An introduction to finite volumes for gas dynamics

François Dubois

Conservatoire National des Arts et Métiers
15 rue Marat, F-78 210 Saint Cyr l'Ecole, France.
and
Centre National de la Recherche Scientifique
Laboratoire ASCI, bât. 506, BP 167, F-91 403 Orsay Cedex.

October 2000 □

Summary

We propose an elementary introduction to the finite volume method in the context of gas dynamics conservation laws. Our approach is founded on the advection equation, the exact integration of the associated Cauchy problem, and the so-called upwind scheme in one space dimension. It is then extended in three directions : hyperbolic linear systems and particularily the system of acoustics, gas dynamics with the help of the Roe matrix and two space dimensions by following the approach proposed by Van Leer. A special emphasis on boundary conditions is proposed all along the text.

AMS Subject Classification: 35L40, 35L60, 35L65, 35Q35, 76N15.

Keywords: Advection, Characteristics, Roe matrix, Van Leer method.

□ Research report CNAM-IAT n°342-2000. Published with the title "An Introduction to Finite Volumes Methods" in *Encyclopedia Of Life Support Systems* (EOLSS, Unesco), Mathematical Sciences, Computational Methods and Algorithms, Vladimir V. Shaidurov and Olivier Pironneau Editors, volume 2, p. 36-105, 2009. Present edition 21 January 2011.

Contents

• *Acknowledgments.* The author thanks Alexandre Gault, listener at the spring 2000 "lectures in computational acoustics" at the Conservatoire National des Arts et Métiers (Paris, France), for providing his personal manuscript notes.

1) Advection equation and method of characteristics.

1.1 Advection equation.

• We consider a given real number $a > 0$ and we wish to solve the so-called advection equation of unknown function $u(x, t)$:

(1.1.1) $$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0 , \qquad t \geq 0 , \quad x \in \mathbb{R} .$$

We first look to the homogeneity coherence of the different terms of equation (1.1.1). On one hand, the ratio $\frac{\partial u}{\partial t}$ is homogeneous to the dimension $[u]$ of function $u(\bullet, \bullet)$ divided by the dimension $[t]$ of the time and we have : $\frac{\partial u}{\partial t} \sim \frac{[u]}{[t]}$. On the other hand the expression $a \frac{\partial u}{\partial x}$ is homogeneous to the dimension $[a]$ of scalar a multiplied by the ratio $\frac{[u]}{[x]}$ and we have $a \frac{\partial u}{\partial x} \sim [a] \frac{[u]}{[x]}$. From equation (1.1.1), the two previous terms $\frac{\partial u}{\partial t}$ and $a \frac{\partial u}{\partial x}$ have the same dimension and we deduce from the previous formulae the equality : $\frac{1}{[t]} \sim \frac{[a]}{[x]}$. Then we have established that the constant a is homogeneous to a **celerity** :

(1.1.2) $$[a] \sim \frac{[x]}{[t]} .$$

• The Cauchy problem for the model equation (1.1.1) is composed by the equation (1.1.1) itself and the following initial condition :

(1.1.3) $$u(x, 0) = u_0(x) , \qquad x \in \mathbb{R} ,$$

where $\mathbb{R} \ni x \longmapsto u_0(x) \in \mathbb{R}$ is some given function. We observe that the solution of equation (1.1.1) is constant along the characteristic (straight) lines that satisfy the differential equation

(1.1.4) $$\frac{\mathrm{d}x}{\mathrm{d}t} = a .$$

Proposition 1.1. The solution is constant along the characteristic lines. Let $0 \leq \lambda \leq t$ be some given parameter and $u(\bullet, \bullet)$ a solution of equation (1.1.1). Then function $u(\bullet, \bullet)$ is constant along the characteristic lines, *i.e.*

(1.1.5) $$u(x - a\lambda, t - \lambda) = u(x, t) , \qquad \forall \; x, t, \lambda .$$

• The **proof of Proposition 1.1** is obtained as follows. We consider a fixed point (x, t) in space-time $\mathbb{R} \times [0, +\infty[$ and the auxiliary function $[0, t] \ni \lambda \longmapsto v(\lambda) = u(x - a\lambda, t - \lambda)$. We have, due to the usual chain rule for derivation of operators :

$$\frac{\mathrm{d}v}{\mathrm{d}\lambda} = \Big[(-a) \frac{\partial u}{\partial x} - \frac{\partial u}{\partial t} \Big] (x - a\lambda, t - \lambda) = 0$$ if function $u(\bullet, \bullet)$ is solution of the advection equation (1.1.1). Then $v(\lambda)$ does not depend on variable λ and we have in particular $v(\lambda) = v(0)$, which exactly expresses the relation

(1.1.5). We have in particular for $\lambda = t$: $u(x,\, t) \,=\, u(x - at,\, 0) \,=\, u_0(x - at)$ as illustrated on Figure 1.1. □

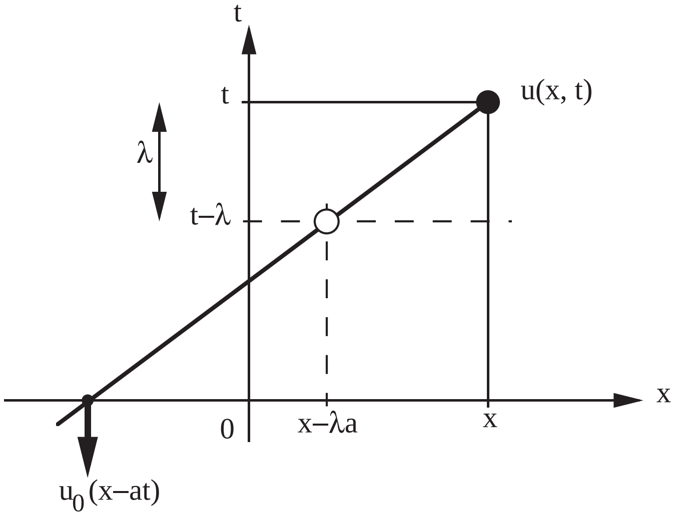

Figure 1.1. The solution $u(x,t)$ of the advection equation is constant along the characteristic lines.

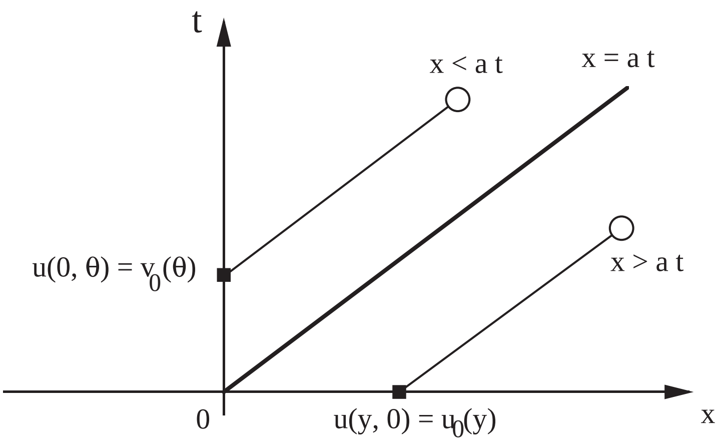

Figure 1.2. Initial-boundary value problem for the advection equation.

1.2 Initial-boundary value problems for the advection equation.

• The second step is concerned by the so-called initial-boundary value problem considered for $x > 0$ and $t > 0$ with some given initial condition $u_0(x)$ for $t = 0$ and a boundary condition $v_0(t)$ for $x = 0$:

(1.2.1) $$\frac{\partial u}{\partial t} + a\, \frac{\partial u}{\partial x} \,=\, 0\,, \qquad t > 0\,, \quad x > 0\,, \qquad \text{(equation)}$$

(1.2.2) $$u(x,\, 0) \,=\, u_0(x)\,, \qquad x > 0\,, \qquad \text{(initial condition)}$$

(1.2.3) $\qquad u(0,\, t) \;=\; v_0(t)\,, \qquad\qquad t > 0\,, \qquad\qquad$ (boundary condition).

Proposition 1.2. Advection in the quadrant $x > 0$ and $t > 0$.
We suppose that $a > 0$. Then the solution of the advection equation (1.2.1) with the initial condition (1.2.2) and the boundary condition (1.2.3) is given by the relations

(1.2.4) $\qquad u(x,\, t) \;=\; u_0(x - at)\,, \qquad x - at \,>\, 0$

(1.2.5) $\qquad u(x,\, t) \;=\; v_0\Big(t - \dfrac{x}{a}\Big)\,, \qquad x - at \,<\, 0\,.$

The initial condition $u_0(\bullet)$ is advected towards space-time point $(x,\, t)$ when $x - at > 0$ and the boundary condition $v_0(\bullet)$ is activated for $x - at < 0$.

• **Proof of Proposition 1.2.**
In order to solve the problem (1.2.1)-(1.2.3), we use the method of characteristics. We fix a point $(x,\, t)$ of space-time domain that satisfies $x > 0,\, t > 0$ and we go upstream in time with the help of the characteristic line that goes through this point (see Figure 1.2) :

(1.2.6) $\qquad x(\lambda) \,=\, x - a\lambda\,, \qquad t(\lambda) \,=\, t - \lambda\,.$

• First case : $x - at \,>\, 0$. When we take the particular value $\lambda = t$ in the previous relation (1.2.6), the particular point $y \,=\, x(t) \,=\, x - at$ on the axis of abscissa is strictly positive then the initial condition $u_0(y)$ is well defined. The solution $u(\bullet,\, \bullet)$ is constant on the characteristic line (see Proposition 1.1) that contains this particular point. Then relation (1.2.4) is established.

• Second case : $x - at \,<\, 0$. We consider the particular value $\lambda = \frac{x}{a}$ inside the expression (1.2.6). Then the corresponding foot of the characteristic belongs to the time axis : $\theta \,=\, t - \lambda = t - \frac{x}{a}$ and $\theta > 0$ due to the inequalities $x < at$ and $a > 0$. The solution is constant along the characteristic line going through this point and the relation (1.2.5) is established. □

• In the particular case where datum $u_0(x)$ is identically equal to zero, *i.e.*

(1.2.7) $\qquad u_0(x) \;=\; 0\,, \qquad\qquad x > 0\,,$

and if the boundary condition $v_0(t)$ is sinusoïdal for time positive to fix the ideas,

(1.2.8) $\qquad v_0(t) \;=\; \sin(\omega t)\,, \qquad t > 0\,,$

the solution of the advection equation in the domain $x > 0,\, t > 0$ via the relations (1.2.4) and (1.2.5) can be considered with the two following view points.

(i) We take a snap shot of the solution $u(\bullet,\, \bullet)$ at a fixed time $T > 0$. We consider the partial function $[0,\, +\infty[\; \ni x \longmapsto u(x,\, T) \in \mathbb{R}$ and taking into account the relations (1.2.4), (1.2.5), (1.2.7) and (1.2.8), we have

(1.2.9) $$u(x,\, T) \;=\; \begin{cases} \sin\Big[\omega\big(T - \dfrac{x}{a}\big)\Big]\,, & x \,<\, aT \\ 0\,, & x \,>\, aT\,. \end{cases}$$

and this function is illustrated on Figure 1.3.

(ii) We fix a particular position X in space and we look, as time is increasing, to the solution $u(\bullet, \bullet)$ at this particular point. We show on Figure 1.4 the function $[0, +\infty[\ni t \longmapsto u(X, t) \in \mathbb{R}$ and taking into account the relations (1.2.4), (1.2.5), (1.2.7) and (1.2.8), we have

$$ u(x, T) = \begin{cases} 0 , & t < \frac{X}{a} \\ \sin\Big[\omega\big(T - \frac{x}{a}\big)\Big] , & t > \frac{X}{a} \end{cases} . \tag{1.2.10} $$

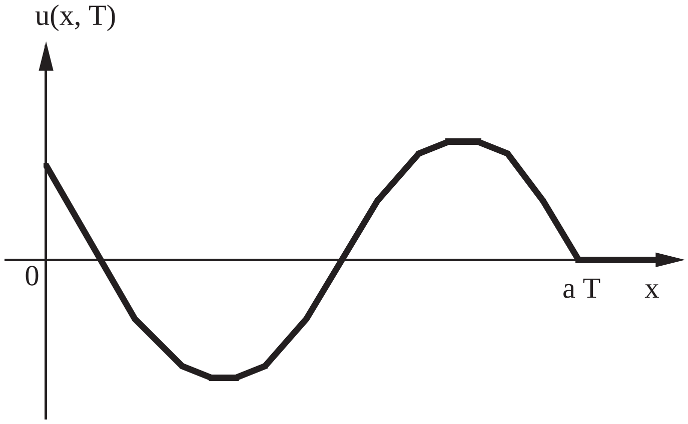

Figure 1.3. Snap shot of the solution of the advection equation at time $t = T$.

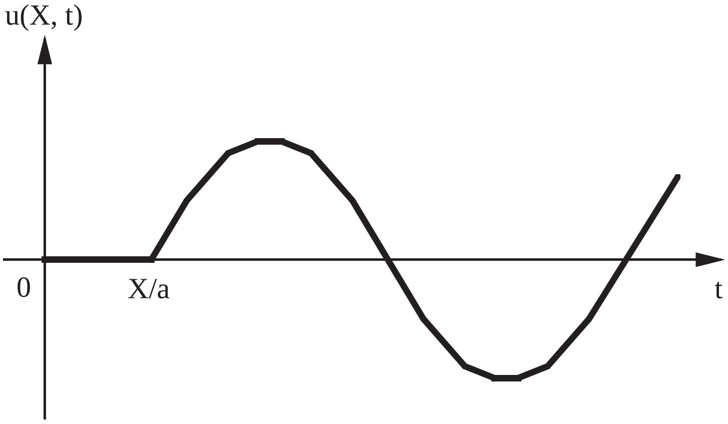

Figure 1.4. Evolution of the solution at the particular point $x = X$.

1.3 Inflow and outflow for the advection equation.

• We still suppose that celerity a is positive and we consider the resolution of the advection (1.2.1) in the space-time domain

$$ 0 < x < L , \qquad t > 0 . \tag{1.3.1} $$

The relations (1.2.4) and (1.2.5) can still be applied because the proof of Proposition 1.2 remains unchanged in this particular case. As a consequence of the previous property, we remark that **no boundary condition** is necessary at the particular position $x = L$ for solving the advection problem in the space-time domain defined in relations (1.3.1). The initial condition (1.2.2) has simply to be restricted in domain $]0, L[$:

(1.3.2) $$u(x, 0) = u_0(x), \qquad 0 < x < L,$$

and the boundary condition (1.2.3) at $x = 0$ remains unchanged :

(1.3.3) $$u(0, t) = v_0(t), \qquad t > 0.$$

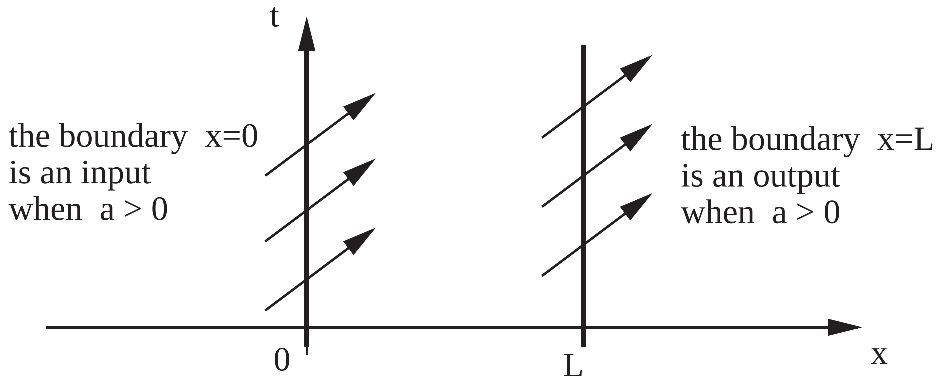

Figure 1.5. Initial-boundary value problem for the advection equation with $a > 0$ in the domain $0 < x < L$ and $t > 0$.

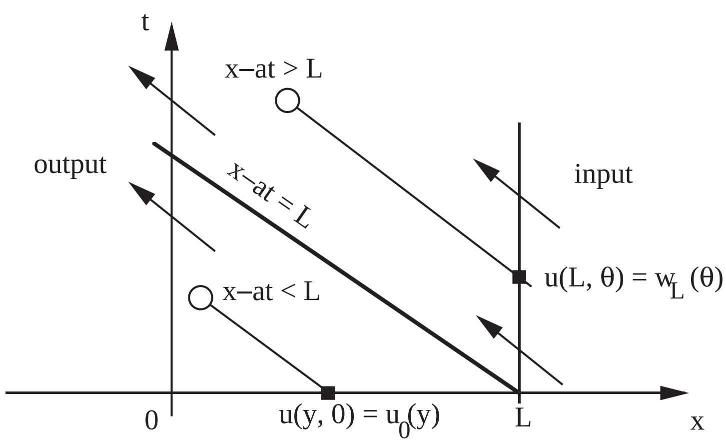

Figure 1.6. Initial-boundary value problem for the advection equation with $a < 0$ in the domain $0 < x < L$ and $t > 0$.

• The difference between point $x = 0$ and point $x = L$ for the resolution of the advection equation in space-time domain (1.3.1) is due to the fact that we choose an orientation of the characteristic lines $x - at =$ constant associated to an **increase** for the time direction. With this choice of time direction, the characteristic lines **enter** inside the space-time domain (1.3.1) at $x = 0$ and they go outside at $x = L$. The boundary condition (1.3.3) is given at the input of the domain (see Figure 1.5) and at $x = L$, there is a free output from space time domain (1.3.1), without necessity to specify any numerical boundary condition.

• If we change the sign of celerity a, *i.e.* if we suppose now

(1.3.4) $\qquad a < 0$,

the above analysis remains unchanged, but the algebraic relations (1.2.4) and (1.2.5) have to be modified (see Figure 1.6). We still start from relation (1.1.5) that expresses that the solution of the advection equation (1.1.1) is constant along the characteristics lines. The foot of the characteristic line that contains the particular point $(x,\, t)$ in space-time is either the point $(y = x - at,\, 0)$ if $x - at < L$, either the point $\big(L,\, \theta = t - \frac{1}{a}(x - L)\big)$ if $x - at > L$. In the first case, we have $y > 0$ and $\theta < 0$ then the initial condition (1.3.2) is advected inside the domain (1.3.1) and we have :

(1.3.5) $\qquad u(x,\, t) \;=\; u_0(x - at)\,, \qquad\qquad x - at \,<\, L\,.$

• On the contrary, if $x - at > L$, we have $y > L$ and $\theta > 0$ then the boundary condition at $x = L$ that takes now the expression

(1.3.6) $\qquad u(L,\, t) \;=\; w_L(t)\,, \qquad t > 0\,,$

is advected inside the domain of study and we have :

(1.3.7) $\qquad u(x,\, t) \;=\; w_L\Big(t + \dfrac{L}{a} - \dfrac{x}{a}\Big)\,, \qquad x - at \,>\, L\,.$

We have established the following

Proposition 1.3. Advection in the domain $0 < x < L$, $a < 0$.

Under the hypothesis (1.3.4), the resolution of the advection equation (1.2.1) in the space-time domain (1.3.1) conducts to a **well posed** problem when we introduce the initial condition (1.3.2) on the interval $]0,\, L[$ and the boundary condition (1.3.6) at the **input** region located at $x = L$, without any boundary condition at the output located at $x = 0$. The solution of Problem (1.2.1), (1.3.2) and (1.3.6) is given by the relations (1.3.5) and (1.3.7).

2) Finite volumes for linear hyperbolic systems.

2.1 Linear advection.

• We still study the advection equation parameterized by some celerity $a > 0$:

(2.1.1) $\qquad \dfrac{\partial W}{\partial t} + \dfrac{\partial}{\partial x}\big(a\, W\big) \;=\; 0\,, \qquad t > 0\,, \qquad x \in \mathbb{R}\,,$

and we search a discrete version of this mathematical model. For doing this, we introduce a space step $\Delta x > 0$ and a space grid composed by points x_j whose coordinates are multiples of this space step Δx, *id est*

(2.1.2) $\qquad x_j \,=\, j\,\Delta x\,, \qquad j \in \mathbb{Z}\,.$

For a finite domain, $]0,\,L[$ to fix the ideas, the above grid is limited to integer values j such that

(2.1.3) $\qquad 0 \,\leq\, j \,\leq\, J \,=\, \dfrac{L}{\Delta x}$

and the vertices $(x_j)_{0\leq j\leq J}$ are usually used in the context of the finite difference method. The intervals $K_{j+1/2} \,=\,]x_j,\, x_{j+1}[$ between two vertices can be considered as finite elements (or finite volumes in our study) and they cover the entire domain $]0,\,L[$:

(2.1.4) $\qquad [0,\,L] \,=\, \displaystyle\bigcup_{0\leq j\leq J-1} [x_j\,,\,x_{j+1}]\,,$

as proposed in the general context of meshes (see *e.g.* Ciarlet [Ci78]). We introduce also a time step $\Delta t > 0$ and the discrete time values at integer multiples of the above quantum :

(2.1.5) $\qquad t^n \,=\, n\,\Delta t\,, \qquad n \in \mathbb{N}\,.$

We consider now a space-time volume $V_{j+1/2}^{n+1/2}$ obtained by cartesian product of the two intervals $]x_j\,,\,x_{j+1}[$ and $]t^n,\,t^{n+1}[$ (see Figure 2.1) :

(2.1.6) $\qquad V_{j+1/2}^{n+1/2} \,=\,]x_j\,,\,x_{j+1}[\,\times\,]t^n,\,t^{n+1}[\,.$

• The finite volume scheme consists simply in integrating the advection equation (2.1.1) inside the space-time domain $V_{j+1/2}^{n+1/2}$ introduced previously :

(2.1.7) $\qquad \displaystyle\int_{V_{j+1/2}^{n+1/2}} \Big[\frac{\partial W}{\partial t} \,+\, \frac{\partial}{\partial x}(a\,W) \Big]\,{\rm d}x\,{\rm d}t \,=\, 0\,, \quad 0\leq j\leq J\,, \qquad n\geq 0\,.$

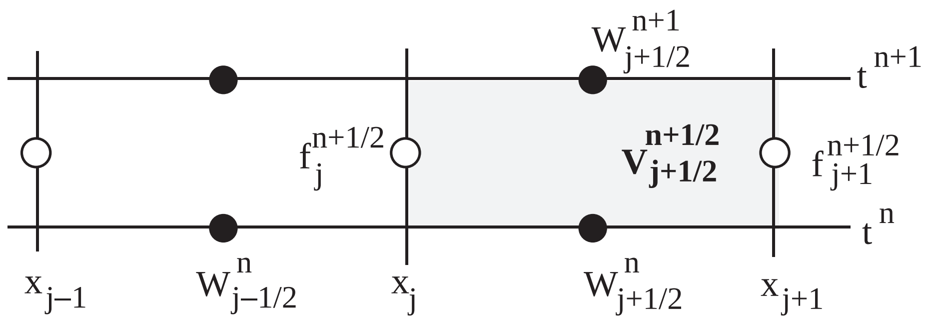

Figure 2.1. Space-time grid for the finite volume method.

Proposition 2.1. Finite volume scheme.

Let $\mathbb{R} \times [0,\,+\infty[\, \ni (x,\,t) \longmapsto W(x,\,t) \in \mathbb{R}$ be a solution of the advection equation (2.1.1). We introduce the space mean value $W_{j+1/2}^n$ of this solution $W(\bullet,\,\bullet)$ in the cell $K_{j+1/2}$:

(2.1.8) $$W_{j+1/2}^{n} \,=\, \frac{1}{\mid K_{j+1/2} \mid} \int_{x_j}^{x_{j+1}} W(x,\, t^n)\, \mathrm{d}x$$

and the time mean value $f_j^{n+1/2}$ of the so-called flux $a\, W(\bullet,\, \bullet)$ at the space position x_j and between discrete times t^n and t^{n+1} :

(2.1.9) $$f_j^{n+1/2} \,=\, \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} (a\, W)(x_j,\, t)\, \mathrm{d}t\,.$$

Then we have the following constitutive relation of finite volumes schemes :

(2.1.10) $$\frac{1}{\Delta t}\big(W_{j+1/2}^{n+1} - W_{j+1/2}^{n}\big) \,+\, \frac{1}{\Delta x}\big(f_{j+1}^{n+1/2} - f_j^{n+1/2}\big) \,=\, 0\,.$$

This numerical modelling characterizes the so-called finite volume method which has been developed thanks to the work of S. Godunov [Go59], Godunov et al [GZIKP79], Patankar [Pa80], Harten, Lax and Van Leer [HLV83] or Faille, Gallouët and Herbin [FGH91] among others.

• The **proof of Proposition 2.1** consists in a precise evaluation of the left hand side of equality (2.1.7). We use Fubini rule for the computation of double integrals and we begin by integrating in time for the $\frac{\partial}{\partial t}$ term :

$$\int_{V_{j+1/2}^{n+1/2}} \frac{\partial W}{\partial t}\, \mathrm{d}x\, \mathrm{d}t \,=\, \int_{x_j}^{x_{j+1}} \Big[\int_{t^n}^{t^{n+1}} \frac{\partial W}{\partial t}(x,\, t)\, \mathrm{d}t \Big]\, \mathrm{d}x$$

$$= \int_{x_j}^{x_{j+1}} \Big[W(x,\, t^{n+1}) - W(x,\, t^n) \Big]\, \mathrm{d}x \,=\, \Delta x \Big[W_{j+1/2}^{n+1} - W_{j+1/2}^{n} \Big]$$

due to the definition (2.1.8). We proceed in an analogous way with the $\frac{\partial}{\partial x}$ term and begin now the Fubini procedure by integrating in space ; we have

$$\int_{V_{j+1/2}^{n+1/2}} \frac{\partial}{\partial x}\big(a\, W\big)\, \mathrm{d}x\, \mathrm{d}t \,=\, \int_{t^n}^{t^{n+1}} \Big[\int_{x_j}^{x_{j+1}} \frac{\partial}{\partial x}\, \big(a\, W\big)(x,\, t)\, \mathrm{d}x \Big]\, \mathrm{d}t$$

$$= \int_{t^n}^{t^{n+1}} \Big[\big(a\, W\big)(x_{j+1},\, t) - \big(a\, W\big)(x_j,\, t) \Big]\, \mathrm{d}t \,=\, \Delta t\, \Big[f_{j+1}^{n+1/2} - f_j^{n+1/2} \Big]$$

according to the definition (2.1.9). We add the two previous results, use identity (2.1.7) and divide by $\Delta t\, \Delta x$. We obtain exactly the relation (2.1.10). □

• The relation (2.1.10) is a very general form for the evolution of the mean values $W_{j+1/2}$ between two time steps. The increment $\big(W_{j+1/2}^{n+1} - W_{j+1/2}^{n}\big)$ is, after correction by a multiplicative factor, equilibrated by the flux difference $\big(f_{j+1}^{n+1/2} - f_j^{n+1/2}\big)$. The idea of a finite volume scheme is to consider now that the algebraic object $W_{j+1/2}$ is nomore the mean value of the exact solution but an **approximation** of this mean value. Then the relation (2.1.10) proposes a numerical scheme for the discrete evolution of the approximated mean values $W_{j+1/2}\,,\; j=0,\cdots,\,J-1$. Nevertheless, the numerical scheme is **not** entirely defined by the relation (2.1.10). Starting from mean values at the initial time step, *i.e.*

$$W^0_{j+1/2} \;=\; \frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} W_0(x)\,\mathrm{d}x\;, \qquad j=0,\cdots,\,J-1\,, \tag{2.1.11}$$

we are able to increment the time step with relation (2.1.10) only if all the fluxes $f_j^{n+1/2}\,,\; j=0,\cdots,\,J$ have been *a priori* first determined as a functional of the previous values. In a very general way, we say that the finite volume scheme (2.1.10) is an **explicit scheme** if each flux $f_j^{n+1/2}$ is a given function Ψ_j of the mean values $\big(W^n_{k+1/2}\big)_{k=1,\cdots,\,J-1}$ at the preceding time step number n :

$$f_j^{n+1/2} \;=\; \Psi_j\big(\{W^n_{k+1/2},\,k=0,\cdots,\,J-1\}\big)\,, \qquad j=0,\,\cdots,\,J-1\,. \tag{2.1.12}$$

The function Ψ_j is called the **local numerical flux function** at point x_j and, joined with the evolution equation (2.1.10), its choice determines the numerical scheme.

• A natural hypothesis claims that we have **translation invariance** for the evaluation of the flux if we move the discrete data in the same way ; in other words, the numerical flux function Ψ_j only depends on the p first neighbors of the interface x_j. Then the explicit numerical flux is a given function Φ of the p first neighbors and we have :

$$f_j^{n+1/2} \;=\; \Phi\big(W^n_{j+1/2-p},\,\cdots,\,W^n_{j-1/2},\,W^n_{j+1/2},\,\cdots,\,W^n_{j+1/2+p-1}\big)\,. \tag{2.1.13}$$

A very important particular case is one of a two-point scheme for the evaluation of the numerical flux. We have in this particular case :

$$f_j^{n+1/2} \;=\; \Phi\big(W^n_{j-1/2},\,W^n_{j+1/2}\big)\,. \tag{2.1.14}$$

With this particular choice, the numerical scheme for incrementing in time of the mean values takes the form :

$$\left\{\begin{array}{l} \dfrac{1}{\Delta t}\big(W^{n+1}_{j+1/2}-W^n_{j+1/2}\big)\;+ \\ \qquad\qquad +\;\dfrac{1}{\Delta x}\Big(\Phi\big(W^n_{j+1/2},\,W^n_{j+3/2}\big)-\Phi\big(W^n_{j-1/2},\,W^n_{j+1/2}\big)\Big)\;=\;0\,. \end{array}\right. \tag{2.1.15}$$

It is also a three-point finite difference scheme. The finite volume scheme (2.1.10) (2.1.13) is said to be **consistent** with the advection equation (2.1.1) when the numerical flux function Φ satisfies the condition

$$\Phi\big(W,\,\cdots,\,W,\,W,\,\cdots,\,W\big) \;=\; a\,W\;, \qquad \forall\,W\in\mathbb{R}\,. \tag{2.1.16}$$

• The crucial question is how to choose a numerical finite volume scheme. The simplest choice consists in a two point explicit scheme such that the finite difference scheme is identical to the upstream-centered scheme (see *e.g.* Richtmyer-Morton [RM67]). It takes the following expressions :

$$(2.1.17) \qquad \frac{1}{\Delta t}\big(W_{j+1/2}^{n+1} - W_{j+1/2}^{n}\big) \;+\; a\Big(W_{j+1/2}^{n} - W_{j-1/2}^{n}\Big) \;=\; 0\,, \qquad a > 0$$

$$(2.1.18) \qquad \frac{1}{\Delta t}\big(W_{j+1/2}^{n+1} - W_{j+1/2}^{n}\big) \;+\; a\Big(W_{j+3/2}^{n} - W_{j+1/2}^{n}\Big) \;=\; 0\,, \qquad a < 0\,.$$

The corresponding flux function is called the **first order upstream-centered flux**, is simply given by the following relations :

$$(2.1.19) \qquad \Phi(W_l,\, W_r) \;=\; \begin{cases} a\, W_l\,, & a > 0 \\ a\, W_r\,, & a < 0\,. \end{cases}$$

When this flux function acts at a given point x_j of the mesh, we have :

$$(2.1.20) \qquad f_j^{n+1/2} \;=\; \Phi\big(W_{j-1/2}^{n},\, W_{j+1/2}^{n}\big) \;=\; \begin{cases} a\, W_{j-1/2}^{n}\,, & a > 0 \\ a\, W_{j+1/2}^{n}\,, & a < 0\,. \end{cases}$$

If $a > 0,$ the exact solution of the advection equation propagates the information from the left to the right ; the flux at the interface x_j is issued from the cell at the left of the interface and this cell at the number $j-1/2$. If $a < 0,$ the propagation of the information with the advection equation is from right to left ; the interface flux at the abscissa x_j is due to the control volume on the right, *i.e.* with number $j+1/2$ as depicted on Figure 2.2.

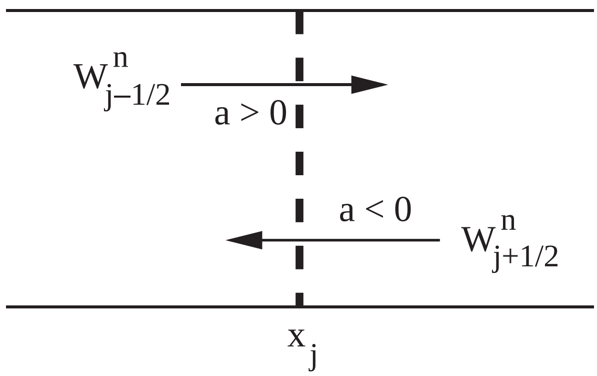

Figure 2.2. Upwinding of the information for the advection equation.

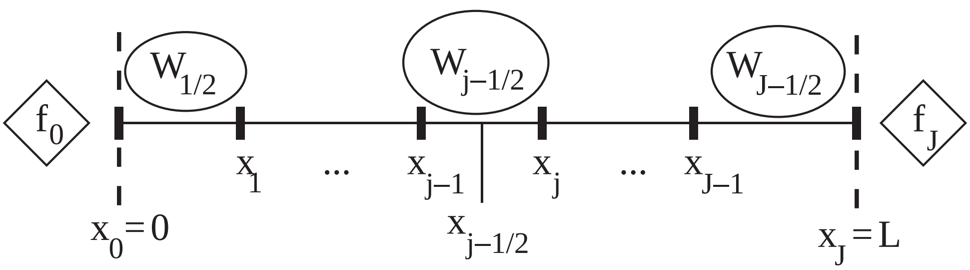

Figure 2.3. Notations for the one-dimensional finite volume method.

• Recall that practical use of the upwind finite volume scheme like (2.1.17) when $a > 0$ or (2.1.18) if $a < 0$ is restricted to the usual Courant-Friedrichs-Lewy stability condition :

(2.1.21) $$a \, \frac{\Delta t}{\Delta x} \, \leq \, 1$$

as developed *e.g.* in the book of Richtmyer and Morton [RM67].

2.2 Numerical flux boundary conditions

• In this section, we focus on the problem of the numerical boundary conditions. Recall that we study the advection equation in the space domain $[0,\, L]$:

(2.2.1) $$0 \leq x \leq L$$

and $J = \frac{L}{\Delta x} \in \mathbb{N}$ control cells (or finite elements) have been used to define a mesh :

(2.2.2) $$J \, \Delta x \, = \, L \, .$$

Note that the $j^{\rm th}$ cell is exactly the interval $]x_{j-1},\, x_j[$ and it is centered at point $x_{j-1/2}$ as shown on Figure 2.3.

• At time step $n\,\Delta t\,,$ the discrete field is entirely known and is composed of all the values $W^n_{j-1/2}$ for $j = 1, \cdots, J\,.$ With a flux function $\Phi(\bullet, \bullet)$ as proposed at relation (2.1.14), we observe that the two boundary fluxes $f_0^{n+1/2}$ and $f_J^{n+1/2}$ are **not** a priori defined because states $W^n_{-1/2}$ or $W^n_{J+1/2}$ does not exist. The situation is more complex with numerical fluxes that use four points or more as proposed in (2.1.13) and will not be detailed in this section. Even if the formula giving the numerical flux at the boundaries has to be specifically studied, the finite volume scheme remains defined by the relation (2.1.10) and we have for the two cells encountering the boundary :

(2.2.3) $$\frac{1}{\Delta t}\big(W^{n+1}_{1/2} - W^n_{1/2}\big) \, + \, \frac{1}{\Delta x}\big(f_1^{n+1/2} - f_0^{n+1/2}\big) \qquad = \, 0 \,,$$

(2.2.4) $$\frac{1}{\Delta t}\big(W^{n+1}_{J-1/2} - W^n_{J-1/2}\big) \, + \, \frac{1}{\Delta x}\big(f_J^{n+1/2} - f_{J-1}^{n+1/2}\big) \, = \, 0 \,.$$

• The question is now to adapt the relation (1.2.14) in order to determine the two **boundary fluxes** $f_0^{n+1/2}$ at the left of the domain and $f_J^{n+1/2}$ at the right. For the advection equation with celerity $a > 0,$ we have observed in the first section that some boundary condition $v_0(t)$ has to be assigned at $x = 0$ and it is not the case for $x = L.$ It is therefore natural to take into account this information at the input of the domain and to set :

(2.2.5) $$f_0^{n+1/2} \, = \, \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} a \, v_0(t) \, {\rm d}t$$

or simply

(2.2.6) $$f_0^{n+1/2} \,=\, a\, v_0\big((n+\tfrac{1}{2})\Delta t\big)\,, \qquad a > 0\,,$$

if function $t \longmapsto v_0(t)$ has a slow time variation at the scale defined by the time step. At the output $x = L$, no numerical datum has to be assigned to set correctly the continuous mathematical problem. We must maintain this property if we wish the numerical method to follow the mathematical physics as efficiently as possible. A simple boundary flux is associated with the previous numerical upwind scheme. For $x = x_J = L$ and $a > 0$, we observe that the upwind scheme (2.1.20) is simply written as :

(2.2.7) $$f_J^{n+1/2} \,=\, a\, W_{J-1/2}^n\,, \qquad a > 0\,,$$

and this relation (2.2.7) defines a **first order extrapolated** boundary flux.

• The roles are reversed when $a < 0$. The abscissa $x = 0$ corresponds to an output for the advection equation and the right boundary $x = L$ is an input where a time field $t \longmapsto w_L(t)$ is given. In the first case, the upwind scheme (2.1.20) can be applied without modification :

(2.2.8) $$f_0^{n+1/2} \,=\, a\, W_{1/2}^n\,, \qquad a < 0\,,$$

and it corresponds to a first order extrapolation of the internal data $\big\{W_{j-1/2}^n,\ j = 1, \cdots, J\big\}$ at the boundary at time step $n\Delta t$. For $x = L$, the boundary flux $f_J^{n+1/2}$ uses the given information between the two time steps :

(2.2.9) $$f_J^{n+1/2} \,=\, a\, w_L\big((n+\tfrac{1}{2})\Delta t\big)\,, \qquad a < 0\,.$$

Proposition 2.2.

Flux boundary conditions for the advection equation.

When we approach the advection equation (2.1.1) with the finite volume method, the numerical boundary conditions induces a choice for the two boundary fluxes $f_0^{n+1/2}$ and $f_J^{n+1/2}$. When $a > 0$, the boundary condition $v_0(t)$ at the input can be introduced into the boundary with the relation (2.2.6) and the free output at the right can be treated with an extrapolation of the type (2.2.7). When $a < 0$, the free output at the left of the domain can be taken into account with the help of relation (2.2.8) whereas the input condition $w_L(t)$ at the right can be introduced thanks to relation (2.2.9).

2.3 A model system with two equations

• Let $a > 0$ and $b > 0$ be two positive real number. We study in this section a model problem that is composed by the juxtaposition of an advection equation with celerity a and an advection with celerity $-b$. We explicit the associated algebra :

(2.3.1) $$\frac{\partial u}{\partial t} \,+\, a\,\frac{\partial u}{\partial x} \,=\, 0\,, \qquad t > 0\,, \qquad x \in \mathbb{R}\,,$$

$$(2.3.2) \qquad \frac{\partial v}{\partial t} - b\,\frac{\partial v}{\partial x} = 0\,, \qquad t > 0\,, \qquad x \in \mathbb{R}\,.$$

We associate the two equations (2.3.1) and (2.3.2) and consider a unique problem with a vector field as unknown. We set :

$$(2.3.3) \qquad \varphi = \begin{pmatrix} u \\ v \end{pmatrix}$$

and the set of equations (2.3.1)-(2.3.2) can naturally be written as a system :

$$(2.3.4) \qquad \frac{\partial \varphi}{\partial t} + \begin{pmatrix} a & 0 \\ 0 & -b \end{pmatrix} \frac{\partial \varphi}{\partial x} = 0\,.$$

By introducing the **flux function** $F(\varphi)$ according to the relation

$$(2.3.5) \qquad F(\varphi) = \begin{pmatrix} a\,u \\ -b\,v \end{pmatrix}$$

the system (2.3.4) takes the general conservative form :

$$(2.3.6) \qquad \frac{\partial \varphi}{\partial t} + \frac{\partial}{\partial x}\big(F(\varphi)\big) = 0\,.$$

• The approximation of system (2.3.6) with a grid parameterized by a space step Δx and a time step Δt is conducted exactly as in the case of the advection equation. The following property is a straightforward generalization of Proposition 2.1. We left the proof to the reader.

Proposition 2.3. Finite volume scheme.

Let $\mathbb{R} \times [0,\, +\infty[\,\ni (x,\, t) \longmapsto \varphi(x,\, t) \in \mathbb{R} \times \mathbb{R}$ be a solution of the linear conservation law (2.3.6). We define the space mean value $\varphi^n_{j+1/2}$ of this solution $\varphi(\bullet,\, \bullet)$ in the cell $K_{j+1/2}$:

$$(2.3.7) \qquad \varphi^n_{j+1/2} = \frac{1}{\mid K_{j+1/2} \mid} \int_{x_j}^{x_{j+1}} \varphi(x,\, t^n)\, \mathrm{d}x$$

and the time mean value $f_j^{n+1/2}$ of the flux function introduced in (2.3.5) at the space position x_j between discrete times t^n and t^{n+1} :

$$(2.3.8) \qquad f_j^{n+1/2} = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} F\big(\varphi(x_j,\, t)\big)\, \mathrm{d}t\,.$$

We have the following relation that characterizes the finite volumes schemes :

$$(2.3.9) \qquad \frac{1}{\Delta t}\big(\varphi^{n+1}_{j+1/2} - \varphi^n_{j+1/2}\big) + \frac{1}{\Delta x}\big(f^{n+1/2}_{j+1} - f^{n+1/2}_j\big) = 0\,.$$

• We have now to propose a precise numerical flux function analogous to the relation (2.1.12) to transform the conservation property (2.3.9) into a finite volume numerical scheme able to propagate the discrete values $\varphi^n_{j+1/2}$ up to the discrete time t^{n+1}. For internal interfaces $x_j,\, j = 1, \cdots, J-1$, it is natural to apply the upwinding scheme (2.1.20) with a left upwinding for the first equation and a right upwinding for the equation (2.3.2). Figure 2.4 illustrates the associated algebra :

(2.3.10) $$f_j^{n+1/2} \;=\; \Phi\big(\varphi_{j-1/2}^n,\, \varphi_{j+1/2}^n\big) \;=\; \begin{pmatrix} a\, u_{j-1/2}^n \\ -b\, v_{j+1/2}^n \end{pmatrix} , \quad j = 1, \cdots, J-1 .$$

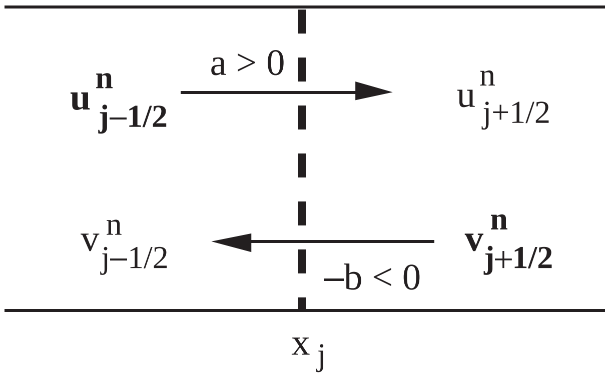

Figure 2.4. Interface upwind numerical flux for a model problem with two equations.

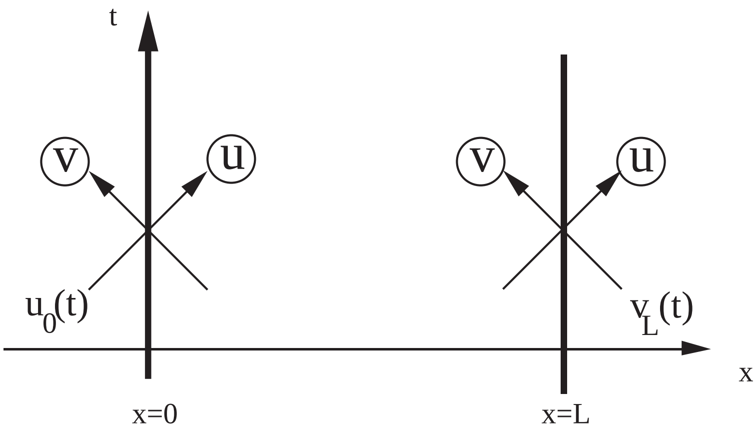

Figure 2.5. Boundary conditions for a model problem with two equations.

• At the left boundary $x = 0$, we have an input for the variable u and we suppose given the associated datum $[0, +\infty[\, \ni t \longmapsto u_0(t) \in \mathbb{R}$:

(2.3.11) $$u(0,\, t) \;=\; u_0(t) \,, \qquad t > 0$$

whereas it is an output for the v variable. By association of relations (2.2.6) and (2.2.8), we obtain

(2.3.12) $$f_0^{n+1/2} \;=\; \begin{pmatrix} a\, u_0\big((n+\frac{1}{2})\Delta t\big) \\ -b\, v_{1/2}^n \end{pmatrix} .$$

At the other boundary of the interval $]0,\, L[$, we have an output for the first variable u and an input for the second one, and an associated boundary condition $[0, +\infty[\, \ni t \longmapsto v_L(t) \in \mathbb{R}$ is supposed to have been given :

(2.3.13) $$v(L,\, t) \;=\; v_L(t) \,, \qquad t > 0$$

as illustrated on Figure 2.5. The numerical flux at the right is evaluated by association of the relations (2.2.7) and (2.2.9) :

$$f_L^{n+1/2} \;=\; \begin{pmatrix} a\, u_{J-1/2}^n \\ -b\, v_L\big((n+\tfrac{1}{2})\Delta t\big) \end{pmatrix} . \qquad (2.3.14)$$

2.4 Unidimensional linear acoustics

• We consider a gas in a pipe of uniform section at normal conditions of temperature and pressure. The reference density is denoted by ρ_0 and the reference pressure is named p_0. The **sound celerity** c_0 of this gas satisfies the relation

$$c_0 \;=\; \sqrt{\frac{\gamma p_0}{\rho_0}} \qquad (2.4.1)$$

with $\gamma = 1.4$ as proved *e.g.* in the book of Landau and Lifchitz [LL54]. A sound wave is a small perturbation of this reference state. The differences of density, pressure and velocity fields are denoted respectively by ρ, p and u. The hypothesis of a small perturbation implies that the entropy of the reference state is maintained for all the time evolution and in consequence, it is easy to establish the following relation between the perturbations of density and pressure :

$$p \;=\; c_0^2\, \rho\,. \qquad (2.4.2)$$

• The conservation of mass leads to a first order linear conservation law :

$$\frac{\partial \rho}{\partial t} \;+\; \rho_0\, \frac{\partial u}{\partial x} \;=\; 0 \qquad (2.4.3)$$

and the conservation of momentum links the time evolution of velocity with the spatial gradient of pressure :

$$\rho_0\, \frac{\partial u}{\partial t} \;+\; \frac{\partial p}{\partial x} \;=\; 0\,. \qquad (2.4.4)$$

We introduce the vector $W = \begin{pmatrix} p \\ u \end{pmatrix}$ of unknowns. Then the equations (2.4.3) and (2.4.4) can be written as a linear hyperbolic system of conservation laws :

$$\frac{\partial W}{\partial t} \;+\; A\, \frac{\partial W}{\partial x} \;=\; 0 \qquad (2.4.5)$$

with

$$A \;=\; \begin{pmatrix} 0 & \rho_0\, c_0^2 \\ \frac{1}{\rho_0} & 0 \end{pmatrix} . \qquad (2.4.6)$$

• When we consider the eigenvalues and eigenvectors of matrix A, it is natural to introduce the **characteristic variables** defined respectively by

$$\varphi_+ \;=\; p \,+\, \rho_0\, c_0\, u \qquad (2.4.7)$$

$$\varphi_- \;=\; p \,-\, \rho_0\, c_0\, u \qquad (2.4.8)$$

and the quantity $\rho_0\, c_0$ is named the **acoustic impedance**. We have from the relations (2.4.3) and (2.4.4) :

$$\begin{aligned}
\frac{\partial \varphi_+}{\partial t} + c_0 \frac{\partial \varphi_+}{\partial x} \;&=\; \Big(\frac{\partial p}{\partial t} + \rho_0\, c_0 \, \frac{\partial u}{\partial t} \Big) \;+\; \Big(c_0 \, \frac{\partial p}{\partial x} + \rho_0\, c_0^2 \, \frac{\partial u}{\partial x} \Big) \\
&=\; c_0^2 \, \Big(\frac{\partial \rho}{\partial t} + \rho_0 \, \frac{\partial u}{\partial x} \Big) \;+\; c_0 \, \Big(\rho_0 \, \frac{\partial u}{\partial t} + \frac{\partial p}{\partial x} \Big) \;=\; 0 \,, \\
\frac{\partial \varphi_-}{\partial t} - c_0 \frac{\partial \varphi_-}{\partial x} \;&=\; \Big(\frac{\partial p}{\partial t} - \rho_0\, c_0 \, \frac{\partial u}{\partial t} \Big) \;-\; c_0 \, \Big(\frac{\partial p}{\partial x} - \rho_0\, c_0 \, \frac{\partial u}{\partial x} \Big) \\
&=\; c_0^2 \, \Big(\frac{\partial \rho}{\partial t} + \rho_0 \, \frac{\partial u}{\partial x} \Big) \;-\; c_0 \, \Big(\rho_0 \, \frac{\partial u}{\partial t} + \frac{\partial p}{\partial x} \Big) \;=\; 0 \,,
\end{aligned}$$

and we recover a system of the type (2.3.4) studied previously :

(2.4.9) $$\frac{\partial}{\partial t} \begin{pmatrix} \varphi_- \\ \varphi_+ \end{pmatrix} \;+\; \begin{pmatrix} -c_0 & 0 \\ 0 & c_0 \end{pmatrix} \frac{\partial}{\partial t} \begin{pmatrix} \varphi_- \\ \varphi_+ \end{pmatrix} \;=\; 0 \,.$$

• A typically physical problem is the following : a given acoustic pressure wave $[0, +\infty[\,\ni t \longmapsto \Pi(t) > 0$ is injected at the left $x = 0$ of the pipe and the waves go away freely at the right boundary $x = L$. At $t = 0$, the velocity and pressure of the fluid are given :

(2.4.10) $\quad u(x,\, 0) \;=\; u_0(x) \,, \qquad 0 < x < L$

(2.4.11) $\quad p(x,\, 0) \;=\; p_0(x) \,, \qquad 0 < x < L \,.$

From a mathematical viewpoint, the boundary conditions have to respect the dynamics of this system of acoustic equations written in diagonal form (2.4.9) : the variable φ_+ must be given at $x = 0$ and the variable φ_- at the abscissa $x = L$. From (2.4.7) and (2.4.8), we determine the pressure as a function of the two characteristics variables φ_+ and φ_- :

(2.4.12) $\quad p \;=\; \dfrac{1}{2} \, (\varphi_+ + \varphi_-)$

and if the pressure is imposed at $x = 0$, the relation (2.4.12) can be written under the form :

(2.4.13) $\quad \varphi_+(0,\, t) \;=\; -\, \varphi_-(0,\, t) \,+\, 2\, \Pi(t) \,, \qquad x = 0 \,, \quad t > 0 \,,$

that makes in evidence a reflection operator : the input variable φ_+ is a given affine function of the output variable φ_-. At the other boundary $x = L$, the notion of free output expresses that the waves that go outside of the domain of study have no reflection at the boundary. When $x = L$, the characteristic variable φ_+ is going outside and there is no boundary condition for this variable. We have to express also that this wave has no influence on the characteristic φ_- that wish to go inside the domain $]0,\, L[$. In other terms, the input value φ_- is independent of the variable φ_+ and also of time. We have in consequence

(2.4.14) $\quad \dfrac{\partial}{\partial t} \varphi_-(L,\, t) \;=\; 0 \,.$

We have established

Proposition 2.4. Boundary conditions for acoustic problem.
The mathematical boundary conditions associated with the datum of a given acoustic pressure wave $[0, +\infty[\ni t \longmapsto \Pi(t) > 0$ at the left of the domain $]0, L[$ admits the expression (2.4.13) and a condition of free output of the waves at the right boundary $x = L$ can be expressed by the relation (2.4.14).

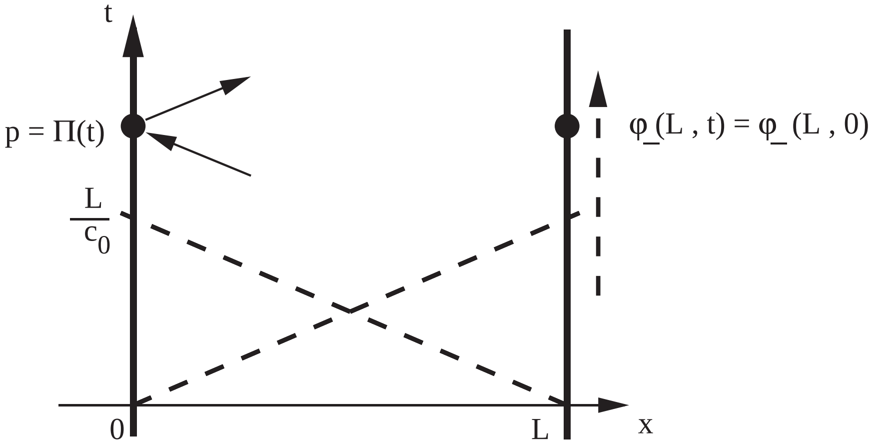

Figure 2.6. Solution of the acoustic equations in one space dimension for a model problem with two equations

• The above acoustic problem associated with the first order partial differential equations (2.4.3) (2.4.4), the initial conditions (2.4.10) (2.4.11) and the boundary conditions (2.4.13) (2.4.14) is illustrated on Figure 2.6. The initial conditions are active in the beginning of the evolution in time $(t \leq \frac{L}{c_0})$ and have a trace for higher times due to the boundary conditon (2.4.13), that gives, due to (2.4.8) and (2.4.13) :

$$(2.4.15) \qquad \varphi_-(x,\, t) \,\equiv\, p(x,\, t) - \rho_0\, c_0\, u(x,\, t) \,=\, p_0(L) - \rho_0\, c_0\, u_0(L)\,, \quad t \geq \frac{L}{c_0}\,.$$

On the other hand, the inflow boundary condition (2.4.12) and the second row of matrix equation (2.4.9) implies :

$$(2.4.16) \qquad \left\{ \begin{array}{l} \varphi_+(x,\, t) \equiv p(x,\, t) + \rho_0\, c_0\, u(x,\, t) \,= \\ \qquad\qquad = \, 2\, \Pi\big(t - \dfrac{x}{c_0}\big) - \varphi_-\big(0,\, t - \dfrac{x}{c_0}\big)\,, \quad t \geq \dfrac{L}{c_0}\,. \end{array} \right.$$

We deduce from the relations (2.4.15) (2.4.16) joined with the definitions (2.4.7) and (2.4.8) :

$$(2.4.17) \qquad p(x,\, t) \,=\, \Pi\big(t - \tfrac{x}{c_0}\big)\,, \qquad 0 \leq x \leq L\,, \quad t \geq \tfrac{L}{c_0}$$

$$(2.4.18) \qquad u(x,\, t) \,=\, u_0(L) + \tfrac{1}{\rho_0\, c_0} \Big(\Pi\big(t - \tfrac{x}{c_0}\big) - p_0(L) \Big)\,, \quad 0 \leq x \leq L\,, \quad t \geq \tfrac{L}{c_0}\,.$$

• We turn now to the numerical finite volume scheme. We have to determine the internal fluxes $f_j^{n+1/2}$, $j = 1, \cdots, J-1$ and the boundary fluxes

$f_0^{n+1/2}$ and $f_J^{n+1/2}$. Recall first that the physical flux $F(W)$ function for the acoustic equation (2.4.5) is equal to

$$F(W) \;=\; \begin{pmatrix} \rho_0\, c_0^2\, u \\ \frac{1}{\rho_0}\, p \end{pmatrix} \qquad \text{with} \qquad W \;=\; \begin{pmatrix} p \\ u \end{pmatrix} . \tag{2.4.19}$$

Proposition 2.5. Upwind scheme for computational acoustics.
The extension of the upwind finite volume scheme (2.3.10), (2.3.12) and (2.3.14) is determined by the following relations :

$$f_j^{n+1/2} \;=\; \begin{pmatrix} \frac{\rho_0\, c_0^2}{2}\, \big(u_{j-1/2}^n + u_{j+1/2}^n\big) \;-\; \frac{c_0}{2}\, \big(p_{j+1/2}^n - p_{j-1/2}^n\big) \\ \frac{1}{2\,\rho_0}\, \big(p_{j-1/2}^n + p_{j+1/2}^n\big) \;-\; \frac{c_0}{2}\, \big(u_{j+1/2}^n - u_{j-1/2}^n\big) \end{pmatrix} \tag{2.4.20}$$

for the internal fluxes, *i.e.* for indexes j that satisfy $1 \leq j \leq J-1$. The two boundary fluxes follow the following relations :

$$f_0^{n+1/2} \;=\; \begin{pmatrix} \rho_0\, c_0^2\, u_{1/2}^n \;+\; c_0\, \Big(\Pi\big((n+\tfrac{1}{2})\Delta t\big) - p_{1/2}^n\Big) \\ \frac{1}{\rho_0}\, \Pi\big((n+\tfrac{1}{2})\Delta t\big) \end{pmatrix} \tag{2.4.21}$$

$$f_J^{n+1/2} \;=\; \begin{pmatrix} \frac{\rho_0\, c_0^2}{2}\, \big(u_{J-1/2}^n + u_{J-1/2}^0\big) \;-\; \frac{c_0}{2}\, \big(p_{J-1/2}^0 - p_{J-1/2}^n\big) \\ \frac{1}{2\,\rho_0}\, \big(p_{J-1/2}^n + p_{J-1/2}^0\big) \;-\; \frac{c_0}{2}\, \big(u_{J-1/2}^0 - u_{J-1/2}^n\big) \end{pmatrix} . \tag{2.4.22}$$

• The internal fluxes are determined with the scheme (2.3.10) applied with the diagonal form of relation (2.4.9). We have

$$\varphi_{+,\,j}^{n+1/2} \;=\; \varphi_{+,\,j-1/2}^n \;\equiv\; p_{j-1/2}^n \;+\; \rho_0\, c_0\; u_{j-1/2}^n \tag{2.4.23}$$

$$\varphi_{-,\,j}^{n+1/2} \;=\; \varphi_{-,\,j+1/2}^n \;\equiv\; p_{j+1/2}^n \;-\; \rho_0\, c_0\; u_{j+1/2}^n \tag{2.4.24}$$

then the relation (2.4.20) is established.
The left boundary flux uses the extension of relation (2.3.12). We first determine the characteristic variables on the left boundary according to relation (2.4.13)

$$\varphi_{+,\,0}^{n+1/2} \;=\; 2\Pi\big((n+\frac{1}{2})\Delta t\big) \;-\; \varphi_{-,\,0}^{n+1/2} \tag{2.4.25}$$

and use a first order extrapolation of the outgoing characteristic variable :

$$\varphi_{-,\,0}^{n+1/2} \;=\; \varphi_{-,\,1/2}^n \;\equiv\; p_{1/2}^n \;-\; \rho_0\, c_0\; u_{1/2}^n\,. \tag{2.4.26}$$

Then we solve the system (2.4.25) (2.4.26) and find finally the relation (2.4.21).
The process is analogous for the right boundary. The input datum is imposed according to the relation (2.4.14) :

$$\varphi_{-,\,J}^{n+1/2} \;=\; \varphi_{-,\,J}^0 \;\equiv\; p_0(L) - \rho_0\, c_0\, u_0(L) \;\approx\; p_{J-1/2}^0 - \rho_0\, c_0\; u_{J-1/2}^0 \tag{2.4.27}$$

and the output characteristic variable is extrapolated from the interior of the domain :

$$\varphi_{+,\,J}^{n+1/2} \;=\; \varphi_{+,\,J-1/2}^n \;\equiv\; p_{J-1/2}^n \;+\; \rho_0\, c_0\; u_{J-1/2}^n\,. \tag{2.4.28}$$

The relation (2.4.22) follows after two steps of elementary algebra. □

• We remark that both relations (2.4.20) and (2.4.22) are identical, except that the boundary state $W_0(L) \approx W^0_{J-1/2}$ has replaced the right state $W^n_{j+1/2}$. Moreover the flux boundary condition (2.4.21) that involves the pressure is a natural discretization of the exact characteristic solution (2.4.17) (2.4.18) at $x = 0$.

2.5 Characteristic variables.

• We suppose now to fix the ideas that the unknown vector $W(\bullet, \bullet)$

(2.5.1) $$[0, L] \times [0, +\infty[\ni (x, t) \longmapsto W(x, t) \in \mathbb{R}^3$$

has three real components w_1, w_2 and w_3. We suppose also that the function $W(\bullet, \bullet)$ is solution of a conservation law of the type

(2.5.2) $$\frac{\partial W}{\partial t} + \frac{\partial}{\partial x} F(W) = 0$$

where the flux $F(W)$ is a **linear** function of vector W :

(2.5.3) $$F(W) = A \bullet W$$

and A is a 3 by 3 diagonalizable real matrix.

• We first detail the fact that matrix A is a diagonalizable matrix. There exists three non null real vectors r_1, r_2, r_3 and three real scalars λ_1, λ_2, λ_3 in such a way that

(2.5.4) $$A \bullet r_j = \lambda_j \, r_j \,, \qquad j = 1, 2, 3.$$

¿From a matricial viewpoint, we denote by $R_{k\,j}$ the k^0 component of the eigenvector r_j, *i.e.*

(2.5.5) $$r_j = \begin{pmatrix} R_{1\,j} \\ R_{2\,j} \\ R_{3\,j} \end{pmatrix} \equiv \begin{pmatrix} (r_j)_1 \\ (r_j)_2 \\ (r_j)_3 \end{pmatrix}$$

and we introduce the 3 by 3 matrix R composed by the scalars $R_{k\,j}$. The vector r_j is the k^0 column of matrix R. The relation (2.5.4) can also be written as

(2.5.6) $$A \bullet R = R \bullet \Lambda \,,$$

and Λ is the diagonal matrix whose diagonal terms are equal to the eigenvalues λ_j :

(2.5.7) $$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} .$$

• We consider now **two** distinct bases for linear space $\mathbb{R}^3$: on one hand the canonical basis $(e_j)_{j=1,2,3}$ defined by

(2.5.8) $$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} , \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} , \qquad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

where the vector W admits the natural decomposition introduced above :

(2.5.9) $\quad W \,=\, \sum_{k=1}^{k=3} w_k \, e_k \,,$

and on the other hand the basis of $\mathbb{R}^3$ composed by the eigenvectors $(r_j)_{j=1,\,2,\,3}$. In the latter, the vector W can be decomposed with a formula of the type

(2.5.10) $\quad W \,=\, \sum_{j=1}^{j=3} \varphi_j \, r_j$

and the scalar φ_j define the **characteristic variables** associated with the system (2.5.2) (2.5.3). The link between the relations (2.5.9) and (2.5.10) is classical : we consider the components $R_{k\,j}$ of vector r_j inside the canonical basis and we get from the relation (2.5.5) :

(2.5.11) $\quad w_k \,=\, \sum_{j=1}^{j=3} \varphi_j \, R_{k\,j} \,.$

Then the relation (2.5.11) can be re-written under a matricial form :

(2.5.12) $\quad W \,=\, R \bullet \varphi \,.$

• The relation (2.5.12) proposes to change the unknown function, *i.e.* to replace the research of $W(x,\,t) \in \mathbb{R}^3$ by the equivalent research of the characteristic vector $\varphi(x,\,t) \in \mathbb{R}^3$ and defined by :

(2.5.13) $\quad \varphi \,=\, R^{-1} \bullet W \,.$

Proposition 2.6. Characteristic variables satisfy advection equations.
The vector $[0,\,L] \times [0,\,+\infty[\,\ni (x,\,t) \longmapsto \varphi(x,\,t) \in \mathbb{R}^3$ of characteristic variables satisfy the matrix equation

(2.5.14) $$\frac{\partial \varphi}{\partial t} + \Lambda \bullet \frac{\partial \varphi}{\partial x} \,=\, 0$$

that takes also the equivalent scalar form :

(2.5.15) $$\frac{\partial \varphi_j}{\partial t} + \lambda_j \, \frac{\partial \varphi_j}{\partial x} \,=\, 0 \,, \qquad j = 1,\,2,\,3 \,.$$

• We have from (2.5.2), (2.5.3), (2.5.6) and (2.5.12) :

$$\frac{\partial W}{\partial t} + A \, \frac{\partial W}{\partial x} \,=\, R \bullet \frac{\partial \varphi}{\partial t} + A \bullet R \bullet \frac{\partial \varphi}{\partial x} \,=\, R \bullet \Big(\frac{\partial \varphi}{\partial t} + R^{-1} \bullet A \bullet R \bullet \frac{\partial \varphi}{\partial x} \Big)$$
$$= R \bullet \Big(\frac{\partial \varphi}{\partial t} + \Lambda \bullet \frac{\partial \varphi}{\partial x} \Big) \,=\, 0 \,,$$

and since the matrix R is invertible, we deduce from the previous calculus the relation (2.5.14). The relation (2.5.15) is an immediate consequence of (2.5.14) and (2.5.7). $\square$

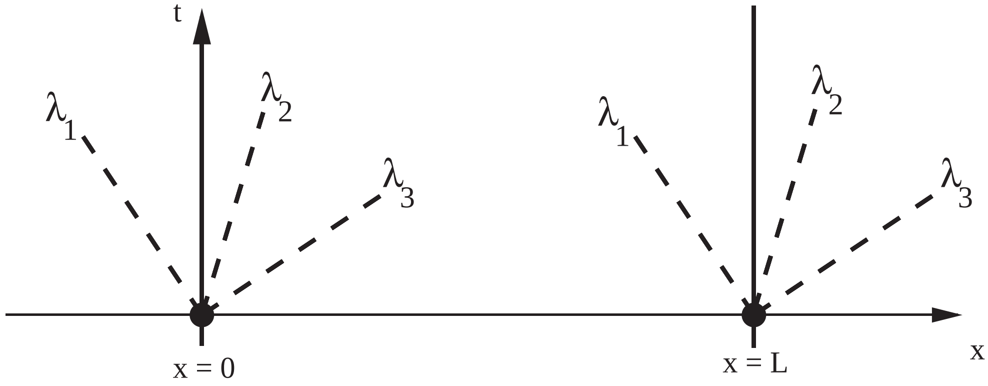

Figure 2.7. Linear hyperbolic system with three equations and eigenvalues satisfying $\lambda_1 < 0 < \lambda_2 < \lambda_3$.

• To fix the ideas, we suppose that the eigenvalues λ_j of matrix A are distinct, enumerated with an increasing order and with distinct signs as illustrated on Figure 2.7 :

(2.5.16) $\qquad \lambda_1 < 0 < \lambda_2 < \lambda_3 .$

The propagation of the first variable φ_1 goes from right to left (because $\lambda_1 < 0$) with celerity $|\lambda_1|$, the second characteristic variable φ_2 from left to right with celerity λ_2 and the same property holds for variable φ_3 with eigenvalue λ_3.

• A set of well posed boundary conditions is a consequence of the diagonal form (2.5.15) of the equations and of the particular choice (2.5.16) for the signs. The directions associated with eigenvalues λ_2 and λ_3 are ingoing at $x=0$ and we have to give some boundary condition for φ_2 and φ_3 at this point :

(2.5.17) $\qquad \varphi_2(x=0,\, t) \;=\; \beta_0(t)$

(2.5.18) $\qquad \varphi_3(x=0,\, t) \;=\; \gamma_0(t) .$

The direction associated with the eigenvalue λ_1 is ingoing at the abscisssa $x=L$, and this condition imposes to have some datum concerning φ_1 at this particular point :

(2.5.19) $\qquad \varphi_1(x=L,\, t) \;=\; \alpha_L(t) .$

The previous boundary conditions (2.5.17) to (2.5.19) define a well posed problem. Nevertheless, the introduction of physically relevant boundary conditions (as a pressure condition as seen in the previous section) requires a more general formulation of the boundary condition. In the linear case, the stability study developed by Kreiss [Kr70] shows that the ingoing characteristic can be an affine function of the outgoing characteristic through a **reflection operator** at the boundary. We can explicit the former with the above example.

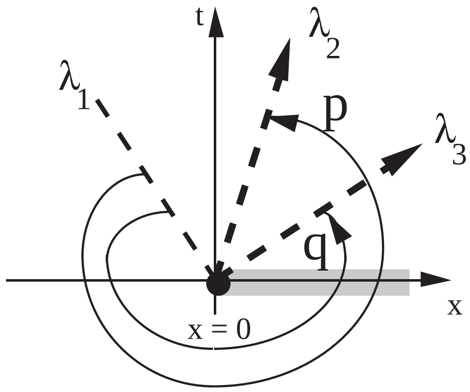

Figure 2.8. Reflection operator at $x = 0$.

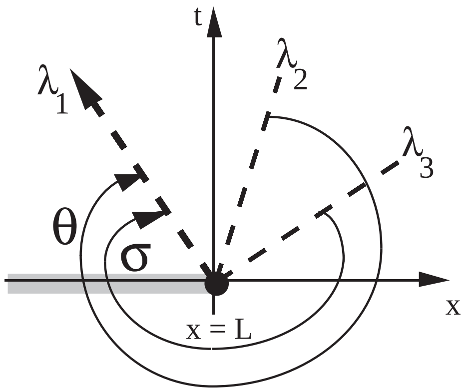

Figure 2.9. Reflection operator at $x = L$.

• At $x=0$, the first characteristic is outgoing and the two last ones are going inside the domain of study. Then we can replace the conditions (2.5.17) and (2.5.18) by the following ones :

(2.5.20) $\quad \varphi_2(x=0,\, t) \;=\; \beta_0(t) \;+\; p(t)\, \varphi_1(x=0,\, t)$

(2.5.21) $\quad \varphi_3(x=0,\, t) \;=\; \gamma_0(t) \;+\; q(t)\, \varphi_1(x=0,\, t)$,

where $t \longmapsto p(t)$ and $t \longmapsto q(t)$ are given fixed real functions of time. The conditions (2.5.20) and (2.5.21) are illustrated on Figure 2.8. We can also write them

(2.5.22) $\quad \varphi^{in}(x,\, t) \;=\; g(t) \;+\; S(t) \bullet \varphi^{out}(x,\, t)$, $\quad x$ point on the boundary,

with $\varphi^{in} = \begin{pmatrix} \varphi_2 \\ \varphi_3 \end{pmatrix}$, $g(t) = \begin{pmatrix} \beta_0(t) \\ \gamma_0(t) \end{pmatrix}$, $S(t) = \begin{pmatrix} p(t) \\ q(t) \end{pmatrix}$, $\varphi^{out} = \varphi_1$

when $x = 0$.

• When $x=L$, the relation (2.5.19) is replaced by a more general one

(2.5.23) $\quad \varphi_1(x=L,\, t) \;=\; \alpha_L(t) \;+\; \theta(t)\; \varphi_2(x=L,\, t) \;+\; \sigma(t)\; \varphi_3(x=L,\, t)$

illustrated on Figure 2.9 and including an affine component of the outgoing characteristic variables. The boundary condition (2.5.23) takes again a form of the type (2.5.22) with this time the following relations : $\varphi^{in} = \varphi_1$, $g(t) = \alpha_L(t)$, $S(t) = \big(\theta(t)\ \sigma(t)\big)$, $\varphi^{out} = \begin{pmatrix} \varphi_2 \\ \varphi_3 \end{pmatrix}$ when $x = L$.

2.6 A family of model systems with three equations

• We still study a 3 by 3 linear hyperbolic system of the type (2.5.2) (2.5.3) with the condition (2.5.16) to fix a particular example. We suggest in this section to explicit a way for evaluation of the numerical flux $f_j^{n+1/2}$ that is the key point for the discrete evolution in time of the mean values $W_{j+1/2}$:

(2.6.1) $$\frac{1}{\Delta t}\Big(W_{j+1/2}^{n+1} - W_{j+1/2}^{n}\Big) + \frac{1}{\Delta x}\Big(f_{j+1}^{n+1/2} - f_j^{n+1/2}\Big) = 0\,.$$

The internal fluxes $\big(f_j^{n+1/2}\big)_{j=1,\cdots,\,J-1}$ are evaluated with the help of a two-point numerical flux function $\Phi(\bullet,\,\bullet)$:

(2.6.2) $$f_j^{n+1/2} = \Phi(\,W_{j-1/2}^{n},\,W_{j+1/2}^{n})$$

and the boundary fluxes $f_0^{n+1/2}$ and $f_J^{n+1/2}$ are detailed in a forthcoming sub-section.

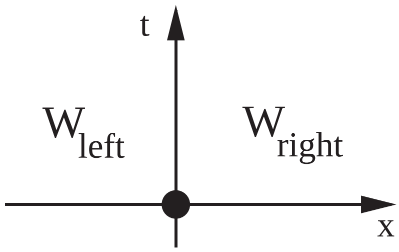

Figure 2.10. Discontinuity at the interface between two cells.

• We change the notations and wish to determine the numerical flux $\Phi(W_l,\ W_r)$ for $W_l = W_{\rm left}$ and $W_r = W_{\rm right}$ given respectively at the left and at the right of the interface (see Figure 2.10). When we consider the advection equation (and in that case the variables W_l and W_r are real numbers) the relation (2.1.19) gives the result : $\Phi(W_l,\,W_r) = a\,W_l$ when $a > 0$ and $\Phi(W_l,\,W_r) = a\,W_r$ when $a < 0$. We have to generalize this study when the field $W(\bullet,\,\bullet)$ is three-dimensional. We first decompose the vector $\Phi(W_l,\ W_r)$ with the basis r_j of eigenvectors and introduce its (scalar) components $\psi_j(W_l,\,W_r)$:

(2.6.3) $$\Phi(W_l,\,W_r) = \sum_{j=1}^{j=3} \psi_j(W_l,\,W_r)\, r_j$$

i.e.

(2.6.4) $$\Phi_k(W_l,\,W_r) = \sum_{j=1}^{j=3} R_{k\,j}\, \psi_j(W_l,\,W_r)\,.$$

For $j = 1$, we have $\lambda_1 < 0$ then the numerical scheme has to be upwinded in the right direction :

(2.6.5) $\psi_1(W_l,\, W_r) \;=\; \lambda_1\, \varphi_{1,\,r}$

whereas for $j=2$ or $j=3$, we have $\lambda_2 > 0$ and $\lambda_3 > 0$ and the scheme must be upwinded to the left. It comes

(2.6.6) $\psi_2(W_l,\, W_r) \;=\; \lambda_2\, \varphi_{2,\,l}\,, \qquad \psi_3(W_l,\, W_r) \;=\; \lambda_3\, \varphi_{3,\,l}\,.$

In consequence of the relations (2.6.3) to (2.6.6), the numerical flux function $\Phi(\bullet,\,\bullet)$ can be written globally :

(2.6.7) $\Phi(W_l,\, W_r) \;=\; \lambda_1\, \varphi_{1,\,r}\, r_1 \;+\; \lambda_2\, \varphi_{2,\,l}\, r_2 \;+\; \lambda_3\, \varphi_{3,\,l}\, r_3\,,$

or in an equivalent way with introducing the Cartesian components :

(2.6.8) $\Phi_k(W_l,\, W_r) \;=\; \lambda_1\, \varphi_{1,\,r}\, R_{k\,1} \;+\; \lambda_2\, \varphi_{2,\,l}\, R_{k\,2} + \lambda_3\, \varphi_{3,\,l}\, R_{k\,3}\,, \quad k=1,\,2,\,3\,.$

• We can also re-write the relation (2.6.8) for the particular interface x_j :

(2.6.9) $W_l \;=\; W_{\rm left} \;=\; W^n_{j-1/2}\,, \qquad W_r \;=\; W_{\rm right} \;=\; W^n_{j+1/2}\,.$

We first decompose the vector W on the eigenvectors of matrix A as in (2.5.11) :

(2.6.10) $\big(W^n_{j+1/2}\big)_k \;=\; \sum_{i=1}^{i=3} \varphi^n_{i,\,j+1/2}\, R_{k\,i}\,, \quad k=1,\,2,\,3\,, \quad j=1,\,\cdots,\,J{-}1\,,$

then we introduce the component number k of the flux $f_j^{n+1/2}$, *i.e.* $(f_j^{n+1/2})_k = \Phi_k(W^n_{j-1/2},\, W^n_{j+1/2})$ at the interface x_j :

(2.6.11) $\big(f_j^{n+1/2}\big)_k \;=\; \lambda_1\, \varphi^n_{1,\,j+1/2}\, R_{k\,1} + \lambda_2\, \varphi^n_{2,\,j-1/2}\, R_{k\,2} + \lambda_3\, \varphi^n_{3,\,j-1/2}\, R_{k\,3}\,.$

• We detail in this sub-section the determination of the numerical flux $f_0^{n+1/2}$ at the boundary $x=0$. We first recall that the continuous boundary conditions at this point take the form given in (2.5.20) (2.5.21). The idea is to try to apply the upwind scheme (2.6.11) at the particular vertex $j=0$: $f_0^{n+1/2} = \lambda_1\, \varphi^n_{1,\,1/2}\, r_1 + \; + \lambda_2\, \varphi^n_{2,\,-1/2}\, r_2 \;+\; \lambda_3\, \varphi^n_{3,\,-1/2}\, r_3$ and then to replace the characteristic values $\varphi^n_{2,\,-1/2}$ and $\varphi^n_{3,\,-1/2}$ (that are not defined on the mesh) by their values evaluated after a rough discretization of relations (2.5.20) and (2.5.21) : $\varphi^n_{2,\,-1/2} \;=\; \beta_0^{n+1/2} + \; + p^{n+1/2}\, \varphi^n_{1,\,1/2}\,, \qquad \varphi^n_{3,\,-1/2} \;=\; \gamma_0^{n+1/2} \;+\; q^{n+1/2}\, \varphi^n_{1,\,1/2}\,.$ We obtain in consequence the following expression for the **boundary flux** at $x=0$:

(2.6.12) $$f_0^{n+1/2} \;=\; \left\{ \begin{array}{l} \lambda_1\, \varphi^n_{1,\,1/2}\, r_1 \;+\; \lambda_2 \left(\beta_0^{n+1/2} \;+\; p^{n+1/2}\, \varphi^n_{1,\,1/2} \right) r_2 \,+ \\ \qquad\quad + \lambda_3 \left(\gamma_0^{n+1/2} \;+\; q^{n+1/2}\, \varphi^n_{1,\,1/2} \right) r_3 \end{array} \right.$$

or in an equivalent way :

(2.6.13) $$f_0^{n+1/2} \;=\; \left\{ \begin{array}{l} \varphi^n_{1,\,1/2} \left(\lambda_1\, r_1 \;+\; \lambda_2\, p^{n+1/2}\, r_2 \;+\; \lambda_3\, q^{n+1/2}\, r_3 \right) \,+ \\ \qquad\quad +\; \lambda_2\, \beta_0^{n+1/2}\, r_2 \;+\; \lambda_3\, \gamma_0^{n+1/2}\, r_3\,. \end{array} \right.$$

• The determination of the **boundary flux** $f_J^{n+1/2}$ can be conducted in the same way. Starting from the expression of the upwind scheme (2.6.11) when $j=J$, *i.e.* formally $f_J^{n+1/2} \;=\; \lambda_1\, \varphi^n_{1,\,J+1/2}\, r_1 \;+\; \lambda_2\, \varphi^n_{2,\,J-1/2}\, r_2 \;+$

$\lambda_3\,\varphi^n_{3,\,J-1/2}\,r_3$, we replace the first characteristic variable that appears external of the domain by its value given by the boundary condition (2.5.23) : $\varphi^n_{1,\,J+1/2} = \alpha_L^{n+1/2} + \theta^{n+1/2}\,\varphi^n_{2,\,J-1/2} + \,+\sigma^{n+1/2}\,\varphi^n_{3,\,J-1/2}$. We deduce :

$$f_J^{n+1/2} \;=\; \left\{ \begin{array}{l} \lambda_1 \left(\alpha_L^{n+1/2} + \theta^{n+1/2}\,\varphi^n_{2,\,J-1/2} + \sigma^{n+1/2}\,\varphi^n_{3,\,J-1/2} \right) r_1 \\ \qquad\qquad + \lambda_2\,\varphi^n_{2,\,J-1/2}\,r_2 + \lambda_3\,\varphi^n_{3,\,J-1/2}\,r_3 \end{array} \right. \qquad (2.6.14)$$

or in an equivalent manner :

$$f_J^{n+1/2} \;=\; \left\{ \begin{array}{l} \lambda_1\,\alpha_L^{n+1/2}\,r_1 + \varphi^n_{2,\,J-1/2} \left(\lambda_1\,\theta^{n+1/2}\,r_1 + \lambda_2\,r_2 \right) + \\ \qquad\qquad + \varphi^n_{3,\,J-1/2} \left(\lambda_1\,\sigma^{n+1/2}\,r_1 + \lambda_3\,r_3 \right) . \end{array} \right. \qquad (2.6.15)$$

2.7 First order upwind-centered finite volumes

• We consider now a general system of conservation laws

$$\frac{\partial W}{\partial t} + \frac{\partial}{\partial x} F(W) \;=\; 0 \qquad (2.7.1)$$

with an unknown vector $W(\bullet,\,\bullet)$ that belongs to linear space $\mathbb{R}^m$:

$$[0,\,L] \times [0,\,+\infty[\; \ni (x,\,t) \longmapsto W(x,\,t) \in \mathbb{R}^m \qquad (2.7.2)$$

and a **linear** flux function $F(\bullet)$

$$F(W) \;=\; A \bullet W \qquad (2.7.3)$$

associated with a diagonalizable matrix A with eigenvalues λ_j and eigenvectors r_j

$$A \bullet r_j \;=\; \lambda_j\,r_j \,, \qquad j = 1,\,2,\,\cdots,\,m\,. \qquad (2.7.4)$$

Introducing the $m \times m$ matrix R as in relation (2.5.5) and the diagonal matrix Λ of eigenvalues as in relation (2.5.7), we have :

$$A \bullet R \;=\; R \bullet \Lambda\,. \qquad (2.7.5)$$

• We propose here to determine a first order upwind flux $\Phi(W_l,\,W_r)$ between the two states $W_{\text{left}} = W_l$ and $W_{\text{right}} = W_r$ that generalizes the relation (2.6.7) when we have not done any hypothesis of the type (2.5.16) concerning the sign of the eigenvalues λ_j. We decompose any state W on the basis of space $\mathbb{R}^m$ characterized by the eigenvectors r_j :

$$W \;=\; \sum_{j=1}^{j=m} \varphi_j\,r_j \,, \quad W_l \;=\; \sum_{j=1}^{j=m} \varphi_{j,\,l}\,r_j \,, \quad W_r \;=\; \sum_{j=1}^{j=m} \varphi_{j,\,r}\,r_j \,, \qquad (2.7.6)$$

and due to the structure introduced at Proposition 2.6, we obtain an advection equation for the j^o characteristic variable φ_j :

$$\frac{\partial \varphi_j}{\partial t} + \lambda_j\,\frac{\partial \varphi_j}{\partial x} \;=\; 0 \,, \qquad j = 1,\,2,\,\cdots,\,m\,. \qquad (2.7.7)$$

Therefore it is natural to introduce the components $\psi_j(W_l,\,W_r)$ of the numerical flux on the basis of the eigenvectors :

$$\Phi(W_l,\,W_r) \;=\; \sum_{j=1}^{j=3} \psi_j(W_l,\,W_r)\,r_j \qquad (2.7.8)$$

and the first order upwind finite volume scheme is defined by the way we evaluate the coefficient $\psi_j(W_l,\, W_r)$ with the upwind scheme associated with the advection equation (2.7.7) :

$$(2.7.9) \qquad \psi_j(W_l,\, W_r) \;=\; \begin{cases} \lambda_j\ \varphi_{j,\,l} & \text{if } \lambda_j > 0 \\ \lambda_j\ \varphi_{j,\,r} & \text{if } \lambda_j < 0\,. \end{cases}$$

• For any real number μ, we introduce the positive part μ^+ and the negative part μ^- by the relations

$$(2.7.10) \qquad \mu^+ \;=\; \begin{cases} \mu & \text{if } \mu \geq 0 \\ 0 & \text{if } \mu \leq 0 \end{cases}, \qquad \mu^- \;=\; \begin{cases} 0 & \text{if } \mu \geq 0 \\ \mu & \text{if } \mu \leq 0\,. \end{cases}$$

We remark that we have

$$(2.7.11) \qquad \mu \;\equiv\; \mu^+ + \mu^- \,, \qquad \forall\, \mu \in \mathbb{R}$$

$$(2.7.12) \qquad |\,\mu\,| \equiv\; \mu^+ - \mu^- \,, \qquad \forall\, \mu \in \mathbb{R}\,.$$

We introduce also the absolute value $|\,\Lambda\,|$ of the diagonal matrix Λ by the condition :

$$(2.7.13) \qquad |\Lambda| \equiv |\,\mathrm{diag}\big(\lambda_1, \cdots, \lambda_m\big)| = \;\mathrm{diag}\big(|\lambda_1|, \cdots, |\lambda_m|\big)$$

and due to the relation (2.7.5), the absolute value $|\,A\,|$ of the matrix A is defined by :

$$(2.7.14) \qquad |A| = \; R\bullet\; |\Lambda|\; \bullet R^{-1}\,.$$

Proposition 2.7. Three expressions of the upwind first order scheme.
Let $\Phi(W_l,\, W_r)$ the upwind flux defined by the relations (2.7.8) and (2.7.9). Then we have the three equivalent expressions :

$$(2.7.15) \qquad \Phi(W_l,\, W_r) \;=\; F(W_l) + \textstyle\sum_{j=1}^{j=m} \lambda_j^- \,\big(\varphi_{j,\,r} - \varphi_{j,\,l}\big)\, r_j$$

$$(2.7.16) \qquad \Phi(W_l,\, W_r) \;=\; F(W_r) - \textstyle\sum_{j=1}^{j=m} \lambda_j^+ \,\big(\varphi_{j,\,r} - \varphi_{j,\,l}\big)\, r_j$$

$$(2.7.17) \qquad \Phi(W_l,\, W_r) \;=\; \tfrac{1}{2}\big(F(W_l) + F(W_r)\big) - \tfrac{1}{2}\,|A|\,\bullet(W_r - W_l)\,.$$

• We write the relation (2.7.9) under the form :

$$(2.7.18) \qquad \psi_j(W_l,\, W_r) \;=\; \lambda_j^+\,\varphi_{j,\,l} \,+\, \lambda_j^-\,\varphi_{j,\,r}$$

and we have :

$$\begin{aligned} \Phi(W_l,\, W_r) \;&=\; \textstyle\sum_{j=1}^{j=m} \big(\lambda_j^+\,\varphi_{j,\,l} + \lambda_j^-\,\varphi_{j,\,r}\big)\, r_j \\ &=\; \textstyle\sum_{j=1}^{j=m} \big((\lambda_j - \lambda_j^-)\,\varphi_{j,\,l} + \lambda_j^-\,\varphi_{j,\,r}\big)\, r_j \qquad \text{due to (2.7.11)} \\ &=\; \textstyle\sum_{j=1}^{j=m} \lambda_j\,\varphi_{j,\,l}\, r_j + \sum_{j=1}^{j=m} \lambda_j^-\,\big(\varphi_{j,\,r} - \varphi_{j,\,l}\big)\, r_j \end{aligned}$$

and the relation (2.7.15) is established. In an analogous way, we have :

$$\begin{aligned} \Phi(W_l,\, W_r) \;&=\; \textstyle\sum_{j=1}^{j=m} \big(\lambda_j^+\,\varphi_{j,\,l} + \lambda_j^-\,\varphi_{j,\,r}\big)\, r_j \\ &=\; \textstyle\sum_{j=1}^{j=m} \big(\lambda_j^+\,\varphi_{j,\,l} + (\lambda_j - \lambda_j^+)\,\varphi_{j,\,r}\big)\, r_j \qquad \text{due to (2.7.11)} \end{aligned}$$

$$\Phi(W_l,\, W_r) \;=\; \textstyle\sum_{j=1}^{j=m} \lambda_j\,\varphi_{j,\,r}\, r_j - \sum_{j=1}^{j=m} \lambda_j^+\,\big(\varphi_{j,\,r} - \varphi_{j,\,l}\big)\, r_j$$

and the relation (2.7.16) holds. We remark that

$$\begin{aligned} |A| \bullet (W_r - W_l) &= R \bullet |\Lambda| \bullet R^{-1} \bullet R \bullet (\varphi_r - \varphi_l) \quad \text{due to (2.7.14) and (2.5.12)} \\ &= R \bullet |\Lambda| \bullet (\varphi_r - \varphi_l) \\ &= \sum_{k=1}^{k=m} \sum_{j=1}^{j=m} R_{k\,j} \, |\lambda_j| \, (\varphi_{j,\,r} - \varphi_{j,\,l}) \, e_k \quad \text{then} \end{aligned}$$

$$|A| \bullet (W_r - W_l) = \sum_{j=1}^{j=m} |\lambda_j| \, (\varphi_{j,\,r} - \varphi_{j,\,l}) \, r_j \, . \qquad (2.7.19)$$

We add the previous results (2.7.15) with (2.5.16), and we divide by two. We obtain :

$$\begin{aligned} \Phi(W_l, W_r) &= \tfrac{1}{2} \big(F(W_l) + F(W_r) \big) - \tfrac{1}{2} \sum_{j=1}^{j=m} \big(\lambda_j^+ - \lambda_j^- \big) \big(\varphi_{j,\,r} - \varphi_{j,\,l} \big) \, r_j \\ &= \tfrac{1}{2} \big(F(W_l) + F(W_r) \big) - \tfrac{1}{2} \sum_{j=1}^{j=m} |\lambda_j| \, \big(\varphi_{j,\,r} - \varphi_{j,\,l} \big) \, r_j \quad \text{due to (2.7.12)} \\ &= \tfrac{1}{2} \big(F(W_l) + F(W_r) \big) - \tfrac{1}{2} \, |A| \bullet (W_r - W_l) \end{aligned}$$

due to the relation (2.7.19). Then the relation (2.7.17) is established and the proposition 2.7 is proven. □

3) Gas dynamics with the Roe method.

3.1 Nonlinear acoustics in one space dimension.

• We propose here to describe quickly a physical problem that comes from the theoretical modelling of trombone, detailed for instance in the work of Hirschberg *et al* [HGMW96] or in our study [MD99] with R. Msallam. In a first approximation, the duct of a trombone is a long cylinder with a constant section and the acoustic waves propagate only in the longitudinal direction. We can use a one-dimensional description of the geometry (see Figure 3.1) and in what follows, the trombone is modelled by a real space variable x that ranges from $x=0$ at the input to $x=L$ at the output.

• At the input $x=0$, a given non-stationary pressure wave $t \longmapsto \Pi(t)$ is emitted ; this wave is a perturbation of the ambiant pressure p_0 of the air :

$$|\Pi(t) - p_0| << p_0 \, , \qquad t > 0 \, . \qquad (3.1.1)$$

At the output $x = L$, the waves go outside without any reflection due to the presence of a pavilion and the boundary condition is a "free output" and a **non-reflecting boundary condition** has to be used. At the initial time $t=0$, we can consider that the air satisfies the usual conditions of pressure $p(x, 0) \equiv p_0$, temperature $T(x, 0) \equiv T_0$ and density $\rho(x, 0) \equiv \rho_0$. We study in this section a finite volume method able to treat nonlinearities in the acoustic modelling and based on the characteristic decompositions developed in the previous section.

Figure 3.1. Long unidimensional pipe for the modelling of a trombone.

3.2 Linearization of the gas dynamics equations.

• We study a perfect gas subjected to a motion with variable velocity in space and time. We have noticed that the primitive unknowns of this problem are the scalar fields that characterize the thermodynamics of the gas, *i.e.* density ρ, internal energy e, temperature T, and pressure p. In what follows, we suppose that the gas is a polytropic perfect gas ; it has constant specific heats at constant volume C_v and at constant pressure C_p. These two quantities do not depend on any thermodynamic variable like temperature or pressure ; we denote by γ their ratio :

(3.2.1) $$\gamma \,=\, \frac{C_p}{C_v} \, (= \text{ constant}) \, .$$

We suppose that the gas satisfies the law of perfect gas that can be written with the following form :

(3.2.2) $$p \,=\, (\gamma - 1) \, \rho \, e \, .$$

As usual, internal energy and temperature are linked together by the Joule-Thomson relation :

(3.2.3) $$e \,=\, C_v \, T \, .$$

• In the formalism proposed by Euler during the 18^{th} century, the motion is described with the help of an unknown vector field u which is a function of space x and time t :

(3.2.4) $$u \,=\, u(x,\, t) \, .$$

In the following, we will suppose that space x has only one dimemsion ($x \in \mathbb{R}$). We have four unknown functions (density, velocity, pressure and internal energy) linked together by the state law (3.2.2). In consequence, we need three complementary equations in order to define a unique solution of the problem. The general laws of Physics assume that mass, momentum and total energy are conserved quantities, at least in the context of classical physics associated to the paradigm of invariance for the Galileo group of space-time transformations (see *e.g.* Landau and Lifchitz [LL54]). When we write the conservation of mass, momentum and energy inside an infinitesimal volume $\mathrm{d}x$ advected with celerity $u(x,\, t)$, which is exactly the mean velocity of particules that compose the gas, it is classical [LL54] to write the fundamental conservation laws of Physics with the help of divergence operators :

$$(3.2.5)\qquad \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) = 0$$

$$(3.2.6)\qquad \frac{\partial}{\partial t}(\rho u) + \frac{\partial}{\partial x}(\rho u^2 + p) = 0$$

$$(3.2.7)\qquad \frac{\partial}{\partial t}\Big(\frac{1}{2}\rho u^2 + \rho e\Big) + \frac{\partial}{\partial x}\Big(\big(\frac{1}{2}\rho u^2 + \frac{p}{\gamma-1}\big) u + p u\Big) = 0 .$$

• We introduce the specific total energy E by unity of volume

$$(3.2.8)\qquad E = \tfrac{1}{2}u^2 + e ,$$

the sound celerity c following the classical expression :

$$(3.2.9)\qquad c = \sqrt{\frac{\gamma p}{\rho}} ,$$

and total enthalpy H defined according to

$$(3.2.10)\qquad H \equiv E + \tfrac{p}{\rho} = \tfrac{1}{2}u^2 + \tfrac{1}{\gamma-1}c^2 .$$

The vector W is therefore composed by the "conservative variables" or more precisely by the "conserved variables" :

$$(3.2.11)\qquad W = \big(\rho , \rho u , \rho E\big)^{\rm t} \equiv \big(\rho , q , \epsilon\big)^{\rm t} .$$

The conservation laws (3.2.5)-(3.2.7) take the following general form of a system of conservation laws :

$$(3.2.12)\qquad \frac{\partial W}{\partial t} + \frac{\partial}{\partial x} F(W) = 0$$

where the flux vector $W \longmapsto F(W)$ satisfies the following algebraic expression :

$$(3.2.13)\qquad F(W) = \big(\rho u , \rho u^2 + p , \rho u H\big)^{\rm t}$$

that can be explicited as a *true* function of state vector W, on one hand with the pressure law $P(W)$ computed with (3.2.2), (3.2.8) and (3.2.11) :

$$(3.2.14)\qquad P(W) = (\gamma-1)\Big(\epsilon - \frac{q^2}{2\rho}\Big)$$

and on the other hand with an explicit use of the conserved variables ρ, q and ϵ. We obtain :

$$(3.2.15)\qquad F(W) = \Big(q , \frac{q^2}{\rho} + P(W) , \frac{q\,\epsilon}{\rho} + P(W)\,\frac{q}{\rho}\Big) .$$

Proposition 3.1. Jacobian matrix of gas dynamics.

• The Jacobian matrix $\mathrm{d}F(W)$ of the flux function $W \longmapsto F(W)$ for the Euler equations of the gas dynamics admits the following expression :

$$(3.2.16)\qquad \mathrm{d}F(W) = \begin{pmatrix} 0 & 1 & 0 \\ (\gamma-1)H - u^2 - c^2 & (3-\gamma)u & \gamma-1 \\ (\gamma-2)uH - u c^2 & H - (\gamma-1)u^2 & \gamma u \end{pmatrix} .$$

• The matrix $\mathrm{d}F(W)$ is diagonalizable ; the eigenvalues $\lambda_j(W)$ satisfy the relations

(3.2.17) $\qquad \lambda_1(W) \equiv u - c \quad < \quad \lambda_2(W) \equiv u \quad < \quad \lambda_3(W) \equiv u + c \,.$

and the associated eigenvectors $r_j(W)$ are proportional to the following ones :

(3.2.18) $$r_1(W) = \begin{pmatrix} 1 \\ u - c \\ H - u\,c \end{pmatrix} , \quad r_2(W) = \begin{pmatrix} 1 \\ u \\ \frac{1}{2} u^2 \end{pmatrix} , \quad r_3(W) = \begin{pmatrix} 1 \\ u + c \\ H + u\,c \end{pmatrix} .$$

• We first differentiate the pressure law $W \longmapsto P(W)$ given in (3.2.14) :

(3.2.19) $$\frac{\partial P}{\partial \rho} = \frac{\gamma - 1}{2} u^2 = (\gamma - 1) H - c^2 , \quad \frac{\partial P}{\partial q} = -(\gamma - 1)\, u , \quad \frac{\partial P}{\partial \epsilon} = (\gamma - 1)$$

and the second row of the matrix (3.2.16) is a direct consequence of the relations $\frac{\partial}{\partial \rho}\big(\frac{q^2}{\rho}\big) = -u^2$ and $\frac{\partial}{\partial q}\Big(\frac{q^2}{\rho}\Big) = 2\,u \,.$

• The calculus of the third row of matrix in (3.2.16) demands first evaluation of the gradient of $\rho\, u\, E = u\, \epsilon$ relatively to the state W. We get

(3.2.20) $$\frac{\partial}{\partial \rho}\Big(\frac{q\,\epsilon}{\rho}\Big) = -u\, E , \quad \frac{\partial}{\partial q}\Big(\frac{q\,\epsilon}{\rho}\Big) = E , \quad \frac{\partial}{\partial \epsilon}\Big(\frac{q\,\epsilon}{\rho}\Big) = u \,.$$

We have also $\frac{\partial}{\partial W}(P\,u) = \frac{\partial P}{\partial W}\, u + p\, \frac{\partial}{\partial W}\big(\frac{q}{\rho}\big)$ then we deduce from (3.2.19) and the following expressions for the gradient of velocity $\frac{\partial}{\partial \rho}\big(\frac{q}{\rho}\big) = -\frac{u}{\rho}$ and $\frac{\partial}{\partial q}\Big(\frac{q}{\rho}\Big) = \frac{1}{\rho}$:

(3.2.21) $$\begin{cases} \dfrac{\partial}{\partial \rho}\Big(\dfrac{P\,q}{\rho}\Big) = \dfrac{\gamma - 1}{2} u^3 - \dfrac{u\,p}{\rho} , \\ \dfrac{\partial}{\partial q}\Big(\dfrac{P\,q}{\rho}\Big) = -(\gamma - 1)\, u^2 + \dfrac{p}{\rho} , \quad \dfrac{\partial}{\partial \epsilon}\Big(\dfrac{P\,q}{\rho}\Big) = (\gamma - 1)\, u \,. \end{cases}$$

We add the relations (3.2.20) and (3.2.21) ; then the third row of matrix (3.2.16) admits the following expression : $\big(\frac{\gamma-1}{2} u^3 - u\,H \,,\; H - (\gamma - 1)\, u^2 \,,\; \gamma\, u \big)$ and this result is exactly the third row of the right hand side of (3.2.16) when we take into account the relation (3.2.10) between H, u^2 and c^2. The relations (3.2.17) and (3.2.18) are elementary to satisfy ; they express simply the three relations :

(3.2.22) $\qquad {\rm d}F(W) \bullet r_j(W) = \lambda_j(W)\, r_j(W) \,, \quad j = 1, 2, 3$

and Proposition 3.1 is established. □

• We keep into memory the following expression of the Jacobian matrix ${\rm d}F(W)$:

(3.2.23) $${\rm d}F(W) = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\gamma-3}{2} u^2 & (3 - \gamma)\, u & \gamma - 1 \\ \frac{\gamma-1}{2} u^3 - u\,H & H - (\gamma - 1)\, u^2 & \gamma u \end{pmatrix}$$

that needs only the datum of velocity u and total enthalpy H of the state W.

3.3 Roe matrix.

• We consider two states $W_{\rm left} \equiv W_l$ and $W_{\rm right} \equiv W_r$ relatively to the gas dynamics, *i.e.* they both belong to space $\mathbb{R}^3$ and have an expression of the form (3.2.11). By **definition**, a Roe matrix $A(W_l, W_r)$ between these two states is a 3 by 3 matrix that satisfy the three following properties :

(3.3.1) $\quad A(W_l, W_r)$ is a diagonalizable matrix on the field $\mathbb{R}$ of real numbers

(3.3.2) $\quad A(W, W) \,=\, {\rm d}F(W)$

(3.3.3) $\quad F(W_r) - F(W_l) \,=\, A(W_l, W_r) \bullet (W_r - W_l)\,.$

In his original article, P. Roe [Roe81] has proposed a very simple algebraic way to construct a Roe matrix for the dynamics of polytropic gas. We propose it in the following Proposition.

Proposition 3.2. Algebraic construction of a Roe matrix [Roe81].
Let W_l and W_r be two states for gas dynamics, defined by their densities ρ_l and ρ_r, their velocities u_l and u_r and their total enthalpies H_l and H_r. We introduce an **intermediate state** $W^*(W_l, W_r)$ by its density ρ^*, its velocity u^* and its total enthalpy H^* according to the following relations :

(3.3.4) $$\rho^* \,=\, \sqrt{\rho_l\,\rho_r}$$

(3.3.5) $$u^* \,=\, \frac{\sqrt{\rho_l}\,u_l \,+\, \sqrt{\rho_r}\,u_r}{\sqrt{\rho_l} \,+\, \sqrt{\rho_r}}$$

(3.3.6) $$H^* \,=\, \frac{\sqrt{\rho_l}\,H_l \,+\, \sqrt{\rho_r}\,H_r}{\sqrt{\rho_l} \,+\, \sqrt{\rho_r}}\,.$$

Then the matrix $A(W_l, W_r)$ defined as the Jacobian matrix of the flux for the intermediate state $W^*(W_l, W_r)$, *i.e.*

(3.3.7) $$A(W_l, W_r) \,=\, {\rm d}F\big(W^*(W_l, W_r)\big)$$

is a Roe matrix.

• Due to the expression (3.2.23) of the Jacobian matrix of gas dynamics, we remark that the formula (3.3.4) giving the density ρ^* is not necessary for the determination of the matrix ${\rm d}F(W^*(W_l, W_r))$ and an entire family of states $W^*(W_l, W_r)$ define a Roe matrix according to the relations (3.3.5), (3.3.6) and (3.3.7). Nevertheless, we keep this definition of density ρ^* by convenience and simplicity for future algebraic expressions. The proof of Proposition 3.2 needs some algebraic developments. We begin by the following technical lemma.

Proposition 3.3.
Under the hypotheses of Proposition 3.2, we have the following relations :

(3.3.8) $$(u^*)^2\,(\rho_r - \rho_l) \,-\, 2\,u^*\,(\rho_r\,u_r - \rho_l\,u_l) \,+\, (\rho_r\,u_r^2 - \rho_l\,u_l^2) \,=\, 0$$

(3.3.9) $$\left\{ \begin{array}{l} -u^*\,H^*\,(\rho_r - \rho_l) \,+\, H^*\,(\rho_r\,u_r - \rho_l\,u_l) \,+\, u^*\,(\rho_r\,H_r - \rho_l\,H_l) \,= \\ \qquad\qquad = \, \rho_r\,u_r\,H_r - \rho_l\,u_l\,H_l\,. \end{array} \right.$$

• We first evaluate the left hand side of relation (3.3.8) :

$$(u^*)^2\,(\rho_r-\rho_l)\;-\;2\,u^*\,(\rho_r\,u_r-\rho_l\,u_l)\;+\;(\rho_r\,u_r^2-\rho_l\,u_l^2)\;=$$
$$=\;u^*(\sqrt{\rho_r}-\sqrt{\rho_l})\,(\sqrt{\rho_r}\,u_r+\sqrt{\rho_l}\,u_l)-2\,u^*(\rho_r\,u_r-\rho_l\,u_l)+(\rho_r\,u_r^2-\rho_l\,u_l^2)$$
$$=\;u^*\,\big(\sqrt{\rho_l}\,(\sqrt{\rho_l}+\sqrt{\rho_r})\,u_l\;-\;\sqrt{\rho_r}\,(\sqrt{\rho_l}+\sqrt{\rho_r})\,u_r\big)\;+\;(\rho_r\,u_r^2-\rho_l\,u_l^2)$$
$$=\;(\sqrt{\rho_l}\,u_l+\sqrt{\rho_r}\,u_r)\,(\sqrt{\rho_l}\,u_l-\sqrt{\rho_r}\,u_r)\;+\;(\rho_r\,u_r^2-\rho_l\,u_l^2)$$
$$=\;0\qquad\text{and the relation (3.3.8) is established.}$$

• We work on the left hand side of (3.3.9) as follows :

$$-u^*\,H^*\,(\rho_r-\rho_l)\;+\;H^*\,(\rho_r\,u_r-\rho_l\,u_l)\;+\;u^*\,(\rho_r\,H_r-\rho_l\,H_l)\;=$$
$$=\;-u^*(\sqrt{\rho_r}-\sqrt{\rho_l})\,(\sqrt{\rho_l}\,H_l+\sqrt{\rho_r}\,H_r)\;+\;H^*\,(\rho_r\,u_r-\rho_l\,u_l)\;+\;u^*\,(\rho_r\,H_r-\rho_l\,H_l)$$
$$=\;\sqrt{\rho_l\,\rho_r}\,u^*\,(H_r-H_l)\;+\;H^*\,(\rho_r\,u_r-\rho_l\,u_l)$$
$$=\;\frac{\sqrt{\rho_l}\,\sqrt{\rho_r}\,(\sqrt{\rho_l}\,u_l+\sqrt{\rho_r}\,u_r)\,(-H_l+H_r)+(-\rho_l\,u_l+\rho_r\,u_r)\,(\sqrt{\rho_l}\,H_l+\sqrt{\rho_r}\,H_r)}{\sqrt{\rho_l}+\sqrt{\rho_r}}$$
$$=\;\frac{1}{\sqrt{\rho_l}+\sqrt{\rho_r}}\,\big[-\rho_l\,(\sqrt{\rho_l}+\sqrt{\rho_r})\,u_l\,H_l\;+\;\rho_r\,(\sqrt{\rho_l}+\sqrt{\rho_r})\,u_r\,H_r\,\big]$$
$$=\;\rho_r\,u_r\,H_r-\rho_l\,u_l\,H_l$$

and the proposition 3.3 is established. □

• The **proof of Proposition 3.2** consists in satisfying the three hypotheses that define a Roe matrix. First, due to the fact that the relation (3.3.7) defines the matrix $A(W_l,\,W_r)$ as a Jacobian of some state, this matrix is diagonalizable with real elements due to the result of Proposition 3.1 and the first property (3.3.1) is satisfied. The second property (3.3.2) is a simple consequence of the fact that if $W_l\,=\,W_r\,=\,W,$ then we have from the relations (3.3.4) to (3.3.6) : $W^*(W_l,\,W_r)\,=\,W$ and the property results from (3.3.7).

• The third property (3.3.3) needs more work. We remark that the first row of this matricial relation is clear. For the second row, we have :

Second row of matrix $A(W_l,\,W_r)\bullet(W_r-W_l)\;=$

$$=\;\frac{\gamma-3}{2}\,(u^*)^2\,(\rho_r-\rho_l)\;+\;(3-\gamma)\,u^*\,(\rho_r\,u_r-\rho_l\,u_l)\;+\;(\gamma-1)\,(\rho_r\,E_r-\rho_l\,E_l)$$
$$=\;\frac{\gamma-3}{2}\,\big[(u^*)^2\,(\rho_r-\rho_l)\;-2\,u^*\,(\rho_r\,u_r-\rho_l\,u_l)\big]\;+\;\frac{\gamma-1}{2}\,(\rho_r\,u_r^2-\rho_l\,u_l^2)\;+\;(p_r-p_l)$$
$$=\;(\rho_r\,u_r^2-\rho_l\,u_l^2)\;+\;(p_r-p_l)\qquad\text{due to (3.3.8)}$$

$=$ second row of the flux difference $F(W_r)-F(W_l)\,.$

• We have also, in consequence of (3.2.23),

Third row of matrix $A(W_l,\,W_r)\bullet(W_r-W_l)\;=$

$$=\;u^*\,\Big(\frac{\gamma-1}{2}\,(u^*)^2-H^*\Big)\,(\rho_r-\rho_l)\;+\;(H^*-(\gamma-1)\,(u^*)^2)\,(\rho_r\,u_r-\rho_l\,u_l)\,+$$
$$+\,\gamma\,u^*\,(\rho_r\,E_r-\rho_l\,E_l)$$
$$=\;\frac{\gamma-1}{2}\,u^*\,\big[(u^*)^2\,(\rho_r-\rho_l)\;-\;2\,u^*\,(\rho_r\,u_r-\rho_l\,u_l)\big]\,+$$

$$+ \big[-H^* \, u^* \, (\rho_r - \rho_l) \, + \, H^* \, (\rho_r \, u_r - \rho_l \, u_l) \big] \, + \, \gamma \, u^* \, (\rho_r \, E_r - \rho_l \, E_l)$$

$$= \frac{\gamma-1}{2} \, u^* \, \big[(u^*)^2 \, (\rho_r - \rho_l) \, - \, 2 \, u^* \, (\rho_r \, u_r - \rho_l \, u_l) \big] \, - \, u^* \, (\rho_r \, H_r - \rho_l \, H_l) \, +$$

$$+ (\rho_r \, u_r \, H_r - \rho_l \, u_l \, H_l) + \gamma \, u^* \, (\rho_r \, E_r - \rho_l \, E_l) \qquad \text{due to (3.3.9)}$$

$$= \frac{\gamma-1}{2} \, u^* \, \big[\, (u^*)^2 \, (\rho_r - \rho_l) \, - \, 2 \, u^* \, (\rho_r \, u_r - \rho_l \, u_l) \, + \, \rho_r \, u_r^2 - \rho_l \, u_l^2 \, \big] +$$

$$+ u^* \, (-\gamma \, \rho_r \, e_r + \gamma \, \rho_l \, e_l \, + \gamma \, \rho_r \, e_r - \gamma \, \rho_l \, e_l) \, + \, \rho_r \, u_r \, H_r - \rho_l \, u_l \, H_l$$

$$= \rho_r \, u_r \, H_r - \rho_l \, u_l \, H_l \qquad \text{due to (3.3.8)}$$

$=$ third row of the flux difference $F(W_r) - F(W_l)$

in the view of relation (3.2.13). The proposition 3.2 is established. □

3.4 Roe flux.

• The principal interest of the Roe matrix is to be able to use all what has been developed for **linear** hyperbolic systems in Section 2. In particular, the following linear hyperbolic system defined with a given Roe matrix $A(W_l, \, W_r)$

(3.4.1) $$\frac{\partial W}{\partial t} \, + \, A(W_l, \, W_r) \bullet \frac{\partial W}{\partial x} \, = \, 0$$

can be treated with the upwind scheme defined at proposition 2.7. We obtain by doing this the following

Proposition 3.4. Three formulae for a flux.

• Let W_l and W_r be two fluid states and W^* the intermediate state defined by the relations (3.3.4) to (3.3.6). The sound celerity c^* of state W^* is defined with the help of relation (3.2.10), *i.e.*

(3.4.2) $$c^* \, = \, \sqrt{(\gamma-1) \Big(H^* - \frac{(u^*)^2}{2} \Big)} \, ,$$

and the eigenvalues λ_j^* of the Roe matrix $A(W_l, \, W_r) \, \equiv \, \mathrm{d}F(W^*(W_l, \, W_r))$ are given by a relation analogous to (3.2.17).

(3.4.3) $\lambda_1^* \equiv u^* - c^* \quad < \quad \lambda_2^* \equiv u^* \quad < \quad \lambda_3^* \equiv u^* + c^* \, .$

The associated eigenvectors $r_j^* \, \equiv \, r_j(W^*)$ are proportional to the following ones :

(3.4.4) $$r_1^* = \begin{pmatrix} 1 \\ u^* - c^* \\ H^* - u^* \, c^* \end{pmatrix} , \quad r_2^* = \begin{pmatrix} 1 \\ u^* \\ \frac{1}{2} (u^*)^2 \end{pmatrix} , \quad r_3^* = \begin{pmatrix} 1 \\ u^* + c^* \\ H^* + u^* \, c^* \end{pmatrix} .$$

• We introduce the decomposition of vector $W_r - W_l$ in the basis r_j^* :

(3.4.5) $W_r - W_l \, = \, \sum_{j=1}^{j=3} \alpha_j \, r_j^* \, .$

The three following relations define a unique numerical flux $\Phi(W_l, \, W_r)$ named the **Roe flux** between the two states W_l and W_r :

(3.4.6) $\Phi(W_l, \, W_r) \, = \, F(W_l) \, + \, \sum_{j=1}^{j=3} (\lambda_j^*)^- \, \alpha_j \, r_j^*$

(3.4.7) $\Phi(W_l, \, W_r) \, = \, F(W_r) \, - \, \sum_{j=1}^{j=3} (\lambda_j^*)^+ \, \alpha_j \, r_j^*$

$$(3.4.8) \qquad \Phi(W_l,\, W_r) \;=\; \tfrac{1}{2}\big(F(W_l) + F(W_r)\big) - \tfrac{1}{2}\; |A(W_l,\, W_r)| \; \bullet (W_r - W_l)\,.$$

• The first non-obvious point is to verify that the relation (3.4.2) defines a real number c^*. We have

$$H^* - \frac{(u^*)^2}{2} \;=\; \frac{\sqrt{\rho_l}\, H_l \,+\, \sqrt{\rho_r}\, H_r}{\sqrt{\rho_l} \,+\, \sqrt{\rho_r}} - \frac{1}{2}\Big(\frac{\sqrt{\rho_l}\, u_l \,+\, \sqrt{\rho_r}\, u_r}{\sqrt{\rho_l} \,+\, \sqrt{\rho_r}} \Big)^2 \;=$$

$$= \frac{1}{(\sqrt{\rho_l} + \sqrt{\rho_r})^2} \Big[(\sqrt{\rho_l} + \sqrt{\rho_r}) \Big(\sqrt{\rho_l}\, \big(\frac{1}{2}\, u_l^2 + \frac{1}{\gamma - 1}\, c_l^2\big) + \sqrt{\rho_r}\, \big(\frac{1}{2}\, u_r^2 + \frac{1}{\gamma - 1}\, c_r^2\big) \Big) - \frac{1}{2} \Big(\rho_l\, u_l^2 \,+\, 2\, \rho^*\, u_l\, u_r \,+\, \rho_r\, u_r^2 \Big) \Big]$$

$$= \frac{1}{(\sqrt{\rho_l} + \sqrt{\rho_r})^2} \Big[\frac{1}{2}\, \big(\rho^*\, u_l^2 - 2\, \rho^*\, u_l\, u_r + \rho^*\, u_r^2\big) \;+\; \frac{\rho_l + \rho^*}{\gamma - 1}\, c_l^2 + \frac{\rho^* + \rho_r}{\gamma - 1}\, c_r^2 \Big]$$

$$= \frac{1}{(\sqrt{\rho_l} + \sqrt{\rho_r})^2} \Big[\frac{1}{2}\, \rho^*\, (u_r - u_l)^2 \;+\; \frac{\rho_l + \rho^*}{\gamma - 1}\, c_l^2 + \frac{\rho^* + \rho_r}{\gamma - 1}\, c_r^2 \Big] \qquad > \; 0$$

and

$$(3.4.9) \qquad c^* \;=\; \frac{\sqrt{\frac{\gamma - 1}{2}\, \rho^*\, (u_r - u_l)^2 \;+\; (\rho_l + \rho^*)\, c_l^2 + (\rho^* + \rho_r)\, c_r^2}}{\sqrt{\rho_l} + \sqrt{\rho_r}}\,.$$

• We make the difference between the right hand sides of (3.4.6) and (3.4.7). We get :

$F(W_r) - F(W_l) \;-\; \sum_{j=1}^{j=3} \big((\lambda_j^*)^+ + (\lambda_j^*)^-\big)\; \alpha_j\; r_j^* \;=$

$= \; A(W_l,\, W_r) \bullet (W_r - W_l) \;-\; \sum_{j=1}^{j=3} \lambda_j^*\; \alpha_j\; r_j^*$ due to (3.3.3) and (2.7.11)

$= \; A(W_l,\, W_r) \bullet \big(\sum_{j=1}^{j=3} \alpha_j\, r_j^*\big) \;-\; \sum_{j=1}^{j=3} \lambda_j^*\; \alpha_j\; r_j^*$ due to (3.4.5)

$= \; 0$ because $A(W_l,\, W_r) \bullet r_j^* \;=\; \lambda_j^*\, r_j^*$ for each integer j.

The proof of relation (3.4.8) is obtained by taking the half sum of (3.4.6) and (3.4.7). It is analogous to the one done for Proposition 2.7. The proof of Proposition 3.4 is completed. □

• We make explicit the parameters α_j introduced in relation (3.4.5) in order to be complete for the implementation of the above formulae on a computer.

Proposition 3.5. New acoustic impedance.

With the notations introduced at Proposition 3.4, and denoting by p_l and p_r the respective pressures of states W_l and W_r, we have the following relations for the scalar components α_j of the state difference $W_r - W_l$ in relation (3.4.5) :

$$(3.4.10) \qquad \alpha_1 \;=\; \frac{1}{2\,(c^*)^2}\, \big[\, (p_r - \rho^*\, c^*\, u_r) - (p_l - \rho^*\, c^*\, u_l) \,\big]$$

$$(3.4.11) \qquad \alpha_2 \;=\; -\frac{1}{(c^*)^2}\, \big[\, (p_r - (c^*)^2\, \rho_r) - (p_l - (c^*)^2\, \rho_l) \,\big]$$

(3.4.12) $$\alpha_3 \;=\; \frac{1}{2\,(c^*)^2}\,\big[\,(p_r+\rho^*\,c^*\,u_r)\,-\,(p_l+\rho^*\,c^*\,u_l)\,\big]$$

with **acoustic impedance** $\rho^*\,c^*$ that is nomore the one $\rho_0\,c_0$ of a reference state as in traditional acoustics but an impedance associated with the Roe intermediate state $W^*(W_l,\,W_r)$ of relations (3.3.4) to (3.3.6).

• We have just to explicit the three components of the relation (3.4.5). It comes :

$$\begin{pmatrix} \rho_r-\rho_l \\ \rho_r\,u_r-\rho_l\,u_l \\ \rho_r\,E_r-\rho_l\,E_l \end{pmatrix} \;=\; \begin{pmatrix} \alpha_1+\alpha_2+\alpha_3 \\ \alpha_1\,(u^*-c^*)\,+\,\alpha_2\,u^*\,+\,\alpha_3\,(u^*+c^*) \\ \alpha_1\,(H^*-u^*\,c^*)\,+\,\alpha_2\frac{(u^*)^2}{2}\,+\,\alpha_3\,(H^*+u^*\,c^*) \end{pmatrix}$$

(3.4.13) $$\alpha_1+\alpha_2+\alpha_3 \;=\; \rho_r-\rho_l\,,$$

and we deduce after multiplying the equation (3.4.13) by $-u^*$ and adding to the second equation of the above matrix equality :

$$\begin{aligned} c^*\,(\alpha_3-\alpha_1) \;&=\; \rho_r\,u_r\,-\,\rho_l\,u_l\,-\,u^*\,(\rho_r-\rho_l) \\ &=\; \rho_r\,u_r\,-\,\rho_l\,u_l\,-\,(\sqrt{\rho_r}-\sqrt{\rho_l})\,(\sqrt{\rho_l}\,u_l+\sqrt{\rho_r}\,u_r)\;, \end{aligned}$$

then

(3.4.14) $$c^*\,(-\alpha_1+\alpha_3) \;=\; \rho^*\,(u_r-u_l)\,.$$

• We deduce from the third equation of relation (3.4.5) :

$$\begin{aligned} \frac{(c^*)^2}{\gamma-1}\,(\alpha_1+\alpha_3) \;&=\; \rho_r\,E_r\,-\,\rho_l\,E_l\,-\,\frac{1}{2}\,(u^*)^2\,(\rho_r-\rho_l)\,-\,u^*\,c^*\,(\alpha_3-\alpha_1) \\ &=\; \frac{1}{\gamma-1}\,(p_r-p_l)+\frac{1}{2}\,(\rho_r\,u_r^2-\rho_l\,u_l^2)+\frac{1}{2}\,(u^*)^2\,(\rho_r-\rho_l)\,-\,u^*\,(\rho_r\,u_r-\rho_l\,u_l) \\ &=\; \frac{1}{\gamma-1}\,(p_r-p_l) \qquad \text{due to (3.3.8). Then we have :} \end{aligned}$$

(3.4.15) $$(c^*)^2\,(\alpha_1+\alpha_3) \;=\; p_r-p_l\,.$$

• The solution of the 2 by 2 linear system with unknowns α_1 and α_3 defined by the relations (3.4.14) and (3.4.15) directly gives the relations (3.4.10) and (3.4.12). The expression (3.4.11) of variable α_2 is a direct consequence of the relations (3.4.10), (3.4.12) and (3.4.13) and Proposition 3.5 is proven. □

Proposition 3.6. An algorithm for the Roe flux.

Let W_l and W_r be two compressible fluid states. The computation of the Roe flux $\Phi(W_l,\,W_r)$ of relations (3.4.6)-(3.4.8) between these two states is summarized by the following points :

• Evaluation of density ρ^*, velocity u^* and total enthalpy H^* of the intermediate state W^* with the relations (3.3.4) to (3.3.6),

• Determination of the sound celerity c^* of the intermediate state W^* from the previous data with the relation (3.4.2),

• Eigenvectors r_j^* of the Roe matrix with the relations (3.4.4),

• Computation of the characteristic variables α_j in (3.4.5) for the difference $W_r - W_l$ with the relations (3.4.10) to (3.4.12),

• Final computation of the Roe flux $\Phi(W_l, W_r)$ with the minimum of work :

(3.4.16) $\Phi(W_l, W_r) = F(W_l)$ if $u^* - c^* \geq 0$

(3.4.17) $\Phi(W_l, W_r) = F(W_l) + (u^* - c^*) \, \alpha_1 \, r_1^*$ if $u^* - c^* \leq 0 < u^*$

(3.4.18) $\Phi(W_l, W_r) = F(W_r) - (u^* + c^*) \, \alpha_3 \, r_3^*$ if $u^* \leq 0 < u^* + c^*$

(3.4.19) $\Phi(W_l, W_r) = F(W_r)$ if $u^* + c^* \leq 0$.

• The proof of the relations (3.4.16)-(3.4.19) is obtained by starting from the expression of the Roe flux given in (3.4.6). We know that $\lambda_1 = u^* - c^*$, $\lambda_2 = u^*$, $\lambda_3 = u^* + c^*$. If $u^* - c^* \geq 0$, then $u^* \geq 0$ and $u^* + c^* \geq 0$, so the relation (3.4.6) reduces to (3.4.16) because, due to (2.7.10), $\mu^- = 0$ if μ is a positive real number. If $u^* - c^* \leq 0 < u^* < u^* + c^*$, the term containing $(\lambda_1^*)^-$ is the only one that contributes in relation (3.4.6) and the relation (3.4.17) is a direct consequence of this remark. If $u^* - c^* < u^* \leq 0 < u^* + c^*$, the term that contains $(\lambda_3^*)^+$ is the only nonzero element among the three inside the relation (3.4.7) and we deduce the relation (3.4.18) from this property. When $u^* - c^* < u^* < u^* + c^* \leq 0$, the term $F(W_r)$ is the only to subsist inside the relation (3.4.7) and the relation (3.4.19) is established. We remark also that the algebraic expression (3.4.11) for α_2 is not necessary for the implementation of the algorithm. □

3.5 Entropy correction.

• The Roe flux replaces the nonlinear waves of the gas dynamics, *i.e.* the rarefactions and the shock waves by **linear** waves that are the **contact discontinuities**. If sufficiently weak shock waves occur for a given discontinuity between two states W_{left} and W_{right}, the Roe flux presented above is a good approximation, but if a rarefaction containing a **sonic** point is present among the nonlinear waves that solves the discontinuity problem between W_{left} and W_{right}, it has been early remarked that for this very particular situation, the Roe flux does not satisfy the entropy condition (see *e.g.* Godlewski and Raviart [GR96]).

• A popular response has been proposed by Harten [Ha83] with a tuning parameter that plays in fact the role of an artificial viscosity and P. Roe himself [Roe85] has proposed a nonparameterized entropy correction for his flux. With G. Mehlman, we have treated the same subject by the introduction of hyperbolic nonlinear models with nonconvex flux functions and have proved a discrete entropy inequality if sufficiently weak nonlinear waves are present in the problem [DM96]. We detail here the modification of the algorithm that we have proposed and tested numerically for various gas dynamics problems.

• We introduce as above two states $W_l \equiv W^0$ and $W_r \equiv W^3$ and the Roe matrix $A(W_l, W_r)$ described in the preceding sub-sections. We have in particular the relation

(3.5.1) $\quad W_r - W_l \;\equiv\; W^3 - W^0 \;=\; \sum_{j=1}^{j=3} \alpha_j\, r_j^*$

and we do not make the confusion between the eigenvalues $\lambda_j(W^0)$ of the left state, $\lambda_j(W^3)$ of the right state, and λ_j^* of the Roe matrix. We introduce the following two intermediate states W^1 and W^2 according to

(3.5.2) $\quad W^1 \;=\; W^0 + \alpha_1\, r_1^*\,, \qquad W^2 \;=\; W^0 + \alpha_1\, r_1^* + \alpha_2\, r_2^* \;=\; W^3 - \alpha_3\, r_3^*$

and illustrated on Figure 3.2. Note that $\lambda_j(W^k)$ is well defined for $j = 1, 2, 3$ and $k = 0, 1, 2, 3$: it is the j^0 eigenvalue of the k^0 intermediate state W^k. We define now the set S of **sonic indices** by the condition that the sign of the j^0 eigenvalue is increasing from negative to positive values accross some j-wave :

(3.5.3) $\quad S \;=\; \big\{\, j \in \{1, 2, , 3\}, \quad \lambda_j(W^{j-1}) \,<\, 0 \,<\, \lambda_j(W^j) \,\big\}\,.$

The modification of the Roe flux is active only for the sonic indices and we introduce a polynomial p_j **of degree 3** by the classical Hermite interpolation conditions

(3.5.4) $$\begin{cases} p_j(0) \,=\, 0\,, & p_j'(0) \,=\, \lambda_j(W^{j-1})\,, \\ p_j(\alpha_j) \,=\, \lambda_j^*\, \alpha_j\,, & p_j'(\alpha_j) \,=\, \lambda_j(W^j)\,, \end{cases} \qquad j \in S$$

that defines explicitely the polynomial $p_j(\bullet)$ by the algebraic relation

(3.5.5) $$\begin{cases} p_j(\xi) \;=\; \dfrac{\lambda_j(W^j) + \lambda_j(W^{j-1}) - 2\,\lambda_j^*}{(\alpha_j)^2}\, \xi^3 \;\; + \\ \qquad + \;\; \dfrac{3\,\lambda_j^* - 2\,\lambda_j(W^{j-1}) - \lambda_j(W^j)}{\alpha_j}\, \xi^2 \;+\; \lambda_j(W^{j-1})\, \xi\,. \end{cases}$$

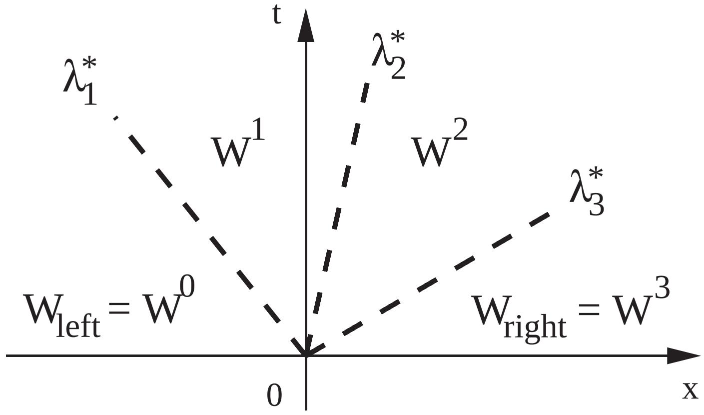

Figure 3.2. Intermediate states for the entropy correction of the Roe upwind scheme.

• With the hypothesis that $j \in S$, it is not difficult to see [DM96] that the polynomial $p_j(\bullet)$ has a unique minimum inside the interval $(0, \alpha_j)$. The argument ξ_j^* of this point of minimum is given according to :

(3.5.6) $$\xi_j^* \,=\, \frac{-\lambda_j(W^{j-1})\;\alpha_j}{\begin{pmatrix} \big(3\,\lambda_j^* - 2\,\lambda_j(W^{j-1}) - \lambda_j(W^j)\big) + \\ + \sqrt{\big(3\,\lambda_j^* - \lambda_j(W^j) - \lambda_j(W^{j-1})\big)^2 - \lambda_j(W^{j-1})\,\lambda_j(W^j)} \end{pmatrix}}\,.$$

Since $p_j(\xi_j^*)$ is the unique minimum of the polynomial $p_j(\bullet)$ on the interval $(0,\,\alpha_j)$, we have $\frac{p_j(\xi_j^*)}{\alpha_j} \leq 0$ and $\frac{p_j(\xi_j^*)}{\alpha_j} \leq \lambda_j^*$. Then the modified flux $\Phi^{\rm modif}(W_l,\;W_r)$ is defined from the Roe flux $\Phi(W_l,\,W_r)$ by the relation

(3.5.7) $$\Phi^{\rm modif}(W_l,\,W_r) = \Phi(W_l,\,W_r) + \sum_{j\in S} \max\Big(\frac{p_j\big(\xi_j^*\big)}{\alpha_j},\,\frac{p_j\big(\xi_j^*\big)}{\alpha_j} - \lambda_j^*\Big)\,\alpha_j\;r_j^*$$

that makes the added numerical viscosity explicit.

3.6 Nonlinear flux boundary conditions.

• At the two extremities $x=0$ and $x=L$ of the pipe, we have to express on one hand the datum of a given nonstationary pressure $\Pi(t)$ at $x=0$ and on the other hand a free output of the waves at $x=L$.

• For the numerical boundary condition for pressure, we follow a general approach founded on the so-called **partial Riemann problem** [Du01] that generalizes to nonlinear hyperbolic systems the reflection operator of relation (2.5.22). For a given discrete time $t^n = n\,\Delta t$, and a given state $W_{1/2}^n \equiv W_r$ in the first cell of the unidimensional mesh, we construct a boundary state $W_0^n \equiv W_l$ that satisfies the boundary constraint

(3.6.1) $$p(W_0^n) \;=\; \Pi^{n+1/2}\,, \quad n \geq 0\,,$$

and moreover, we impose that the state $W_{1/2}^n$ present in the first cell is issued from the boundary state W_0^n with an ingoing 3-wave, *i.e.* we impose the relation

(3.6.2) $$W_{1/2}^n - W_0^n \;=\; \alpha_3\;r_3^*$$

as illustrated on Figure 3.3. This problem has a unique solution, as claims the

Proposition 3.7. Pressure flux boundary condition with Roe matrix.
We consider a left boundary condition associated with a pressure Π and a right datum defined by state W_r. Then there exists a **unique** left state W_l that satisfies the pressure condition

(3.6.3) $$p(W_l) \;=\; \Pi$$

and such that when we construct the Roe intermediate state W^* according to the relations (3.3.4) to (3.3.6), the difference $W_r - W_l$ has only one component over the third eigenvector of the Roe matrix ${\rm d}F(W^*)$, *i.e.* the relation (3.4.5) can been written under the form

(3.6.4) $\quad W_r - W_l \;=\; \alpha_3 \; r_3^* \,.$

The density ρ_l and the velocity u_l of the left state W_l are given according to the relations

(3.6.5) $$\rho_l \;=\; \frac{(\gamma{+}1)\,\Pi \;+\; (\gamma{-}1)\,p_r}{(\gamma{-}1)\,\Pi \;+\; (\gamma{+}1)\,p_r}\,\rho_r$$

(3.6.6) $$u_l \;=\; u_r \;+\; (\Pi - p_r)\,\sqrt{\frac{2}{\rho_r\,\big((\gamma{-}1)\,\Pi \;+\; (\gamma{+}1)\,p_r\big)}}\,.$$

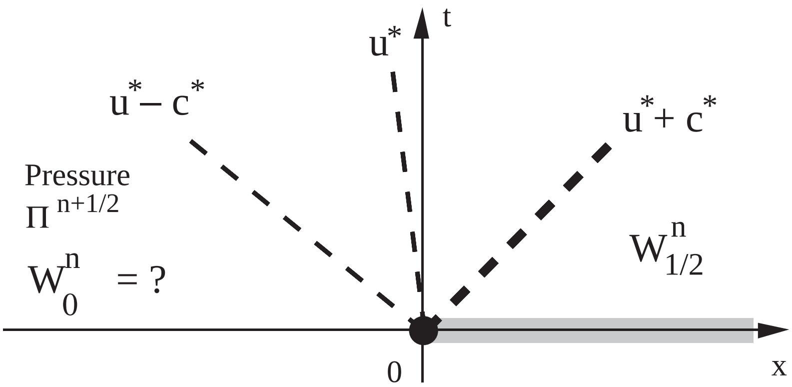

Figure 3.3. Nonlinear boundary condition for given pressure at $x = 0$ with the Roe upwind scheme.

• The **proof of Proposition 3.7** is a consequence precisely of the preceding subsections about the Roe flux. We first remark that the relations (3.4.5) and (3.6.4) are absolutly **identical**. Then we deduce that necessarily $\alpha_1 = \alpha_2 = 0$ and according to the relations (3.4.11) and (3.4.12), we get

(3.6.7) $\quad \Pi - \rho^*\,c^*\,u_l \;=\; p_r - \rho^*\,c^*\,u_r$

(3.6.8) $\quad \Pi - (c^*)^2\,\rho_l \;=\; p_r - (c^*)^2\,\rho_r \,.$

We deduce simply $\;u_r - u_l \;=\; \dfrac{p_r - \Pi}{\rho^*\,c^*} \;=\; \dfrac{c^*\,(\rho_r - \rho_l)}{\rho^*}\;$ and due to the relation (3.4.9), we get :

$$(c^*)^2 = \frac{1}{(\sqrt{\rho_r}+\sqrt{\rho_l})^2}\Big[\,\frac{\gamma{-}1}{2}\,\rho^*\,\Big(\frac{c^*\,(\rho_r-\rho_l)}{\rho^*}\Big)^2 + \gamma\,\Big(\frac{\rho_l+\rho^*}{\rho_l}\,\Pi + \frac{\rho^*+\rho_r}{\rho_r}\,p_r\Big)\Big]\,.$$

Then after multiplication by $(\rho_r - \rho_l)$, we obtain with the help of (3.6.8) :

$$0 \;=\; (p_r - \Pi) \;-\; \frac{\gamma{-}1}{2}\,(p_r - \Pi)\,\frac{(\sqrt{\rho_r}-\sqrt{\rho_l})^2}{\rho^*}$$

$$-\,\gamma\,\frac{\sqrt{\rho_r}-\sqrt{\rho_l}}{\sqrt{\rho_r}+\sqrt{\rho_l}}\,\Big(\frac{\rho_l+\rho^*}{\rho_l}\,\Pi \;+\; \frac{\rho^*+\rho_r}{\rho_r}\,p_r\Big) \qquad = \qquad \gamma\,(p_r - \Pi)$$

$$-\,\frac{\gamma{-}1}{2}\,(p_r - \Pi)\,\Big(\frac{\sqrt{\rho_r}}{\sqrt{\rho_l}} + \frac{\sqrt{\rho_l}}{\sqrt{\rho_r}}\Big) \;-\; \gamma\,\Big(\frac{\sqrt{\rho_r}}{\sqrt{\rho_l}} - 1\Big)\,\Pi \;-\; \gamma\,\Big(1 - \frac{\sqrt{\rho_l}}{\sqrt{\rho_r}}\Big)\,p_r\,.$$

We multiply the previous equality by $\sqrt{\frac{\rho_l}{\rho_r}}$ and we get :

$$-\frac{\gamma-1}{2}\,(p_r-\Pi)\,(\,1+\frac{\rho_l}{\rho_r}\,)\,-\,\gamma\,\Pi\,+\,\gamma\,\frac{\rho_l}{\rho_r}\,p_r \;=\; 0\,,$$

$$id\ est \quad \Big[-\frac{\gamma-1}{2}\,(p_r-\Pi)\,+\,\gamma\,p_r\Big]\,\frac{\rho_l}{\rho_r} \;=\; \gamma\,\Pi\,+\,\frac{\gamma-1}{2}\,(p_r-\Pi)$$

and the relation (3.6.5) is established.

• We deduce from the previous relation :

$$\rho^* \;=\; \sqrt{\frac{(\gamma+1)\,\Pi\,+\,(\gamma-1)\,p_r}{(\gamma-1)\,\Pi\,+\,(\gamma+1)\,p_r}}\;\rho_r\,,\qquad \rho_r-\rho_l \;=\; \frac{2\,(p_r-\Pi)}{(\gamma-1)\,\Pi\,+\,(\gamma+1)\,p_r}\;\rho_r$$

and due to the relation (3.6.8) :

$$(c^*)^2 \;=\; \frac{(\gamma-1)\,\Pi\,+\,(\gamma+1)\,p_r}{2\,\rho_r}\,,\qquad \rho^*\,c^* \;=\; \sqrt{\frac{(\gamma-1)\,\Pi\,+\,(\gamma+1)\,p_r}{2}}\;\sqrt{\rho_r}$$

and the relation (3.6.6) is an easy consequence of the last equality joined with (3.6.7). The proposition 3.7 is established. □

• The determination of a **nonlinear nonreflecting** boundary condition at $x=L$ is still an open mathematical problem. We recommand for deriving a flux boundary condition for such a situation to impose that **no wave** are present at the interaction for the last interface $j=J$. We just write

$$(3.6.9) \qquad f_J^{n+1/2} \;=\; F(W_{J-1/2}^n)\,, \qquad n \geq 0$$

which is equivalent of introducing a right boundary state W_J^n according to the simple relation $W_J^n \,=\, W_{J-1/2}^n$ and then making these two states interacting with the Roe flux : $f_J^{n+1/2} \;=\; \Phi(W_{J-1/2}^n,\, W_J^n)$. This last definition is equivalent to the one proposed in (3.6.9) due to the property (3.3.2) of the Roe matrix.

4) Second order and two space dimensions.

4.1 Towards second order accuracy.

• The finite volume method described in the previous sections is a natural method for the discretization of systems of m conservation laws. It conducts to an explicit scheme in time : the evaluation of the field W^{n+1} at time step $(n+1)\,\Delta t$ needs only the knowledge of the field $W_{j+1/2}^n$ for $j\,=\,0,\cdots,J-1$ at the preceding time step $n\,\Delta t$. This evaluation needs a certain number of auxiliary computations without the resolution of any linear system involving the new field. The method is parameterized by the choice of a numerical flux and a great flexibility can be adopted at this level. We have proposed two fluxes for nonlinear problems related to nonlinear acoustics and gas dynamics, the Roe flux

$\Phi(\bullet,\,\bullet)$ of relations (3.4.6)-(3.4.8) that conduct to a discrete scheme according to the relation

$$f_j^{n+1/2} \,=\, \Phi(W_{j-1/2}^n,\, W_{j+1/2}^n)\,, \qquad (4.1.1)$$

and the modified Roe flux $\Phi^{\rm modif}(\bullet,\,\bullet)$ of relation (3.6.7) that enforces the entropy condition. This explicit version of the finite volume method is submitted to a **stability** condition that can be written as a first approximation for linear cases as :

$$c_0\,\frac{\Delta t}{\Delta x} \,\leq\, 1\,. \qquad (4.1.2)$$

• Nevertheless, the above finite volume method is only **first order accurate**. If we insert an exact solution $W(x,\,t)$ of the conservation law

$$\frac{\partial W}{\partial t} \,+\, \frac{\partial}{\partial x}F(W) \,=\, 0 \qquad (4.1.3)$$

inside the formal expression of the flux (4.1.1), it is easy to see that the finite difference $\frac{1}{\Delta x}(f_{j+1}^{n+1/2} - f_j^{n+1/2})$ is first order accurate :

$$\frac{1}{\Delta x}\big(f_{j+1}^{n+1/2} - f_j^{n+1/2}\big) \,=\, \Big(\frac{\partial F(W)}{\partial x}\Big)_{j+1/2}^{n+1/2} \,+\, {\rm O}\big(\Delta t + \Delta x\big)\,. \qquad (4.1.4)$$

In a similar way, the use of an explicit scheme in time conducts to

$$\frac{1}{\Delta t}\,(W_{j+1/2}^{n+1} - W_{j+1/2}^n) \,+\, \frac{1}{\Delta x}\big(f_{j+1}^{n+1/2} - f_j^{n+1/2}\big) \,=\, 0 \qquad (4.1.5)$$

and maintains this first order accuracy for the finite volume scheme.

• We develop in this section the fact that it is possible to improve the method, *i.e.* to define a method with a relation of the type (4.1.5), and that conduct to a troncation error **of second order** :

$$\left\{ \begin{array}{l} \displaystyle\frac{1}{\Delta t}\,(W_{j+1/2}^{n+1} - W_{j+1/2}^n) \,+\, \frac{1}{\Delta x}\big(f_{j+1}^{n+1/2} - f_j^{n+1/2}\big) \,= \\ \displaystyle\qquad\qquad = \Big(\frac{\partial W}{\partial t} \,+\, \frac{\partial}{\partial x}F(W)\Big)_{j+1/2}^{n+1/2} \,+\, {\rm O}\big(\Delta t^2 + \Delta x^2\big)\,. \end{array} \right. \qquad (4.1.6)$$

The price to pay is to develop flux formulae much more complicated than the simple relation (4.1.1). When the second order precision (4.1.6) is achieved with a stable scheme, the precision is sufficient to develop predictive computations in acoustics and aerodynamics, whereas that is not the case with the initial scheme (4.1.1) (4.1.5).

4.2 The method of lines.

• The simplest way to extend the first order finite volume scheme is first to develop a new vision of the method with emphasis more on abstraction. We have presented a method founded on the integration of the conservation law (4.1.3) inside the space-time domain $V_{j+1/2}^{n+1/2} =]x_j \,,\, x_{j+1}[\times]t^n ,\, t^{n+1}[$ as suggested in (2.1.6). With the method of lines, we just integrate the conservation (4.1.3) in **space** in each control volume $K_{j+1/2} =]x_j \,,\, x_{j+1}[$. It is straightforward to introduce the mean value $W_{j+1/2}(t)$ in this finite element :

(4.2.1) $$W_{j+1/2}(t) \;=\; \frac{1}{\mid K_{j+1/2} \mid} \int_{x_j}^{x_{j+1}} W(x,\, t)\, \mathrm{d}x \,;$$

then we integrate the conservation law (4.1.3) in space in the cell $K_{j+1/2}$ and taking into account the relation $\frac{\mathrm{d}}{\mathrm{d}t} W_{j+1/2}(t) \;=\; \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \frac{\partial W}{\partial t}(x,\, t)\, \mathrm{d}x,$ we get simply

(4.2.2) $$\frac{\mathrm{d}}{\mathrm{d}t} W_{j+1/2}(t) \;+\; \frac{1}{\Delta x} \big[\, f_{j+1}(t) \;-\; f_j(t) \,\big] \;=\; 0$$

with

(4.2.3) $$f_j(t) \;=\; F\big(W(x_j,\, t)\big).$$

• As usual with the finite volume method, a numerical scheme can be obtained from the relations (4.2.2) (4.2.3) by replacing the relation (4.2.3) by some explicit function over the set of all discrete variables introduced for the relation (4.2.1). To fix the ideas, we introduce a **dynamic state vector** $Z(t)$ composed by all the dynamic variables on the finite mesh :

(4.2.4) $$Z(t) \;=\; \big(\, W_{1/2}(t),\, \cdots,\, W_{j+1/2}(t),\, \cdots,\, W_{J-1/2}(t) \,\big) \quad \in \; (\mathbb{R}^m)^J \,.$$

The discretization in space is achieved if we are able to determine the numerical flux $f_j(t)$ with the help of both the dynamic state vector $Z(\bullet)$ and the boundary conditions, that is the input pressure $\Pi(t)$ in the example considered in the last section. As in relation (2.1.12), we introduce a **local numerical flux function** $\Psi_j(\bullet,\, \bullet)$ relative to the vertex x_j :

(4.2.5) $$f_j(t) \;=\; \Psi_j\big(\Pi(t),\, Z(t)\big) \,.$$

We replace the relation (4.2.3) by the numerical approximation (4.2.5) inside the equation (4.2.2) of dynamic evolution of the state variable $W_{j+1/2}(\bullet)$. We obtain the following **ordinary differential equation**

(4.2.6) $$\frac{\mathrm{d}}{\mathrm{d}t} W_{j+1/2}(t) \;+\; \frac{1}{\Delta x} \big[\, \Psi_{j+1}\big(\Pi(t),\, Z(t)\big) \;-\; \Psi_j\big(\Pi(t),\, Z(t)\big) \,\big] \;=\; 0 \,.$$

• The **method of lines** is a semi-discrete version of the finite volume method. It is obtained by integration in space of the conservation law without integration in time. The result is not a numerical scheme but *just* an ordinary differential equation for the dynamic state vector $Z(\bullet)$ described component by component with the equation (4.2.6). The method is parameterized by the local numerical flux functions $\Psi_j\big(\bullet,\, \bullet\big)$ and take the general form of a dynamical system parameterized by the pressure function $t \longmapsto \Pi(t)$:

(4.2.7) $$\frac{\mathrm{d}}{\mathrm{d}t} Z(t) \;=\; G(Z,\, t) \,.$$

The **discrete dynamic function** $(\mathbb{R}^m)^J \times [0, +\infty[\ni (Z, t) \mapsto G(Z, t) \in (\mathbb{R}^m)^J$ is a vector valued expression with J components :

(4.2.8) $G(Z, t) = \big(G_{1/2}(Z, t), \cdots, G_{j+1/2}(Z, t), \cdots, G_{J-1/2}(Z, t) \big) \in (\mathbb{R}^m)^J$

and it is defined from the $(J+1)$ local numerical fluxes $(\Psi_j)_{j=0, \cdots, J}$ with the very simple algebra relative to the finite volume method :

(4.2.9) $$G_{j+1/2}(Z, t) = -\frac{1}{\Delta x}\big(\Psi_{j+1}(\Pi(t), Z) - \Psi_j(\Pi(t), Z) \big) , \; 0 \leq j \leq J-1 .$$

Proposition 4.1. Explicit Euler scheme.
With the choice of the first order scheme in space, that is

(4.2.10) $$\Psi_j\big(\Pi^n, Z^n\big) = \Phi(W^n_{j-1/2}, W^n_{j+1/2}) \qquad \text{if} \quad j = 1, \cdots, J-1,$$

and the first order **explicit** forward Euler scheme for the ordinary differential equation (4.2.7), *id est*

(4.2.11) $$\frac{1}{\Delta t}\,(Z^{n+1} - Z^n) = G(Z^n, t^n) ,$$

we recover the previous first order finite volume scheme

(4.2.12) $$\left\{ \begin{array}{l} \frac{1}{\Delta t}\big(W^{n+1}_{j+1/2} - W^n_{j+1/2}\big) + \frac{1}{\Delta x}\Big(\Phi\big(W^n_{j+1/2}, W^n_{j+3/2}\big) \\ \qquad\qquad - \Phi\big(W^n_{j-1/2}, W^n_{j+1/2}\big)\Big) = 0 \qquad \text{for} \quad j = 1, \cdots, J-2. \end{array} \right.$$

• We write the relation (4.2.10) for the particular control volume $K_{j+1/2}$ and we get : $W^{n+1}_{j+1/2} = W^n_{j+1/2} + \Delta t \; G(z^n, t^n)$ due to (4.2.11), then

$$W^{n+1}_{j+1/2} = W^n_{j+1/2} - \frac{1}{\Delta x}\big(\Psi_{j+1}(\Pi^n, Z^n) - \Psi_j(\Pi^n, Z^n) \big) \qquad \text{due to (4.2.9)}$$

$$= W^n_{j+1/2} - \frac{1}{\Delta x}\big(\Phi(W^n_{j+1/2}, W^n_{j+3/2}) - \Phi(W^n_{j-1/2}, W^n_{j+1/2}) \big) \quad \textit{c.f.} \text{ (4.2.10)}$$

and the relation (4.2.12) is established. □

4.3 The method of Van Leer.

• We turn now to the construction of a second order accurate version of the finite volume method as proposed initially with the "Multidimensional Upwind-centered Scheme for Conservation Laws" of B. Van Leer [VL79]. The fundamental idea of this scheme is the reconstruction of a function $\mathbb{R} \ni x \longmapsto W(x) \in \mathbb{R}$ from his mean values $W_{j+1/2}$ in each cell $K_{j+1/2}$. The reconstructed function is regular inside each control volume $K_{j+1/2}$ and is discontinuous at the interfaces x_l between two control volumes. The application to the finite volume method replaces the scheme (4.1.1) by the same Roe flux interaction $\Phi(\bullet, \bullet)$ considered for the two extrapolated data W^-_j and W^+_j on each side of the boundary :

(4.3.1) $$f_j = \Phi(W^-_j, W^+_j) .$$

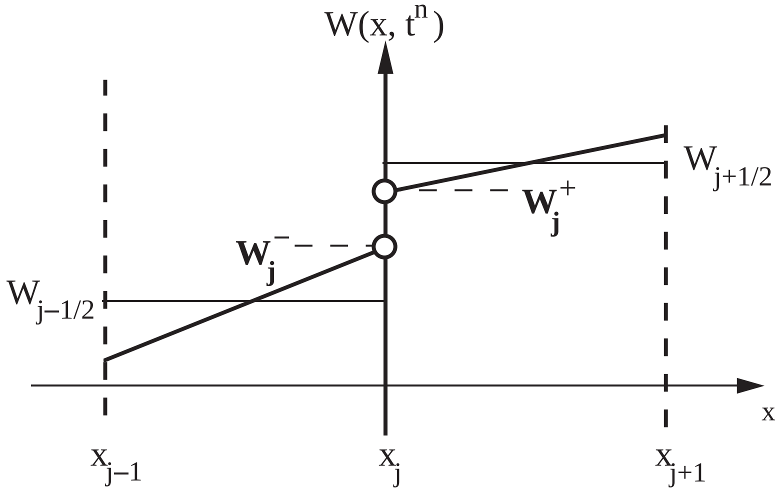

Figure 4.1. First order and second order interpolation at the interface x_j.

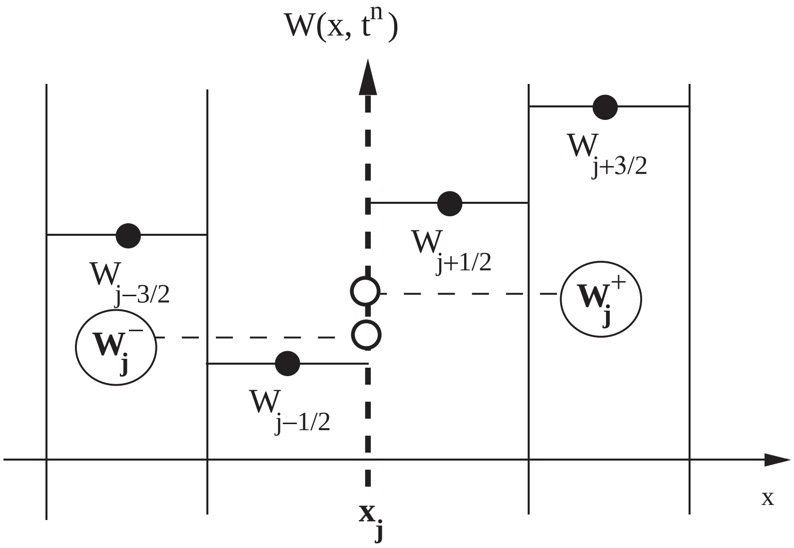

Figure 4.2. Construction of the nonlinear interpolated values W_j^- and W_j^+ with the Van Leer method.

• The simplest case is illustrated on Figure 4.1. It imposes simply the reconstructed function $W(x)$ to be **constant** in each interval :

(4.3.2) $\quad W(x) \,\equiv\, W_{j+1/2}\,, \qquad x_j \,<\, x \,<\, x_{j+1}\,.$

The two limit values on each side of the interface x_j are the following ones : $W_j^- \,=\, W_{j-1/2}\,,\quad W_j^+ \,=\, W_{j+1/2}$ and the explicit version of this finite volume scheme is the standard first order numerical flux (4.1.1) as seen at Proposition 4.1. In the context of the method of lines, we obtain :

(4.3.3) $\quad f_j \,=\, \Phi(W_{j-1/2},\, W_{j+1/2})\,.$

• The second order accurate Muscl method consists first in restricting the methodology to a **scalar** field $W(\bullet)$ and to construct an **affine** function in each interval $K_{j+1/2}$ instead of a constant function as in (4.3.2). We set

(4.3.4) $$W(x) \equiv W_{j+1/2} + p_{j+1/2}\,\big(x - x_{j+1/2}\big)\,, \qquad x_j < x < x_{j+1}\,.$$

The simplest choice for a slope is the one of the **centered** scheme :

(4.3.5) $$p_{j+1/2} = \frac{1}{\Delta x}(W_{j+3/2} - W_{j-1/2})\,,$$

and due to (4.3.4) and (4.3.5), the extrapolated values W_j^- and W_j^+ on each side of the interface located at the position x_j are the following ones :

(4.3.6) $$W_j^- = W_{j-1/2} + \frac{1}{4}\,(W_{j+1/2} - W_{j-3/2})$$

(4.3.7) $$W_j^+ = W_{j+1/2} - \frac{1}{4}\,(W_{j+3/2} - W_{j-1/2})\,.$$

• The choice of a numerical flux given according to the relations

(4.3.8) $$f_j = \Phi\big(W_{j-1/2} + \frac{1}{4}(W_{j+1/2} - W_{j-3/2})\,,\, W_{j+1/2} - \frac{1}{4}(W_{j+3/2} - W_{j-1/2})\big)$$

lead to an **unstable** scheme when we consider the particular case of the advection equation with the first order explicit scheme in time.

Proposition 4.2. Linear Muscl scheme is unstable.
We apply the linear Muscl approach for the advection equation. Then the numerical scheme obtained by association of (4.3.8) and the upwind scheme (2.1.19) conducts to the following explicit first order scheme :

(4.3.9) $$\left\{ \begin{array}{l} W_{j+1/2}^{n+1} - W_{j+1/2}^{n} + \dfrac{a\,\Delta t}{\Delta x}\Big(\big(W_{j+1/2}^{n} + \dfrac{1}{4}\,(W_{j+3/2}^{n} - W_{j-1/2}^{n}\big) \\ \qquad\qquad - \big(W_{j-1/2}^{n} + \dfrac{1}{4}\,(W_{j+1/2}^{n} - W_{j-3/2}^{n})\big)\Big) = 0\,. \end{array} \right.$$

This scheme is unstable for each $\Delta t > 0$.

• Due to the expression (2.1.19) of the upwind scheme, the discrete first order in time advection equation can be written :

$$W_{j+1/2}^{n+1} - W_{j+1/2}^{n} + \frac{a\,\Delta t}{\Delta x}\,\big(W_{j+1}^{-,\,n} - W_{j}^{-,\,n}\big) = 0$$

and the expression (4.3.9) is a consequence of the left extrapolation (4.3.6). For the study of stability, we introduce a profile of the type $W_{j+1/2}^n = \mathrm{e}^{(i\,k\,(j+1/2)\,\Delta x)}$ with a wave number k. The scheme (4.3.9) can be written as

$$W_{j+1/2}^{n+1} = g\big(k\,\Delta x,\, \frac{a\,\Delta t}{\Delta x}\big)\,W_{j+1/2}^{n}$$

with an amplification coefficient $g(\xi,\,\sigma)$ $\,(\xi = k\,\Delta t,\ \sigma = \frac{a\,\Delta t}{\Delta x})$ given simply by the expression

$$g(\xi,\,\sigma) = 1 - \sigma\,\big(1 - \mathrm{e}^{-i\,\xi}\big) - \frac{\sigma}{4}\,\big(\mathrm{e}^{i\,\xi} - 1 - \mathrm{e}^{-i\,\xi} + \mathrm{e}^{-2\,i\,\xi}\big)\,, \qquad \text{then}$$

$$g(\xi,\,\sigma) \;=\; 1 - \sigma\,(\,1 - \cos\xi\,) - \frac{\sigma}{4}\,(\,-1 + \cos 2\xi\,) - i\,\sigma\,\Big(\sin\xi + \frac{1}{2}\sin\xi - \frac{1}{4}\sin 2\xi\Big)$$

$$\;=\; 1 - \sigma\,\Big(1 - \cos\xi + \frac{1}{4}\,(\,2\cos^2\xi - 2\,)\Big) - i\,\sigma\,\Big(\frac{3}{2}\sin\xi - \frac{1}{2}\sin\xi\cos\xi\Big)\,,$$

$$(4.3.10)\qquad g(\xi,\,\sigma) \;=\; 1 - \frac{\sigma}{2}\,(\,1 - \cos\xi\,)^2 - i\,\frac{\sigma}{2}\sin\xi\,(\,3 - \cos\xi\,)\,.$$

For ξ arbitrarily small, we deduce from (4.3.10) : $|\,g(\xi,\,\sigma)\,|^2 = 1 + \sigma^2\,\xi^2 + \mathrm{O}\big(\xi^4\big)$ which establishes the instability for all $\sigma \neq 0$. □

• The above remark motivates the introduction of so-called **slope limiters**, intensively studied during the period 1980-90 after the pioneering work of Van Leer [VL77]. The idea is to search an interpolation W_j^- of the field $W(\bullet)$ at the left of the point x_j from the neighbouring mean values $W_{j-3/2}$, $W_{j-1/2}$ and $W_{j+1/2}$ and by left-right invariance of the procedure, to construct an interpolated value W_j^+ from the first right neighbours $W_{j-1/2}$, $W_{j+1/2}$ and $W_{j+3/2}$ as suggested on Figure 4.2. We replace the relations (4.3.6) and (4.3.7) by a **nonlinear interpolation** parameterized by a slope limiter $\mathbb{R} \ni r \longmapsto \varphi(r) \in \mathbb{R}$:

$$(4.3.11)\qquad W_j^- \;=\; W_{j-1/2} + \frac{1}{2}\,\varphi\Big(\frac{W_{j-1/2} - W_{j-3/2}}{W_{j+1/2} - W_{j-1/2}}\Big)\,\big(W_{j+1/2} - W_{j-1/2}\big)$$

$$(4.3.12)\qquad W_j^+ \;=\; W_{j+1/2} - \frac{1}{2}\,\varphi\Big(\frac{W_{j+3/2} - W_{j+1/2}}{W_{j+1/2} - W_{j-1/2}}\Big)\,\big(W_{j+1/2} - W_{j-1/2}\big)\,.$$

We remark that the relations (4.3.6) and (4.3.7) are a particular case of the general nonlinear relations (4.3.11) and (4.3.12) with the particular choice $\varphi(r) = \frac{1}{2}\,(1 + r)$. The limiter function satisfies very often the functional relation

$$(4.3.13)\qquad \varphi(r) \;\equiv\; r\,\varphi\Big(\frac{1}{r}\Big)\,,\qquad r > 0\,.$$

Among all the possible choices, we have adopted for fluid mechanics [DM92] the so-called STS-limiter defined by the relations

$$(4.3.14)\qquad \varphi^{STS}(r) \;=\; \begin{cases} 0\,, & r \leq 0 \\ \dfrac{3}{2}\,r\,, & 0 \leq r \leq \dfrac{1}{2} \\ \dfrac{1+r}{2}\,, & \dfrac{1}{2} \leq r \leq 2 \\ \dfrac{3}{2}\,, & r \geq 2\,, \end{cases}$$

and illustrated on Figure 4.3.

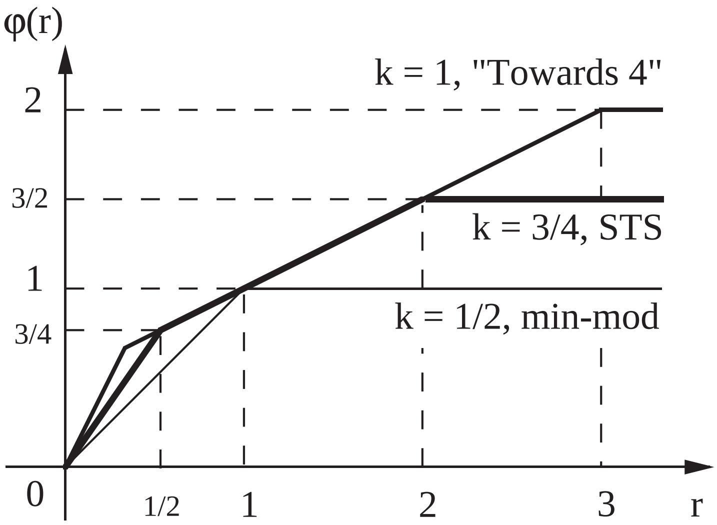

Figure 4.3. Examples of limiter functions that can be easily extended to unstructured meshes.

4.4 Second order accurate finite volume method for fluid problems.

• We detail in this section a generalization for unstructured meshes of the Muscl scheme proposed by Van Leer [VL79]. At one space dimension on a uniform mesh, it is classical to consider a scalar field z among the primitive variables, *id est*

(4.4.1) $$z \in \{ \rho , u , v , p \} \qquad \text{(primitive variables)}$$

and instead of computing the interface flux with relation (4.1.1), to first construct two interface states W_S^- and W_S^+ on each side of the interface S. Then the flux is evaluated by the decomposition of the discontinuity :

(4.4.2) $$f_S = \Phi\big(W_S^- , W_S^+\big) , \quad S \in \big\{ x_1, \cdots , x_{J-1} \big\} .$$

This nonlinear interpolation is done with a slope limiter $\varphi(\bullet)$ that operates on each variable proposed in (4.4.1) and we have typically when a left-right invariance is assumed [Du91] :

(4.4.3) $$z_S^- = z_{j-1/2} + \frac{1}{2} \varphi\bigg(\frac{z_{j-1/2} - z_{j-3/2}}{z_{j+1/2} - z_{j-1/2}}\bigg) \big(z_{j+1/2} - z_{j-1/2}\big) , \quad S = x_j$$

(4.4.4) $$z_S^+ = z_{j+1/2} - \frac{1}{2} \varphi\bigg(\frac{z_{j+3/2} - z_{j+1/2}}{z_{j+1/2} - z_{j-1/2}}\bigg) \big(z_{j+1/2} - z_{j-1/2}\big) , \quad S = x_j .$$

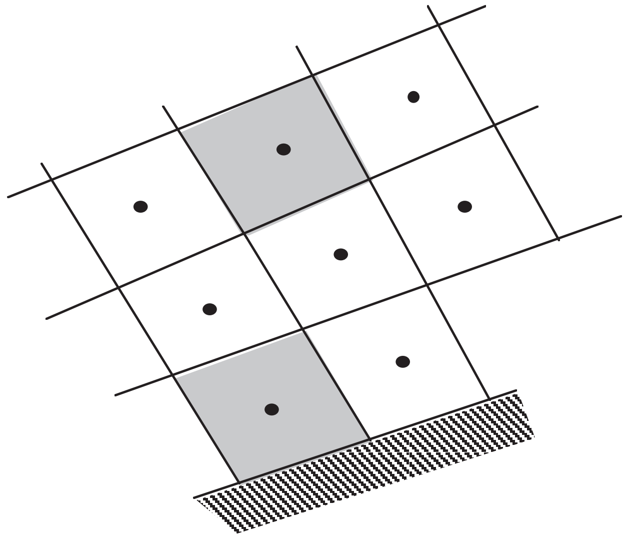

Figure 4.4. Structured Cartesian mesh.
The control volumes are exactly the elements of mesh $\mathcal{T}$.

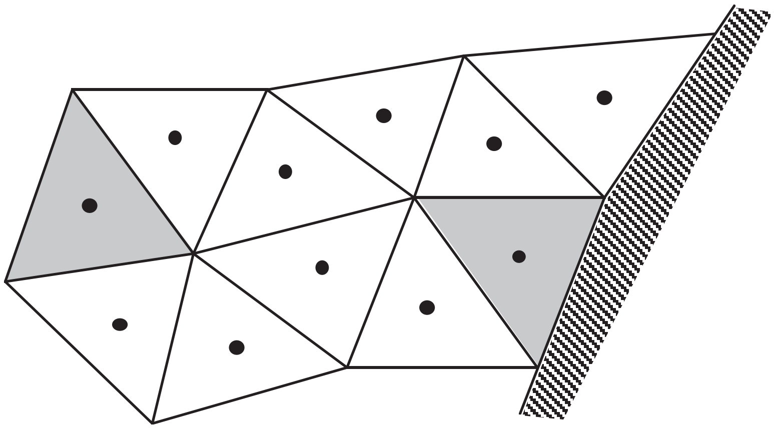

Figure 4.5. Unstructured mesh composed by triangular elements.
The control volumes are exactly the elements of the mesh.

• We focus now on the use of unstructured meshes for the extension to second order accuracy of the finite volume method. As in the one-dimensional case, the domain of study is decomposed into finite elements (or control volumes) $K \in \mathcal{E}_{\mathcal{T}}$ than can be structured in a Cartesian way (Figure 4.4) or with a cellular complex as in Figure 4.5. In both cases, the intersection of two finite elements define an interface $f \in \mathcal{F}_{\mathcal{T}}$. We denote by $\mathbf{n}_f$ the normal at the interface f that separates a left control volume $K_l(f)$ and a right control volume $K_r(f)$. The ordinary differential equation (4.2.6) is replaced by a multidimensional version :

$$(4.4.5) \qquad | K | \; \frac{\mathrm{d}W_K}{\mathrm{d}t} + \sum_{f \subset \partial K} | f | \; \Phi\big(W_K \, , \, \mathbf{n}_f \, , \, W_{K_r(f)}\big) \; = \; 0 \, , \qquad K \in \mathcal{E}_{\mathcal{T}} \, .$$

For internal interfaces, the function $\Phi(\bullet\,, \mathbf{n}_f\,, \bullet)$ is equal *e.g.* to the Roe flux between states $W_{K_l(f)}$ and $W_{K_r(f)}$ in the one-dimensional direction along normal $\mathbf{n}_f$ in order to take into account the invariance by rotation of the equations of gas dynamics (see [GR96]).

• We consider now a finite element K internal to the domain. The extension to second order accuracy of the finite volume scheme consists in replacing the arguments $W_{K_l(f)}$ and $W_{K_r(f)}$ in relation (4.4.5) by nonlinear extrapolations $W_{K_l(f),\,f}$ and $W_{K_r(f),\,f}$ on each side of the boundary of state data and evaluated as described in what follows. We first introduce the set $\mathcal{N}(K)$ of neighbouring cells of given finite element $K \in \mathcal{E}_\mathcal{T}$, as illustrated on Figure 4.6 :

(4.4.6) $$\mathcal{N}(K) \,=\, \big\{ L \in \mathcal{E}_\mathcal{T}, \quad \exists\, f \in \mathcal{F}_\mathcal{T}, \quad f \,\subset\, \partial K \cap \partial L \big\}\,.$$

For $L \in \mathcal{N}(K)$, we suppose by convention that the normal $\mathbf{n}_f$ to the face $f \subset \partial K \cap \partial L$ is external to the element K *id est* $K_r(f) = K$, $K_l(f) = L$. We introduce also the point $y_{K,\,f}$ on the interface $f \subset \partial K$ that links the barycenters x_K and $x_{K_r(f)}$:

(4.4.7) $$\left\{ \begin{array}{l} y_{K,\,f} \;\equiv\; (1-\theta_{K,\,f})\, x_K \;+\; \theta_{K,\,f}\, x_{K_r(f)}\,, \qquad y_{K,\,f} \in f\,, \\ \qquad\qquad f \subset \partial K\,, \qquad K \text{ finite element internal to mesh } \mathcal{T}\,. \end{array} \right.$$

Then, following Pollet [Po88], for z equal to one scalar variable of the family :

(4.4.8) $$z \,\in\, \{\rho\,, \rho\, u\,, \rho\, v\,, p\}$$

we evaluate a mean value $\overline{z_{K,\,f}}$ on the interface f :

(4.4.9) $$\overline{z_{K,\,f}} \;=\; (1-\theta_{K,\,f})\, z_K \;+\; \theta_{K,\,f}\, z_{K_r(f)}$$

and the gradient $\nabla z(K)$ of field $z(\bullet)$ in volume K with a Green formula :

(4.4.10) $$\nabla z(K) \;=\; \frac{1}{\mid K \mid} \int_{\partial K} \overline{z}\, \mathbf{n}\, \mathrm{d}\gamma \;=\; \frac{1}{\mid K \mid} \sum_{f \subset \partial K} \mid f \mid\, \overline{z_{K,\,f}}\, \mathbf{n}_f\,, \;\; K \in \mathcal{E}_\mathcal{T}\,.$$

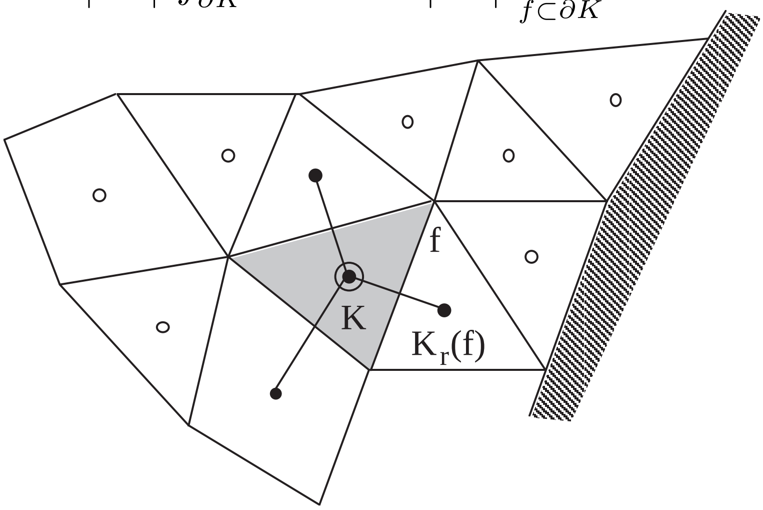

Figure 4.6. Cellular complex mesh with triangles and quadrangles. Three neighbouring cells are necessary to determine the gradient in triangle K and to limit eventually its variation.

• An ideal extrapolation of field $z(\bullet)$ at the interface f would be :

(4.4.11) $$z_{K,\,f} \;=\; z_K \,+\, \nabla z(K) \bullet \big(y_{K,\,f} - x_K \big)$$

but the corresponding scheme is unstable as explicited at Proposition 4.2. When the variation $\nabla z(K) \bullet \big(y_{K,\,f} - x_K \big)$ is very important, it has to be "limited" as first suggested by Van Leer [VL77]. For doing this in a very general way, we introduce the minimum $m_K(z)$ and the maximum $M_K(z)$ of field $z(\bullet)$ in the neighbouring cells :

(4.4.12) $$m_K(z) \;=\; \min \big\{ \, z_L \,, \quad L \in \mathcal{N}(K) \, \big\}$$

(4.4.13) $$M_K(z) \;=\; \max \big\{ \, z_L \,, \quad L \in \mathcal{N}(K) \, \big\} \,.$$

If the value z_K is extremum among the neighbouring ones, *i.e.* if $z_K \,\leq\, m_K(z)$, or $z_K \,\geq\, M_K(z)$, we impose that the interpolated value $z_{K,\,f}$ is equal to the cell value z_K :

(4.4.14) $$z_{K,\,f} \;=\; z_K \qquad \text{if} \quad z_K \,\leq\, m_K(z) \quad \text{or} \quad z_K \,\geq\, M_K(z) \,, \quad f \subset \partial K \,.$$

When on the contrary z_K lies inside the interval $[m_K(z),\, M_K(z)]$, we impose that the variation $z_{K,\,f} - z_K$ is limited by some coefficient k $(0 \,\leq\, k \,\leq\, 1)$ multiplied by the variations $z_K - m_K(z)$ and $M_K(z) - z_K$. We introduce a nonlinear extrapolation of the field $z(\bullet)$ between center x_K and boundary face $y_{K,\,f}\, (f \subset \partial K)$:

(4.4.15) $$z_{K,\,f} \;=\; z_K \,+\, \alpha_K(z) \, \nabla z(K) \bullet \big(y_{K,\,f} - x_K \big) \,, \qquad f \subset \partial K$$

with a **limiting coefficient** $\alpha_K(z)$ satisfying the following conditions :

(4.4.16) $$\left\{ \begin{array}{l} 0 \,\leq\, \alpha_K(z) \,\leq\, 1 \,, \quad z(\bullet) \text{ scalar field defined in (4.4.8)}, \quad K \in \mathcal{E}_{\mathcal{T}} \\ k \, (z_K - m_K(z)) \,\leq\, \alpha_K(z) \, \nabla z(K) \bullet \big(y_{K,\,f} - x_K \big) \,\leq\, k \, (M_K(z) - z_K) \\ \qquad\qquad\qquad\qquad\qquad\qquad \forall \, f \subset \partial K \,, \quad K \in \mathcal{E}_{\mathcal{T}} \,. \end{array} \right.$$

Then $\alpha_K(z)$ is chosen as large as possible and less than or equal to 1 in order to satisfy the constraints (4.4.16) :

(4.4.17) $$\alpha_K(z) \;=\; \min \bigg[\, 1 \,, \, k \, \frac{\min\big(\, M_K(z) - z_K \,, \, z_K - m_K(z) \, \big)}{\max \big\{ \, | \, \nabla z(K) \bullet (y_{K,\,f} - x_K) \, | \,, \, f \subset \partial K \, \big\}} \, \bigg] \,.$$

• In the one dimensional case with a regular mesh, it is an exercice to re-write the extrapolation (4.4.15) under the usual form (4.4.3) in the context of finite differences. In this particular case, some limiter functions $r \longmapsto \varphi_k(r)$ associated with particular parameters k are shown on Figure 4.3. For $k = 1$, we recover the initial limiter proposed by Van Leer in the fourth paper of the family "Towards the ultimate finite difference scheme..." [VL77] ; for this reason, we have named it the "Towards 4" limiter (see Figure 4.3). When $k = \frac{1}{2}$ we obtain the "min-mod" limiter proposed by Harten [Ha83]. The intermediate value $k = \frac{3}{4}$ is a good compromise between the "nearly unstable" choice $k = 1$ and the "too compressive" min-mod choice. We have named it STS (see also (4.3.14)) and it has been chosen for our Euler computations in [DM92].

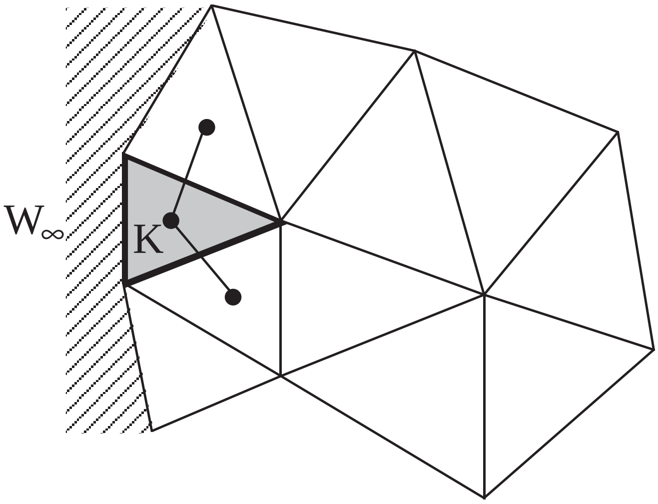

Figure 4.7. Slope limitation at a fluid boundary.

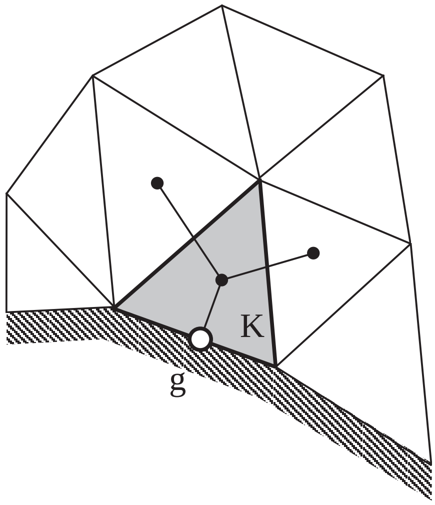

Figure 4.8. Slope limitation at a solid boundary.

• We explain now the way the preceding scheme is adapted near the boundary. We first consider a **fluid boundary**. When K is a finite element with some face $g \subset \partial K$ lying on the boundary, we still define the set $\mathcal{N}(K)$ of neighbouring cells by the relation (4.4.6) as shown on Figure 4.7. The number of neighbouring cells is just less important in this case. Then points $y_{K,\,f}$ are introduced by relation (4.4.7) if face f does not lie on the boundary and by taking the barycenter of face g if it is lying on the boundary. The only difference is the way the values $\overline{z_{K,\,g}}$ are extrapolated for the face g that is on the boundary ; we set

$$\text{(4.4.18)} \qquad \left\{ \begin{array}{l} \overline{z_{K,\,g}} \;=\; z_K \;, \quad g \subset \partial K \;, \\ \qquad\qquad\qquad g \text{ face lying on the boundary of the domain.} \end{array} \right.$$

When values $\overline{z_{K,\,f}}$ are determined for all the faces $f \subset \partial K$, the gradient

$\nabla z(K)$, the minimal $m_K(z)$ and maximal $M_K(z)$ values among the neighbouring cells are still determined with the relations (4.4.10), (4.4.12) and (4.4.13) respectively. The constraints (4.4.16) remain unchanged except that no limitation process is due to the faces lying on the boundary. In a precise way, we set :

$$(4.4.19)\quad \left\{ \begin{array}{l} \alpha_K(z) \;= \\ \qquad \min\Big[1, \dfrac{k \quad \min\big(M_K(z)-z_K\,,\, z_K-m_K(z)\big)}{\max\big\{|\nabla z(K)\bullet(y_{K,\,f}-x_K)|,\; f\subset\partial K,\; K_r(f)\in\mathcal{N}(K)\big\}}\Big]. \end{array}\right.$$

Then the interpolated values $z_{K,\,f}$ for all the faces $f\subset\partial K$ are again predicted with the help of relation (4.4.15).

• For a **rigid wall**, the limitation process is a little modified, as presented at Figure 4.8. We first introduce the limit face g inside the set of neighbouring cells :

$$(4.4.20)\qquad \mathcal{N}(K) \;=\; \begin{array}{l}\big\{L\in\mathcal{E}_{\mathcal{T}},\; \exists f\subset\partial K\cap\partial L\big\}\;\cup \\ \qquad\qquad \big\{g\in\mathcal{F}_{\mathcal{T}},\; g\subset\partial K,\; g \text{ on the boundary}\big\}.\end{array}$$

For the face(s) $g\subset\partial K$ lying on the solid boundary, we determine preliminary values $\overline{z_{K,\,g}}$ by taking in consideration at this level the impenetrability boundary condition $\mathbf{u}\bullet\mathbf{n}_g = 0$. We introduce the two components n_g^x and n_g^y of the normal $\mathbf{n}_g$ at the boundary and we set, in coherence with variables (4.4.8) :

$$(4.4.21)\qquad \left\{\begin{array}{l} \overline{\rho_{K,\,g}} \;=\; \rho_K \\ \overline{\rho_{K,\,g}}\;\overline{u_{K,\,g}} \;=\; \rho_K\,\big(u_K-(\mathbf{u}_K\bullet\mathbf{n}_g)\,n_g^x\big) \\ \overline{\rho_{K,\,g}}\;\overline{v_{K,\,g}} \;=\; \rho_K\,\big(v_K-(\mathbf{u}_K\bullet\mathbf{n}_g)\,n_g^y\big) \\ \overline{p_{K,\,g}} \;=\; p_K\,. \end{array}\right.$$

We consider also these values for the limitation algorithm. We define "external values" z_L for $L=g$ and face g lying on the boundary as equal to the ones defined in relation (4.4.21) :

$$(4.4.22)\qquad z_g \equiv \overline{z_{K,\,g}}\,,\quad z(\bullet) \text{ field defined in (4.4.21)},\;\; g\subset\partial K \text{ on the boundary.}$$

Then the extrapolation algorithm that conducts to relation (4.4.15) for extrapolated values $z_{K,\,f}$ is used as in the internal case.

4.5 Explicit Runge-Kutta integration with respect to time.

• When all values $z_{K,\,f}$ are known for all control volumes $K\in\mathcal{E}_{\mathcal{T}}$, all faces $f\subset K$ and all fields $z(\bullet)$ defined at relation (4.4.8), extrapolated states $W_{K,\,f}$ are naturally defined by going back to the conservative variables. Then we introduce these states as arguments of the flux function $\Phi(\bullet\,,\mathbf{n}_f\,,\bullet)$ and obtain by this way a new system of ordinary differential equations :

$$(4.5.1)\quad |K|\;\frac{\mathrm{d}W_K}{\mathrm{d}t} + \sum_{f\subset\partial K} |f|\;\Phi\big(W_{K,\,f}\,,\mathbf{n}_f\,,W_{K_r(f),\,f}\big) \;=\; 0\,,\qquad K\in\mathcal{E}_{\mathcal{T}}\,.$$

• The numerical integration of such kind of system is done with a Runge-Kutta scheme as presented in [CDV92]. We have used with success in [DM92] the Heun scheme of second order accuracy for discrete integration of (4.5.1) between time steps $n\,\Delta t$ and $(n+1)\,\Delta t$:

$$\frac{|K|}{\Delta t}\Big(\widetilde{W_K}-W_K^n\Big)+\sum_{f\subset\partial K}|f|\;\Phi\Big(W_{K,\,f}^n\,,\,\mathbf{n}_f\,,\,W_{K_r(f),\,f}^n\Big)\;=\;0\,,\;K\in\mathcal{E}_\mathcal{T}\tag{4.5.2}$$

$$\frac{|K|}{\Delta t}\Big(\widetilde{\widetilde{W_K}}-\widetilde{W_K}\Big)+\sum_{f\subset\partial K}|f|\;\Phi\Big(\widetilde{W}_{K,\,f}\,,\,\mathbf{n}_f\,,\,\widetilde{W}_{K_r(f),\,f}\Big)\;=\;0\,,\;K\in\mathcal{E}_\mathcal{T}\tag{4.5.3}$$

$$W_K^{n+1}\;=\;\frac{1}{2}\Big(\widetilde{\widetilde{W_K}}\;+\;W_K^n\Big)\,,\qquad K\in\mathcal{E}_\mathcal{T}\,.\tag{4.5.4}$$

5) References.

[CDV92] D. Chargy, F. Dubois, J.P. Vila. Méthodes numériques pour le calcul d'écoulements compressibles, applications industrielles, *cours de l'Institut pour la Promotion des Sciences de l'Ingénieur*, Paris, septembre 1992.

[Ci78] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems,* North Holland, Amsterdam, 1978.

[DM92] F. Dubois, O. Michaux. Solution of the Euler Equations Around a Double Ellipsoïdal Shape Using Unstructured Meshes and Including Real Gas Effects, *Workshop on Hypersonic Flows for Reentry Problems*, (Désidéri-Glowinski-Périaux Editors), Springer Verlag, vol. II, p. 358-373, 1992.

[DM96] F. Dubois, G. Mehlman. A non-parameterized entropy correction for Roe's approximate Riemann solver, *Numerische Mathematik*, vol 73, p. 169-208, 1996.

[Du91] F. Dubois. Nonlinear Interpolation and Total Variation Diminishing Schemes, *Third International Conference on Hyperbolic Problems*, (Engquist-Gustafsson Editors), Chartwell-Bratt, p. 351-359, 1991. See also hal-00493555.

[Du01] F. Dubois. Partial Riemann problem, Boundary conditions and gas dynamics, in *Artificial Boundary Conditions, with Applications to Computational Fluid Dynamics Problems*, (L. Halpern and L. Tourette Eds), Nova Science Publishers, p. 16-77, 2001. See also hal-00555600.

[FGH91] I. Faille, T. Gallouët, R. Herbin. Les Mathématiciens découvrent les Volumes Finis, *Matapli*, n° 23, p. 37-48, octobre 1991.

[GR96] E. Godlewski, P.A. Raviart. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*, Applied Mathematical Sciences, vol. 118, Springer, New York, 1996.

[Go59] S.K. Godunov. A Difference Method for the Numerical Computation of Discontinuous Solutions of the Equations of Fluid Dynamics, *Math. Sbornik*, vol. 47, p. 271-290, 1959.

[GZIKP79] S.K. Godunov, A. Zabrodine, M. Ivanov, A. Kraiko, G. Prokopov. *Résolution numérique des problèmes multidimensionnels de la dynamique des gaz,* Editions de Moscou, 1979.

[Ha83] A. Harten. High Resolution Schemes for Hyperbolic Conservation Laws, *Journal of Computational Physics,* vol. 49, p. 357-393, 1983.

[HLV83] A. Harten, P.D. Lax, B. Van Leer. On Upstream Differencing and Godunov-type Schemes for Hyperbolic Conservation Laws, *SIAM Review,* vol. 25, n° 1, p. 35-61, January 1983.

[HGMW96] A. Hirschberg, J. Gilbert, R. Msallam, A.P.J. Wijnands. Shock waves in trombones, *J. Acoust. Soc. Am.*, vol. 99, n°3, p. 1754-1758, 1996.

[Kr70] H.O. Kreiss. Initial Boundary Value Problems for Hyperbolic Systems, *Comm. Pure Applied Math.*, vol. 23, p. 277-298, 1970.

[LL54] L. Landau, E. Lifchitz. *Fluid Mechanics*, Pergamon Press, 1954.

[MD99] R. Msallam, F. Dubois. Mathematical model for coupling a quasi-unidimensional perfect flow with an acoustic boundary layer, *Research report CNAM-IAT* n° 326/99, 1999. See also hal-00491417.

[Pa80] S.V. Patankar. *Numerical Heat Transfer and Fluid Flow*, Hemisphere publishing, 1980.

[Po88] M. Pollet. Méthodes de calcul relatives aux interfaces missiles-propulseurs, *Internal report*, Aerospatiale Les Mureaux, 1988.

[Roe81] P. Roe. Approximate Riemann Solvers, Parameter Vectors and Difference Schemes, *Journal of Computational Physics,* vol. 43, p. 357-372, 1981.

[Roe85] P. Roe. Some contributions to the Modelling of Discontinuous Flows, in *Lectures in Applied Mathematics*, vol. 22, (Engquist, Osher, Sommerville Eds), AMS, p. 163-193, 1985.

[RM67] R.D. Richtmyer, K.W. Morton. *Difference Methods for Initial-Value Problems*, Interscience Publishing, J. Wiley & Sons, New York, 1967.

[VL77] B. Van Leer. Towards the Ultimate Conservative Difference Scheme IV. A New Approach to Numerical Convection, *Journal of Computational Physics,* vol. 23, p. 276-299, 1977.

[VL79] B. Van Leer. Towards the Ultimate Conservative Difference Scheme V. A Second Order Sequel to Godunov's Method, *Journal of Computational Physics,* vol 32, n^0 1, p. 101-136, 1979.