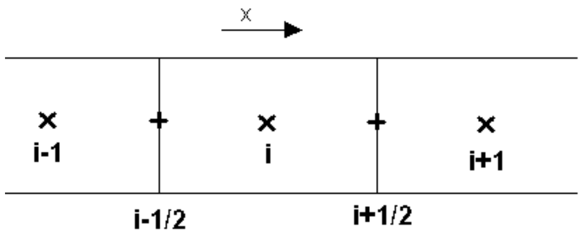

Fig. 4: Unidimensional uniform mesh.

duct contains a grid at each end that filters the tracer while letting pass the fluid. Boundary condition i is called open boundary condition and boundary condition ii is equivalent to a fully closed boundary condition. The particularity of the open boundary condition is that we must impose a tracer concentration of $P_{bL}(t)$ on the left end of the duct, and a tracer concentration of $P_{bR}(t)$ on the right end, be it Dirichelet or Neumann conditions. On the other hand, the main interest of the closed boundary condition is to test if the numerical methods conserve the tracer's total mass inside the domain as expected.

3.1 The mesh to the problem

Given that the fluid flow is uniform and along a closed square-shaped pipeline, the best way to describe the geometry of the system is by cutting N uniform slices of volume area V, where the leftmost slice is indexed 1 and the rightmost slice is indexed N. This sliced square-shaped pipeline defines our mesh of uniform cells of length Δx, width Δy and height Δz such that $\Delta x\,\Delta y\,\Delta z = V$. A one-dimensional projection of the mesh is illustrated in figure 4.

4 The numerical models

Considering u, Δx and Δt constants throughout the problem, let us define $Cr \equiv \frac{u\,\Delta t}{\Delta x}$, $Dif \equiv \frac{K\,\Delta t}{\Delta x^2}$, $CrL \equiv \frac{\|Cr\|+Cr}{2}$, $CrR \equiv \frac{\|Cr\|-Cr}{2}$. As a note, Cr is known in the literature as the Courant number and $Pec \equiv \frac{\|Cr\|}{Dif}$ is known as the Péclet number. They are both adimensional numbers and are relevant to characterize the positivity and stability of the numerical schemes. In particular, it can be seen from inequation 31 that a Péclet number below or equal to 2 is necessary to avoid negative oscillations in the central difference scheme.

$$CrL = \begin{cases} Cr & \text{if } u \geq 0 \\ 0 & \text{if } u < 0 \end{cases},$$

$$CrR = \begin{cases} 0 & \text{if } u \geq 0 \\ -Cr & \text{if } u < 0 \end{cases}.$$

Let us define the concentration of the tracer imposed at the boundaries by:

$$P_{bL} = \begin{cases} f_L & \text{if Dirichelet} \\ P_1 & \text{if Neumann} \end{cases},$$

$$P_{bR} = \begin{cases} f_R & \text{if Dirichelet} \\ P_N & \text{if Neumann} \end{cases}.$$

The stability and positivity criteria introduced below are the simultaneous combination of inequations 31, 32 and 33 applied to this case.

4.1 Upwind explicit

$$\begin{aligned} P_i(t+\Delta t) = &(CrR + Dif)\ P_{i+1}(t) \\ &+ (1 - CrR - CrL - 2\,Dif)\ P_i(t) \\ &+ (CrL + Dif)\ P_{i-1}(t), \end{aligned}$$

stability criteria:

$$0 \leq \|Cr\| + 2\,Dif \leq 1.$$

closed boundary condition:

$$\begin{aligned} P_1(t+\Delta t) = &(CrR + Dif)\ P_2(t) \\ &+ (1 - \cancel{\mathrm{CrR}} - CrL - \cancel{2}\,Dif)\ P_1(t) \\ &+ (\cancel{\mathrm{CrL}} + \cancel{\mathrm{Dif}})\ P_{bL}(t), \end{aligned}$$

$$\begin{aligned} P_N(t+\Delta t) = &(\cancel{\mathrm{CrR}} + \cancel{\mathrm{Dif}})\ P_{bR}(t) \\ &+ (1 - CrR - \cancel{\mathrm{CrL}} - \cancel{2}\,Dif)\ P_N(t) \\ &+ (CrL + Dif)\ P_{N-1}(t), \end{aligned}$$

open boundary condition:

$$\begin{aligned} P_1(t+\Delta t) = &(CrR + Dif)\ P_2(t) \\ &+ (1 - CrR - CrL - 2\,Dif)\ P_1(t) \\ &+ (CrL + Dif)\ P_{bL}(t), \end{aligned}$$

$$
\begin{aligned}
P_N(t+\Delta t) = {} & (CrR + Dif)\ P_{bR}(t) \\
& + (1 - CrR - CrL - 2\,Dif)\ P_N(t) \\
& + (CrL + Dif)\ P_{N-1}(t).
\end{aligned}
$$

4.2 Central differences explicit

$$
\begin{aligned}
P_i(t+\Delta t) = {} & (-Cr/2 + Dif)\ P_{i+1}(t) \\
& + (1 - Cr/2 + Cr/2 - 2\,Dif)\ P_i(t) \\
& + (Cr/2 + Dif)\ P_{i-1}(t),
\end{aligned}
$$

stability and positivity criteria:

$$0 \leq \|Cr\| \leq 2\,Dif \leq 1.$$

closed boundary condition:

$$
\begin{aligned}
P_1(t+\Delta t) = {} & (-Cr/2 + Dif)\ P_2(t) \\
& + (1 - Cr/2 + \cancel{\mathrm{Cr}/2} - \cancel{2}\,Dif)\ P_1(t) \\
& + (\cancel{\mathrm{Cr}/2} + \cancel{\mathrm{Dif}})\ P_{bL}(t),
\end{aligned}
$$

$$
\begin{aligned}
P_N(t+\Delta t) = {} & (-\cancel{\mathrm{Cr}/2} + \cancel{\mathrm{Dif}})\ P_{bR}(t) \\
& + (1 - \cancel{\mathrm{Cr}/2} + Cr/2 - \cancel{2}\,Dif)\ P_N(t) \\
& + (Cr/2 + Dif)\ P_{N-1}(t),
\end{aligned}
$$

open boundary condition:

$$
\begin{aligned}
P_1(t+\Delta t) = {} & (-Cr/2 + Dif)\ P_2(t) \\
& + (1 - Cr/2 + Cr/2 - 2\,Dif)\ P_1(t) \\
& + (Cr/2 + Dif)\ P_{bL}(t),
\end{aligned}
$$

$$
\begin{aligned}
P_N(t+\Delta t) = {} & (-Cr/2 + Dif)\ P_{bR}(t) \\
& + (1 - Cr/2 + Cr/2 - 2\,Dif)\ P_N(t) \\
& + (Cr/2 + Dif)\ P_{N-1}(t),
\end{aligned}
$$

4.3 Upwind implicit

$$
\begin{aligned}
&(-CrR - Dif)\; P_{i+1}(t + \Delta t) \\
&+ (1 + CrR + CrL + 2\,Dif)\; P_i(t + \Delta t) \\
&+ (-CrL - Dif)\; P_{i-1}(t + \Delta t) \\
&= P_i(t),
\end{aligned}
$$

unconditionally stable:

$$0 \le \|Cr\| + 2\,Dif \le +\infty.$$

closed boundary condition:

$$
\begin{aligned}
&(-CrR - Dif)\; P_2(t + \Delta t) \\
&+ (1 + \cancel{\mathrm{CrR}} + CrL + \cancel{2}\,Dif)\; P_1(t + \Delta t) \\
&= P_1(t) - (-\cancel{\mathrm{CrL}} - \cancel{\mathrm{Dif}})\; P_{bL}(t),
\end{aligned}
$$

$$
\begin{aligned}
&+ (1 + CrR + \cancel{\mathrm{CrL}} + \cancel{2}\,Dif)\; P_N(t + \Delta t) \\
&+ (-CrL - Dif)\; P_{N-1}(t + \Delta t) \\
&= P_N(t) - (-\cancel{\mathrm{CrR}} - \cancel{\mathrm{Dif}})\; P_{bR}(t),
\end{aligned}
$$

open boundary condition:

$$
\begin{aligned}
&(-CrR - Dif)\; P_2(t + \Delta t) \\
&+ (1 + CrR + CrL + 2\,Dif)\; P_1(t + \Delta t) \\
&= P_1(t) - (-CrL - Dif)\; P_{bL}(t),
\end{aligned}
$$

$$
\begin{aligned}
&(1 + CrR + CrL + 2\,Dif)\; P_N(t + \Delta t) \\
&+ (-CrL - Dif)\; P_{N-1}(t + \Delta t) \\
&= P_N(t) - (-CrR - Dif)\; P_{bR}(t),
\end{aligned}
$$

4.4 Central differences implicit

$$
\begin{aligned}
&(Cr/2 - Dif)\; P_{i+1}(t + \Delta t) \\
&+ (1 + Cr/2 - Cr/2 + 2\,Dif)\; P_i(t + \Delta t) \\
&+ (-Cr/2 - Dif)\; P_{i-1}(t + \Delta t) \\
&= P_i(t),
\end{aligned}
$$

positivity criteria:

$$0 \leq \|Cr\| \leq 2\,Dif \leq +\infty.$$

closed boundary condition:

$$\begin{aligned}
&(Cr/2 - Dif)\; P_2(t+\Delta t)\\
&+ (1 + Cr/2 - \cancel{\mathrm{Cr}/2} + \cancel{2}\,Dif)\; P_1(t+\Delta t)\\
&= P_1(t) - (-\cancel{\mathrm{Cr}/2} - \cancel{\mathrm{Dif}})\; P_{bL}(t),
\end{aligned}$$

$$\begin{aligned}
&+ (1 + \cancel{\mathrm{Cr}/2} - Cr/2 + \cancel{2}\,Dif)\; P_N(t+\Delta t)\\
&+ (-Cr/2 - Dif)\; P_{N-1}(t+\Delta t)\\
&= P_N(t) - (\cancel{\mathrm{Cr}/2} - \cancel{\mathrm{Dif}})\; P_{bR}(t),
\end{aligned}$$

open boundary condition:

$$\begin{aligned}
&(Cr/2 - Dif)\; P_2(t+\Delta t)\\
&+ (1 + Cr/2 - Cr/2 + 2\,Dif)\; P_1(t+\Delta t)\\
&= P_1(t) - (-Cr/2 - Dif)\; P_{bL}(t),
\end{aligned}$$

$$\begin{aligned}
&(1 + Cr/2 - Cr/2 + 2\,Dif)\; P_N(t+\Delta t)\\
&+ (-Cr/2 - Dif)\; P_{N-1}(t+\Delta t)\\
&= P_N(t) - (Cr/2 - Dif)\; P_{bR}(t).
\end{aligned}$$

4.5 Hybrid schemes

$$\alpha = \begin{cases} 1 & \text{upwind} \\ 0 & \text{central differences} \end{cases},$$

$$\beta = \begin{cases} 0 & \text{explicit} \\ 1 & \text{implicit} \end{cases},$$

$$\text{Crank-Nicholson}: \quad \begin{cases} \beta = 0.5 & \text{i.e. hybrid} \\ \alpha = 0 & \text{i.e. central differences} \end{cases},$$

$$\begin{aligned}
&\beta\,((1-\alpha)\,Cr/2 - \alpha CrR - Dif)\; P_{i+1}(t+\Delta t)\\
&+ (1 + \beta\,(\alpha\,(CrR + CrL) + 2\,Dif))\; P_i(t+\Delta t)\\
&+ \beta\,(-\,(1-\alpha)\,Cr/2 - \alpha CrL - Dif)\; P_{i-1}(t+\Delta t)\\
&=\\
&(1-\beta)\,(-\,(1-\alpha)\,Cr/2 + \alpha CrR + Dif)\; P_{i+1}(t)\\
&+ \big(1 - (1-\beta)\,\big(\alpha\,(CrR + CrL) + 2\,Dif\big)\big)\; P_i(t)\\
&+ (1-\beta)\,((1-\alpha)\,Cr/2 + \alpha CrL + Dif)\; P_{i-1}(t),
\end{aligned}$$

stability and positivity criteria:

$$\forall_{\alpha,\beta\in[0\;\;1]},\; -(1-\alpha)\,\|Cr\| + 2\,Dif \begin{cases} \geq 0 \\ \leq \frac{2}{\beta} \\ \leq -\|Cr\| + \frac{1}{1-\beta} \end{cases},$$

closed boundary condition:

$$\begin{aligned}
&\beta\left((1-\alpha)\,Cr/2 - \alpha CrR - Dif\right)\; P_2(t+\Delta t)\\
&+\left(1+\beta\left((1-\alpha)\,(Cr/2 - \cancel{\mathrm{Cr}/2}) + \alpha\,(\cancel{\mathrm{CrR}} + CrL) + \cancel{2}\,Dif\right)\right)\; P_1(t+\Delta t)\\
&=\\
&(1-\beta)\,(-(1-\alpha)\,Cr/2 + \alpha CrR + Dif)\; P_2(t)\\
&+\Big(1-(1-\beta)\,\big((1-\alpha)\,(-Cr/2 + \cancel{\mathrm{Cr}/2}) + \alpha\,(\cancel{\mathrm{CrR}} + CrL) + \cancel{2}\,Dif\big)\Big)\; P_1(t)\\
&+(1-\beta)\,((1-\alpha)\,\cancel{\mathrm{Cr}/2} + \alpha\cancel{\mathrm{CrL}} + \cancel{\mathrm{Dif}})\; P_{bL}(t)\\
&-\beta\,(-(1-\alpha)\,\cancel{\mathrm{Cr}/2} - \alpha\cancel{\mathrm{CrL}} - \cancel{\mathrm{Dif}})\; P_{bL}(t),
\end{aligned}$$

$$\begin{aligned}
&\left(1+\beta\left((1-\alpha)\,(\cancel{\mathrm{Cr}/2} - Cr/2) + \alpha\,(CrR + \cancel{\mathrm{CrL}}) + \cancel{2}\,Dif\right)\right)\; P_N(t+\Delta t)\\
&+\beta\,(-(1-\alpha)\,Cr/2 - \alpha CrL - Dif)\; P_{N-1}(t+\Delta t)\\
&=\\
&-\beta\,((1-\alpha)\,\cancel{\mathrm{Cr}/2} - \alpha\cancel{\mathrm{CrR}} - \cancel{\mathrm{Dif}})\; P_{bR}(t)\\
&+(1-\beta)\,(-(1-\alpha)\,\cancel{\mathrm{Cr}/2} + \alpha\cancel{\mathrm{CrR}} + \cancel{\mathrm{Dif}})\; P_{bR}(t)\\
&+\Big(1-(1-\beta)\,\big((1-\alpha)\,(-\cancel{\mathrm{Cr}/2} + Cr/2) + \alpha\,(CrR + \cancel{\mathrm{CrL}}) + \cancel{2}\,Dif\big)\Big)\; P_N(t)\\
&+(1-\beta)\,((1-\alpha)\,Cr/2 + \alpha CrL + Dif)\; P_{N-1}(t),
\end{aligned}$$

open boundary condition:

$$\begin{aligned}
&\beta\left((1-\alpha)\,Cr/2 - \alpha CrR - Dif\right)\; P_2(t+\Delta t)\\
&+\left(1+\beta\left(\alpha\,(CrR + CrL) + 2\,Dif\right)\right)\; P_1(t+\Delta t)\\
&=\\
&(1-\beta)\,(-(1-\alpha)\,Cr/2 + \alpha CrR + Dif)\; P_2(t)\\
&+\big(1-(1-\beta)\,\big(\alpha\,(CrR + CrL) + 2\,Dif\big)\big)\; P_1(t)\\
&+(1-\beta)\,((1-\alpha)\,Cr/2 + \alpha CrL + Dif)\; P_{bL}(t)\\
&-\beta\,(-(1-\alpha)\,Cr/2 - \alpha CrL - Dif)\; P_{bL}(t),
\end{aligned}$$

$$
\begin{aligned}
&(1+\beta\,(\alpha\,(CrR+CrL)+2\,Dif))\;P_N(t+\Delta t)\\
&+\beta\,(-\,(1-\alpha)\,Cr/2-\alpha CrL-Dif)\;P_{N+1}(t+\Delta t)\\
&=\\
&-\,\beta\,((1-\alpha)\,Cr/2-\alpha CrR-Dif)\;P_{bR}(t)\\
&(1-\beta)\,(-\,(1-\alpha)\,Cr/2+\alpha CrR+Dif)\;P_{bR}(t)\\
&+\big(1-(1-\beta)\,\big(\alpha\,(CrR+CrL)+2\,Dif\big)\big)\;P_N(t)\\
&+(1-\beta)\,((1-\alpha)\,Cr/2+\alpha CrL+Dif)\;P_{N+1}(t).
\end{aligned}
$$

4.6 Characterization of the numerical methods

The implicit schemes are the most robust as they are unconditionally stable, but they also produce the most numerical diffusion. The explicit schemes are the less robust. The central differences method may, additionally, violate positivity (either with the explicit, either with the implicit schemes) but has a higher precision (second-order) and produces less numerical diffusion. The upwind method never violates positivity, but shows a lot of numerical diffusion and has less precision than the central differences. The hybrid method is more robust than the explicit one, is more resilient to positivity violation than central differences, and produces less numerical diffusion than the implicit methods.

5 Adding sources and sinks to build an ecological model

The previous section limited to describe numerical methods suitable for advective and diffusive processes only. However, generic tracers may also have source/sink and/or growth/decay terms. Besides involving more terms in the equation, the stability conditions of the numerical scheme also change. This section proposes to describe mathematically a mass conservative ecological system composed by three variables of state describing a predador-prey-nutrients system, that are transported in a 1D-channel. The stability conditions of the numerical implementation are also considered.

5.1 the mathematical model of a prey-predator-nutrients system

The mathematical model will be described in two parts. The first part will contain only the ecological modeling terms, acting in a single control-volume (without any advective or diffusive process). The second part will add the advective and diffusive terms to the equations.

The equations describing the ecological model composed by three variables, predator, prey and nutrients (let them be zooplankton, phytoplankton and nutrients) evolving in a single control-volume devoid of any advective or diffusive processes are

$$\begin{cases} \frac{dP_{\mathrm{zoo}}}{dt} & = e_{\mathrm{h}}\, k_{\mathrm{h}}\, P_{\mathrm{phy}}\, P_{\mathrm{zoo}} - k_{\mathrm{mz}}\, P_{\mathrm{zoo}}, \\ \\ \frac{dP_{\mathrm{phy}}}{dt} & = k_{\mathrm{g}}\, P_{\mathrm{nut}}\, P_{\mathrm{phy}} - k_{\mathrm{h}}\, P_{\mathrm{zoo}}\, P_{\mathrm{phy}}, \\ \\ \frac{dP_{\mathrm{nut}}}{dt} & = -k_{\mathrm{g}}\, P_{\mathrm{phy}}\, P_{\mathrm{nut}} + (1-e_{\mathrm{h}})\, k_{\mathrm{h}}\, P_{\mathrm{phy}}\, P_{\mathrm{zoo}} + k_{\mathrm{mz}}\, P_{\mathrm{zoo}}, \end{cases} \tag{35}$$

where P_{zoo}, P_{phy} and P_{nut} are, respectively, the concentrations of zooplankton, phytoplankton and nutrients. k_{h} is the assimilation rate of phytoplankton by zooplankton, k_{g} is the grazing rate of phytoplankton and k_{mz} is the mortality rate of zookplankton. e_{h} is the zooplankton's efficiency rate of assimilation of phytoplankton (phytoplankton grazes with 100 percent of efficiency). All the remains of dead zooplankton and dead phytoplankton are wholly decomposed into nutrients by bacteria. Growth terms come with a plus sign, and depletion terms come with a minus sign. The system of equations 35 is globally conservative, as the sum of the initial masses of the three properties is preserved over time.

By substituting the ecological model in equations 35 in the source and sink terms of equation 8, the full ecological model system of differential equations writes

$$\begin{cases} \frac{\partial P_{\mathrm{zoo}}}{\partial t} & = -\nabla \cdot (\mathbf{v}\, P_{\mathrm{zoo}}) + \nabla \cdot (K_{\mathrm{zoo}}\, \nabla P_{\mathrm{zoo}}) + \lambda_{\mathrm{zoo}}\, P_{\mathrm{zoo}}, \\ \frac{\partial P_{\mathrm{phy}}}{\partial t} & = -\nabla \cdot (\mathbf{v}\, P_{\mathrm{phy}}) + \nabla \cdot (K_{\mathrm{phy}}\, \nabla P_{\mathrm{phy}}) + \lambda_{\mathrm{phy}}\, P_{\mathrm{phy}}, \\ \frac{\partial P_{\mathrm{nut}}}{\partial t} & = -\nabla \cdot (\mathbf{v}\, P_{\mathrm{nut}}) + \nabla \cdot (K_{\mathrm{nut}}\, \nabla P_{\mathrm{nut}}) - \lambda_{\mathrm{nut}}\, P_{\mathrm{nut}} + \mathrm{Sc}_{\mathrm{nut}}, \end{cases} \tag{36}$$

where

$$\begin{cases} \lambda_{\mathrm{zoo}} & \equiv e_{\mathrm{h}}\, k_{\mathrm{h}}\, P_{\mathrm{phy}} - k_{\mathrm{mz}}, \\ \lambda_{\mathrm{phy}} & \equiv k_{\mathrm{g}}\, P_{\mathrm{nut}} - k_{\mathrm{h}}\, P_{\mathrm{zoo}}, \\ \lambda_{\mathrm{nut}} & \equiv k_{\mathrm{g}}\, P_{\mathrm{phy}}, \\ \mathrm{Sc}_{\mathrm{nut}} & \equiv \left((1-e_{\mathrm{h}})\, k_{\mathrm{h}}\, P_{\mathrm{phy}} + k_{\mathrm{mz}}\right)\, P_{\mathrm{zoo}}. \end{cases} \tag{37}$$

In particular, λ_{zoo} and λ_{phy} are growth or decay coefficients of the zooplankton and phytoplankton, respectively. Their sign can change over time. λ_{nut} is always positive and is a decay coefficient in the nutrients equation. Finally, $\mathrm{Sc}_{\mathrm{nut}}$ is a source term in the nutrients equation. Growth and source terms tend to stabilize the numerical solution of equations 36, using the explicit method, whereas decay terms may reduce the stability interval when using the explicit method.

5.2 The discretization of ecological properties

Discretizing equations 36 follows the same approach that led to the discretized equations 30, except that this time growth/decay and source/sink terms appear in the equations i.e.

$$\begin{aligned} P_i(t+\Delta t) = A_i\,P_{i-1}(t) + (1 - B_i + \Delta t\,\lambda_i)P_i(t) + C_i\,P_{i+1}(t) \\ +\Delta t\;(\mathrm{Sc}_i - \mathrm{Sk}_i)\,. \end{aligned} \tag{38}$$

To meet the positivity (eq. 31) and stability (eq. 32) criteria, the following constraints must be followed for equations 38,

$$\forall i \begin{cases} 0 & \leq \quad A_i \leq 1, \\ 0 & \leq \quad B_i - \Delta t\,\lambda_i \leq 1, \\ 0 & \leq \quad C_i \leq 1. \end{cases} \tag{39}$$

However, even if the constraints in equations 39 are met, the numerical solution may still undershoot (i.e. yield a negative concentration) if the sink term is predominant over the source term, i.e.

$$\text{If} \quad \exists i, \quad \mathrm{Sc}_i < \mathrm{Sk}_i \quad \text{then the model may undershoot.} \tag{40}$$

In the ecological model presented in equations 36 and 37 there are no sink terms. Thus, the model will never undershoot on account of criteria given in equation 40. However the stability criteria provided in equations 39 still require to be met in order to obtain a stable solution with the explicit method. The implicit method still remains unconditionally stable (though non-positive schemes, like the central-differences method, may yield too much spurious negative oscillations).

The implicit method is straightforward from equation 38, by analogy with equations 30 and 34:

$$\begin{aligned} -A_i\,P_{i-1}(t+\Delta t) + (1 + B_i - \Delta t\,\lambda_i)P_i(t+\Delta t) - C_i\,P_{i+1}(t+\Delta t) \\ = P_i(t) + \Delta t\;(\mathrm{Sc}_i - \mathrm{Sk}_i)\,. \end{aligned} \tag{41}$$

References

[1] Chapra, S. C. Surface water-quality modeling. Mcgraw-hill Series In Water Resources And Environmental Engineering (1997).

[2] Collado, F. J. Reynolds transport theorem for a two-phase flow. Applied Physics Letters 90, 2 (2007), 024101+.

[3] Ferziger, J. H., and Perić, M. Computational methods for fluid dynamics. Springer, Berlin, 1999.

[4] Fletcher, C. A. J., and Srinivas, K. Computational Techniques for Fluid Dynamics 1. Springer, 1991.

[5] Kantha, L. H., and Clayson, C. A. Numerical models of oceans and oceanic processes. Academic Press, 2000.