Chapter 6

Finite Difference Methods for Hyperbolic Systems

6.1. Discretization of time and space

6.1.1. *Discretization for one-dimensional problems*

Hyperbolic systems of conservation laws can seldom be solved analytically when real-world problems are dealt with. This is because the initial and boundary conditions and the geometry of most real-world problems cannot be described analytically in a simple way. Most engineering applications therefore involve the solution of approximations of the governing equations. The solution of the approximate equations, that is easier to derive, can be expected to be "reasonably" close to that of the real equations provided that a number of criteria are satisfied (see Appendix B for basic notions in numerical analysis). The operation by which the original equations are approximated is called discretization. In the finite difference approach, presented in this chapter, time and space are treated identically in the discretization process. Other approaches exist, as shown in Chapters 7 and 8.

The discretization process consists of transforming the originally continuous time and space coordinates into discrete variables (see Figure 6.1). A set of points (called the computational points) is defined in time and space by the modeler. The solution is calculated at these points. The governing equations are approximated using the differences between the known and unknown values of the computational solution at the predefined points. Denoting by U_i^n the solution at the computational point x_i at the computational time t^n, the following difference may be seen as a

Wave Propagation in Fluids: Models and Numerical Techniques, Second Edition Vincent Guinot
© 2010 ISTE Ltd. Published 2010 by ISTE Ltd.

"good" approximation of the derivative $\partial U / \partial x$ over the interval $[x_i, x_{i+1}]$ at the time t^n.

$$D_x = \frac{U_{i+1}^n - U_i^n}{x_{i+1} - x_i} \qquad [6.1]$$

Equation [6.1] is not the only possible approximation for the derivative $\partial U / \partial x$. Many alternative formulations may be proposed. The accuracy of the numerical solution depends on the accuracy with which the governing equations are approximated (see section B.1 in Appendix B). Approximation methods where the differences between the point values are used to estimate the derivatives are referred to as finite difference methods.

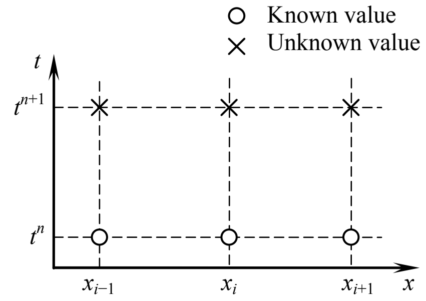

Figure 6.1. *Discretization of time and space in the one-dimensional case*

In what follows, the distance between the points i and $i + 1$ is denoted by $\Delta x_{i+1/2}$. It is often referred to as the grid spacing, or cell width. The difference between two successive computational times t^n and t^{n+1}, also called the computational time step, is usually denoted by Δt. The time t^n is usually referred to as "the time level n".

6.1.2. *Multidimensional discretization*

Several options are available for the discretization of multidimensional equations. Two main types of grid are distinguished (Figure 6.2):

1) Structured grids. The computational points of a structured grid are arranged along rows and columns. A computational point is located using two indices (i, j) in a two-dimensional space. The index i generally indicates the location of the point along the x-coordinate, while the index j indicates its location along the y-coordinate. A computational point is located using a triple index (i, j, k) in a three-

dimensional space. The value of the variable U at the point (i, j, k) at the time level n is usually denoted by $U^{n}_{i,j,k}$. The family of structured grid is divided into two subcategories:

1.1) Cartesian, or rectangular grids. The computational points form the intersection between two families of orthogonal, straight lines (Figure 6.2a).

1.2) Curvilinear grids. The computational points are located at the intersections between two families of curves that do not necessarily intersect at straight angles (Figure 6.2b). While curvilinear grids allow real-world geometries to be described more accurately, they require more work from the modeler in that it is the modeler's responsibility to define the curvature of the lines and the locations at which the computational points are to be distributed along the lines.

2) Unstructured grids. Such grids do not use the arrangement in lines and columns used by structured grids. The computational points are placed and linked to each other at the modeler's convenience, mostly based on the geometrical constraints imposed by the problem to be solved (Figure 6.2c).

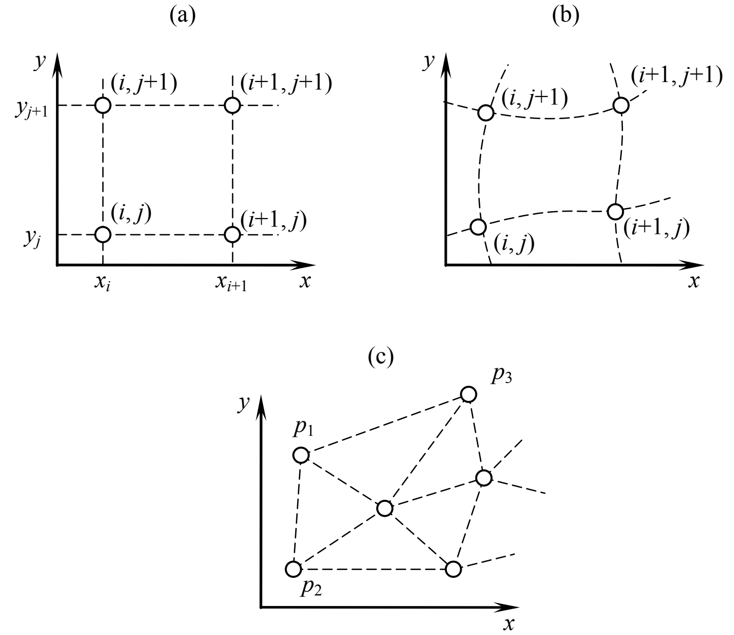

Figure 6.2. *The various types of grid: structured, Cartesian (a), structured, curvilinear (b), unstructured (c). Sketch for a two-dimensional problem*

6.1.3. *Explicit schemes, implicit schemes*

All scalar hyperbolic laws and hyperbolic systems of conservation laws contain derivatives with respect to time and space and a source term. The time derivative is that of the conserved variable, while the space derivative is that of the flux or the conserved variable, depending on whether the conservation form or the non-conservation form is being discretized.

When the space derivative and the source term are discretized using only the known values of U at the time level n, the discretization is said to be explicit because the unknown solution at the time level $n + 1$ can be computed directly from the known values at the time level n.

When the space derivative and/or the source term are discretized using the unknown values of U at the time level $n + 1$, the estimate of U^{n+1} is based on expressions that involve U^{n+1} itself. In such a case, the discretization is said to be implicit because U^{n+1} is defined as a function of itself.

Explicit schemes are easily programmed and maintained. However, they are subjected to a so-called stability constraint that yields a restriction in the range of permissible computational time steps (see section B.2 in Appendix B). The computational time step must remain smaller than a threshold value $\Delta t_{\max}$ above which the numerical solution becomes unstable. Explicit methods usually lead to small computational time steps, thus increasing the number of calculations and leading to time-consuming simulations.

As shown in section B.2 of Appendix B, the so-called Courant number Cr (also called CFL number) is a key factor to the stability of explicit numerical methods. The Courant number expresses the ratio of the area covered by a wave of speed λ during the interval Δt to the total area of the grid cells. Most explicit methods are stable only when the absolute value of the Courant number is smaller than one.

Implicit schemes are not subjected to stability constraints. Most of them are indeed unconditionally stable, regardless of the Courant number. This makes them popular for industrial use because they allow larger computational time steps to be used, thus restricting the number of time steps and leading to faster applications. It should be remembered however that a fast method is not necessarily an accurate one.

6.2. The method of characteristics (MOC)

6.2.1. *MOC for scalar hyperbolic laws*

6.2.1.1. *Principle of the method*

The conservation form [1.1] of a scalar hyperbolic conservation law is recalled:

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}=S$$

As shown in Chapter 1, equation [1.1] may be rewritten in the characteristic form [1.27], recalled here:

$$\frac{\mathrm{d}U}{\mathrm{d}t}=S' \qquad \text{for } \frac{\mathrm{d}x}{\mathrm{d}t}=\lambda$$

where λ and S' are given as in equations [1.21], recalled here:

$$\left.\begin{aligned}\lambda &= \frac{\partial F}{\partial U}\\ S' &= S-\left(\frac{\partial F}{\partial x}\right)_{U=\mathrm{Const}}\end{aligned}\right\}$$

Equation [1.27] is solved numerically by discretizing time and space as illustrated in Figure 6.3. The calculation of the unknown value U_i^{n+1} at the time level $n + 1$ is detailed hereafter.

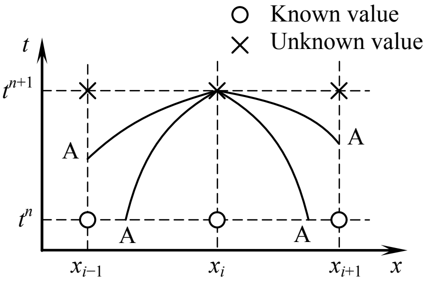

Figure 6.3. *Definition sketch for the MOC*

The principle of the MOC is the following. Assume first that the shape of the characteristic passing at (x_i, t^{n+1}) can be determined in the phase space. The intersection between the characteristic and the grid lines, also called the foot of the characteristic, is denoted by A. The various possible locations of the foot of the characteristic are illustrated in Figure 6.3. Integrating the characteristic form [1.27] between A and the point (x_i, t^{n+1}) leads to:

$$U_i^{n+1} = U_\mathrm{A} + \int_\mathrm{A}^{(x_i,t^{n+1})} S'(x,t)\mathrm{d}t, \qquad \frac{\mathrm{d}x}{\mathrm{d}t} = \lambda \tag{6.2}$$

The unknown value U_i^{n+1} can be calculated provided that the following two terms are estimated: (i) the value U_A of U at the foot of the characteristic and (ii) the integral of the source term between A and M. Classically, polynomial interpolation formulae are used. The most widely used are the first- and second-order and the Hermite polynomial-based interpolation introduced by Holly and Preissmann [HOL 77]. Interpolation issues are dealt with in the next three sections.

6.2.1.2. *Interpolation at the foot of the characteristic: first-order formula*

The value U_A of U at the foot of the characteristic is interpolated from the known values at the computational points. Consider first the case of the linear, or first-order interpolation, where U is estimated using the equation of a straight line. Two cases may be distinguished (Figure 6.4).

1) Point A is located between the points (i–1, n) and (i, n). The following first-order interpolation formula is used:

$$U_\mathrm{A} = \frac{x_i - x_\mathrm{A}}{\Delta x_{i-1/2}} U_{i-1}^n + \frac{x_\mathrm{A} - x_{i-1/2}}{\Delta x_{i-1/2}} U_i^n \tag{6.3}$$

Defining the cell size and the average wave speed as:

$$\left.\begin{aligned} \Delta x_{i-1/2} &= x_i - x_{i-1} \\ \lambda_i^{n+1/2} &= \frac{x_i - x_\mathrm{A}}{\Delta t} \end{aligned}\right\} \tag{6.4}$$

leads to:

$$U_\mathrm{A} = \frac{\lambda_i^{n+1/2}\Delta t}{\Delta x_{i+1/2}} U_{i-1}^n + \frac{\Delta x_{i-1/2} - \lambda_i^{n+1/2}\Delta t}{\Delta x_{i-1/2}} U_i^n \tag{6.5}$$

where $\lambda_i^{n+1/2}$ represents the average wave speed of the characteristic that passes at the point $(i, n+1)$ between the time levels n and $n+1$. The expression [6.5] is simplified by introducing the Courant number, that represents the fraction of the cell covered by the characteristic over the time step Δt:

$$\mathrm{Cr} = \frac{\lambda_i^{n+1/2}\Delta t}{\Delta x_{i-1/2}} \qquad [6.6]$$

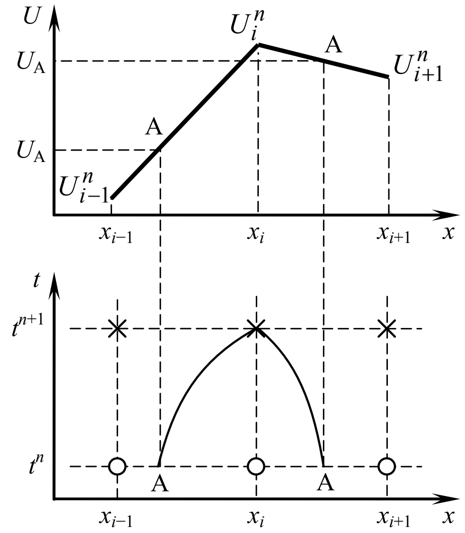

Figure 6.4. *Interpolation at the foot of the characteristic. Sketch in the physical space (top) and in the phase space (bottom) for a first-order linear interpolation*

Note that Cr is between 0 and 1 when A is located between the points $(i-1, n)$ and (i, n). Substituting definition [6.6] into equation [6.5] leads to the following expression:

$$U_A = \mathrm{Cr}U_{i-1}^n + (1-\mathrm{Cr})U_i^n, \qquad 0 \le \mathrm{Cr} \le 1 \qquad [6.7]$$

2) The point A is located between the points (i, n) and $(i+1, n)$. It is easy to check that the following formula applies:

$$U_A = -\mathrm{Cr}U_{i+1}^n + (1+\mathrm{Cr})U_i^n, \qquad -1 \le \mathrm{Cr} \le 0 \qquad [6.8]$$

where the Courant number Cr is defined as:

$$\mathrm{Cr} = \frac{\lambda_i^{n+1/2}\Delta t}{\Delta x_{i+1/2}} \qquad [6.9]$$

3) Point A is located between $(i-1, n)$ and $(i-1, n+1)$. By definition, the Courant number is larger than unity. The following interpolation is used:

$$U_\mathrm{A} = \frac{t^{n+1} - t_\mathrm{A}}{\Delta t} U_{i-1}^n + \frac{t_\mathrm{A} - t^n}{\Delta t} U_{i-1}^{n+1} \qquad [6.10]$$

By definition of the average wave speed and the Courant number, we have:

$$\left.\begin{aligned} t^{n+1} - t_\mathrm{A} &= \frac{\Delta x_{i-1/2}}{\lambda_i^{n+1/2}} = \frac{\Delta t}{\mathrm{Cr}} \\ t_\mathrm{A} - t^n &= \Delta - t\,\frac{\Delta x_{i-1/2}}{\lambda_i^{n+1/2}} = \frac{\mathrm{Cr}-1}{\mathrm{Cr}}\Delta t \end{aligned}\right\} \qquad [6.11]$$

Substituting equations [6.11] into equation [6.10] leads to:

$$U_\mathrm{A} = \frac{1}{\mathrm{Cr}} U_{i-1}^n + \frac{\mathrm{Cr}-1}{\mathrm{Cr}} U_{i-1}^{n+1}, \qquad \mathrm{Cr} \ge 1 \qquad [6.12]$$

4) The point A is located between $(i+1, n)$ and $(i+1, n+1)$. By definition, the Courant number is smaller than -1. The following interpolation formula is used:

$$U_\mathrm{A} = -\frac{1}{\mathrm{Cr}} U_{i+1}^n + \frac{\mathrm{Cr}+1}{\mathrm{Cr}} U_{i+1}^{n+1}, \qquad \mathrm{Cr} \le -1 \qquad [6.13]$$

The formulae above can be summarized as follows:

$$U_\mathrm{A} = \begin{cases} \dfrac{1}{\mathrm{Cr}} U_{i-1}^n + \dfrac{\mathrm{Cr}-1}{\mathrm{Cr}} U_{i-1}^{n+1} & \text{if } \mathrm{Cr} \ge 1 \\ \mathrm{Cr}\,U_{i-1}^n + (1-\mathrm{Cr})U_i^n & \text{if } 0 \le \mathrm{Cr} \le 1 \\ (1+\mathrm{Cr})U_i^n - \mathrm{Cr}\,U_{i+1}^n & \text{if } -1 \le \mathrm{Cr} \le 0 \\ -\dfrac{1}{\mathrm{Cr}} U_{i+1}^n + \dfrac{\mathrm{Cr}+1}{\mathrm{Cr}} U_{i+1}^{n+1} & \text{if } \mathrm{Cr} \le -1 \end{cases} \qquad [6.14]$$

where the Courant number is defined as:

$$\mathrm{Cr} = \begin{cases} \dfrac{\lambda_i^{n+1/2}\Delta t}{\Delta x_{i-1/2}} & \text{if } \lambda_i^{n+1/2} \geq 0 \\ \dfrac{\lambda_i^{n+1/2}\Delta t}{\Delta x_{i+1/2}} & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \qquad [6.15]$$

Note that the first and fourth formulae [6.14] are implicit because they involve the unknown values U_{i-1}^{n+1} and U_{i+1}^{n+1}. In contrast, the second and third formulae [6.14] are explicit because they allow U_i^{n+1} to be calculated directly from the known values at the time level *n*.

The calculation of the Courant number Cr requires that $\lambda_i^{n+1/2}$ be estimated. A number of possible formulae are proposed hereafter (the list is non-exhaustive):

$$\begin{aligned} &\lambda_i^{n+1/2} \approx \lambda_i^n && \text{(explicit without interpolation)} \\ &\lambda_i^{n+1/2} \approx \frac{\Delta x_{i+1/2}\lambda_{i-1}^n + \Delta x_{i-1/2}\lambda_{i+1}^n}{\Delta x_{i-1/2} + \Delta x_{i+1/2}} && \text{(explicit with interpolation)} \\ &\lambda_i^{n+1/2} \approx \lambda_i^{n+1} && \text{(implicit without interpolation)} \\ &\lambda_i^{n+1/2} \approx \frac{\Delta x_{i+1/2}\lambda_{i-1}^{n+1} + \Delta x_{i-1/2}\lambda_{i+1}^{n+1}}{\Delta x_{i-1/2} + \Delta x_{i+1/2}} && \text{(implicit with interpolation)} \end{aligned} \qquad [6.16]$$

The last two formulae are said to be implicit because they involve the unknown values of λ at the time level $n + 1$. Since λ is a function of *U* in the general case, the determination of $\lambda_i^{n+1/2}$ in an implicit method is iterative.

Note that when the Courant number is an integer, the foot of the characteristic is located at a grid point. The interpolation formulae are exact and the numerical solution is the analytical solution. Also note that formulae [6.14] are defined over different intervals that coincide only at Cr = – 1, Cr = 0 and Cr = + 1. Both the first and second formula [6.14] give the analytical solution U_{i-1}^n when Cr tends to 1 by a lower or upper value; both the second and third formulae [6.14] yield the analytical solution U_i^n when the Courant number tends to zero; eventually, both the third and fourth formulae [6.14] give the solution U_{i+1}^n when the Courant number tends to – 1 by a lower or upper value. This continuous behavior of the interpolation formula is referred to as a “continuous switch” in the literature.

The linear interpolation procedures introduce a numerical effect known as "numerical diffusion" (see section B.1.3). Numerical diffusion may be eliminated using higher-order interpolations that use more computational points. A second-order interpolation technique is presented in the next section.

6.2.1.3. *Interpolation at the foot of the characteristic: second-order formula*

This section deals with a second-order interpolation method for U_A (Figure 6.5):

$$U_A = U_i^n + (x_A - x_i)^2 a_i + (x_A - x_i) b_i \qquad [6.17]$$

Equation [6.17] is valid when point A is located between the points $(i-1, n)$ and $(i+1, n)$. It is applicable for Courant numbers ranging from -1 to $+1$. The coefficients a_i and b_i are determined by stating that function [6.17] should coincide with the value U_{i-1}^n for $x_A = x_{i-1}$ and with U_{i+1}^n for $x_A = x_{i+1}$:

$$\left.\begin{aligned} U_{i-1}^n &= U_i^n + \Delta x_{i-1/2}^2 a_i - \Delta x_{i-1/2} b_i \\ U_{i+1}^n &= U_i^n + \Delta x_{i+1/2}^2 a_i + \Delta x_{i+1/2} b_i \end{aligned}\right\} \qquad [6.18]$$

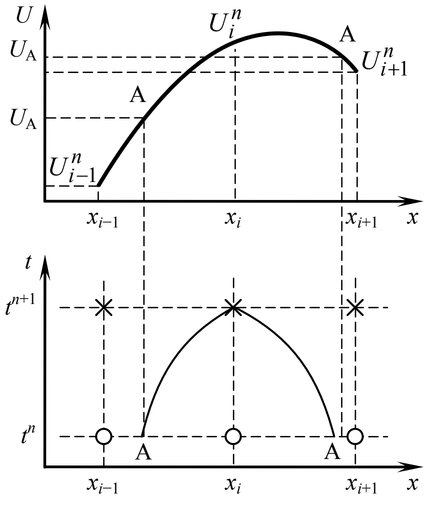

Figure 6.5. *Second-order interpolation. Sketch in the physical space (top) and in the phase space (bottom)*

Solving equations [6.18] for a_i and b_i leads to:

$$\left.\begin{aligned} a_i &= \frac{\Delta x_{i+1/2}U_{i-1}^n - (\Delta x_{i-1/2} + \Delta x_{i+1/2})U_i^n + \Delta x_{i-1/2}U_{i+1}^n}{(\Delta x_{i-1/2} + \Delta x_{i+1/2})\Delta x_{i-1/2}\Delta x_{i+1/2}} \\ b_i &= \frac{-\Delta x_{i+1/2}^2 U_{i-1}^n + (\Delta x_{i+1/2}^2 - \Delta x_{i-1/2}^2)U_i^n + \Delta x_{i+1/2}^2 U_{i+1}^n}{(\Delta x_{i-1/2} + \Delta x_{i+1/2})\Delta x_{i-1/2}\Delta x_{i+1/2}} \end{aligned}\right\} \qquad [6.19]$$

When the grid is regular, $\Delta x_{i-1/2} = \Delta x_{i+1/2} = \Delta x$ and equations [6.19] simplify into:

$$\left.\begin{aligned} a_i &= \frac{U_{i-1}^n - 2U_i^n + U_{i+1}^n}{2\Delta x^2} \\ b_i &= \frac{U_{i+1}^n - U_{i-1}^n}{2\Delta x} \end{aligned}\right\} \qquad [6.20]$$

Substituting equation [6.20] into equation [6.17], using the relationship $x_A = x_i - \text{Cr}\,\Delta x$ leads to:

$$U_A = (\text{Cr}+1)\frac{\text{Cr}}{2}U_{i-1}^n + (1-\text{Cr}^2)U_i^n + (\text{Cr}-1)\frac{\text{Cr}}{2}U_{i+1}^n \qquad [6.21]$$

Note that for Cr = – 1, Cr = 0 or Cr = + 1, the foot of the characteristic coincides with a grid point and the interpolation is exact. The analytical solution is obtained.

6.2.1.4. *Estimation of the source term*

The integral of the source term must be estimated in equation [6.2]. Four options are proposed hereafter:

$$\int_A^{(x_i,t^{n+1})} S'(x,t)\mathrm{d}t \approx \begin{cases} S'(U_i^n)\Delta t \\ S'(U_A)\Delta t \\ [S'(U_A) + S'(U_i^{n+1/2})]\Delta t/2 \\ S'(U_i^{n+1})\Delta t \end{cases} \qquad [6.22]$$

The first two options are explicit because they use only known values of U, whether they are point values at the computational points or interpolated values. The remaining two expressions are implicit because they require the knowledge of the unknown value U_i^{n+1}. The solution is determined iteratively in the general case.

6.2.1.5. *Treatment of boundary conditions*

Assume first that the wave speed $\lambda_1^{n+1/2}$ at the first computational point (i.e. at the left-hand boundary of the domain) is positive. Then the characteristic $\mathrm{d}x/\mathrm{d}t = \lambda$ enters the computational domain and the value of U at $i = 1$ cannot be computed from the internal points. It must be supplied in the form of a boundary condition (see Figure 6.6). Applying equations [6.14] with $i = 1$ in the case of a positive Courant number gives:

$$U_2^{n+1} = \begin{cases} \dfrac{1}{\mathrm{Cr}} U_1^n + \dfrac{\mathrm{Cr}-1}{\mathrm{Cr}} U_1^{n+1} & \text{if } \mathrm{Cr} \geq 1 \\ \mathrm{Cr}U_1^n + (1-\mathrm{Cr})U_2^n & \text{if } 0 \leq \mathrm{Cr} \leq 1 \end{cases} \qquad [6.23]$$

Conversely, if the wave speed $\lambda_M^{n+1/2}$ at the right-hand boundary of the computational domain is negative, the value of U at $i = M$ must be supplied in the form of a boundary condition. Equations [6.14] lead to:

$$U_{M-1}^{n+1} = \begin{cases} (1+\mathrm{Cr})U_{M-1}^n - \mathrm{Cr}U_M^n & \text{if } -1 \leq \mathrm{Cr} \leq 0 \\ -\dfrac{1}{\mathrm{Cr}} U_M^n + \dfrac{\mathrm{Cr}+1}{\mathrm{Cr}} U_M^{n+1} & \text{if } \mathrm{Cr} \leq -1 \end{cases} \qquad [6.24]$$

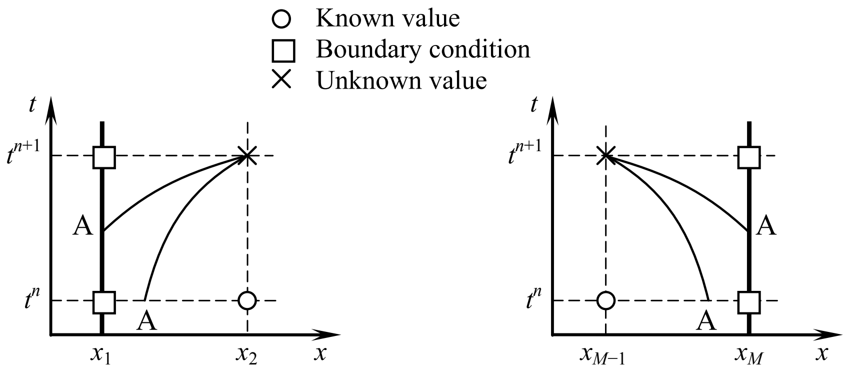

Figure 6.6. *Treatment of boundary conditions. The boundary is indicated by a bold line*

6.2.2. *The MOC for hyperbolic systems of conservation laws*

6.2.2.1. *Principle of the method*

The generalization of the MOC to hyperbolic systems of conservation laws is best known as the Courant-Isaacson-Rees (CIR) scheme [COU 52]. The CIR scheme can be viewed as a direct application of the scalar MOC to the Riemann invariants (Figure 6.7). The purpose is to solve the conservation form [2.2], recalled hereafter:

$$\frac{\partial \mathrm{U}}{\partial t}+\frac{\partial \mathrm{F}}{\partial x}=\mathrm{S}$$

As shown in Chapter 2, equation [2.2] can be written in characteristic form as in equation [2.25], recalled hereafter:

$$\frac{\mathrm{d}W_p}{\mathrm{d}t}=S_p^{"} \qquad \text{for}\,\frac{\mathrm{d}x}{\mathrm{d}t}=\lambda^{(p)} \qquad \forall p=1,2,\ldots,m$$

where S" = K^{-1}S' and the wave speeds $\lambda^{(p)}$ are the eigenvalues of the Jacobian matrix A of F with respect to U and K is the matrix formed by the eigenvectors of A (see section 2.1.3 for the details of the developments).

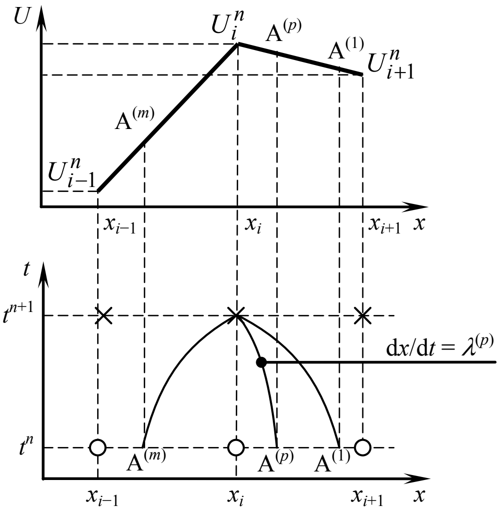

Figure 6.7. *Interpolation at the feet of the characteristics. Definition sketch in the physical space (top) and in the phase space (bottom) for a linear interpolation*

The foot of the *p*th characteristic passing at the point (*i*, *n* + 1) is denoted by A$^{(p)}$ (Figure 6.7). Integrating relationship [2.25] along the characteristic line A$^{(p)}$M leads to:

$$(W_p)_i^{n+1} = (W_p)_{\mathrm{A}^{(p)}} + \int_{\mathrm{A}^{(p)}}^{(x_i,t^{n+1})} S_p''(x,t)\mathrm{d}t, \frac{\mathrm{d}x}{\mathrm{d}t} = \lambda^{(p)}, p = 1,2,\ldots,m \qquad [6.25]$$

Note that this formulation is identical to equation [6.2] for the scalar problem. Two options are available for the calculation of the *p*th Riemann invariant.

1) Interpolate W_p at the foot of the *p*th characteristic from the known values at the computational points:

$$(W_p)_{\mathrm{A}^{(p)}} = \begin{cases} \dfrac{1}{\mathrm{Cr}^{(p)}}(W_p)_{i-1}^n + \dfrac{\mathrm{Cr}^{(p)}-1}{\mathrm{Cr}^{(p)}}(W_p)_{i-1}^{n+1} & \text{if } \mathrm{Cr}^{(p)} \geq 1 \\ \mathrm{Cr}^{(p)}(W_p)_{i-1}^n + (1-\mathrm{Cr}^{(p)})(W_p)_i^{n+1} & \text{if } 0 \leq \mathrm{Cr}^{(p)} \leq 1 \\ (1+\mathrm{Cr}^{(p)})(W_p)_i^{n+1} - \mathrm{Cr}^{(p)}(W_p)_{i+1}^n & \text{if } -1 \leq \mathrm{Cr}^{(p)} \leq 0 \\ -\dfrac{1}{\mathrm{Cr}^{(p)}}(W_p)_{i+1}^n + \dfrac{\mathrm{Cr}^{(p)}+1}{\mathrm{Cr}^{(p)}}(W_p)_{i+1}^{n+1} & \text{if } \mathrm{Cr}^{(p)} \leq -1 \end{cases} \qquad [6.26]$$

where $(W_p)_i^n$ is the value of the *p*th Riemann invariant calculated from the value of U_i^n and Cr$^{(p)}$ is the Courant number for the *p*th wave:

$$\mathrm{Cr}^{(p)} = \begin{cases} \dfrac{\lambda^{(p)}{}_i^{n+1/2}\Delta t}{\Delta x_{i-1/2}} & \text{if } \lambda^{(p)}{}_i^{n+1/2} \geq 0 \\ \dfrac{\lambda^{(p)}{}_i^{n+1/2}\Delta t}{\Delta x_{i+1/2}} & \text{if } \lambda^{(p)}{}_i^{n+1/2} \leq 0 \end{cases} \qquad [6.27]$$

2) Calculate W_p from the interpolated value of U at the foot of the characteristic:

$$(W_p)_{\mathrm{A}^{(p)}} = W_p(\mathrm{U}_{\mathrm{A}^{(p)}}) \qquad [6.28]$$

where $\mathrm{U}_{\mathrm{A}^{(p)}}$ is the estimate of U at the foot $\mathrm{A}^{(p)}$ of the *p*th characteristic:

$$\mathrm{U}_{\mathrm{A}^{(p)}} = \begin{cases} \dfrac{1}{\mathrm{Cr}^{(p)}}\mathrm{U}_{i-1}^{n} + \dfrac{\mathrm{Cr}^{(p)}-1}{\mathrm{Cr}^{(p)}}\mathrm{U}_{i-1}^{n+1} & \text{if } \mathrm{Cr}^{(p)} \geq 1 \\ \mathrm{Cr}^{(p)}\mathrm{U}_{i-1}^{n} + (1-\mathrm{Cr}^{(p)})\mathrm{U}_{i}^{n+1} & \text{if } 0 \leq \mathrm{Cr}^{(p)} \leq 1 \\ (1+\mathrm{Cr}^{(p)})\mathrm{U}_{i}^{n+1} - \mathrm{Cr}^{(p)}\mathrm{U}_{i+1}^{n} & \text{if } -1 \leq \mathrm{Cr}^{(p)} \leq 0 \\ -\dfrac{1}{\mathrm{Cr}^{(p)}}\mathrm{U}_{i+1}^{n} + \dfrac{\mathrm{Cr}^{(p)}+1}{\mathrm{Cr}^{(p)}}\mathrm{U}_{i+1}^{n+1} & \text{if } \mathrm{Cr}^{(p)} \leq -1 \end{cases} \qquad [6.29]$$

When the governing equations are linear (as is the case with the water hammer equations), the Riemann invariants are linear combinations of the components of U and both options give the same result.

The wave speeds and the source term may be estimated using any of the formulae [6.16] and [6.22]. The number of boundary conditions to be provided at each boundary of the domain is equal to the number of characteristics that enter the domain.

The practical implementation of the CIR scheme for Courant numbers smaller than one is straightforward. Indeed, the case |Cr| ≤ 1 uses only the second and third formulae [6.26] or [6.29], where only known values of U are required for the calculation of the solution. In contrast, when the absolute value of the Courant number is larger than one, an implicit formulation must be used, leading to a dependence between the unknown value of U at two neighboring points. The nonlinear dependence between the wave speed and the conserved variable U makes the procedure time-consuming in the general case.

6.2.2.2. *Application example: the water hammer equations*

This section deals with the application of the CIR scheme to the water hammer equations. Expression [2.79] is recalled:

$$\left.\begin{aligned} &\frac{\mathrm{d}p}{\mathrm{d}t} - \frac{\rho c}{A}\frac{dQ}{\mathrm{d}t} = \left(k|u|u + \rho g A \sin\theta\right)\frac{c}{A} && \text{for } \frac{\mathrm{d}x}{\mathrm{d}t} = -c \\ &\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\rho c}{A}\frac{dQ}{\mathrm{d}t} = \left(-k|u|u - \rho g A \sin\theta\right)\frac{c}{A} && \text{for } \frac{\mathrm{d}x}{\mathrm{d}t} = c \end{aligned}\right\}$$

For the sake of clarity, the source term is assumed to be zero hereafter. Equations [2.79] simplify into equations [2.83] where S" is set to zero:

$$\left.\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\left(p-\frac{\rho c}{A}Q\right)=0 \qquad &&\text{for}\,\frac{\mathrm{d}x}{\mathrm{d}t}=-c \\ &\frac{\mathrm{d}}{\mathrm{d}t}\left(p-\frac{\rho c}{A}Q\right)=0 \qquad &&\text{for}\,\frac{\mathrm{d}x}{\mathrm{d}t}=c \end{aligned}\right\} \qquad [6.30]$$

Integrating equations [6.30] along the characteristics dx/dt = – c and dx/dt = + c yields the following algebraic system:

$$\left.\begin{aligned} p_i^{n+1}-\frac{\rho c}{A}Q_i^{n+1} &= p_{\mathrm{A}^{(1)}}-\frac{\rho c}{A}Q_{\mathrm{A}^{(1)}} \\ p_i^{n+1}+\frac{\rho c}{A}Q_i^{n+1} &= p_{\mathrm{A}^{(2)}}+\frac{\rho c}{A}Q_{\mathrm{A}^{(2)}} \end{aligned}\right\} \qquad [6.31]$$

where A$^{(1)}$ and A$^{(2)}$ are the feet of the first and second characteristic passing at (i, n + 1) respectively (Figure 6.8).

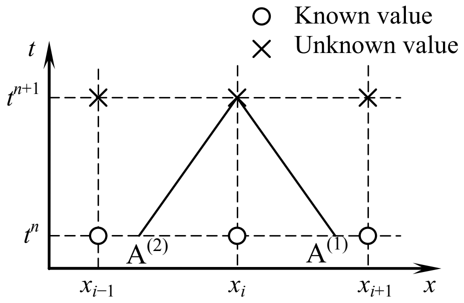

Figure 6.8. *Application of the CIR scheme to the water hammer equations*

Solving equations [6.31] for p and Q leads to:

$$\left.\begin{aligned} p_i^{n+1} &= \frac{p_{\mathrm{A}^{(1)}}+p_{\mathrm{A}^{(2)}}}{2}+\frac{\rho c}{2A}\left(Q_{\mathrm{A}^{(2)}}-Q_{\mathrm{A}^{(1)}}\right) \\ Q_i^{n+1} &= \frac{Q_{\mathrm{A}^{(1)}}+Q_{\mathrm{A}^{(2)}}}{2}\,p_{\mathrm{A}^{(2)}}+\frac{A}{2\rho c}\left(p_{\mathrm{A}^{(1)}}-p_{\mathrm{A}^{(2)}}\right) \end{aligned}\right\} \qquad [6.32]$$

The value of *p* and *Q* at the feet of the characteristics may be interpolated indifferently from U and W. It is easy to check that the linear dependence between U and W leads to identical formulations for [6.26] and [6.29]. In the case of a regular cell size Δx, the following formula is obtained:

$$\left.\begin{aligned} p_{\mathrm{A}^{(1)}} &= (1+\mathrm{Cr}^{(1)})p_i^n - \mathrm{Cr}^{(1)}p_{i+1}^n \\ Q_{\mathrm{A}^{(1)}} &= (1+\mathrm{Cr}^{(1)})Q_i^n - \mathrm{Cr}^{(1)}Q_{i+1}^n \\ p_{\mathrm{A}^{(2)}} &= \mathrm{Cr}^{(2)}p_{i-1}^n + (1-\mathrm{Cr}^{(2)})p_i^n \\ Q_{\mathrm{A}^{(2)}} &= \mathrm{Cr}^{(2)}Q_{i-1}^n + (1-\mathrm{Cr}^{(2)})Q_i^n \end{aligned}\right\} \qquad [6.33]$$

Substituting [6.33] into [6.32], noting that $\mathrm{Cr}^{(2)} = -\,\mathrm{Cr}^{(1)} = \mathrm{Cr}$, yields

$$\left.\begin{aligned} p_i^{n+1} &= \frac{\mathrm{Cr}\ p_{i-1}^n + 2(1-\mathrm{Cr})p_i^n + \mathrm{Cr}\ p_{i-1}^n}{2} + \mathrm{Cr}\,\frac{\rho c}{2A}(Q_{i-1}^n - Q_{i+1}^n) \\ Q_i^{n+1} &= \frac{\mathrm{Cr}\ Q_{i-1}^n + 2(1-\mathrm{Cr})Q_i^n + \mathrm{Cr}\ Q_{i-1}^n}{2} + \mathrm{Cr}\,\frac{A}{2\rho c}(p_{i-1}^n - p_{i+1}^n) \end{aligned}\right\} \qquad [6.34]$$

where Cr is given by

$$\mathrm{Cr} = -\mathrm{Cr}^{(1)} = \mathrm{Cr}^{(2)} = \frac{c\Delta t}{\Delta x} \qquad [6.35]$$

Many software packages for water hammer simulation use the fact that the speed of sound *c* is constant in pipes with homogenous material and geometrical properties. The time step and/or cell size are adjusted in such a way that the Courant number is equal to one over the computational domain. The need for an interpolation procedure is then eliminated and the solution is exact. Equations [6.34] simplify into:

$$\left.\begin{aligned} p_i^{n+1} &= \frac{p_{i-1}^n + p_{i-1}^n}{2} + \frac{\rho c}{2A}(Q_{i-1}^n - Q_{i+1}^n) \\ Q_i^{n+1} &= \frac{Q_{i-1}^n + Q_{i-1}^n}{2} + \frac{A}{2\rho c}(p_{i-1}^n - p_{i+1}^n) \end{aligned}\right\} \qquad [6.36]$$

The reduced number of computations makes the particular application [6.36] approximately three times as fast as the general form [6.34].

6.2.3. *Application examples*

6.2.3.1. *The linear advection equation*

A major drawback of the MOC is that the equations are not solved in conservation form. Due to this, mass and/or momentum conservation may not always be guaranteed. The non-conservation form [1.48] of the linear advection equation, recalled hereafter, is solved:

$$\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} = 0$$

The flow velocity u is assumed to be uniform over the computational domain. The parameters of the problem are given in Table 6.1. Such parameters are typical for the advection of a contaminant in a river. The initial condition is a top hat function, the width of which is 5 km. The calculation is carried out over an irregular computational grid. The cell size is 1 km over the entire domain, except between x = 10 km and x = 14 km, where Δx = 4 km. Since u is uniform, the Courant number is not. It is equal to one all over the domain, except between x = 10 km and x = 14 km, where it is equal to 1/4.

Symbol	Meaning	Value
A	River cross-sectional area	1,000 m^2
C_i^0	Initial concentration	1 g/l for 1 km $\le x \le$ 5 km, 0 g/l otherwise
C_G	Concentration at the left-hand boundary	0 g/l
L	Length of the domain	30 km
u	Flow velocity	1 km/hr
Δt	Computational time step	1 h
Δx	Cell size	4 km for 10 km $\le x \le$14 km, 1 km otherwise

Table 6.1. *Physical and numerical parameters for the solution of the advection equation over an irregular grid*

The numerical solution computed by the first-order MOC is compared to the analytical solution in Figure 6.9. The concentration signal reaches the zone Δx = 4 km at t = 4 hr. Up to this time, the numerical solution is identical to the analytical solution because the concentration signal is transported with a Courant number equal to one. At t = 5 hr, the signal enters the zone Δx = 4 km. Owing to the interpolation between x_{11} = 10 km and x_{12} = 14 km, the solution is smoothed out and spreads artificially over the cell. Owing to the interpolation between the zero value at x_{11} and the non-zero value at x_{12}, the concentration at $x = x_{12}$ does not return to zero, even after an infinite time.

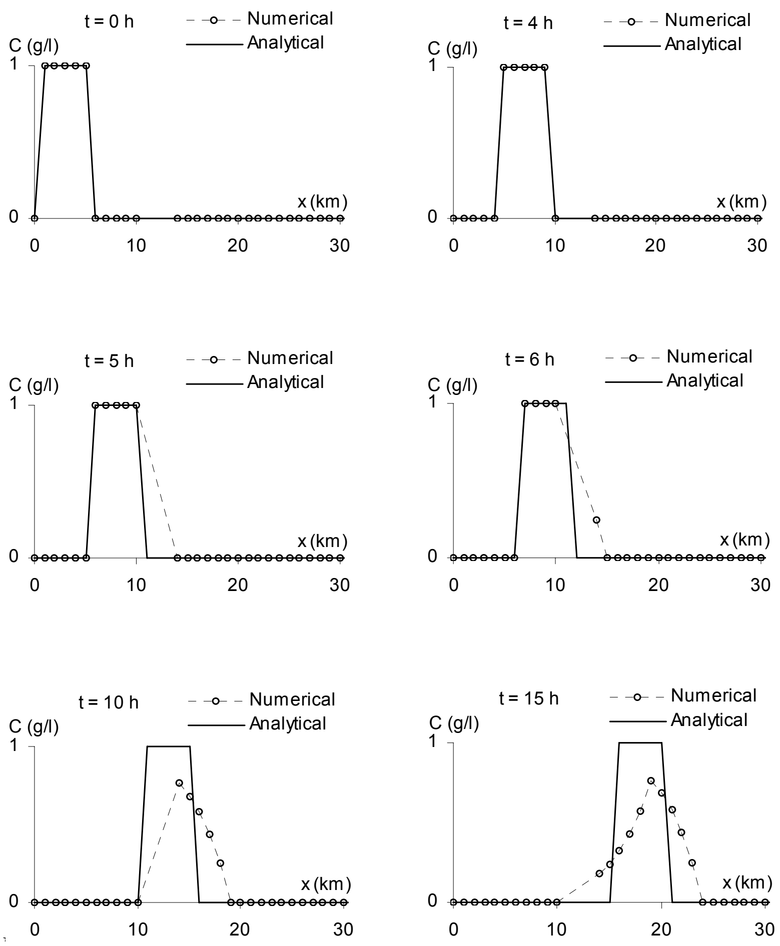

Figure 6.9. *Advection of a square concentration signal. Comparison between the analytical solution and the numerical solution obtained using the first-order MOC on an irregular grid*

The total mass of contaminant in the computational domain is plotted as a function of time in Figure 6.10. The mass is calculated as the integral of the concentration profile:

$$M_T(t^n) = \left[\sum_{i=1}^{M-1} (C_i^n + C_{i+1}^n)\Delta x_{i+1/2} + C_M^n \Delta x_{M-1/2}\right]\frac{A}{2} \qquad [6.37]$$

The total mass of contaminant is not constant, which indicates that the conservation properties of the numerical solution are violated. The artificial spreading of the contaminant front at $t = 5$ hr results in a simultaneous increase in the total mass of contaminant. When the contaminant leaves the zone $\Delta x = 4$ km, the total mass suddenly decreases below its initial value. The initial value is recovered asymptotically as the contaminant travels downstream. However, it is never totally recovered, even for infinite times.

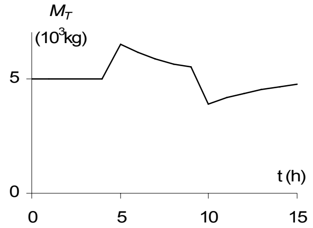

Figure 6.10. *Advection of a square concentration signal. Total mass of contaminant in the domain as computed by the first-order MOC on an irregular grid*

6.2.3.2. *The inviscid Burgers equation*

The characteristic form [1.68] of the inviscid Burgers equation is solved. A Riemann problem, the parameters of which are given in Table 6.2, is solved using the first-order MOC.

Symbol	Meaning	Value
L	Length of the domain	20 m
u_i^0	Initial speed	1 m/s for $x \le 5$ m, 0 m/s otherwise
Δt	Computational time step	1 s
Δx	Cell size	1 m

Table 6.2. *Physical and numerical parameters for the solution of the inviscid Burgers equation using the first-order MOC*

The following options are used for the estimate of λ:

$$\lambda_i^{n+1/2} = \begin{cases} u_i^n & \text{(option 1)} \\ (u_{i-1}^n + u_i^n)/2 & \text{(option 2)} \\ (u_{i-1}^n + u_{i+1}^n)/2 & \text{(option 3)} \end{cases} \qquad [6.38]$$

The numerical solution at $t = 20$ s is compared to the analytical solution for each of the three options in Figure 6.11. Note that the analytical solution is a shock that propagates at a speed given by the average value of the speeds on both sides of the discontinuity.

Option 1 gives a zero value for λ. The shock does not move.

Option 2 is based on the analytical formula for the shock speed. The numerical solution moves at the right speed, but the front is subjected to numerical diffusion. This is because the front moves at a speed $c_S = 0.5$ m/s, which corresponds to Cr = 1/2. As shown in section B.2.5, the phase portrait of the first-order MOC shows that numerical diffusion is maximum for Cr = 1/2, hence the smoothing in the neighborhood of the front.

Option 3 leads to an underestimated shock speed. The numerical profile is slower than the analytical profile.

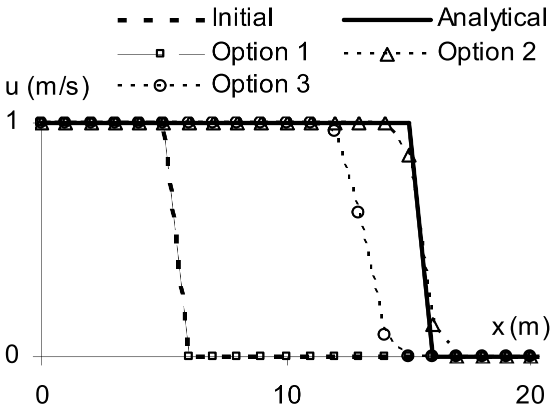

Figure 6.11. *Analytical solution and numerical profiles computed by the first-order MOC for the three calculation options [6.38]*

As shown by the two examples above, the MOC may lead to conservation problems when the Courant number is not uniform over the computational domain. This is because the MOC does not solve the governing equations in conservation form. This is one of the reasons why the method is seldom used in modern computational software packages (with the exception of the water hammer equations that are based on a constant wave speed).

6.3. Upwind schemes for scalar laws

6.3.1. *The explicit upwind scheme (non-conservative version)*

Upwind schemes aim to solve the conservation form [1.1] recalled here:

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} = S$$

The derivative of U with respect to time is estimated as:

$$\frac{\partial U}{\partial t} \approx \frac{U_i^{n+1} - U_i^n}{\Delta t} \tag{6.39}$$

The scheme is said to be "upwind" because the derivative of F with respect to space is estimated using the computational point located upstream of the point i (Figure 6.12).

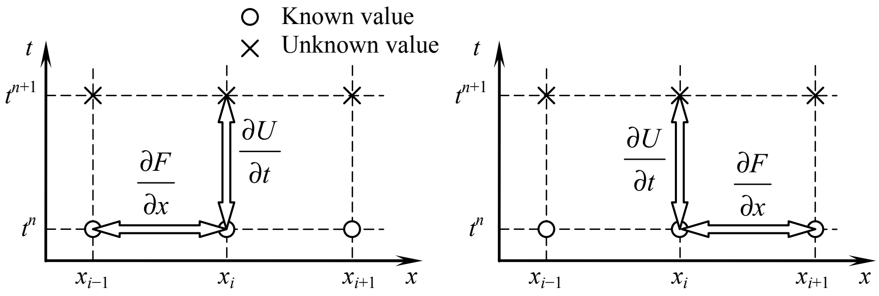
○ Known value
× Unknown value
t
t^{n+1}
t^n
$\frac{\partial F}{\partial x}$
$\frac{\partial U}{\partial t}$
x_{i-1}
x_i
x_{i+1}
x
t
t^{n+1}
t^n
$\frac{\partial U}{\partial t}$
$\frac{\partial F}{\partial x}$
x_{i-1}
x_i
x_{i+1}
x

Figure 6.12. *Definition sketch for the explicit upwind scheme. Sketch for a positive wave speed (left), for a negative wave speed (right)*

In the explicit approach, the flux F and the source term S are estimated using the known values of U:

$$\left.\begin{aligned} &\frac{\partial F}{\partial x} \approx \begin{cases} \dfrac{F_i^n - F_{i-1}^n}{\Delta x_{i-1/2}} & \text{if } \lambda_i^{n+1/2} \geq 0 \\ \dfrac{F_{i+1}^n - F_i^n}{\Delta x_{i+1/2}} & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \\ &S \approx S_i^n \end{aligned}\right\} \tag{6.40}$$

where $F_i^n = F(U_i^n)$ and $S_i^n = S(U_i^n)$. The numerical solution is stable if the absolute value of the Courant number is smaller than or equal to one. The Courant number is defined as in equation [6.15], recalled hereafter:

$$\mathrm{Cr} = \begin{cases} \dfrac{\lambda_i^{n+1/2}\Delta t}{\Delta x_{i-1/2}} & \text{if } \lambda_i^{n+1/2} \geq 0 \\ \dfrac{\lambda_i^{n+1/2}\Delta t}{\Delta x_{i+1/2}} & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases}$$

Substituting approximations [6.39–40] into equation [1.1] leads to:

$$U_i^{n+1} = \begin{cases} U_i^n + \dfrac{\Delta t}{\Delta x_{i-1/2}}(F_{i-1}^n - F_i^n) + \Delta t S_i^n & \text{if } \lambda_i^{n+1/2} \geq 0 \\ U_i^n + \dfrac{\Delta t}{\Delta x_{i+1/2}}(F_i^n - F_{i+1}^n) + \Delta t S_i^n & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \qquad [6.41]$$

In the particular case of the linear advection equation, $F = uC$. It is easy to check that equations [6.41] simplify into the same expression as the second and third equations [6.14]:

$$U_i^{n+1} = \begin{cases} \mathrm{Cr}U_{i-1}^n + (1-\mathrm{Cr})U_{i-1}^n & \text{if } \lambda_i^{n+1/2} \geq 0 \\ -\mathrm{Cr}U_{i+1}^n + (1+\mathrm{Cr})U_i^n & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \qquad [6.42]$$

In other words, the explicit upwind scheme is equivalent to the MOC when the absolute value of the Courant number is smaller than one. It is stable only when the Courant number lies within the range [– 1, + 1].

6.3.2. *The implicit upwind scheme (non-conservative version)*

In this scheme the time derivative is estimated as in equation [6.39], recalled here:

$$\frac{\partial U}{\partial t} \approx \frac{U_i^{n+1} - U_i^n}{\Delta t}$$

while the space derivative of the flux is estimated using the unknown values at the time level $n + 1$:

$$\left.\begin{aligned} \frac{\partial F}{\partial x} &\approx \begin{cases} \dfrac{F_i^{n+1} - F_{i-1}^{n+1}}{\Delta x_{i-1/2}} & \text{if } \lambda_i^{n+1/2} \geq 0 \\ \dfrac{F_{i+1}^{n+1} - F_i^{n+1}}{\Delta x_{i+1/2}} & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \\ S &\approx S_i^{n+1} \end{aligned}\right\} \qquad [6.43]$$

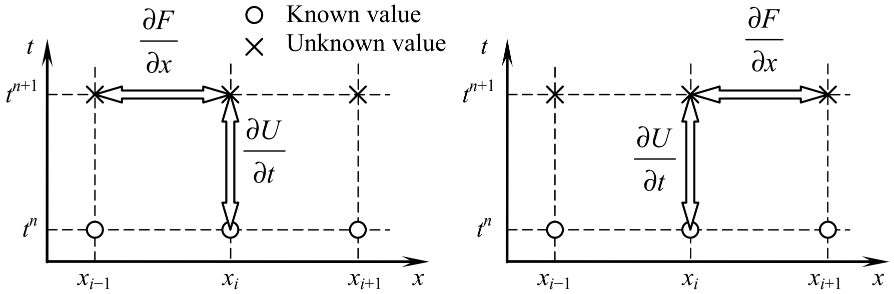
$\frac{\partial F}{\partial x}$
○ Known value
× Unknown value
t
t^{n+1}
$\frac{\partial U}{\partial t}$
t^n
x_{i-1}
x_i
x_{i+1}
x
t
t^{n+1}
$\frac{\partial F}{\partial x}$
$\frac{\partial U}{\partial t}$
t^n
x_{i-1}
x_i
x_{i+1}
x

Figure 6.13. *Definition sketch for the implicit upwind scheme. Sketch for a positive wave speed (left), for a negative wave speed (right)*

Substituting approximations [6.39] and [6.43] into equation [1.1] leads to:

$$U_i^{n+1} = \begin{cases} U_i^n + \dfrac{\Delta t}{\Delta x_{i-1/2}}(F_{i-1}^{n+1} - F_i^{n+1}) + S_i^n & \text{if } \lambda_i^{n+1/2} \geq 0 \\ U_i^n + \dfrac{\Delta t}{\Delta x_{i+1/2}}(F_i^{n+1} - F_{i+1}^{n+1}) + S_i^n & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \qquad [6.44]$$

Applying the implicit scheme to the particular case of the advection equation leads to the following formula that is not equivalent to the MOC:

$$U_i^{n+1} = \begin{cases} \dfrac{\mathrm{Cr}}{1+\mathrm{Cr}} U_{i-1}^{n+1} + \dfrac{1}{1+\mathrm{Cr}} U_i^n & \text{if } \lambda_i^{n+1/2} \geq 0 \\ \dfrac{-\mathrm{Cr}}{1-\mathrm{Cr}} U_{i+1}^{n+1} + \dfrac{1}{1-\mathrm{Cr}} U_i^n & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \qquad [6.45]$$

A consistency analysis (see section B.1) shows that the implicit upwind scheme is more diffusive than the first-order MOC for Courant numbers larger than one. The numerical solution is stable for all values of the Courant number.

6.3.3. *Conservative versions of the implicit upwind scheme*

The explicit upwind scheme presented in section 6.3.1 leads to the same formulation as the first-order MOC when applied to the advection equation. Consequently, conservation is not guaranteed when the scheme is applied with irregular grids. Conservation can be restored via a minor modification in the estimate of the time derivative. This is done by attaching a control volume to each computational point. The control volume is delineated by interfaces located at mid-distance between two adjacent computational points (Figure 6.14).

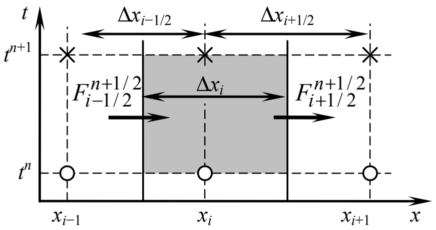

Figure 6.14. *Defining control volumes for the upwind scheme*

A mass balance over the control volume attached to the point i gives:

$$\Delta x_i U_i^{n+1} = \Delta x_i U_i^n + (F_{i-1/2}^{n+1/2} - F_{i+1/2}^{n+1/2}) + \Delta t S_i^{n+1/2} \qquad [6.46]$$

where $F_{i-1/2}^{n+1/2}$ and $F_{i+1/2}^{n+1/2}$ are the average value of the fluxes across the left- and right-hand interface of the control volume respectively and $S_{i+1/2}^{n+1/2}$ is the average value of the source term over the control volume. The size of the control volume is given by $\Delta x_i = (\Delta x_{i-1/2} + \Delta x_{i+1/2})/2$. Note that in equation [6.46] the quantity U_i^n is not a point value but the average value of U over the control volume at the time level n. This approach is to be put in parallel with the finite volume approach described in Chapter 7.

In the conservative upwind approach, the flux is estimated at a given interface using the variable in the cell located immediately upstream of the interface. For instance, the flux at the interface $i - 1/2$ is computed using the cell $i - 1$ if the wave speed is positive. It is computed using the cell i if the wave speed is negative. In the same way, the flux at the interface $i + 1/2$ is computed using the cell i or $i + 1$ for a positive and negative wave speed respectively. These formulations can be summarized as follows:

$$\left.\begin{array}{l} F_{i-1/2}^{n+1/2} \approx \begin{cases} F_{i-1}^{n} & \text{if } \lambda_i^{n+1/2} \geq 0 \\ F_i^n & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \\ F_{i+1/2}^{n+1/2} \approx \begin{cases} F_i^n & \text{if } \lambda_i^{n+1/2} \geq 0 \\ F_{i+1}^n & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \end{array}\right\} \qquad [6.47]$$

Substituting equations [6.47] into equation [6.46] leads to:

$$U_i^{n+1} = \begin{cases} U_i^n + 2\dfrac{F_{i-1}^n - F_i^n}{\Delta x_{i-1/2} + \Delta x_{i+1/2}}\Delta t + \Delta t S_i^n & \text{if } \lambda_i^{n+1/2} \geq 0 \\ U_i^n + 2\dfrac{F_i^n - F_{i+1}^n}{\Delta x_{i-1/2} + \Delta x_{i+1/2}}\Delta t + \Delta t S_i^n & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \qquad [6.48]$$

In the implicit version of the scheme, the flux and the source term are calculated using the unknown values at the time level $n + 1$:

$$U_i^{n+1} = \begin{cases} U_i^n + 2\dfrac{F_{i-1}^{n+1} - F_i^{n+1}}{\Delta x_{i-1/2} + \Delta x_{i+1/2}}\Delta t + \Delta t S_i^{n+1} & \text{if } \lambda_i^{n+1/2} \geq 0 \\ U_i^n + 2\dfrac{F_i^{n+1} - F_{i+1}^{n+1}}{\Delta x_{i-1/2} + \Delta x_{i+1/2}}\Delta t + \Delta t S_i^{n+1} & \text{if } \lambda_i^{n+1/2} \leq 0 \end{cases} \qquad [6.49]$$

The total amount of U contained within the control volume is conserved. Indeed, the mass is defined as:

$$M_T^n = \sum_{i=1}^{M} \Delta x_i U_i^n \qquad [6.50]$$

The flux $F_{i+1/2}^{n+1/2}$ leaving the control volume i across the interface $i + 1/2$ enters the control volume $i + 1/2$ through the same interface. Therefore, no mass is gained or lost during the time step.

This conservative version of the finite difference approach is sometimes referred to as "finite differences with control volume".

6.3.4. *Application examples*

The conservative scheme is applied to the same test cases as in section 6.2.3, with the same test parameters (Tables 6.1 and 6.2).

The results of the first test, which deals with the linear advection of a concentration profile, are illustrated by Figures 6.15–16.

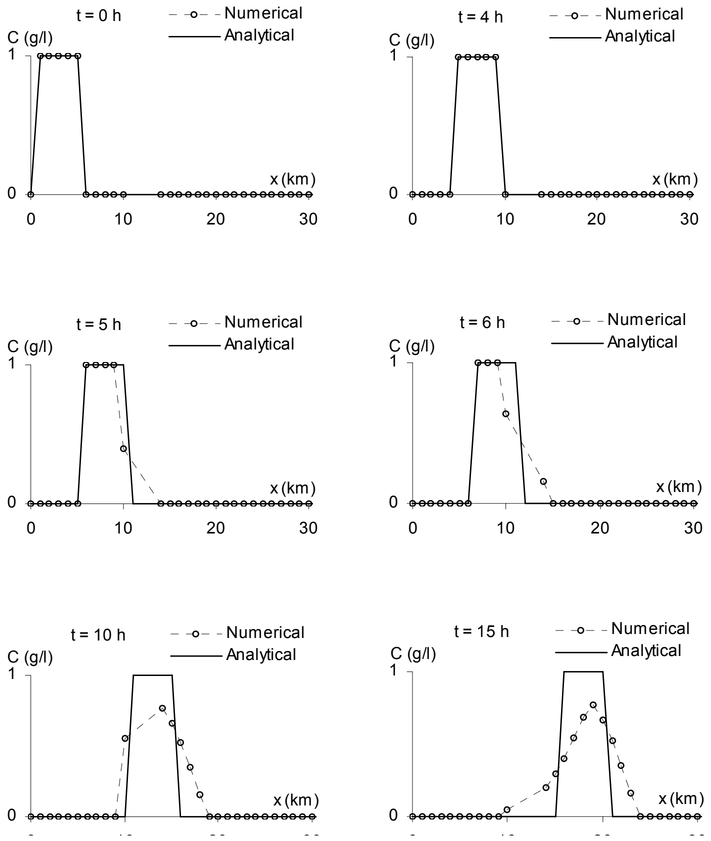

Figure 6.15. *Pure advection of a square concentration signal. Analytical solution and numerical solution calculated by the explicit upwind scheme on an irregular grid*

In contrast with the first-order MOC, the integral of the numerical profile obtained using the conservative upwind scheme is the same at all times. This is confirmed by Figure 6.16, which shows the variation in the total mass of contaminant in the domain. The mass is constant in the limit of the precision of the computer. It must be stressed however that the conservative upwind scheme remains diffusive, which leads to an underestimation of the peak concentration compared to the analytical solution.

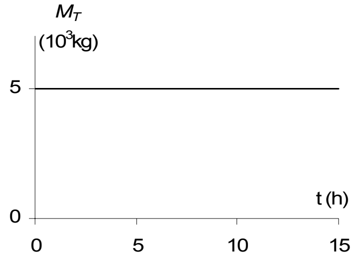

Figure 6.16. *Pure advection of a square concentration signal calculated by the conservative, explicit upwind scheme. Total mass of contaminant in the domain as a function of time*

The second test, that deals with a Riemann problem for the inviscid Burgers equation, is applied to the conservation form [1.69]. In contrast with the application to the first-order MOC in section 6.2.3, the three options [6.38] give the same result. Conservation being ensured intrinsically by the scheme, the front propagates at the correct speed (Figure 6.17). Nevertheless, the scheme remains diffusive and the shape of the front is altered by numerical smoothing.

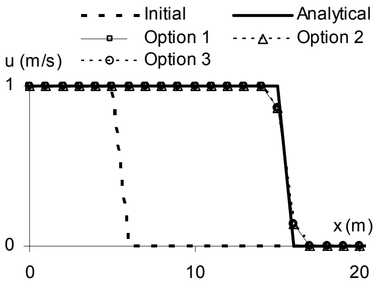

Figure 6.17. *Analytical solution and numerical profiles computed by the conservative, explicit upwind scheme for the three calculation options [6.38]*

6.4. The Preissmann scheme

6.4.1. *Formulation*

The Preissmann scheme [PRE 61a-c] is a conservative scheme. It is used in a number of commercially available packages for the simulation of open channel hydraulics such as flows in rivers, urban drainage and sewer systems. The discretization follows the same rule for a hyperbolic system as for a scalar, conservation law. The scheme uses the four computational points that define the corners of a "box" in the phase space (Figure 6.18), hence the name of "box scheme".

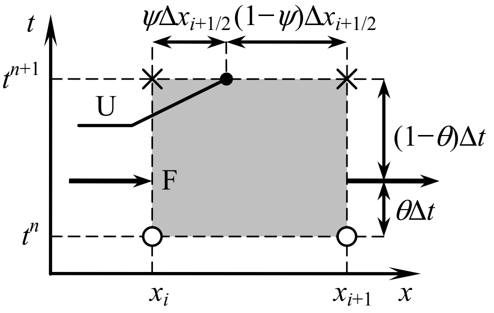

Figure 6.18. *Definition sketch for the control volume in the Preissmann scheme*

A balance over the control volume materialized by the gray-shaded area in Figure 6.18 gives:

$$\frac{U_{i+1/2}^{n+1} - U_{i+1/2}^{n}}{\Delta t} + \frac{F_{i+1}^{n+1/2} - F_i^{n+1/2}}{\Delta t} = S_{i+1/2}^{n+1/2} \qquad [6.51]$$

where $U_{i+1/2}^{n}$ is the average value of U over the segment $[x_i, x_{i+1}]$ at the time level n, $F_i^{n+1/2}$ is the average of the flux F at the point i over the time step and $S_{i+1/2}^{n+1/2}$ is the average value of the source term S over the segment $[x_i, x_{i+1}]$ between the time levels n and $n + 1$. The following estimates are used for $U_{i+1/2}^{n}$ and $F_i^{n+1/2}$:

$$\left.\begin{aligned} U_{i+1/2}^{n} &\approx (1-\psi)U_i^n + \psi\ U_{i+1}^n \\ F_i^{n+1/2} &\approx (1-\theta)F_i^n + \theta\ F_i^{n+1} \end{aligned}\right\} \qquad [6.52]$$

where θ and ψ are parameters ranging from 0 to 1. θ is usually called the time-centering, or implicitation, parameter. Most practical implementations of the scheme use the value $\psi = 1/2$. The same weight is then given to the waves with positive and negative wave speeds.

In this case the numerical solution is unconditionally stable for $\theta \geq 1/2$, while it is unconditionally unstable for $\theta < 1/2$. Substituting equations [6.52] into equation [6.51] gives:

$$\frac{(1-\psi)(\mathrm{U}_i^{n+1}-\mathrm{U}_i^n)+\psi(\mathrm{U}_{i+1}^{n+1}-\mathrm{U}_{i+1}^n)}{\Delta t}+\frac{(1-\theta)(\mathrm{F}_{i+1}^n-\mathrm{F}_i^n)+\theta(\mathrm{F}_{i+1}^{n+1}-\mathrm{F}_i^{n+1})}{\Delta x_{i+1/2}}=S_{i+1/2}^{n+1/2} \qquad [6.53]$$

Multiplying by Δt and rearranging leads to:

$$\begin{aligned}&(1-\psi)\mathrm{U}_i^{n+1}+\psi\mathrm{U}_{i+1}^{n+1}+\frac{\theta\Delta t}{\Delta x_{i+1/2}}(\mathrm{F}_{i+1}^{n+1}-\mathrm{F}_i^{n+1})=\\&(1-\psi)\mathrm{U}_i^n+\psi\mathrm{U}_{i+1}^n+\frac{(1-\theta)\Delta t}{\Delta x_{i+1/2}}(\mathrm{F}_{i+1}^n-\mathrm{F}_i^n)+\Delta t S_{i+1/2}^{n+1/2}\end{aligned} \qquad [6.54]$$

Equation [6.54] is a vector equation, that is, a system of m scalar equations. In the general case, the nonlinearity of F with respect to U makes the solution of the system [6.54] computationally intensive. In a number of approaches [CUN 80], solving a linearized version of the system makes the solution process easier:

$$\left.\begin{aligned}&\mathrm{U}_i^{n+1}=\mathrm{U}_i^n+\Delta\mathrm{U}_i\\&\mathrm{F}_i^{n+1}=\mathrm{F}_i^n+\frac{\partial\mathrm{F}}{\partial\mathrm{U}}\Delta\mathrm{U}_i=\mathrm{F}_i^n+\mathrm{A}_i^{n+1/2}\Delta\mathrm{U}_i\end{aligned}\right\} \qquad [6.55]$$

where the matrix $\mathrm{A}_i^{n+1/2}$ is an approximation of the average value of the Jacobian matrix $\mathrm{A}=\partial\mathrm{F}/\partial\mathrm{U}$ between the time levels n and $n+1$. How this matrix should be approximated is the subject of section 6.4.2. Equation [6.53] can be rewritten as:

$$\begin{aligned}&\frac{(1-\psi)\Delta\mathrm{U}_i+\psi\Delta\mathrm{U}_{i+1}}{\Delta t}+\frac{\theta(\mathrm{A}_{i+1}^{n+1/2}\Delta\mathrm{U}_{i+1}-\mathrm{A}_i^{n+1/2}\Delta\mathrm{U}_i)}{\Delta x_{i+1/2}}=\\&\mathrm{S}_{i+1/2}^{n+1/2}-\frac{\mathrm{F}_{i+1}^n-\mathrm{F}_i^n}{\Delta x_{i+1/2}}\end{aligned} \qquad [6.56]$$

Simplifying by Δt leads to:

$$\left[(1-\psi)\mathrm{I}-\frac{\theta\Delta t\mathrm{A}_i^{n+1/2}}{\Delta x_{i+1/2}}\right]\Delta\mathrm{U}_i+\left[\psi\;\mathrm{I}+\frac{\theta\Delta t\mathrm{A}_{i+1}^{n+1/2}}{\Delta x_{i+1/2}}\right]\Delta\mathrm{U}_{i+1}= \\ \Delta t\mathrm{S}_{i+1/2}^{n+1/2}-\Delta t\frac{\mathrm{F}_{i+1}^n-\mathrm{F}_i^n}{\Delta x_{i+1/2}} \tag{6.57}$$

where I is the identity matrix. System [6.57] is linear and can be solved using standard matrix inversion techniques. When a scalar equation is to be solved, the matrix A becomes the wave speed λ and equation [6.57] simplifies into:

$$(1-\psi-\theta\mathrm{Cr}_i^{n+1/2})\Delta U_i+(\psi+\theta\mathrm{Cr}_{i+1}^{n+1/2})\Delta U_{i+1}= \\ \Delta tS_{i+1/2}^{n+1/2}-\Delta t\frac{F_{i+1}^n-F_i^n}{\Delta x_{i+1/2}} \tag{6.58}$$

where the average Courant numbers at the points i and $i+1$ are given by:

$$\left.\begin{aligned}\mathrm{Cr}_i^{n+1/2}&=\frac{\lambda_i^n+\lambda_i^{n+1}}{2}\frac{\Delta t}{\Delta x_{i+1/2}}\\ \mathrm{Cr}_{i+1}^{n+1/2}&=\frac{\lambda_{i+1}^n+\lambda_{i+1}^{n+1}}{2}\frac{\Delta t}{\Delta x_{i+1/2}}\end{aligned}\right\} \tag{6.59}$$

6.4.2. *Estimation of nonlinear terms – algorithmic aspects*

The Jacobian matrix A and the source term S being functions of U, the average values $\mathrm{A}_i^{n+1/2}$ and $\mathrm{S}_{i+1/2}^{n+1/2}$ over the time steps necessarily depend on the (unknown) value of U at the time level $n+1$. The following expressions are used in practice:

$$\left.\begin{aligned}\mathrm{S}_{i+1/2}^{n+1/2}&=(1-\theta)[(1-\psi)\mathrm{S}_i^n+\psi\mathrm{S}_{i+1}^n]+\theta[(1-\psi)\mathrm{S}_i^{n+1}+\psi\mathrm{S}_{i+1}^{n+1}]\\ \mathrm{A}_i^{n+1/2}&=(1-\theta)\mathrm{A}_i^n+\theta\mathrm{A}_i^{n+1}\end{aligned}\right\} \tag{6.60}$$

The calculation procedure is iterative. It consists of the following steps:

1) Initialize $\mathrm{A}_i^{n+1/2}$ and $\mathrm{S}_{i+1/2}^{n+1/2}$ using the values at the beginning of the time step:

$$\left.\begin{array}{l} \mathrm{S}_{i+1/2}^{n+1/2} \approx (1-\psi)\mathrm{S}_i^n + \psi \mathrm{S}_{i+1}^n \\ \mathrm{A}_i^{n+1/2} = \mathrm{A}_i^n \end{array}\right\} \qquad [6.61]$$

2) Solve system [6.57] using estimates [6.61].

3) Update $\mathrm{A}_i^{n+1/2}$ and $\mathrm{S}_{i+1/2}^{n+1/2}$ using equation [6.60].

Steps 2–3 must be repeated until convergence is achieved. In practical applications such as the Saint Venant equations, only a few iterations are needed.

The Preissmann scheme has the drawback that it cannot be used in a straightforward manner in applications where the speed of the waves change sign over the computational domain [MES 97]. In such a case, instability may occur. A new version of the scheme, based on an approach similar to flux splitting (see section 6.7), has been proposed in [JOH 02].

6.4.3. *Numerical applications*

This chapter details the application of the Preissmann scheme to the test cases presented in sections 6.2.3 and 6.3.4, with the difference that the linear advection equation is solved on a regular grid (see Table 6.3). Using different values for the time step leads to different values of the Courant number. In this test, the influence of θ on the accuracy of the numerical solution is investigated.

The Preissmann scheme is applied to the linear advection equation by defining *U* and *F* as $U = AC$, $F = AuC$ in equation [6.58] and dividing by *A*.

Symbol	Meaning	Value
A	River cross-sectional area	1,000 m^2
C_i^0	Initial concentration	1 g/l for 1 km $\le x \le$ 5 km, 0 g/l otherwise
C_G	Concentration at the left-hand boundary	0 g/l
L	Length of the domain	30 km
u	Flow velocity	1 km/hr
Δt	Computational time step	0.5 hr, 1 hr, 2 hr
Δx	Cell size	1 km

Table 6.3. *Physical and numerical parameters for the numerical solution of the linear advection equation on a regular grid*

$$
\begin{aligned}
(\psi + \theta\,\mathrm{Cr})C_{i+1}^{n+1} &= (\psi - 1 + \theta\,\mathrm{Cr})C_i^{n+1} + \\
&\quad [1 - \psi + (1-\theta)\mathrm{Cr}]C_i^n + [\psi - (1-\theta)\mathrm{Cr}]C_{i+1}^n
\end{aligned}
\qquad [6.62]
$$

Equation [6.62] is a recurrence relationship between the unknown values of C at the points i and $i+1$. In the case of a positive flow velocity, the computational domain is swept from left to right by rewriting equation [6.62] as:

$$
\begin{aligned}
C_{i+1}^{n+1} &= \frac{\psi - 1 + \theta\,\mathrm{Cr}}{\psi + \theta\,\mathrm{Cr}} C_i^{n+1} + \frac{1 - \psi + (1-\theta)\mathrm{Cr}}{\psi + \theta\,\mathrm{Cr}} C_i^n \\
&\quad + \frac{\psi - (1-\theta)\mathrm{Cr}}{\psi + \theta\,\mathrm{Cr}} C_{i+1}^n
\end{aligned}
\qquad [6.63]
$$

Figure 6.19 illustrates the behavior of the numerical solution after 10 hours for various values of the numerical parameter θ. Note that:

– for $\theta = \psi = 1/2$ the scheme is dispersive if the Courant number is different from one. The numerical dispersion is reflected by the oscillations in the computed profile. When the Courant number is smaller than one the oscillations propagate faster than the analytical solution. For Courant numbers larger than one the oscillations propagate more slowly than the analytical solution;

– increasing the value of θ induces numerical diffusion, which contributes to dampening the oscillations. Numerical dispersion still occurs but its effects are "hidden" by those of numerical diffusion. Using $\theta = 0.7$ with a Courant number $\mathrm{Cr} = 1/2$ allows the oscillations to be almost completely eliminated (Figure 6.19).

In the second test the Preissmann scheme is applied to the Riemann problem for the inviscid Burgers equation. The parameters of the test case are given in Table 6.4.

Symbol	Meaning	Value
L	Length of the domain	30 m
u_i^0	Initial flow velocity	2 m/s for $x \le 5$ m, 1 m/s otherwise
Δt	Computational time step	0.5 s, 1 s
Δx	Cell width	1 m

Table 6.4. *Physical and numerical parameters for the solution of the inviscid Burgers equation*

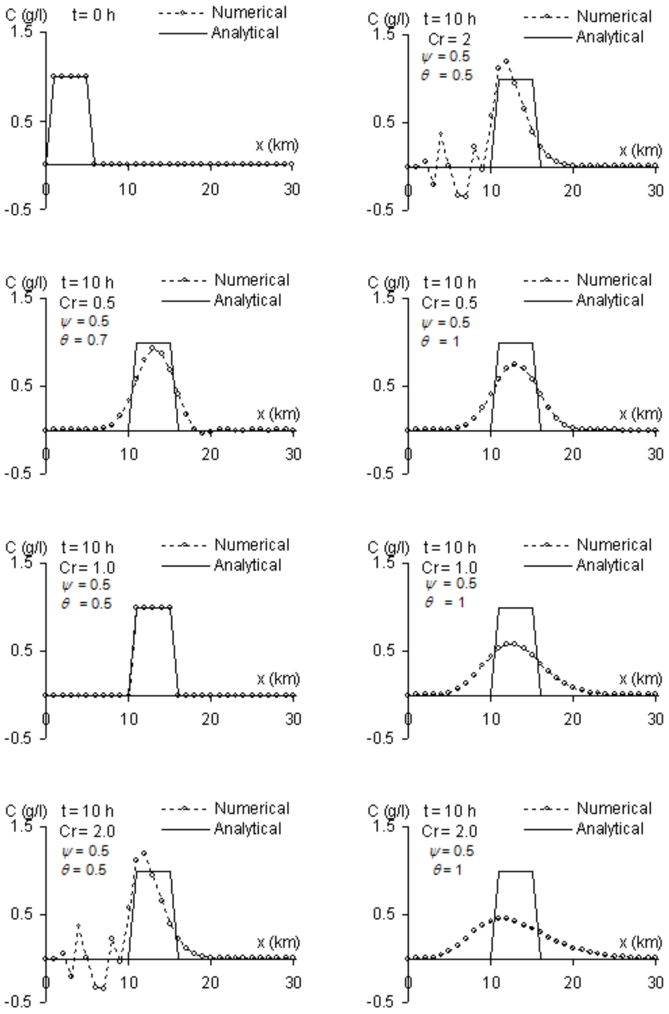

Figure 6.19. *Pure advection of a square concentration signal. Comparison between the numerical and analytical solutions for various values of the parameter θ and the Courant number*

The linearized version [6.58] of the scheme is applied with $U = u$ and $F = u^2/2$:

$$\frac{1-\psi-\theta\,\mathrm{Cr}_i^{n+1/2}}{\Delta t}\Delta u_i+\frac{\psi+\theta\,\mathrm{Cr}_{i+1}^{n+1/2}}{\Delta t}\Delta u_{i+1}=\frac{(u_i^n)^2-(u_{i+1}^n)^2}{2\Delta x_{i+1/2}}\qquad[6.64]$$

where the following expressions are used for the Courant number:

$$\left.\begin{aligned}\mathrm{Cr}_i^{n+1/2} &= \frac{u_i^n + u_i^{n+1}}{2}\frac{\Delta t}{\Delta x_{i+1/2}}\\ \mathrm{Cr}_{i+1}^{n+1/2} &= \frac{u_{i+1}^n + u_{i+1}^{n+1}}{2}\frac{\Delta t}{\Delta x_{i+1/2}}\end{aligned}\right\} \qquad [6.65]$$

For a positive flow velocity the wave speed is positive and the domain is swept directly from the left-hand boundary to the right-hand boundary:

$$\Delta u_{i+1} = \frac{\Delta t \dfrac{(u_i^n)^2 - (u_{i+1}^n)^2}{2\Delta x_{i+1/2}} - (1 - \psi - \theta\,\mathrm{Cr}_i^{n+1/2})\Delta u_i}{(\psi + \theta\,\mathrm{Cr}_{i+1}^{n+1/2})} \qquad [6.66]$$

which allows u_{i+1}^{n+1} to be calculated as:

$$u_{i+1}^{n+1} = u_{i+1}^n + \Delta u_{i+1} \qquad [6.67]$$

The analytical solution is a shock, the speed of which is the average of the speeds on both sides, i.e. 1.5 m/s. The numerical profile obtained at $t = 10$ s for various values of the time step and the parameter θ are compared to the analytical profile in Figure 6.20.

Four iterations are used for the determination of the Courant number. However, 2 iterations would have been sufficient in that the values obtained using 2 iterations and those obtained using 4 iterations differ by less than 1%.

As in the linear case, using $\theta = 1/2$ eliminates numerical diffusion and the effect of numerical dispersion becomes clearly visible. When the Courant number is larger than one (as is the case for $\Delta t = 1$ s) the oscillations appear behind the shock. When the Courant number is smaller than one (as is the case for $\Delta t = 0.5$ s) the oscillations appear ahead of the shock. Increasing θ even by a slight amount leads to a dramatic damping of the oscillations. For $\theta = 0.7$ the oscillations are almost absent from the profile. They disappear completely for $\theta = 1$.

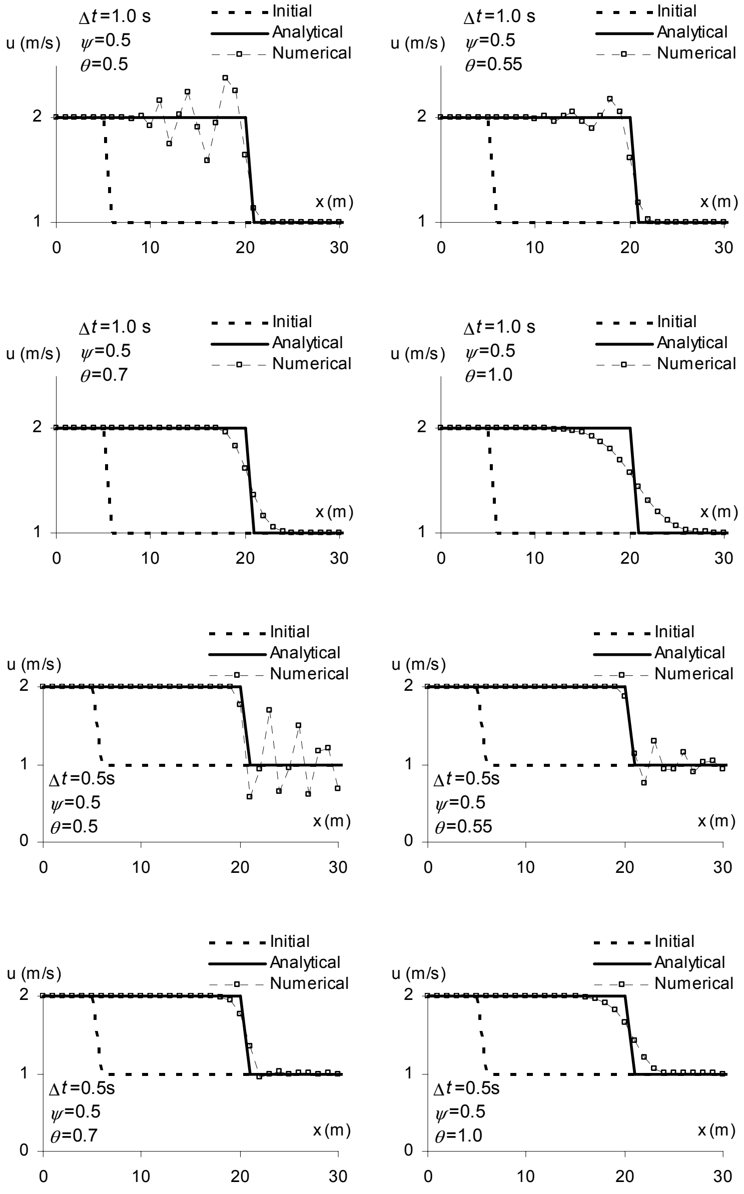

Figure 6.20. *Riemann problem for the inviscid Burgers equation with the data in Table 6.4. Analytical and numerical solution at t = 10 s*

The Preissmann scheme may lead to unstable solutions when applied to problems where the direction of the wave changes locally or becomes locally zero. This is the case with the Riemann problem specified in Table 6.2. The wave speed in the right state of the initial profile is zero. Adding numerical diffusion via the implicitation parameter θ does not always lead to a total damping of the oscillations. This is one of the reasons why the Preissmann scheme, although frequently applied to free surface flow simulations, remains restricted to the simulation of subcritical regimes.

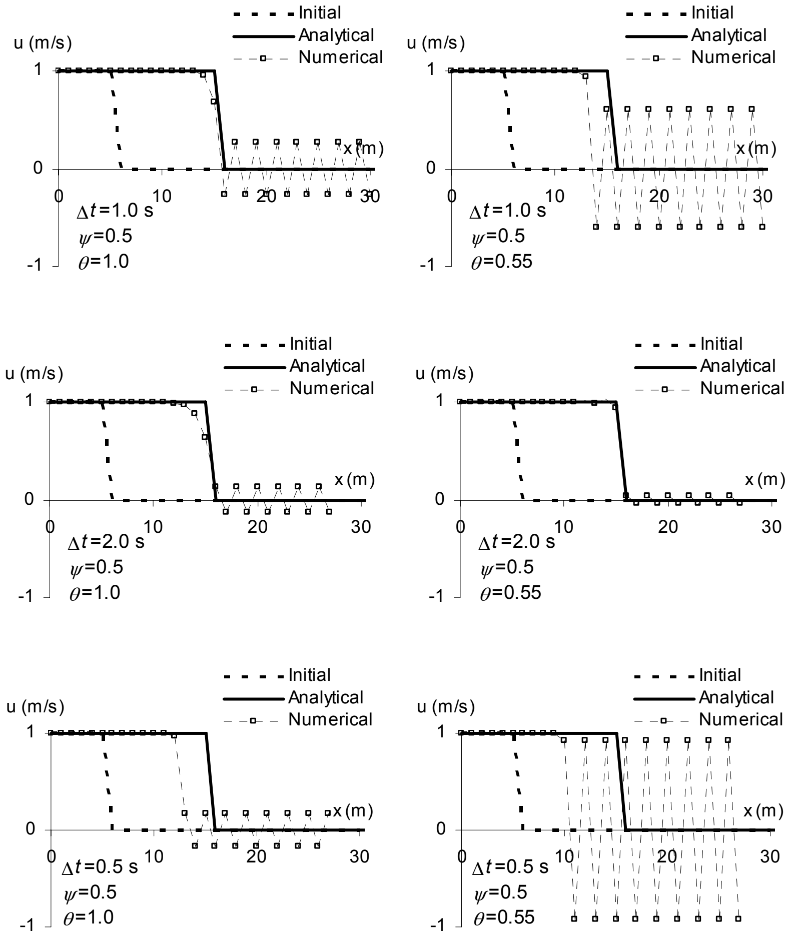

Figure 6.21. *Riemann problem for the inviscid Burgers equation with the parameters given in Table 6.2. Analytical and numerical solution at t = 10 s*

6.5. Centered schemes

6.5.1. *The Crank-Nicholson scheme*

The MOC and the upwind schemes presented in sections 6.2 and 6.3 are sensitive to the direction of the waves. As a consequence, a test should be carried out at each computational point in order to determine the direction from which the information comes and which computational points must be used. Moreover, characteristics-based and upwind schemes are considered to be too diffusive (see section B.1 for detailed considerations on numerical diffusion) in a number of computational applications, among which are turbulent flow simulations. Centered schemes aim to eliminate this drawback.

The derivative of the conserved variable with respect to time is estimated as in equation [6.39]:

$$\frac{\partial \mathrm{U}}{\partial t} \approx \frac{\mathrm{U}_i^{n+1} - \mathrm{U}_i^n}{\Delta t}$$

The derivative of the flux with respect to space is estimated as:

$$\frac{\partial \mathrm{F}}{\partial x} \approx \frac{1}{2}\frac{\mathrm{F}_{i+1}^n - \mathrm{F}_{i-1}^n}{2\Delta x} + \frac{1}{2}\frac{\mathrm{F}_{i+1}^{n+1} - \mathrm{F}_{i-1}^{n+1}}{2\Delta x} \qquad [6.68]$$

Note that the estimate of the space derivative is symmetric with respect to the point *i*, hence the term “centered scheme”. Substituting equations [6.39] and [6.68] into equation [2.1] leads to:

$$\mathrm{U}_i^{n+1} + \frac{\Delta t}{4\Delta x}(\mathrm{F}_{i+1}^{n+1} - \mathrm{F}_{i-1}^{n+1}) = \mathrm{U}_i^n + \frac{\Delta t}{4\Delta x}(\mathrm{F}_{i+1}^n - \mathrm{F}_{i-1}^n) \qquad [6.69]$$

A domain with M computational points (including the boundary points) allows $M - 2$ equations [6.69] to be written. The values of U at the remaining points $i = 1$ and $i = M$ cannot be computed using equation [6.69] because the points 0 and $M + 1$ do not exist. Since the vector U has m components, $2m$ unknown values are to be determined. Except in the particular case where all the characteristics enter the computational domain across both boundaries, the boundary conditions do not allow the solution to be determined uniquely. Additional conditions, such as zero gradient conditions or fixed values, must be prescribed at the points $i = 1$ and $i = M$. Another possibility is to discretize the governing equations at the first and last points in the domain using schemes that involve only two adjacent points in space, such as the CIR scheme or the Preissmann scheme.

If the flux function is nonlinear, system [6.69] is a nonlinear system. Its solution may be computationally demanding. The computational effort can be reduced to some extent by linearizing the system as proposed in equations [6.55], recalled hereafter:

$$\left.\begin{aligned} U_i^{n+1} &= U_i^n + \Delta U_i \\ F_i^{n+1} &= F_i^n + \frac{\partial F}{\partial U}\Delta U_i = F_i^n + A_i^{n+1/2}\Delta U_i \end{aligned}\right\}$$

Substituting equations [6.55] into equation [6.69] leads to:

$$\Delta U_i + \frac{\Delta t}{4\Delta x} A_i^{n+1/2}(A_{i+1}^{n+1/2}\Delta U_{i+1} - A_{i-1}^{n+1/2}\Delta U_{i-1}) = 0 \qquad [6.70]$$

where the Jacobian matrices $A_{i\pm 1}^{n+1/2}$ are estimated as explained in section 6.4.2.

6.5.2. *Centered schemes with Runge-Kutta time stepping*

The Crank-Nicholson scheme presented in the previous section is an implicit scheme. This implies an iterative linearization of the equations, followed by the solution of a system of algebraic equations in the form [6.70]. The question thus arises of the possibility to develop explicit, centered schemes so as to preserve the non-dissipative character of the centered formulation, while making the computational procedure simpler. The following discretization may be seen as a good candidate for the central discretization of the conservation form [2.2] or the non-conservation form [2.5]:

$$\left.\begin{aligned} \frac{\partial U}{\partial t} &\approx \frac{U_i^{n+1} - U_i^n}{\Delta t} \\ \frac{\partial F}{\partial x} &\approx \frac{F_{i+1}^n - F_{i-1}^n}{2\Delta x} \\ \frac{\partial U}{\partial x} &\approx \frac{U_{i+1}^n - U_{i-1}^n}{2\Delta x} \end{aligned}\right\} \qquad [6.71]$$

Applying the discretization above to the linear advection equation [1.48] leads to the following numerical scheme:

$$C_i^{n+1} = C_i^n + \frac{u\Delta t}{2\Delta x}(C_{i-1}^n - C_{i+1}^n) \qquad [6.72]$$

However, a simple stability analysis [VIC 82] reveals that the numerical solution obtained using equation [6.72] is unconditionally unstable. Stable solutions can be obtained only if the derivative with respect to space is integrated using multiple step integration procedures, such as the Runge-Kutta technique. The most widely used options are the second- and fourth-order Runge-Kutta time stepping techniques. The principle of such techniques is outlined hereafter.

Assume that the scheme can be expressed in the form:

$$\mathrm{U}_i^{n+1} = \mathrm{U}_i^n + \mathrm{M}\mathrm{U}_i^n \qquad [6.73]$$

where M is a matrix operator. If equation [6.72] is to be used, M is given by:

$$\mathrm{M} = \frac{u\Delta t}{2\Delta x}(\delta_{-1} + \delta_{+1}) \qquad [6.74]$$

where δ is the shift operator, defined as:

$$\left.\begin{aligned} \delta_{-1}\mathrm{U}_i^n &= \mathrm{U}_{i-1}^n \\ \delta_{+1}\mathrm{U}_i^n &= \mathrm{U}_{i+1}^n \end{aligned}\right\} \qquad [6.75]$$

The formula for the pth-order Runge-Kutta technique is:

$$\mathrm{U}_i^{n+1} = \mathrm{U}_i^n + \sum_{k=1}^{p} \frac{\Delta t^k}{k!}\mathrm{M}^k\mathrm{U}_i^n \qquad [6.76]$$

where the notation M^k indicates that the operator M is to be applied k successive times. For instance, assuming that M is given by equation [6.72], the operator M^2 is defined as:

$$\begin{aligned} \mathrm{M}^2 C_i^n &= \mathrm{M}(\mathrm{M}C_i^n) \\ &= \frac{u\Delta t}{2\Delta x}(\mathrm{M}C_{i-1}^n - \mathrm{M}C_{i+1}^n) \\ &= \frac{u\Delta t}{2\Delta x}\left[\frac{u\Delta t}{2\Delta x}(C_{i-2}^n - C_i^n) - \frac{u\Delta t}{2\Delta x}(C_i^n - C_{i+2}^n)\right] \\ &= \left(\frac{u\Delta t}{2\Delta x}\right)^2 (C_{i-2}^n - 2C_i^n + C_{i+2}^n) \end{aligned} \qquad [6.77]$$

As shown by a linear stability analysis, second-order Runge-Kutta methods do not yield stable solutions. The minimal order for which the numerical solution can be made stable is $p = 3$. The stability domains of the third- and fourth-order Runge-Kutta time stepping methods are given by (see [VIC 82]):

$$\left.\begin{array}{ll} \left|\mathrm{Cr}\right| \le 1.8 & \text{(order 3)} \\ \left|\mathrm{Cr}\right| \le 2.85 & \text{(order 4)} \end{array}\right\} \tag{6.78}$$

The range of stability of the centered scheme with third- and fourth-order Runge-Kutta time stepping is wider than that of classical explicit schemes. In contrast, the computational effort required by Runge-Kutta time stepping schemes is larger than that required by classical schemes in that the operator M must be applied several times over a given time step and the results of the successive applications must be combined linearly as in equation [6.76].

6.6. TVD schemes

6.6.1. *Definitions*

Consider the numerical solution U_i^n of a scalar conservation law over a one-dimensional domain, $i = 1,\ldots, M$. The total variation TV^n of the solution over the domain at the time level n is defined as:

$$\mathrm{TV}(U)^n = \sum_{i=1}^{M-1} \left| U_{i+1}^n - U_i^n \right| \tag{6.79}$$

The total variation serves as a quantitative indicator for the oscillatory character of the solution. A numerical scheme is said to be Total Variation Diminishing (TVD) if it satisfies the following property:

$$\mathrm{TV}(U)^{n+1} \le \mathrm{TV}(U)^n \tag{6.80}$$

Applying a TVD scheme to an initially monotone numerical solution necessarily yields a monotone numerical solution.

A scheme is said to be monotony-preserving if the following conditions hold:

$$\left.\begin{array}{lll} U_i^n \le U_{i-1}^n \quad \forall i & \Rightarrow & U_i^{n+1} \le U_{i-1}^{n+1} \quad \forall i \\ U_i^n \ge U_{i-1}^n \quad \forall i & \Rightarrow & U_i^{n+1} \ge U_{i-1}^{n+1} \quad \forall i \end{array}\right\} \tag{6.81}$$

A TVD scheme is always monotony-preserving.

The monotony property has the advantage that spurious oscillations do not arise in the numerical solution. Monotone schemes are developed with the purpose of minimizing numerical diffusion, while preserving the monotony of the solution.

6.6.2. *General formulation of TVD schemes*

This section focuses on three-point TVD schemes as applied to the linear advection equation. The generalization of such schemes to hyperbolic systems of conservation laws is dealt with in section 6.7. Applying the second-order MOC seen in section 6.2 to the linear advection equation on a regular grid leads to equation [6.21]. In the particular case of a zero source term, equation [6.21] simplifies into:

$$U_i^{n+1} = (\mathrm{Cr}+1)\frac{\mathrm{Cr}}{2}U_{i-1}^n + (1-\mathrm{Cr}^2)U_i^n + (\mathrm{Cr}-1)\frac{\mathrm{Cr}}{2}U_{i+1}^n \qquad [6.82]$$

Equation [6.82] is also known as the Lax-Wendroff scheme [LAX 60]. Equation [6.82] can be obtained from a second-order consistency analysis by specifying a zero numerical diffusion condition. The scheme may also be written as the combination of an upwind scheme and a function of the variations in the conserved variable. Assuming that the advection velocity is positive, equation [6.82] can be rewritten as:

$$\begin{aligned} U_i^{n+1} &= \mathrm{Cr}U_{i-1}^n + (1-\mathrm{Cr})U_i^n \\ &\quad + \frac{(\mathrm{Cr}-1)\mathrm{Cr}}{2}\left[(U_{i+1}^n - U_i^n) - (U_i^n - U_{i-1}^n)\right] \\ &= U_i^n - (U_i^n - U_{i-1}^n)\mathrm{Cr} \\ &\quad + \frac{(\mathrm{Cr}-1)\mathrm{Cr}}{2}\left[(U_{i+1}^n - U_i^n) - (U_i^n - U_{i-1}^n)\right] \end{aligned} \qquad [6.83]$$

Reasoning by symmetry leads to the following expression for a negative advection velocity:

$$\begin{aligned} U_i^{n+1} &= U_i^n + (U_i^n - U_{i+1}^n)\mathrm{Cr} \\ &\quad + \frac{(\mathrm{Cr}+1)\mathrm{Cr}}{2}\left[(U_{i+1}^n - U_i^n) - (U_i^n - U_{i-1}^n)\right] \end{aligned} \qquad [6.84]$$

Equations [6.83] and [6.84] can be rewritten in the following, condensed form:

$$\begin{aligned} U_i^{n+1} &= U_i^n - \frac{|\mathrm{Cr}| + \mathrm{Cr}}{2}(U_i^n - U_{i-1}^n) + \frac{|\mathrm{Cr}| - \mathrm{Cr}}{2}(U_{i+1}^n - U_i^n) \\ &+ \frac{(|\mathrm{Cr}| - 1)|\mathrm{Cr}|}{2}\left[(U_{i+1}^n - U_i^n) - (U_i^n - U_{i-1}^n)\right] \end{aligned} \qquad [6.85]$$

Scheme [6.85] yields oscillatory solutions in the neighborhood of steep fronts. TVD schemes aim to limit the contribution of the variations in U locally when the gradient of the solution becomes too large. To do so, scheme [6.85] is modified into:

$$\begin{aligned} U_i^{n+1} &= U_i^n - \frac{|\mathrm{Cr}| + \mathrm{Cr}}{2}(U_i^n - U_{i-1}^n) + \frac{|\mathrm{Cr}| - \mathrm{Cr}}{2}(U_{i+1}^n - U_i^n) \\ &+ \frac{(|\mathrm{Cr}| - 1)|\mathrm{Cr}|}{2}\left[(U_{i+1}^n - U_i^n)\phi_{i+1/2}^n - (U_i^n - U_{i-1}^n)\phi_{i-1/2}^n\right] \end{aligned} \qquad [6.86]$$

where the so-called limiting function ϕ, also referred to as a limiter, is a real number between 0 and 1. The value of ϕ depends on the regularity (or smoothness) of the profile. If the profile is regular (or smooth) enough the limiter is set to unity and scheme [6.86] is equivalent to the original Lax-Wendroff scheme. When the solution becomes irregular, with sudden changes in the local slope of the solution profile, ϕ is decreased, which is equivalent to adding numerical diffusion. In the particular case $\phi = 0$, equation [6.86] is equivalent to the upwind scheme. The smoothness of the solution is characterized by a monotony indicator θ defined as:

$$\theta_{i+1/2}^n = \begin{cases} \dfrac{U_i^n - U_{i-1}^n}{U_{i+1}^n - U_i^n} & \text{if Cr} \geq 0 \\[2ex] \dfrac{U_{i+2}^n - U_{i+1}^n}{U_{i+1}^n - U_i^n} & \text{if Cr} \leq 0 \end{cases} \qquad [6.87]$$

The monotony indicator is the ratio of the slope of the profile upstream of the point i to the slope of the profile downstream of it (Figure 6.22). The indicator is positive when the solution is monotone and negative otherwise. It is zero if the slope of the profile is zero upstream of the point i. It is infinite if the profile is horizontal downstream of the point i. The various TVD schemes proposed in the literature use different formulations for the limiter $\phi(\theta)$. The conditions that must be fulfilled by the function $\phi(\theta)$ are detailed in the next section.

Note that a number of schemes proposed in the literature before the appearance of TVD schemes were found later to be particular cases of the general formulation

[6.86], provided that $\phi(\theta)$ is defined appropriately. This is the case with the schemes listed in Table 6.5. Note that not all of these schemes are TVD.

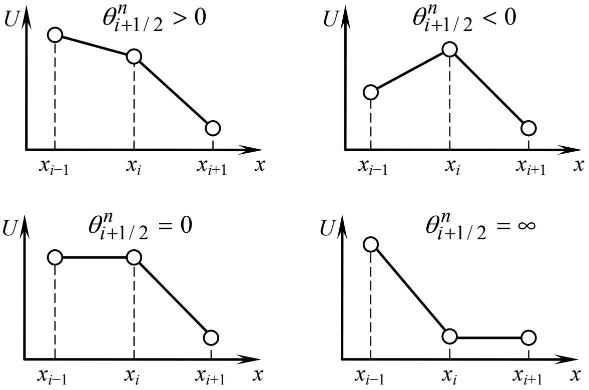

Figure 6.22. *Monotony indicator for various possible configurations (sketched here only for a positive wave speed)*

Scheme	Equivalent limiter
Upwind scheme	$\phi(\theta) = 0$
Lax-Wendroff scheme	$\phi(\theta) = 1$
Beam and Warming scheme	$\phi(\theta) = \theta$
Fromm scheme	$\phi(\theta) = (1+\theta)/2$

Table 6.5. *Expression of the limiter for classical schemes*

6.6.3. *Harten's and Sweby's criteria*

The necessary conditions for a scheme to be TVD were derived by Harten [HAR 83a, HAR 84]. Harten's analysis focuses on schemes in the form:

$$U_i^{n+1} = U_i^n - (U_i^n - U_{i-1}^n)a_{i-1/2}^n + (U_{i+1}^n - U_i^n)b_{i+1/2}^n \qquad [6.88]$$

Such schemes are TVD provided that *a* and *b* satisfy the following conditions:

$$\left.\begin{array}{r} a_{i-1/2}^n \geq 0 \\ b_{i+1/2}^n \geq 0 \\ a_{i+1/2}^n + b_{i+1/2}^n \leq 1 \end{array}\right\} \qquad [6.89]$$

Note that equation [6.86] can be written in the form [6.88] by defining *a* and *b* as:

$$\left.\begin{aligned} a_{i-1/2}^{n} &= \mathrm{Cr} + \frac{(\mathrm{Cr}-1)\mathrm{Cr}}{2}\left[\phi(\theta_{i-1/2}^{n}) - \frac{\phi(\theta_{i+1/2}^{n})}{\theta_{i+1/2}^{n}}\right] \\ b_{i+1/2}^{n} &= 0 \end{aligned}\right\} \qquad [6.90]$$

Also note that definitions [6.90] are given only for a positive wave speed. Substituting equations [6.90] into equations [6.89] leads to:

$$0 \le \mathrm{Cr} + \frac{(\mathrm{Cr}-1)\mathrm{Cr}}{2}\left[\phi(\theta_{i-1/2}^{n}) - \frac{\phi(\theta_{i+1/2}^{n})}{\theta_{i+1/2}^{n}}\right] \le 1 \qquad [6.91]$$

Since the Courant number is assumed to be smaller than one for the sake of stability, equation [6.91] becomes:

$$\left.\begin{aligned} 0 \le \phi(\theta) \le 2 \\ 0 \le \phi(\theta) \le 2\theta \end{aligned}\right\} \text{ for } \theta \ge 0 \qquad [6.92]$$

If θ is negative the limiter ϕ is set to zero and the scheme becomes locally equivalent to the TVD, explicit upwind scheme. This leads to the final set of conditions for θ:

$$\left.\begin{aligned} &\phi(0) = 0 && \text{for } \theta \le 0 \\ &0 \le \phi(\theta) \le \min\bmod(2, 2\theta) && \text{for } \theta \ge 0 \end{aligned}\right\} \qquad [6.93]$$

Figure 6.23 provides a representation of the TVD region in the (θ, ϕ) space. The TVD region is the gray-shaded area in the figure. Note that the Lax-Wendroff and the explicit upwind scheme are represented by the straight lines $\phi = 1$ and $\phi = 0$ respectively.

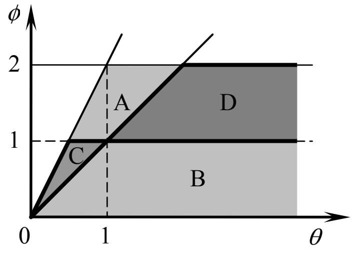

Figure 6.23. *Representation of the TVD region in the (θ, ϕ) space. General TVD region (light gray-shaded area), Sweby's TVD region (dark gray-shaded area)*

Limiters that belong to region A exhibit a compressive behavior that tends to make the fronts artificially steeper, while limiter functions located in region B induce an artificial smoothing of the solution via numerical diffusion. Sweby [SWE 84] suggested that optimal limiting is achieved in regions B and C in Figure 6.23, hence the following criteria for ϕ:

$$\left.\begin{array}{ll} \phi(\theta) = 0 & \text{for } \theta \le 0 \\ \theta \le \phi(\theta) \le \min(2\theta,1) & \text{for } \theta \in [0,1] \\ 1 \le \phi(\theta) \le \min(2\theta,2) & \text{for } \theta \ge 1 \end{array}\right\} \qquad [6.94]$$

A number of classical limiters are presented in section 6.6.4.

6.6.4. *Classical limiters*

A number of limiters have been proposed in the literature. Well-known examples are the minmod, MC, Superbee and Van Leer limiters. The corresponding formulae are given in Table 6.6. A graphical representation of the limiting functions in the (θ, ϕ) space is provided in Figure 6.24. A limiter, that is very close to the MC limiter, is also proposed for the Lax-Wendroff scheme. The minmod limiter, that is the most diffusive limiter among the limiters presented here, follows the lower bound of Sweby's TVD region. The Superbee limiter, that follows the upper bound of Sweby's TVD region, may prove to be overcompressive in some cases. The MC limiter follows the minimum between of Beam and Warming's scheme [WAR 76], Fromm's scheme [FRO 68] and the maximum permissible value $\phi = 2$. Van Leer's limiter is a rational function of θ.

Limiter / scheme	Formula	Figure
Upwind scheme	$\phi(\theta) = 0$	6.24a
Lax-Wendroff scheme	$\phi(\theta) = 1$	6.24a
Monotone Lax-Wendroff scheme	$\phi(\theta) = \max[0, \min(1,2\theta)]$	6.24b
Minmod limiter	$\phi(\theta) = \max[0, \min(1, \theta)]$	6.24c
Superbee limiter	$\phi(\theta) = \max[0, \min(1,2\theta), \min(2, \theta)]$	6.24d
MC limiter	$\phi(\theta) = \max\{0, \min[(1+\theta)/2, 2, 2\theta]\}$	6.24e
Van Leer's limiter	$\phi(\theta) = (\theta + \lvert\theta\rvert)/(1+\lvert\theta\rvert)$	6.24f

Table 6.6. *Expression of the limiter ϕ for various classical schemes and limiters available from the literature. All limiters take a zero value for negative values of θ*

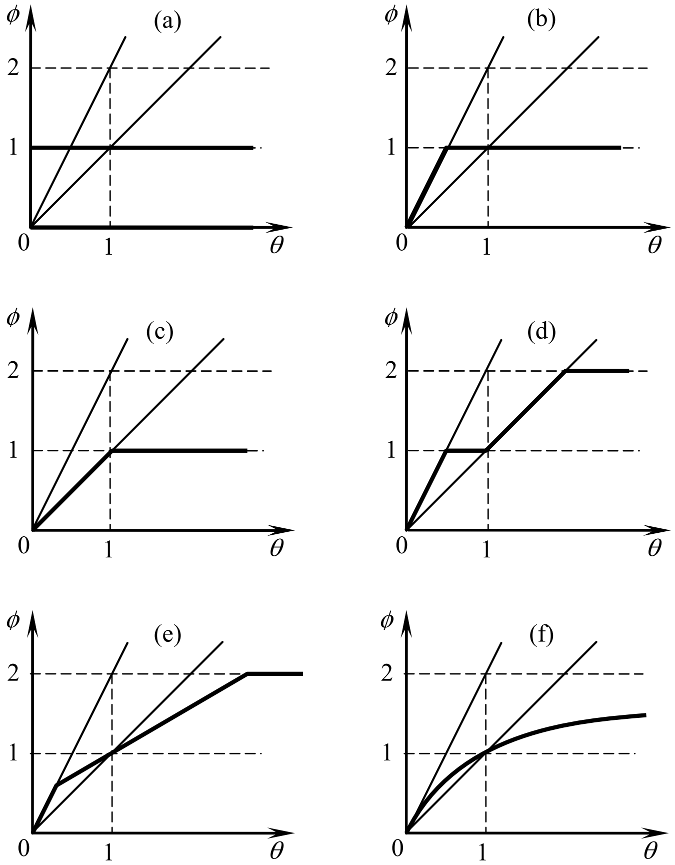

Figure 6.24. *Representation of classical schemes and limiters in the (θ, ϕ) space. Upwind scheme and Lax-Wendroff scheme (a), Monotone Lax-Wendroff (b), minmod limiter (c), Superbee limiter (d),MC limiter (e), Van Leer's limiter (f)*

6.6.5. *Computational example*

The performance of the various schemes and limiters are illustrated by an application to the linear advection equation. The pure advection of a square concentration profile in a uniform velocity field is simulated by solving equation [1.48] numerically. The parameters of the test case are given in Table 6.7. The Courant number is equal to 0.5, which is the configuration where the effect of the limiter (or the absence of limiter) is maximum because the quantity (Cr – 1) Cr, that is the coefficient of the limiting function, is maximum (see e.g. equation [6.86] or equation [6.90]). Figure 6.25 shows the numerical solutions computed at $t = 50$ s using a 15 m wide initial, square concentration profile.

Symbol	Meaning	Value
u	Advection velocity	1 m/s
Δt	Computational time step	0.5 s
Δx	Cell width	1 m

Table 6.7. *Pure advection of a square concentration profile in a uniform velocity field. Parameters of the test case*

The performance of the schemes with limiters is intermediate between that of the first-order, diffusive upwind scheme and the second-order, dispersive Lax-Wendroff scheme. The amplitude and the total variation of the solutions at $t = 50$ s is shown in Table 6.8. The non-monotone character of the Lax-Wendroff scheme is illustrated by the increase in the total variation between $t = 0$ s and $t = 50$ s.

The compressive Superbee limiter is the only one that allows the amplitude of the signal to be preserved after 100 time steps. The MC and Van Leer's limiter give very similar results. The monotone Lax-Wendroff scheme yields an asymmetrical solution, with a larger amplitude and total variation than the minmod limiter.

Solution method	Amplitude at $t = 50$ s	Total variation at $t = 50$ s
Analytical solution	1	2
Upwind scheme	0.837	1.673
Lax-Wendroff scheme	1.373	2.990
Monotone Lax-Wendroff scheme	0.996	1.992
Minmod limiter	0.981	1.962
Superbee limiter	1.000	1.999
MC limiter	0.999	1.998
Van Leer's limiter	0.998	1.997

Table 6.8. *Amplitude and total variation of the numerical solution at $t = 50$ s*

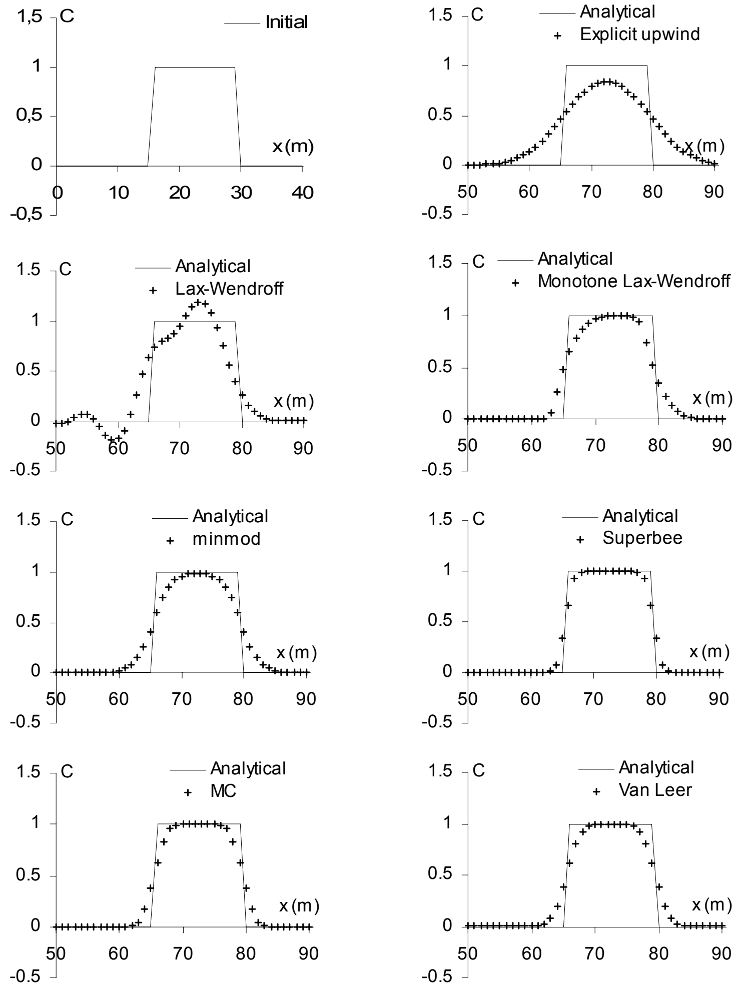

Figure 6.25. *Pure advection of a square concentration profile (arbitrary units). Analytical and numerical solutions after 100 time steps with Cr = 1/2*

6.7. The flux splitting technique

6.7.1. *Principle of the approach*

The flux splitting technique [STE 81], also known as the flux vector splitting technique or flux difference splitting technique, allows the TVD approach to be generalized to hyperbolic systems of conservation laws. The discretization can be made conservative provided that a number of conditions are satisfied. The flux splitting technique is applicable to all upwind schemes, the formulation of which depends on the direction in which the waves propagate. In hyperbolic systems of conservation laws, the various waves propagate at different speeds and in different directions. The purpose of the flux splitting technique is precisely to account for the propagation direction of each of the waves in the formulation of the upwind terms. The flux splitting technique has been applied to a variety of hyperbolic systems. The original publications focused on the equations of gas dynamics [STE 81, DIC 85]. An application to the one-dimensional Saint Venant equations can be found in [ALC 92, HUB 00].The purpose is to solve the non-conservation form [2.5]. The discretization can be made conservative as explained in section 6.7.2. For the sake of clarity, the source term is assumed to be zero hereafter. The non-conservation form [2.5] then simplifies into:

$$\frac{\partial \mathrm{U}}{\partial t} + \mathrm{A}\frac{\partial \mathrm{U}}{\partial x} = 0 \qquad [6.95]$$

The derivative of U with respect to time is classically discretized as:

$$\frac{\partial \mathrm{U}}{\partial t} \approx \frac{\mathrm{U}_i^{n+1} - \mathrm{U}_i^n}{\Delta t} \qquad [6.96]$$

while the derivative with respect to space is discretized as:

$$\mathrm{A}\frac{\partial \mathrm{U}}{\partial x} \approx \mathrm{A}_{i-1/2}^{+}\left(\frac{\partial \mathrm{U}}{\partial x}\right)_{i-1/2}^{n+1/2} + \mathrm{A}_{i+1/2}^{-}\left(\frac{\partial \mathrm{U}}{\partial x}\right)_{i+1/2}^{n+1/2} \qquad [6.97]$$

where the terms $(\partial \mathrm{U}/\partial x)_{i-1/2}^{n+1/2}$ and $(\partial \mathrm{U}/\partial x)_{i+1/2}^{n+1/2}$ are estimated over the intervals $[i-1, i]$ and $[i, i+1]$ respectively. The estimates may be obtained using a first-order, upwind formulation or one of the more complex TVD formulations seen in section 6.6. The superscript $n + 1/2$ indicates that the terms are estimated between the time levels n and $n + 1$. The estimate may be purely explicit, purely implicit or semi-implicit as in the Preissmann scheme. The flux splitting technique consists of estimating the matrices $\mathrm{A}_{i-1/2}^{+}$ and $\mathrm{A}_{i+1/2}^{-}$ so as to preserve the upwind character of the formulation, that is, by eliminating all the waves that do not propagate in the direction of the point i. This is achieved using the diagonal form [2.22]:

$$\frac{\partial \mathrm{W}}{\partial t} + \Lambda\frac{\partial \mathrm{W}}{\partial x} = 0 \qquad [6.98]$$

where Λ is the diagonal matrix formed by the eigenvalues of A and W is the vector formed by the Riemann invariants. The matrix Λ is written as the sum of two matrices Λ^- and Λ^+ that contain only the negative and positive eigenvalues of A respectively:

$$\frac{\partial \mathrm{W}}{\partial t} + (\Lambda^- + \Lambda^+)\frac{\partial \mathrm{W}}{\partial x} = 0 \qquad [6.99]$$

where Λ^- and Λ^+ are defined as:

$$\Lambda^- = \begin{bmatrix} \lambda^{(1)} & & & & & \\ & \ddots & & & 0 & \\ & & \lambda^{(r)} & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{bmatrix}, \quad \lambda^{(1)} < \lambda^{(2)} < \cdots < \lambda^{(r)} \le 0$$

$$\Lambda^+ = \begin{bmatrix} 0 & & & & & \\ & \ddots & & & 0 & \\ & & 0 & & & \\ & & & \lambda^{(r+1)} & & \\ & 0 & & & \ddots & \\ & & & & & \lambda^{(m)} \end{bmatrix}, \quad 0 < \lambda^{(r+1)} < \cdots < \lambda^{(m)} \qquad [6.100]$$

Equation [6.99] is discretized as follows:

$$\Lambda \frac{\partial \mathrm{W}}{\partial x} \approx \Lambda^-_{i+1/2}\left(\frac{\partial \mathrm{W}}{\partial x}\right)^{n+1/2}_{i+1/2} + \Lambda^+_{i-1/2}\left(\frac{\partial \mathrm{W}}{\partial x}\right)^{n+1/2}_{i-1/2} \qquad [6.101]$$

where the terms $(\partial \mathrm{W}/\partial x)^{n+1/2}_{i-1/2}$ and $(\partial \mathrm{W}/\partial x)^{n+1/2}_{i+1/2}$ are estimated over the intervals $[i-1, i]$ and $[i, i+1]$ respectively. Multiplying equation [6.101] by the matrix K, introducing the product $\mathrm{K}^{-1}\,\mathrm{K} = \mathrm{I}$ leads to:

$$\begin{aligned} \mathrm{K}\Lambda\mathrm{K}^{-1}\mathrm{K}\frac{\partial \mathrm{W}}{\partial x} &\approx \mathrm{K}\Lambda^-_{i+1/2}\mathrm{K}^{-1}\mathrm{K}\left(\frac{\partial \mathrm{W}}{\partial x}\right)^{n+1/2}_{i+1/2} \\ &\quad + \mathrm{K}\Lambda^+_{i-1/2}\mathrm{K}^{-1}\mathrm{K}\left(\frac{\partial \mathrm{W}}{\partial x}\right)^{n+1/2}_{i-1/2} \end{aligned} \qquad [6.102]$$

By definition (see equations [2.16] and [2.21]):

$$\left.\begin{aligned} \mathrm{K\Lambda K^{-1}} &= \mathrm{A} \\ \mathrm{K\ dW} &= \mathrm{dU} \end{aligned}\right\} \qquad [6.103]$$

Substituting equations [6.103] into equation [6.102] leads to:

$$\mathrm{A}\frac{\partial \mathrm{U}}{\partial x} \approx \mathrm{K\Lambda}_{i+1/2}^{-}\mathrm{K}^{-1}\left(\frac{\partial \mathrm{U}}{\partial x}\right)_{i+1/2}^{n+1/2} + \mathrm{K\Lambda}_{i-1/2}^{+}\mathrm{K}^{-1}\left(\frac{\partial \mathrm{U}}{\partial x}\right)_{i-1/2}^{n+1/2} \qquad [6.104]$$

Comparing equation [6.104] and [6.97] yields the following expressions for $\mathrm{A}_{i-1/2}^{+}$ and $\mathrm{A}_{i+1/2}^{-}$:

$$\left.\begin{aligned} \mathrm{A}_{i+1/2}^{-} &\approx \mathrm{K\Lambda}_{i+1/2}^{-}\mathrm{K}^{-1} \\ \mathrm{A}_{i-1/2}^{+} &\approx \mathrm{K\Lambda}_{i-1/2}^{+}\mathrm{K}^{-1} \end{aligned}\right\} \qquad [6.105]$$

6.7.2. *Application to classical schemes*

6.7.2.1. *Explicit upwind scheme for the water hammer equations*

The conservation form of the explicit upwind scheme uses the following approximations for the space derivatives:

$$\left.\begin{aligned} \left(\frac{\partial \mathrm{U}}{\partial x}\right)_{i-1/2}^{n+1/2} &\approx \frac{\mathrm{U}_i^n - \mathrm{U}_{i-1}^n}{\Delta x_i} \\ \left(\frac{\partial \mathrm{U}}{\partial x}\right)_{i+1/2}^{n+1/2} &\approx \frac{\mathrm{U}_{i+1}^n - \mathrm{U}_i^n}{\Delta x_i} \end{aligned}\right\} \qquad [6.106]$$

Substituting equations [6.96], [6.104] and [6.106] into equation [6.95] gives:

$$\frac{\mathrm{U}_i^{n+1} - \mathrm{U}_i^n}{\Delta t} + \mathrm{A}_{i-1/2}^{+}\frac{\mathrm{U}_{i-1}^n - \mathrm{U}_i^n}{\Delta x_i} + \mathrm{A}_{i+1/2}^{-}\frac{\mathrm{U}_i^n - \mathrm{U}_{i+1}^n}{\Delta x_i} = 0 \qquad [6.107]$$

which leads to:

$$\mathrm{U}_i^{n+1} = \mathrm{U}_i^n + \frac{\mathrm{A}_{i-1/2}^{+}\Delta t}{\Delta x_i}(\mathrm{U}_{i-1}^n - \mathrm{U}_i^n) + \frac{\mathrm{A}_{i+1/2}^{-}\Delta t}{\Delta x_i}(\mathrm{U}_i^n - \mathrm{U}_{i+1}^n) \qquad [6.108]$$

The expressions of the matrices A, K and K^{-1} derived in section 2.4.3 are recalled hereafter:

$$\mathrm{A} = \begin{bmatrix} 0 & 1 \\ c^2 & 0 \end{bmatrix}, \quad \mathrm{K} = \begin{bmatrix} 1 & 1 \\ -c & c \end{bmatrix}, \quad \mathrm{K}^{-1} = \frac{1}{2c}\begin{bmatrix} c & -1 \\ c & 1 \end{bmatrix} \qquad [6.109]$$

The matrices Λ, Λ^- and Λ^+ are:

$$\Lambda = \begin{bmatrix} -c & 0 \\ 0 & c \end{bmatrix}, \quad \Lambda^- = \begin{bmatrix} -c & 0 \\ 0 & 0 \end{bmatrix}, \quad \Lambda^+ = \begin{bmatrix} 0 & 0 \\ 0 & c \end{bmatrix} \qquad [6.110]$$

Hence the matrices A$^-$ and A$^+$:

$$\mathrm{A}^- = \frac{1}{2}\begin{bmatrix} -c & 1 \\ c^2 & -c \end{bmatrix}, \qquad \mathrm{A}^+ = \frac{1}{2}\begin{bmatrix} c & 1 \\ c^2 & c \end{bmatrix} \qquad [6.111]$$

The variable vector U is defined as:

$$\mathrm{U} = \begin{bmatrix} \rho A \\ \rho Q \end{bmatrix} \qquad [6.112]$$

Substituting equations [6.111–112] into equation [6.108] leads to the following scheme:

$$\left.\begin{aligned} (\rho A)_i^{n+1} &= (\rho A)_i^n + \frac{c\Delta t}{2\Delta x_i}\left[(\rho A)_{i-1}^n - 2(\rho A)_i^n + (\rho A)_{i+1}^n\right] \\ &\quad + \frac{\Delta t}{2\Delta x_i}\left[(\rho Q)_{i-1}^n - (\rho Q)_{i+1}^n\right] \\ (\rho Q)_i^{n+1} &= (\rho Q)_i^n + \frac{c^2\Delta t}{2\Delta x_i}\left[(\rho A)_{i-1}^n - (\rho A)_{i+1}^n\right] \\ &\quad + \frac{c\Delta t}{2\Delta x_i}\left[(\rho Q)_{i-1}^n - 2(\rho Q)_i^n + (\rho Q)_{i+1}^n\right] \end{aligned}\right\} \qquad [6.113]$$

The linear dependence between the mass per unit length and the pressure and the approximation [2.75] are recalled:

$$\left.\begin{aligned} \mathrm{d}(\rho A) &\approx \frac{A}{c^2}\mathrm{d}p \\ \mathrm{d}(\rho u A) &\approx \rho \mathrm{d}Q \end{aligned}\right\} \qquad [6.114]$$

Substituting approximations [6.114] into equations [6.113] leads to:

$$\left.\begin{aligned} p_i^{n+1} &= \frac{\mathrm{Cr}}{2}(p_{i-1}^{n+1} + p_{i+1}^{n+1}) + (1-\mathrm{Cr})p_i^n + \mathrm{Cr}\frac{\rho c}{2A}(Q_{i-1}^{n+1} - Q_{i+1}^{n+1}) \\ Q_i^{n+1} &= \frac{\mathrm{Cr}}{2}(Q_{i-1}^{n+1} + Q_{i+1}^{n+1}) + (1-\mathrm{Cr})Q_i^n + \mathrm{Cr}\frac{A}{2\rho c}(p_{i-1}^{n+1} - p_{i+1}^{n+1}) \end{aligned}\right\} \qquad [6.115]$$

where the Courant number is defined as in equation [6.35], recalled hereafter:

$$\mathrm{Cr} = \frac{c\Delta t}{\Delta x_i}$$

Note that equations [6.115] and [6.34] are equivalent when the grid is regular.

6.7.2.2. *TVD scheme for the water hammer equations*

The general expression for a scalar TVD scheme is provided in section 6.6.2. The generalization of the scheme to vector variables requires the evaluation of the derivatives for positive and negative wave speeds. Remember that TVD schemes can be recast in the following form

$$U_i^{n+1} = \begin{cases} U_i^n - \lambda\Delta t\left(\dfrac{\partial U}{\partial x}\right)_{i-1/2} & \text{if} \quad \lambda \ge 0 \\ U_i^n - \lambda\Delta t\left(\dfrac{\partial U}{\partial x}\right)_{i+1/2} & \text{if} \quad \lambda \le 0 \end{cases} \qquad [6.116]$$

For the sake of clarity, the time index is omitted in the derivatives. The derivatives may be estimated using an explicit formulation (in which case the time level n is used), or an implicit approach (in which case the superscript $n + 1$ should be used), or any intermediate formulation. Comparing equations [6.116] and [6.86], using the definition $\mathrm{Cr} = \lambda\,\Delta t/\Delta x$ leads to:

$$\left.\begin{aligned}\left(\frac{\partial U}{\partial x}\right)_{i-1/2} &\approx \frac{U_i - U_{i-1}}{\Delta x} \\ &\quad - \frac{|\mathrm{Cr}| - 1}{2\Delta x}\left[(U_{i+1} - U_i)\phi^{+}_{i+1/2} - (U_i - U_{i-1})\phi^{+}_{i-1/2}\right] \\ \left(\frac{\partial U}{\partial x}\right)_{i+1/2} &\approx \frac{U_{i+1} - U_i}{\Delta x} \\ &\quad + \frac{|\mathrm{Cr}| - 1}{2\Delta x}\left[(U_{i+1} - U_i)\phi^{-}_{i+1/2} - (U_i - U_{i-1})\phi^{-}_{i-1/2}\right]\end{aligned}\right\} \qquad [6.117]$$

where the limiters used at the interfaces $i - 1/2$ and $i + 1/2$ depend on the sign of the wave propagation speed. The limiters ϕ^- and ϕ^+ are defined for the waves propagating in the direction of negative and positive x respectively:

$$\left.\begin{aligned}\phi^{-}_{i+1/2} &= \phi(\theta^{-}_{i+1/2}) \\ \phi^{+}_{i+1/2} &= \phi(\theta^{+}_{i+1/2})\end{aligned}\right\} \qquad [6.118]$$

where the monotony indicators are defined from equation [6.87] as:

$$\left.\begin{aligned}\theta^{+}_{i+1/2} &= \frac{U^n_i - U^n_{i-1}}{U^n_{i+1} - U^n_i} \\ \theta^{-}_{i+1/2} &= \frac{U^n_{i+2} - U^n_{i+1}}{U^n_{i+1} - U^n_i}\end{aligned}\right\} \qquad [6.119]$$

Equation [6.117] is generalized to hyperbolic systems as follows:

1) In the particular case of the water hammer equations, both waves propagate in opposite directions, the absolute values of both speeds are identical. Therefore, the absolute value of the Courant number attached to both waves is the same. The same value of Cr is applied identically to all the components U_p of U in equation [6.117].

2) Several waves may propagate in the same direction in the general case. Each wave has a specific Courant number. The following options are available:

2.1) Apply the wave speed, the absolute value of which is the smallest. This option minimizes the anti-diffusion.

2.2) Apply equations [6.117] to each of the Riemann invariants individually and apply the following estimate for the derivative of U;

$$\left(\frac{\partial \mathrm{U}}{\partial x}\right)_{i\pm 1/2} = \mathrm{K}_{i\pm 1/2}\left(\frac{\partial \mathrm{W}}{\partial x}\right)_{i\pm 1/2} \qquad [6.120]$$

Option 1) is used in the particular case of the water hammer equations. The limiter is computed for each of the components of the vector U. The limiter function is generalized into a diagonal matrix and equation [6.117] is generalized into:

$$\left.\begin{aligned}\left(\frac{\partial \mathrm{U}}{\partial x}\right)_{i-1/2} &\approx \frac{\mathrm{U}_i - \mathrm{U}_{i-1}}{\Delta x} \\ &\quad - \frac{|\mathrm{Cr}|-1}{2\Delta x}\left[\Phi^{+}_{i+1/2}(\mathrm{U}_{i+1}-\mathrm{U}_i) - \Phi^{+}_{i-1/2}(\mathrm{U}_i - \mathrm{U}_{i-1})\right] \\ \left(\frac{\partial \mathrm{U}}{\partial x}\right)_{i+1/2} &\approx \frac{\mathrm{U}_{i+1} - \mathrm{U}_{i}}{\Delta x} \\ &\quad + \frac{|\mathrm{Cr}|-1}{2\Delta x}\left[\Phi^{-}_{i+1/2}(\mathrm{U}_{i+1}-\mathrm{U}_i) - \Phi^{-}_{i-1/2}(\mathrm{U}_i - \mathrm{U}_{i-1})\right]\end{aligned}\right\} \qquad [6.121]$$

where the coefficients of the limiter matrix are given by:

$$(\Phi_{p,q})^{\pm}_{i+1/2} = \begin{cases} \phi[(\theta_p)^{\pm}_{i+1/2}] & \text{if} \quad p = q \\ 0 & \text{if} \quad p \neq q \end{cases} \qquad [6.122]$$

where θ_p is the monotony indicator for the pth component of U. The superscript + or – indicates that two monotony indicators must be defined: one in the direction of negative waves, another in the direction of positive waves.

6.7.2.3. *Computational example*

Discretization [6.121] is applied to the simulation of the instantaneous failure of a valve in a pipe. The parameters of the test case are given in Table 6.9.

Symbol	Meaning	Value
A	Cross-sectional area of the pipe	1 m^2
c	Sound speed	1,000 m/s
L	Length of the pipe	500 m
p_0	Initial pressure	5.10^5 Pa for x < 250 m, 10^5 Pa otherwise
T	Simulated time	0.15 s
Δt	Computational time step	0.005 s
Δx	Cell width	10 m
ρ	Density of water	1,000 kg/m

Table 6.9. *Sudden failure of a valve. Physical and numerical parameters*

The cell width and the computational time step are chosen such that the Courant number is 0.5, a value for which the effect of profile limiting (when applied) is maximum. The pressure profiles computed at $t = 0.15$ s are represented in Figure 6.26.

Note that the monotone version of the Lax-Wendroff scheme yields an asymmetry in the fronts, which the other limiters do not do. The Superbee limiter remains the most compressive, while the MC and Van Leer limiters yield comparable results.

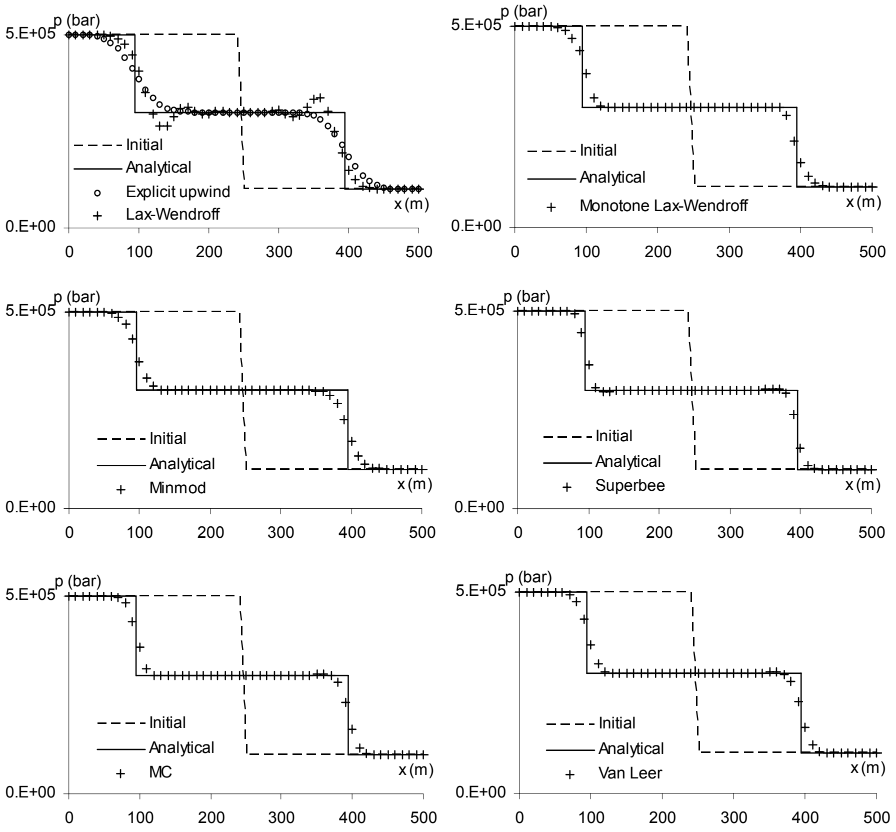

Figure 6.26. *Sudden failure of a valve. Pressure profiles computed at t = 0.15 s by the upwind scheme, the Lax-Wendroff scheme, and TVD schemes with various limiters*

6.8. Conservative discretizations: Roe's matrix

6.8.1. *Rationale and principle of the approach*

A number of techniques presented in the previous sections are based on the discretization of the equations in non-conservation form. This is the case of the Preissmann scheme seen in section 6.4 with the linearization technique [6.55] presented in section 6.4.2. This is also the case of the flux splitting technique seen in section 6.7. Both techniques require that the Jacobian matrix A of the flux F with respect to the conserved variable U be estimated. The computational examples provided in section 6.2.3 show however that the non-conservation form of the equations may give erroneous results when discontinuities appear in the solution. This is because the propagation speed of shocks is extremely sensitive to the formulation used to estimate the wave speed in the non-conservation form. If a wrong estimate is used, conservation may be violated. The approach presented here allows conservation to be preserved even though the equation is not solved in conservation form.

Consider first the conservation form [2.2] without source term. A balance over a control volume centered around the point *i* between the time level *n* and *n* + 1 yields:

$$\mathrm{U}_i^{n+1} = \mathrm{U}_i^{n+1} + \frac{\Delta t}{\Delta x_i}\left(\mathrm{F}_{i-1/2}^{n+1/2} - \mathrm{F}_{i+1/2}^{n+1/2}\right) \qquad [6.123]$$

The linearized Preissmann scheme and flux splitting techniques use discretizations of the non-conservation form that can be written as:

$$\mathrm{U}_i^{n+1} = \mathrm{U}_i^{n+1} + \frac{\Delta t}{\Delta x_i}\left[\mathrm{A}_{i-1/2}^{n+1/2}\Delta\mathrm{U}_{i-1/2}^{n+1/2} - \mathrm{A}_{i+1/2}^{n+1/2}\Delta\mathrm{U}_{i+1/2}^{n+1/2}\right] \qquad [6.124]$$

where ΔU denotes the variation in U between two adjacent computational points:

$$\Delta\mathrm{U}_{i+1/2}^{n+1/2} = \mathrm{U}_{i+1}^{n+1/2} - \mathrm{U}_i^{n+1/2} \qquad [6.125]$$

The superscript *n* + 1/2 indicates that the estimate is carried out between the time levels *n* and *n* + 1. It could be an explicit estimate, a fully implicit estimate or a semi-implicit estimate. Roe [ROE 81] provides a set of necessary conditions that should be satisfied by the estimate of the matrix A for conservation to be guaranteed:

$$\left.\begin{aligned} A_{i-1/2}^{n+1/2} &= A_{i-1/2}^{n+1/2}(U_{i-1}^{n+1/2}, U_i^{n+1/2}) \\ A_{i-1/2}^{n+1/2}(U, U) &= \partial F / \partial U(U) \\ A_{i-1/2}^{n+1/2} \Delta U_{i-1/2}^{n+1/2} &= F_i^{n+1/2} - F_{i-1}^{n+1/2} \\ \exists(K, \Lambda) \neq (0{,}0), &\quad \Lambda = K^{-1} A_{i-1/2}^{n+1/2} K \end{aligned}\right\} \qquad [6.126]$$

The first and second conditions [6.126] express consistency. The third condition enforces conservation. The last condition states the hyperbolic character of the system of conservation laws. Substituting the third equation [6.126] into equation [6.124] leads to equation [6.123] and conservation is guaranteed.

6.8.2. *Expression of Roe's matrix*

6.8.2.1. *Roe's method*

In his publication [ROE 81], Roe provides a general method for the derivation of the matrix A. The method is based on the following reasoning.

The expression of A is a function of the values U_L and U_R of U on the left- and right-hand sides of the discontinuity. Conservation is guaranteed if:

$$A(U_L - U_R) = F_L - F_R \qquad [6.127]$$

A "parameter vector" V is introduced, such that:

$$\left.\begin{aligned} U_L - U_R &= B(V_L - V_R) \\ F_L - F_R &= C(V_L - V_R) \end{aligned}\right\} \qquad [6.128]$$

where B and C are matrices, the expression of which is to be determined from equation [6.128]. Comparing equations [6.127] and [6.128] leads to the following expression for A:

$$A = CB^{-1} \qquad [6.129]$$

Application example: Roe [ROE 81] proposed the following parameter vector for the Euler equations:

$$V = \begin{bmatrix} \rho^{1/2} \\ \rho^{1/2} u \\ \rho^{1/2} H \end{bmatrix} \qquad [6.130]$$

where the enthalpy H is defined as:

$$H = (e + p)\rho \qquad [6.131]$$

The following expression is obtained for the matrix B:

$$\mathrm{B} = \begin{bmatrix} 2\rho^{1/2} & 0 & 0 \\ \rho^{1/2}u & \rho^{1/2} & 0 \\ \rho^{1/2}H/\gamma & (\gamma-1)\rho^{1/2}u/\gamma & \rho^{1/2}/\gamma \end{bmatrix} \qquad [6.132]$$

and the following expression is obtained for C:

$$\mathrm{C} = \begin{bmatrix} \rho^{1/2}u & \rho^{1/2} & 0 \\ (\gamma-1)\rho^{1/2}H/\gamma & (\gamma+1)\rho^{1/2}u/\gamma & (\gamma-1)\rho^{1/2}/\gamma \\ 0 & \rho^{1/2}H & \rho^{1/2}u \end{bmatrix} \qquad [6.133]$$

Note that the eigenvalues and eigenvectors of A may be obtained directly from the expressions of B and C without using equation [6.129]. By definition the eigenvalues of A verify:

$$\left|\mathrm{A} - \lambda \mathrm{I}\right| = 0 \qquad [6.134]$$

Consequently:

$$\left|\mathrm{AB} - \lambda \mathrm{B}\right| = 0 \qquad [6.135]$$

Using equation [6.129], condition [6.135] becomes:

$$\left|\mathrm{C} - \lambda \mathrm{B}\right| = 0 \qquad [6.136]$$

Substituting equations [6.132–133] into equation [6.129] leads to:

$$\left.\begin{aligned} \rho^{1/2} &= \frac{\rho_{\mathrm{L}}^{1/2} + \rho_{\mathrm{L}}^{1/2}}{2} \\ u &= \frac{\rho_{\mathrm{L}}^{1/2}u_{\mathrm{L}} + \rho_{\mathrm{R}}^{1/2}u_{\mathrm{R}}}{\rho_{\mathrm{L}}^{1/2} + \rho_{\mathrm{R}}^{1/2}} \\ H &= \frac{\rho_{\mathrm{L}}^{1/2}H_{\mathrm{L}} + \rho_{\mathrm{R}}^{1/2}H_{\mathrm{R}}}{\rho_{\mathrm{L}}^{1/2} + \rho_{\mathrm{R}}^{1/2}} \end{aligned}\right\} \qquad [6.137]$$

6.8.2.2. *Expression in the base of eigenvectors*

The expression of A may also be obtained by writing the variations ΔU and ΔF as linear combinations of the eigenvectors of A. The so-called wave strengths a_p (p = 1, …, m) are introduced:

$$\left.\begin{aligned}\Delta\mathrm{U} &= \sum_{p=1}^{m}\alpha_p \mathrm{K}^{(p)}\\ \Delta\mathrm{F} &= \sum_{p=1}^{m}\alpha_p \lambda^{(p)}\mathrm{K}^{(p)}\end{aligned}\right\}\qquad [6.138]$$

where the eigenvectors $\mathrm{K}^{(p)}$ and the wave speeds $\lambda^{(p)}$ are estimated from a weighted average between $\mathrm{U_L}$ and $\mathrm{U_R}$. Finding the expressions of the wave strengths allows the relationship between ΔU and ΔF to be determined.

Application example. Consider the Saint Venant equations. U, F and K are defined as:

$$\mathrm{U} = \begin{bmatrix} A \\ Q \end{bmatrix}, \mathrm{F} = \begin{bmatrix} Q \\ uQ + P/\rho \end{bmatrix}, \mathrm{K}^{(1)} = \begin{bmatrix} 1 \\ u - c \end{bmatrix}, \mathrm{K}^{(2)} = \begin{bmatrix} 1 \\ u + c \end{bmatrix}\qquad [6.139]$$

where A is the channel cross-sectional area, c is the speed of the waves in still water, P is the pressure force exerted on the channel cross-section, Q is the liquid discharge, u is the flow velocity and ρ is the density of water. The first equation [6.138] yields the following system:

$$\left.\begin{aligned}\alpha_1 + \alpha_2 &= \Delta A\\ (u-c)\alpha_1 + (u+c)\alpha_2 &= \Delta Q\end{aligned}\right\}\qquad [6.140]$$

Solving equations [6.140] for the wave strengths leads to:

$$\left.\begin{aligned}\alpha_1 &= \frac{u+c}{2c}\Delta A - \frac{\Delta Q}{2c}\\ \alpha_2 &= -\frac{u-c}{2c}\Delta A + \frac{\Delta Q}{2c}\end{aligned}\right\}\qquad [6.141]$$

The second equation [6.138] gives:

$$\left.\begin{aligned}(u-c)\alpha_1 + (u+c)\alpha_2 &= \Delta Q\\ (u-c)^2\alpha_1 + (u+c)^2\alpha_2 &= \Delta(uQ + P/\rho)\end{aligned}\right\}\qquad [6.142]$$

Note that the first equation [6.142] is equivalent to the second equation [6.140]. Solving equations [6.142] for c and u leads to:

$$\left.\begin{aligned} u &= \frac{A_{\mathrm{L}}^{1/2} u_{\mathrm{L}} + A_{\mathrm{R}}^{1/2} u_{\mathrm{R}}}{A_{\mathrm{L}}^{1/2} + A_{\mathrm{R}}^{1/2}} \\ c^2 &= \frac{g}{\rho} \frac{P_{\mathrm{R}} - P_{\mathrm{L}}}{A_{\mathrm{R}} - A_{\mathrm{L}}} \end{aligned}\right\} \qquad [6.143]$$

6.9. Multidimensional problems

6.9.1. *Explicit alternate directions*

Most classical numerical schemes were originally developed in one dimension of space. As well as for historical reasons, the numerical properties of one-dimensional schemes (consistency, solution stability, etc.) are easier to study than those of multidimensional schemes. Alternate directions, sometimes referred to as dimension splitting or time splitting techniques, allow multidimensional problems to be solved using one-dimensional techniques. Alternate directions techniques are analyzed in [STR 68] and [GOU 77]. Such techniques are mainly used on Cartesian grids.

Consider the following multidimensional problem:

$$\frac{\partial \mathrm{U}}{\partial t} + \frac{\partial \mathrm{F}}{\partial x} + \frac{\partial \mathrm{G}}{\partial y} + \frac{\partial \mathrm{H}}{\partial z} = \mathrm{S} \qquad [6.144]$$

where U is the conserved variable, F, G and H are the flux vectors in the x-, y- and z-direction respectively, and S is the source term. The solution U^n at the time level n is assumed to be known at all points of the computational grid. The purpose is to compute the solution U^{n+1} at the time level $n + 1$. The solution may be approximated as:

$$\mathrm{U}^{n+1} = \mathrm{L}_{\Delta t}^{(S)} \mathrm{L}_{\Delta t}^{(z)} \mathrm{L}_{\Delta t}^{(y)} \mathrm{L}_{\Delta t}^{(x)} \mathrm{U}^n \qquad [6.145]$$

where $\mathrm{L}_{\Delta t}^{(x)}$ is the numerical scheme (also referred to as "numerical operator") that solves the conservation part of equation [6.144] in the x-direction over the time step Δt. The conservation part of equation [6.144] in the x-direction is:

$$\frac{\partial \mathrm{U}}{\partial t} + \frac{\partial \mathrm{F}}{\partial x} = 0 \qquad [6.146]$$

and $\mathrm{L}_{\Delta t}^{(x)}$ is the numerical scheme that solves equation [6.146] numerically over the time step Δt via a formula that can be written in the form:

$$\mathrm{U}_i^{n+1} = \mathrm{L}_{\Delta t}^{(x)} \mathrm{U}_i^n \qquad [6.147]$$

By definition of the time derivative, the values of U at the time levels n and $n + 1$ are related by:

$$\begin{aligned} \mathrm{U}_i^{n+1} &\approx \mathrm{U}_i^n + \Delta t \left(\frac{\partial \mathrm{U}}{\partial t} \right)_i^{n+1/2} \\ &\approx \mathrm{U}_i^n - \Delta t \frac{\partial \mathrm{F}}{\partial x} (\mathrm{U}_i^n) \end{aligned} \qquad [6.148]$$

where $\partial \mathrm{F} / \partial x (\mathrm{U}_i^n)$ is the discretized form of $\partial \mathrm{F} / \partial x$ at the computational point i. Comparing equations [6.147] and [6.148] leads to the following definition for the operator $\mathrm{L}_{\Delta t}^{(x)}$:

$$\mathrm{L}_{\Delta t}^{(x)} = \mathrm{I} - \Delta t \frac{\partial \mathrm{F}}{\partial x} \qquad [6.149]$$

where I is the identity matrix. The operators $\mathrm{L}_{\Delta t}^{(y)}$ and $\mathrm{L}_{\Delta t}^{(z)}$ solve the conservation part of the equation in the y- and z-direction respectively and the operator $\mathrm{L}_{\Delta t}^{(S)}$ accounts for the contribution of the source term by solving:

$$\frac{\partial \mathrm{U}}{\partial t} = \mathrm{S} \qquad [6.150]$$

Equation [6.145] describes the following sequence. The conservation part of equation [6.144] is solved in the x-direction. The result is used as a starting point for the solution of the conservation part in the y-direction. The result of this computation is used as an initial state to solve the conservation part of the equation in the z-direction. The result of this step is used as an initial condition in the solution of equation [6.150].

The numerical solution is stable if each of the operators in sequence [6.145] is stable. The advection operators in the x-, y- and z-direction give stable solutions if the absolute value of the Courant number in the corresponding direction of space is smaller than one:

$$\max\left(\left|\mathrm{Cr}_x\right|, \left|\mathrm{Cr}_y\right|, \left|\mathrm{Cr}_z\right|\right) \leq 1 \qquad [6.151]$$

where Cr_x, Cr_y and Cr_z are respectively the Courant numbers in the x-, y- and z-direction. Sequence [6.145] is first-order accurate with respect to time.

When second- and higher-order schemes are used, the first-order time stepping sequence [6.145] may not be accurate enough. The accuracy of the solution may be improved by using the following sequence for the conservation part, as suggested in [STR 68]:

$$\mathrm{U}^{n+1} = \mathrm{L}^{(z)}_{\Delta t/2}\mathrm{L}^{(y)}_{\Delta t/2}\mathrm{L}^{(x)}_{\Delta t}\mathrm{L}^{(y)}_{\Delta t/2}\mathrm{L}^{(z)}_{\Delta t/2}\mathrm{U}^{n} \qquad [6.152]$$

where the subscripts $\Delta t/2$ indicate that the operators are applied over half a time step. Equation [6.152] can be generalized to account for the influence of the source term as:

$$\mathrm{U}^{n+1} = \mathrm{L}^{(S)}_{\Delta t/2}\mathrm{L}^{(z)}_{\Delta t/2}\mathrm{L}^{(y)}_{\Delta t/2}\mathrm{L}^{(x)}_{\Delta t}\mathrm{L}^{(y)}_{\Delta t/2}\mathrm{L}^{(z)}_{\Delta t/2}\mathrm{L}^{(S)}_{\Delta t/2}\mathrm{U}^{n} \qquad [6.153]$$

Since the operators $\mathrm{L}^{(y)}$ and $\mathrm{L}^{(z)}$ are used over half a time step, the stability criterion becomes:

$$\max\left(\left|\mathrm{Cr}_x\right|, \left|\mathrm{Cr}_y/2\right|, \left|\mathrm{Cr}_z/2\right|\right) \le 1 \qquad [6.154]$$

Sequence [6.153] allows spurious effects such as solution anisotropy to be reduced to a large extent compared to the first-order approach [6.145]. However, it is more time-consuming.

6.9.2. *The ADI method*

The Alternate Directions Implicit (ADI) method is iterative. It uses implicit one-dimensional schemes that allow large time steps to be used without making the numerical solution unstable. The schemes being one-dimensional, the linear algebraic systems to be solved remain diagonal and can be solved using fast system inversion techniques. Any implicit discretization of equation [6.144] can be written in the form:

$$\begin{aligned}\mathrm{U}^{n+1} + \frac{\partial \mathrm{F}(\mathrm{U}^{n+1})}{\partial x} + \frac{\partial \mathrm{G}(\mathrm{U}^{n+1})}{\partial y} + \frac{\partial \mathrm{H}(\mathrm{U}^{n+1})}{\partial z}& \\ -\mathrm{S}(\mathrm{U}^{n+1}) = \mathrm{R}(\mathrm{U}^{n})&\end{aligned} \qquad [6.155]$$

where $\mathrm{R}(\mathrm{U}^n)$ contains all the terms that are functions of the known solution at the time level n. Equation [6.155] is solved as follows:

1) The equation is solved in the x-direction, all the remaining terms in the equation being assumed known. In the first iteration, the derivatives of G and H must be "guessed". The simplest method is to use the value at the time level n. The following equation is solved:

$$\mathrm{U}^{n+1,x} + \frac{\partial \mathrm{F}(\mathrm{U}^{n+1,x})}{\partial x} = -\frac{\partial \mathrm{G}(\mathrm{U}^{n})}{\partial y} - \frac{\partial \mathrm{H}(\mathrm{U}^{n})}{\partial z} + \mathrm{S}(\mathrm{U}^{n}) + \mathrm{R}(\mathrm{U}^{n}) \qquad [6.156]$$

where $\mathrm{U}^{n+1,x}$ is the solution of equation [6.156] over the time step Δt.

2) The solution $\mathrm{U}^{n+1,x}$ is used as an initial condition for the equation in the y-direction:

$$\mathrm{U}^{n+1,y} + \frac{\partial \mathrm{G}(\mathrm{U}^{n+1,y})}{\partial y} = -\frac{\partial \mathrm{F}(\mathrm{U}^{n+1,x})}{\partial x} - \frac{\partial \mathrm{H}(\mathrm{U}^{n+1,x})}{\partial z} + \mathrm{S}(\mathrm{U}^{n+1,x}) + \mathrm{R}(\mathrm{U}^{n+1,x}) \qquad [6.157]$$

where $\mathrm{U}^{n+1,y}$ is the solution of equation [6.157] over the time step Δt.

3) The solution $\mathrm{U}^{n+1,y}$ is used as an initial condition for the equation in the z-direction:

$$\mathrm{U}^{n+1,z} + \frac{\partial \mathrm{G}(\mathrm{U}^{n+1,z})}{\partial y} = -\frac{\partial \mathrm{F}(\mathrm{U}^{n+1,y})}{\partial x} - \frac{\partial \mathrm{H}(\mathrm{U}^{n+1,y})}{\partial z} + \mathrm{S}(\mathrm{U}^{n+1,y}) + \mathrm{R}(\mathrm{U}^{n+1,y}) \qquad [6.158]$$

where $\mathrm{U}^{n+1,z}$ is the solution of equation [6.158] over the time step Δt.

4) The solution $\mathrm{U}^{n+1,z}$ is used as an initial condition for the contribution of the source term:

$$\mathrm{U}^{n+1,S} - \mathrm{S}(\mathrm{U}^{n+1,S}) = -\frac{\partial \mathrm{G}(\mathrm{U}^{n+1,z})}{\partial y} - \frac{\partial \mathrm{F}(\mathrm{U}^{n+1,z})}{\partial x} - \frac{\partial \mathrm{H}(\mathrm{U}^{n+1,z})}{\partial z} + \mathrm{R}(\mathrm{U}^{n+1,z}) \qquad [6.159]$$

where $\mathrm{U}^{n+1,S}$ is the solution of equation [6.159] over the time step Δt.

5) The solution $\mathrm{U}^{n+1,S}$ is used as an initial condition to solve the equation in the x-direction:

$$\mathrm{U}^{n+1,x} + \frac{\partial \mathrm{F}(\mathrm{U}^{n+1,x})}{\partial x} = -\frac{\partial \mathrm{G}(\mathrm{U}^{n+1,S})}{\partial y} - \frac{\partial \mathrm{H}(\mathrm{U}^{n+1,S})}{\partial z} + \mathrm{S}(\mathrm{U}^{n+1,S}) + \mathrm{R}(\mathrm{U}^{n+1,S}) \quad [6.160]$$

Steps 2–5, that form a single iteration, must be repeated until convergence is achieved, that is:

– the value of U at a given step (e.g. step 5) between two successive iterations should not differ by more than a given threshold value ε specified by the modeler;

– the values of U between two successive steps within the same iteration should not differ by more than the threshold value ε.

In many industrial implementations of the ADI method, only the first condition is checked. In other applications, the first condition is not checked and the second condition is checked only between two steps (e.g. between steps 2 and 3) instead of the complete sequence. When this is the case, the solution may be abnormally sensitive to the orientation chosen for the main axes of the grid. The ADI method is known to converge slowly when applied to nonlinear systems. For this reason, most industrial implementations of the method use a maximum permissible number of iterations, after which the sequence 2)–5) will be stopped regardless of the convergence of the iterative process. The maximum permissible number of iterations may be left to the modeler's choice or hard-programmed in the software. The fact that the user is not informed of or has no control over the parameters that influence the accuracy of the method may lead to erroneous computational results, thus destroying the predictive power of the simulation.

6.9.3. *Multidimensional schemes*

Multidimensional schemes solve multidimensional problems by treating all the spatial derivatives within one single step. When a two-dimensional problem is to be solved, the operators $\mathrm{L}^{(x)}$ and $\mathrm{L}^{(y)}$ introduced in the previous section are applied as follows:

$$\mathrm{U}^{n+1} = \left[\mathrm{L}_{\Delta t}^{(x)} + \mathrm{L}_{\Delta t}^{(y)} - \mathrm{I}\right]\mathrm{U}^{n} \quad [6.161]$$

where I is the identity matrix. Equation [6.161] is derived as follows. Extending equation [6.149] to the y-direction gives:

$$L_{\Delta t}^{(y)} = I - \Delta t \frac{\partial G}{\partial y} \qquad [6.162]$$

The discretized form of the equation to be solved is:

$$U_i^{n+1} = U_i^n - \Delta t \frac{\partial F}{\partial x}(U_i^n) - \Delta t \frac{\partial G}{\partial y}(U_i^n) \qquad [6.163]$$

It is easy to check that substituting equations [6.149] and [6.162] into equation [6.163] leads to equation [6.161]. Extending the reasoning to three dimensions of space leads to the following formula:

$$U^{n+1} = \left[L_{\Delta t}^{(x)} + L_{\Delta t}^{(y)} + L_{\Delta t}^{(z)} - 2I\right] U^n \qquad [6.164]$$

Since the effects of the operators in each direction of space are added within a single step, the stability constraint attached to multidimensional schemes is usually more restrictive than for alternate direction techniques. Most two-dimensional explicit schemes are subjected to the following constraint:

$$\max\left(\left|\mathrm{Cr}_x\right| + \left|\mathrm{Cr}_y\right|\right) \le 1 \qquad [6.165]$$

the three-dimensional version of which is:

$$\max\left(\left|\mathrm{Cr}_x\right| + \left|\mathrm{Cr}_y\right| + \left|\mathrm{Cr}_z\right|\right) \le 1 \qquad [6.166]$$

6.10. Summary

6.10.1. *What you should remember*

Finite difference methods are based on the discretization of space and time, which allows partial differential equations to be approximated in the form of differences between the values taken by the solution at predefined time and space coordinates. When the unknown value of the solution at the next time level can be expressed only as a function of the known value at the current time step, the method is said to be explicit. A discretization of the partial differential equation that provides a relationship between several unknowns at the next time level is said to be implicit. Explicit methods are subjected to stability constraints, while implicit methods are not in general.

The performance of a numerical method for hyperbolic conservation laws is conditioned by the Courant number, defined as the ratio of the distance covered by the wave over a time step to the distance between two adjacent computational points.

Characteristic-based methods, covered in section 6.2, solve the governing equations in characteristic form. The Riemann invariants are interpolated at the feet of the characteristics from their values at the computational points. Linear interpolations lead to first-order, diffusive numerical schemes. Parabolic interpolations lead to second-order, dispersive schemes. Characteristic-based methods may fail to preserve the conservation properties of the solution, especially in the presence of shocks and when the Courant number is not uniform over the computational domain.

Upwind schemes for scalar laws, presented in section 6.3, may be applied to the conservation and non-conservation form of the equations. In such schemes the derivative with respect to space is discretized using the point located upstream of the point at which the time derivative is discretized.

The Preissmann scheme presented in section 6.4 is a conservative scheme. It is used by a number of commercially available software packages for free-surface flow modeling. The degree of consistency of the discretization can be adjusted in time and space via two weighting parameters ψ and θ. The parameters ψ that control the weighting in space is usually set to 1/2. When this is the case, setting $\theta = 1/2$ leads to a purely dispersive scheme. Increasing θ induces numerical diffusion.

Centered schemes are dealt with in section 6.5. They are not sensitive to the direction in which the waves propagate. They are less diffusive than upwind schemes. The Crank-Nicholson scheme is an implicit scheme. Explicit centered schemes require the use of third- and higher-order Runge-Kutta time integration algorithms for stability reasons. Their range of stability is wider than that of classical explicit schemes.

TVD schemes for scalar laws are presented in section 6.6. They can be seen as weighted combinations between the upwind scheme and the second-order Lax-Wendroff scheme. The weighting between the two schemes is a function of the local variations in the slope of the variable. The contributions of the gradient of the variable are limited using a so-called limiter, for which many formulations have been proposed in the literature.

Flux splitting techniques are presented in section 6.7. Such techniques allow upwind and TVD schemes to be generalized to hyperbolic systems of conservation laws. They use the non-conservation form of the equations. The flux splitting approach consists of separating the Jacobian matrix of the flux into two matrices, one that accounts for the waves propagating in the direction of negative *x*, the other accounting for the waves that propagate in the direction of positive *x*.

The flux splitting technique allows conservation to be preserved provided that the Jacobian matrix satisfies a number of criteria. Roe's linearization technique, presented in section 6.8, provides guidelines for the derivation of the Jacobian matrix.

Several options are available for the treatment of multidimensional problems (see section 6.9). The alternate direction technique consists of solving the governing equations in each direction of space successively. Multidimensional schemes treat all the directions of space within a single step.

6.10.2. *Application exercises*

6.10.2.1. *Exercise 6.1: finite difference methods for scalar laws*

Check the conclusions of Exercises 1.1 to 1.5 using finite difference methods. The following methods are advised:

– a characteristic-based method,

– an upwind scheme (conservative version),

– Preissmann's scheme,

– a TVD scheme.

Indications and searching tips for the solution of this exercise can be found at the following URL: http://vincentguinot.free.fr/waves/exercises.htm.

6.10.2.2. *Exercise 6.2: finite difference methods for hyperbolic systems*

Check the conclusions derived in Exercises 2.2 and 2.5 by solving the equations numerically. The first-order MOC is advised for the sake of simplicity.

Indications and searching tips for the solution of this exercise can be found at the following URL: http://vincentguinot.free.fr/waves/exercises.htm.

6.10.2.3. *Exercise 6.3: finite difference methods for hyperbolic systems*

Implement the numerical schemes used in the examples presented in sections 6.2.3, 6.2.4, 6.4.3, 6.6.5 and 6.7.2.3.

Indications and searching tips for the solution of this exercise can be found at the following URL: http://vincentguinot.free.fr/waves/exercises.htm.