10

Gas Dynamics: The Riemann Problem and Discontinuous Solutions: Application to the Shock Tube Problem

Project Summary

Level of difficulty:	3
Keywords:	Nonlinear hyperbolic systems, Euler equations for gas dynamics, centered schemes: Lax–Wendroff, MacCormack; upwind schemes: Godunov, Roe
Application fields:	Shock tube, supersonic flows

The interest in studying the shock tube problem is threefold. From a fundamental point of view, it offers an interesting framework to introduce some basic notions about nonlinear hyperbolic systems of partial differential equations (PDEs). From a numerical point of view, this problem constitutes, since the exact solution is known, an inevitable and difficult test case for any numerical method dealing with discontinuous solutions. Finally, there is a practical interest, since this model is used to describe real shock tube experimental devices.[1]

10.1 Physical Description of the Shock Tube Problem

The fundamental idea of the shock tube is the following: consider a long one-dimensional (1D) tube, closed at its ends and divided into two equal regions by a thin diaphragm (see Fig. 10.1). Each region is filled with the same gas, but with different thermodynamic parameters (pressure, density, and temperature). The region with the highest pressure is called the *driven*

[1] The first shock tube facility was built in 1899 by Paul Vieille to study the deflagration of explosive charges. Nowadays, shock tubes are currently used as low-cost high-speed wind tunnels, in which a wide variety of aerodynamic or aeroballistic topics are studied: supersonic aircraft flight, gun performance, asteroid impacts, shuttle atmospheric entry, etc.

section of the tube, while the low-pressure part is the *working section*. The gas being initially at rest, the sudden breakdown of the diaphragm generates a high-speed flow, which propagates in the working section (this is the place where the model of a free-flying object, such as a supersonic aircraft, will be placed).

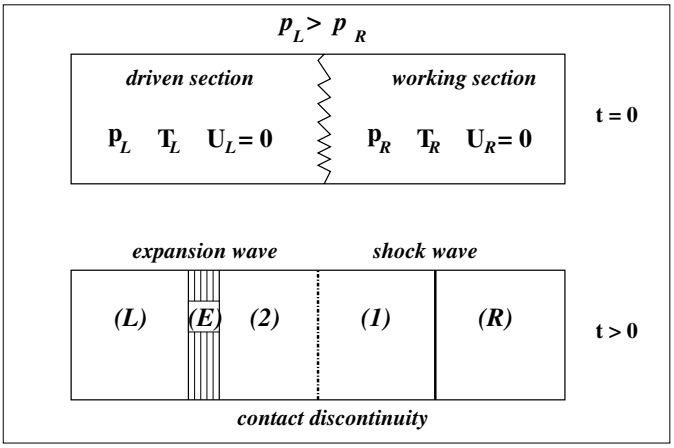

Fig. 10.1. Sketch of the initial configuration of the shock tube ($t = 0$) and waves propagating in the tube after the diaphragm breakdown ($t > 0$).

Let us get into a more detailed analysis of the problem. Consider (Fig. 10.1) that the left part of the tube is the driven section, defined by the pressure p_L, the density ρ_L, the temperature T_L, and the initial velocity $U_L = 0$. Similarly, the parameters of the (right part) working section are $p_R < p_L$, ρ_R, T_R, and $U_R = 0$.

At time $t = 0$ the diaphragm breaks, generating a process that naturally tends to equalize the pressure in the tube. The gas at high pressure expands through an *expansion (or rarefaction) wave* and flows into the working section, pushing the gas of this part. The rarefaction is a continuous process and takes place inside a well-defined region (the expansion fan) that propagates to the left (region (E) in Fig. 10.1); the width of the expansion fan grows in time.

The compression of the low-pressure gas generates a *shock wave* propagating to the right. The expanded gas is separated from the compressed gas by a *contact discontinuity*, which can be regarded as a fictitious membrane traveling to the right at constant speed. At this point of our simplified description, we just note that some of the physical functions defining the flow in the tube ($p(x), \rho(x), T(x)$, and $U(x)$) are discontinuous across the shock wave and the contact discontinuity. These discontinuities, which cause the difficulty of the problem, will be described in greater detail in the following sections.

10.2 Euler Equations of Gas Dynamics

To simplify the mathematical description of the shock tube problem we consider an infinitely long tube (to avoid reflections at the tube ends) and neglect viscous effects in the flow. We also suppose that the diaphragm is completely removed from the flow at $t = 0$. Under these simplifying hypotheses, the compressible flow in the shock tube is described by the one-dimensional (1D) Euler system of PDEs (see, for instance, Hirsch, 1988; LeVeque, 1992)

$$\frac{\partial}{\partial t}\underbrace{\begin{pmatrix}\rho\\ \rho U\\ E\end{pmatrix}}_{W(x,t)}+\frac{\partial}{\partial x}\underbrace{\begin{pmatrix}\rho U\\ \rho U^2+p\\ (E+p)U\end{pmatrix}}_{F(W)}=0, \tag{10.1}$$

where ρ is the density of the gas and E the total energy:

$$E=\frac{p}{\gamma-1}+\frac{\rho}{2}U^2. \tag{10.2}$$

To close this system of equations, we need to write the constitutive law of the gas (or equation of state). Considering the perfect gas model, the equation of state is

$$p=\rho\mathcal{R}T. \tag{10.3}$$

The constants $\mathcal{R}$ and γ characterize the thermodynamic properties of the gas ($\mathcal{R}$ is the universal gas constant divided by the molecular mass and γ is the ratio of specific heat coefficients). It is also useful to define the local speed of sound a, the Mach number M, and the total enthalpy H:

$$a=\sqrt{\gamma\mathcal{R}T}=\sqrt{\gamma\frac{p}{\rho}},\quad M=\frac{U}{a},\quad H=\frac{E+p}{\rho}=\frac{a^2}{\gamma-1}+\frac{1}{2}U^2. \tag{10.4}$$

Considering the column vector of unknowns $W=(\rho,\rho U,E)^t$, the Euler system of equations (10.1) can be written in the following *conservative* form:

$$\frac{\partial W}{\partial t}+\frac{\partial}{\partial x}F(W)=0, \tag{10.5}$$

with the initial condition (we denote by x_0 the abscissa of the diaphragm):

$$W(x,0)=\begin{cases}(\rho_L,\rho_L U_L,E_L),\ x\le x_0,\\ (\rho_R,\rho_R U_R,E_R),\ x>x_0.\end{cases} \tag{10.6}$$

The vector W contains the conserved variables and $F(W)$ the conserved fluxes. Note that with this choice of the vector of unknowns W, the pressure is not an unknown, since it can be derived from (10.2) using the components of W.

The mathematical analysis of the Euler system of PDEs usually considers its *quasilinear* form:[2]

$$\frac{\partial W}{\partial t} + A\frac{\partial W}{\partial x} = 0, \tag{10.7}$$

with the Jacobian matrix

$$A = \frac{\partial F}{\partial W} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma-3)U^2 & (3-\gamma)U & \gamma-1 \\ \frac{1}{2}(\gamma-1)U^3 - UH & H-(\gamma-1)U^2 & \gamma U \end{pmatrix}. \tag{10.8}$$

It is interesting to note that the matrix A satisfies the following remarkable relationship:

$$AW = F(W). \tag{10.9}$$

Furthermore, we can easily calculate its eigenvalues

$$\lambda^0 = U, \quad \lambda^+ = U + a, \quad \lambda^- = U - a, \tag{10.10}$$

and the corresponding eigenvectors

$$v^0 = \begin{pmatrix} 1 \\ U \\ \frac{1}{2}U^2 \end{pmatrix}, \quad v^+ = \begin{pmatrix} 1 \\ U+a \\ H+aU \end{pmatrix}, \quad v^- = \begin{pmatrix} 1 \\ U-a \\ H-aU \end{pmatrix}. \tag{10.11}$$

We conclude that the Jacobian matrix A is diagonalizable, i.e., it can be decomposed as $A = P\Lambda P^{-1}$, where

$$\Lambda = \begin{pmatrix} U-a & 0 & 0 \\ 0 & U & 0 \\ 0 & 0 & U+a \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & 1 \\ U-a & U & U+a \\ H-aU & \frac{1}{2}U^2 & H+aU \end{pmatrix}. \tag{10.12}$$

We can easily verify that

$$P^{-1} = \begin{pmatrix} \frac{1}{2}\left(\alpha_1 + \frac{U}{a}\right) & -\frac{1}{2}\left(\alpha_2 U + \frac{1}{a}\right) & \frac{\alpha_2}{2} \\ 1-\alpha_1 & \alpha_2 U & -\alpha_2 \\ \frac{1}{2}\left(\alpha_1 - \frac{U}{a}\right) & -\frac{1}{2}\left(\alpha_2 U - \frac{1}{a}\right) & \frac{\alpha_2}{2} \end{pmatrix}, \tag{10.13}$$

where $\alpha_1 = (\gamma-1)U^2/(2a^2)$ and $\alpha_2 = (\gamma-1)/a^2$.

Definition 10.1. *The system (10.7) with the matrix A diagonalizable with real eigenvalues is called hyperbolic.*

The hyperbolic character of the system (10.7) has important consequences on the propagation of the information in the flow field. Certain quantities,

[2] The reader who has already explored Chap. 1 of this book may notice that this form is similar to that of the convection equation. The underlying idea is here to generalize the analysis of characteristics in the case of a system of PDEs.

called *invariants*,[3] are transported along particular curves in the plane (x,t), called *characteristics*. From a numerical point of view, this suggests a simple way to calculate the solution in any point $P(x,t)$ by gathering all the information transported through the characteristics starting from P and going back to regions where the solution is already known (imposed by the initial condition, for example).

The general form of the equation defining a characteristic is $dx/dt = \lambda$, where λ is an eigenvalue of the Jacobian matrix A. Since the corresponding invariant r is constant along the characteristic, it satisfies

$$\frac{dr}{dt} = \frac{\partial r}{\partial t} + \frac{\partial r}{\partial x}\frac{dx}{dt} = 0, \quad \text{or} \quad \frac{\partial r}{\partial t} + \lambda \frac{\partial r}{\partial x} = 0. \tag{10.14}$$

In the simplest case of the convection equation $\partial u/\partial t + c\,\partial u/\partial x = 0$, with constant transport velocity c, there exists a single characteristic curve, which is the line $x = c\,t$, and the corresponding invariant is the solution itself, $r = u$ (see also Chap. 1). From (10.10) we infer that the system (10.7) has three distinct characteristics:

$$C^0 : \frac{dx}{dt} = U, \quad C^+ : \frac{dx}{dt} = U + a, \quad C^- : \frac{dx}{dt} = U - a. \tag{10.15}$$

The invariants can be generally expressed as differential relations (see, for instance, Hirsch (1988); Godlewski and Raviart (1996) for details)

$$dr^0 = dp - a^2\, d\rho = 0, \quad dr^+ = dp + \rho a\, dU = 0, \quad dr^- = dp - \rho a\, dU = 0,$$

which have to be integrated along the corresponding characteristic curves. In the case of an isentropic flow[4] we obtain

$$r^0 = p/\rho^\gamma, \quad r^+ = U + \frac{2a}{\gamma - 1}, \quad r^- = U - \frac{2a}{\gamma - 1}. \tag{10.16}$$

The above relations will be used in the following to derive the exact solution of the shock tube problem.

Definition 10.2. *The nonlinear hyperbolic system of PDEs (10.5) and piecewise constant initial condition (10.6) define the Riemann problem.*

---

[3] For a rigorous analysis of hyperbolic systems of PDEs and related definitions (in particular the definition of Riemann invariants), the reader can refer to Hirsch (1988); LeVeque (1992); Godlewski and Raviart (1996).

[4] The entropy variation of a perfect gas during its evolution starting from a reference state A is $s - s_A = C_v \ln\left(\frac{p/p_A}{(\rho/\rho_A)^\gamma}\right)$, where $C_v = \mathcal{R}/(\gamma - 1)$ is the heat coefficient under constant volume. For an isentropic flow, since the entropy remains unchanged ($s = s_A$), we deduce that the ratio p/ρ^γ is constant.

10.2.1 Dimensionless Equations

When building numerical applications we usually prefer to remove physical units from equations and work with dimensionless variables. This simplifies the problem formulation and may reduce computational round-off errors. Physical variables in previous equations are nondimensionalized (or scaled) using a reference state defined by the parameters of the working section:

$$\begin{aligned} \rho^* = \rho/\rho_R,\ U^* = U/a_R,\ a^* = a/a_R,\ T^* = T/(\gamma T_R), \\ p^* = p/(\rho_R a_R^2) = p/(\gamma p_R),\ E^* = E/(\rho_R a_R^2),\ H^* = H/a_R^2. \end{aligned} \quad (10.17)$$

We also nondimensionalize space and time variables as $x^* = x/L$, $t^* = t/(L/a_R)$, where L is the length of the tube.

The Euler equations for the dimensionless variables (denoted by the star superscript) keep the same differential form as previously:

$$\frac{\partial}{\partial t^*} \underbrace{\begin{pmatrix} \rho^* \\ \rho^* U^* \\ E^* \end{pmatrix}}_{W^*(x^*,t^*)} + \frac{\partial}{\partial x^*} \underbrace{\begin{pmatrix} \rho^* U^* \\ \rho^* U^{*2} + p^* \\ (E^* + p^*) U^* \end{pmatrix}}_{F^*(W^*)} = 0. \quad (10.18)$$

Dimensionless total energy E^* and total enthalpy H^* become

$$E^* = \frac{p^*}{\gamma - 1} + \frac{\rho^*}{2} U^{*2}, \quad H^* = \frac{(a^*)^2}{\gamma - 1} + \frac{1}{2} U^{*2}. \quad (10.19)$$

Differences with respect to previous physical equations appear in the equation of state

$$p^* = \rho^* T^*, \quad (10.20)$$

and in the definition of the speed of sound

$$a^* = \sqrt{\gamma \frac{p^*}{\rho^*}} = \sqrt{\gamma T^*}. \quad (10.21)$$

In the interests of simplicity, we drop the star superscript in subsequent equations; only dimensionless variables will be considered in the following sections.

10.2.2 Exact Solution

The exact solution of the shock tube problem follows the physical and mathematical descriptions given in previous sections. The tube is separated (see Fig. 10.1) into four uniform regions, i.e., with constant parameters (pressure, density, temperature, and velocity): the left (L) and right (R) regions (which keep the parameters imposed by the initial condition) and two intermediate regions, denoted by subscripts 1 and 2.

It is important to identify these regions in the (x,t) plane (see Fig. 10.2). All the waves are centered at the initial position of the diaphragm ($t=0, x=x_0$). Since the shock and the contact discontinuity propagate in uniform zones, they have constant velocities and hence are displayed as lines in the (x,t) diagram. The expansion wave extends through the new zone (E), the expansion fan, in which the flow parameters vary continuously (see below). We remember that the shock wave and the contact discontinuity propagate to the right, while the expansion fan moves to the left.

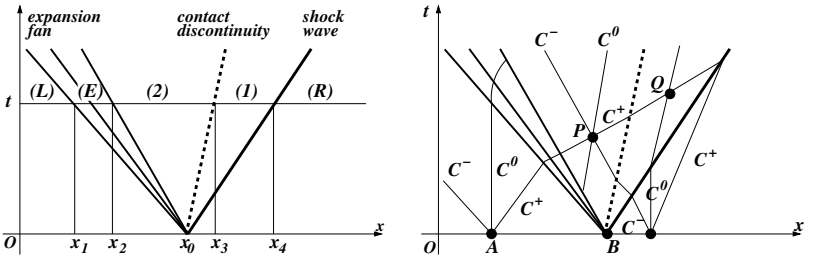

Fig. 10.2. Diagram in the (x,t) plane of the exact solution of the shock tube problem (left). Characteristics used to calculate the exact solution (right).

We start the calculation of the exact solution by writing the dimensionless parameters of the (L) and (R) regions (which are in fact the input parameters for a computer program):

$$\text{Region (R):} \quad \rho_R = 1,\ p_R = 1/\gamma,\ T_R = 1/\gamma,\ a_R = 1,\ U_R = 0, \tag{10.22}$$

$$\text{Region (L):} \quad \rho_L,\ p_L,\ T_L,\ a_L,\ U_L = 0 \text{ (given quantities).} \tag{10.23}$$

We then use the jump relations across the discontinuities and take into account the propagation of the information along the characteristics, as follows:

1. The shock wave implies the discontinuity of all the parameters of the flow. The jump between regions (1) and (R) is described by the Rankine–Hugoniot relations (see, for example, Hirsch (1988)):

$$\frac{p_1}{p_R} = \frac{2\gamma}{\gamma+1} M_s^2 - \frac{\gamma-1}{\gamma+1}, \tag{10.24}$$

$$\frac{\rho_R}{\rho_1} = \frac{2}{\gamma+1}\frac{1}{M_s^2} + \frac{\gamma-1}{\gamma+1}, \tag{10.25}$$

$$U_1 = \frac{2}{\gamma+1}\left(M_s - \frac{1}{M_s}\right), \tag{10.26}$$

where M_s is the Mach number of the shock, defined in physical units as $M_s = U_s^{phys}/a_R^{phys}$. We note that using our scaling, $M_s = U_s$, where U_s is the dimensionless propagation speed of the shock. We remember that U_s is constant.

2. The contact discontinuity is in fact a discontinuity of the density function, the pressure and the velocity being continuous. Hence

$$U_2 = U_1, \quad p_2 = p_1. \tag{10.27}$$

3. We now link the parameters of region (2) to those of region (L). For this purpose, we consider a point P inside the region (2) and draw the characteristics passing through this point (see Fig. 10.2). We notice that only C^0 and C^+ characteristics will cross the expansion fan to search the information in region (L). Using the expressions (10.15) for the invariants r^0 and r^+ and taking into account that $U_L = 0$, we obtain

$$\frac{\rho_2}{\rho_L} = \left(\frac{p_2}{p_L}\right)^{1/\gamma}, \quad U_2 = \frac{2}{\gamma - 1}(a_L - a_2). \tag{10.28}$$

4. Finally, we combine the previous relations to obtain an implicit equation for the unknown M_s. The detailed calculation follows:

$$M_s - \frac{1}{M_s} \overset{(10.26)}{=} \frac{\gamma+1}{2}U_1 \overset{(10.27)}{=} \frac{\gamma+1}{2}U_2 \overset{(10.28)}{=} a_L\frac{\gamma+1}{\gamma-1}\left(1 - \frac{a_2}{a_L}\right).$$

Since

$$\frac{a_2}{a_L} \overset{(10.21)}{=} \left(\frac{p_2}{p_L}\frac{\rho_L}{\rho_2}\right)^{1/2} \overset{(10.28)}{=} \left(\frac{p_2}{p_L}\right)^{\frac{\gamma-1}{2\gamma}} \overset{(10.27)}{=} \left(\frac{p_1}{p_L}\right)^{\frac{\gamma-1}{2\gamma}},$$

we replace p_1/p_L from (10.24) and finally get the following *compatibility equation*:

$$M_s - \frac{1}{M_s} = a_L\frac{\gamma+1}{\gamma-1}\left\{1 - \left[\frac{p_R}{p_L}\left(\frac{2\gamma}{\gamma+1}M_s^2 - \frac{\gamma-1}{\gamma+1}\right)\right]^{\frac{\gamma-1}{2\gamma}}\right\}. \tag{10.29}$$

Once this implicit nonlinear equation is solved (using an iterative Newton method, for example), the value of M_s will be used in previous relations to determine all the parameters of uniform regions (1) and (2).

To complete the exact solution, we need to determine the extent of each region (i.e., calculate the values of the abscissas x_1, x_2, x_3, x_4 in Fig. 10.2) for a given time value t. We proceed as follows:

- The expansion fan (E) is left-bounded by the C^- characteristic starting from the point B, considered to belong to region (L), i.e., the line of slope $dx/dt = -a_L$. The right bound of the expansion fan is the C^- characteristic starting from the same point B, but considered this time to belong to region (2), i.e., the line of slope $dx/dt = U_2 - a_2$. The values of x_1 and x_2 are consequently

$$x_1 = x_0 - a_L t, \quad x_2 = x_0 + (U_2 - a_2)t. \tag{10.30}$$

Consider now a point (x,t) inside the region (E), i.e., $x_1 \leq x \leq x_2$. Since this point belongs to a C^- characteristic starting from B, necessarily $(x - x_0)/t = U - a$. Using the C^+ characteristic coming from region (L), we also get that $a + (\gamma - 1)U/2 = a_L$. Combining these two relations and remembering that the flow is isentropic, we can conclude that the exact solution inside the expansion fan is

$$U = \frac{2}{\gamma+1}\left(a_L + \frac{x - x_0}{t}\right),\ a = a_L - (\gamma-1)\frac{U}{2},\ p = p_L\left(\frac{a}{a_L}\right)^{\frac{2\gamma}{\gamma-1}}. \tag{10.31}$$

- The contact discontinuity is transported at constant velocity $U_2 = U_1$, so

$$x_3 = x_0 + U_2 t. \tag{10.32}$$

- Since the shock wave also propagates at constant dimensionless velocity $U_s = M_s$, we finally obtain

$$x_4 = x_0 + M_s t. \tag{10.33}$$

Remark 10.1. The exact solution $W(x,t)$ of the shock tube problem depends only on the ratio x/t, as one would have expected from the characteristics analysis of the Euler system of PDEs.

Exercise 10.1. Write a MATLAB function to compute the exact solution of the shock tube problem. The definition header of the function will be as follows:

```
function uex=HYP_shock_exact(x,x0,t)
% Input arguments:
% x   vector of abscissas of dimension M
%      x0  the initial position of the diaphragm
%      t   time at which the solution is calculated
% Output arguments:
% uex  vector of dimensions (3,n) containing the solution as
% uex(1,1:M) the density
% uex(2,1:M) the velocity
% uex(3,1:M) the pressure
```

Plot the dimensionless exact solution ($\rho(x)$, $U(x)$ and $p(x)$) at time $t = 0.2$. Consider $x \in [0,1]$, $x_0 = 0.5$, and a regular (equidistant) grid with $M = 81$ computational points. The physical parameters correspond to those used by Sod (see also Hirsch, 1988): $\gamma = 1.4$, $\rho_L = 8$, $p_L = 10/\gamma$.
Hint: define all the physical parameters as global variables; use the MATLAB built-in function `fzero` to solve the compatibility equation (10.29).

The expected result is displayed in Fig. 10.3. This solution was obtained using the MATLAB program presented in Sect. 10.4 at page 232.

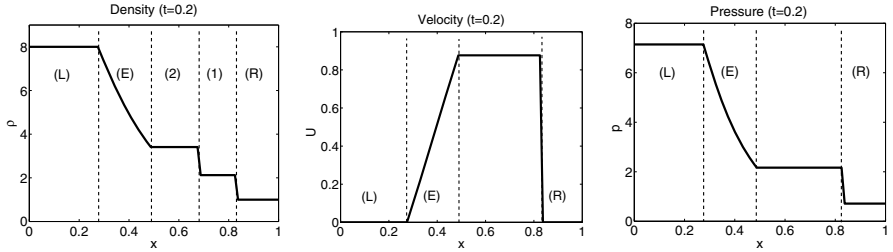

Fig. 10.3. Exact solution of the shock tube problem (Sod's data) at time $t = 0.2$.

10.3 Numerical Solution

The first idea one would have in mind when attempting to numerically solve the Euler system of PDEs (10.18) is to use *elementary* discretization methods discussed in Chap. 1 for scalar PDEs, for example, an Euler or a Runge–Kutta method for the time integration and centered finite differences for the space discretization. We shall see, however, that such methods are not appropriate to compute discontinuous solutions, since they generate nonphysical oscillations. This drawback of the space-centered schemes for computing the shock tube problem will be illustrated using the more sophisticated Lax–Wendroff and MacCormack schemes. We shall also give a quick description of upwind schemes that take into account the hyperbolic character of the system and allow a better numerical solution. Results using Roe's upwind scheme will be discussed at the end.

10.3.1 Lax–Wendroff and MacCormack Centered Schemes

The space-centered schemes were historically the first to be derived to solve hyperbolic systems. The two most popular schemes, the Lax and Wendroff scheme and the MacCormack scheme, are still used in some industrial numerical codes. We shall apply these schemes to solve the Euler system (10.18) written in the conservative form

$$\frac{\partial W}{\partial t} + \frac{\partial}{\partial x} F(W) = 0. \tag{10.34}$$

We use a regular (or equidistant) discretization of the domain of definition of the problem $(x, t) \in [0, 1] \times [0, T]$:

- in space

$$x_j = (j-1)\delta x, \quad \delta x = \frac{1}{M-1}, \quad j = 1, 2, \ldots, M, \tag{10.35}$$

- and in time

$$t^n = (n-1)\delta t, \quad \delta t = \frac{T}{N-1}, \quad n = 1, 2, \ldots, N. \tag{10.36}$$

For both schemes, the numerical solution W_j^{n+1} (at time t_{n+1} and space position x_j) is computed in two steps (a predictor and a corrector step) following the formulas displayed in Fig. 10.4.

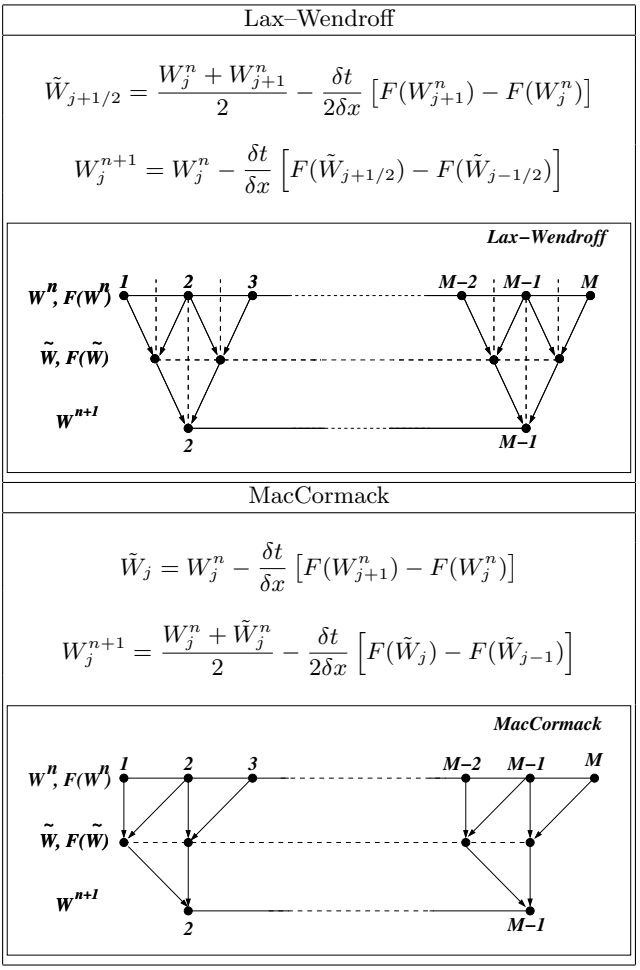

Fig. 10.4. Formulas of Lax–Wendroff and MacCormack space-centered schemes. Schematic representation of their predictor and corrector steps.

We discuss in the following some remarkable features of these schemes.

1. (*Boundary values.*) From the schematic representation of the predictor and corrector steps in Fig. 10.4, we notice that only the components $j = 2, \ldots, (M-1)$ of the solution are calculated. The remaining components for $j = 1$ and $j = M$ need to be prescribed by appropriate boundary conditions.

Since the tube is assumed infinite, we impose $W_1^n = W_L$ and $W_M^n = W_R$ at any time level t_n. Practically, this is equivalent to leaving unchanged the first and last components of the solution vector. Meanwhile, it is obvious that the computation must stop before one of the waves (expansion or shock) hits the boundary.

2. (*Propagation of information.*) The predictor step of the Lax–Wendroff scheme computes an intermediate solution at interfaces $\left(j+\frac{1}{2}\right)$ and $\left(j-\frac{1}{2}\right)$ using forward finite differences. These intermediate values are then used in the centered finite difference scheme of the corrector step.

The MacCormack scheme combines backward differences for the predictor step with forward differences for the corrector step. We can show in fact that the idea behind this scheme is the following Taylor expansion:

$$W_j^{n+1} = W_j^n + \left(\frac{\overline{\partial W}}{\partial t}\right)_j \delta t, \tag{10.37}$$

where

$$\left(\frac{\overline{\partial W}}{\partial t}\right)_j = \frac{1}{2}\left[\left(\frac{\partial W}{\partial t}\right)_j^n + \left(\frac{\partial \tilde{W}}{\partial t}\right)_j\right] = \frac{1}{2}\left[\frac{\tilde{W}_j - W_j^n}{\delta t} - \frac{F(\tilde{W}_j) - F(\tilde{W}_{j-1})}{\delta x}\right]$$

is an approximation of the first derivative in time.

In conclusion, the information is searched on both sides of the computed point j. The information propagation along characteristics is not taken into account, since no distinction is made between upstream and downstream influences. We shall see that this lack of physics in the numerical schemes will generate unwanted (nonphysical) oscillations of the solution.

3. (*Accuracy.*) Both schemes use a three-point stencil $(j-1, j, j+1)$ to reach second-order accuracy in time and space.

4. (*Stability.*) Both schemes are explicit and consequently subject to stability conditions. Similar to the (scalar) convection equation (see Chap. 1), we can write the stability (or CFL[5]) condition in the general form

$$\max_i\{|\lambda_i|\}\cdot\frac{\delta t}{\delta x} \leq 1,$$

where λ_i, $i = 1, 2, 3$, are the eigenvalues of the Jacobian matrix $\partial F/\partial W$, regarded here as propagation speeds of the corresponding characteristic waves ($dx/dt = \lambda$). Using (10.15), we obtain the stability condition

$$(|U| + a)\,\frac{\delta t}{\delta x} \leq 1. \tag{10.38}$$

For numerical applications, this condition is used to compute the time step

$$\delta t = \text{cfl}\cdot\frac{\delta x}{|U| + a}, \quad \text{with} \quad \text{cfl} < 1. \tag{10.39}$$

[5] Courant–Friedrichs–Lewy

Exercise 10.2. For the same physical and numerical parameters as in the previous exercise, compute the numerical solution of the shock tube problem at $t = 0.2$ using Lax–Wendroff and MacCormack centered schemes. Compare to the exact solution and comment on the results. Hints:
• set an array `w(1:3,1:M)` to store the discrete values of the vector $W = (\rho, \rho U, E)^t$ of conservative variables;
• using (10.39) with cfl $= 0.95$, compute the time step in a separate function `function dt = HYP_calc_dt(w,dx,cfl)`;
• write a function to compute $F(W)$;
• use vectorial programming to translate the formulas in Fig. 10.4 into MATLAB program lines (avoid loops!); for example, the predictor step of the Lax–Wendroff scheme will be coded in a single line:

```
wtilde=0.5*(w(:,1:M-1)+w(:,2:M))-0.5*dt/dx*(F(:,2:M)-F(:,1:M-1));
```

• for each scheme, superimpose numerical and exact solutions for (ρ, U, p) as in Fig. 10.5.
A solution of this exercise is proposed in Sect. 10.4 at page 232.

The numerical results of both schemes, displayed in Fig. 10.5, show good accuracy in smooth regions, whereas unwanted oscillations appear at the interfaces between different regions of the solution. The contact discontinuity is also poorly captured. The MacCormack scheme seems to capture the shock discontinuity better, but introduces higher-amplitude oscillations at the end of the expansion wave where the flow is strongly accelerated.

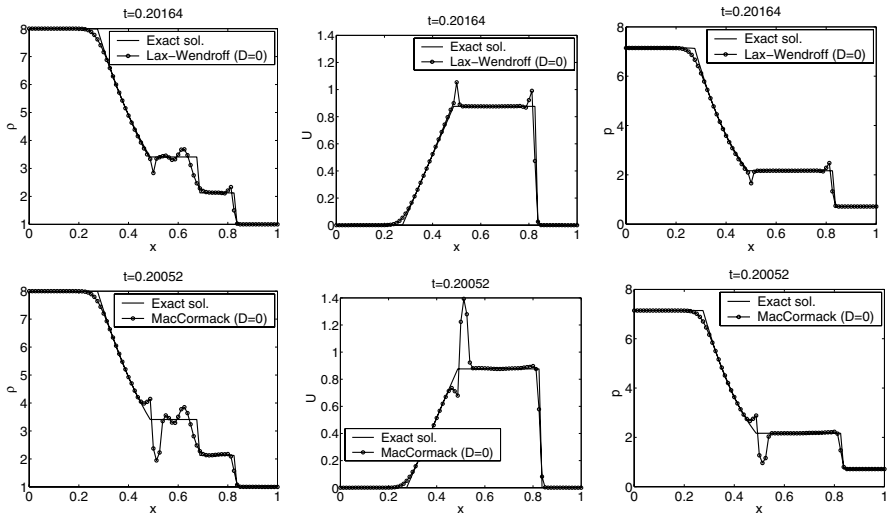

Fig. 10.5. Numerical results for the shock tube problem (Sod's parameters) using centered schemes. Lax–Wendroff scheme (up) and MacCormack scheme (down).

Artificial Dissipation

The oscillations generated by the centered schemes around discontinuities can be damped by adding a supplementary term to the initial equation (10.34):

$$\frac{\partial W}{\partial t} + \frac{\partial}{\partial x} F(W) - \delta x^2 \frac{\partial}{\partial x} \left(D(x) \frac{\partial W}{\partial x} \right) = 0. \tag{10.40}$$

The mathematical form of this term is inspired by the heat equation (discussed in Chap. 1). The idea is to simulate the effects of a physical dissipation (or diffusion) process which is well known to have a smoothing effect[6] on the solution. Since the dissipation term is proportional to the gradient $\partial W/\partial x$ of the solution, the smoothing will be important in regions with sharp gradients (as the shock discontinuity) where numerical oscillations are expected to disappear.

The coefficient $D(x)$, also called *artificial viscosity* by analogy with Navier–Stokes equations (see Chap. 12), has to be positive to ensure a stabilizing effect[7] on the numerical solution. Moreover, its value has to be chosen such that the influence of the artificial term is negligible (i.e., of an order greater than or equal to the truncation error) in the smooth regions of the solution.

Several methods have been proposed to prescribe the artificial viscosity $D(x)$ and to modify classical centered schemes accordingly (see, for instance, Hirsch (1988), Fletcher (1991)). We illustrate the simplest technique, which considers a constant coefficient $D(x) = D$ and writes (10.40) in the conservative form (10.34) with a modified flux $F^*(W)$:

$$\frac{\partial W}{\partial t} + \frac{\partial}{\partial x} F^*(W) = 0, \quad \text{where} \quad F^*(W) = F(W) - D\delta x^2 \frac{\partial W}{\partial x}. \tag{10.41}$$

In order to use the same three-points stencil to define the schemes, the new vector $F^*(W)$ will be discretized

- using backward differences in the predictor step

$$F^*(W_j) = F(W_j) - (D\delta x)(W_j - W_{j-1}), \tag{10.42}$$

- and forward differences in the corrector step

$$F^*(\tilde{W}_j) = F(\tilde{W}_j) - (D\delta x)(\tilde{W}_{j+1} - \tilde{W}_j). \tag{10.43}$$

Exercise 10.3. Modify the previous program by adding an artificial dissipation term to both Lax–Wendroff and MacCormack schemes. Use (10.42)–(10.43) to modify the flux $F(W)$. Discuss the effect of the value of the artificial viscosity D (take $0 \leq D \leq 10$). What is the influence of D on the value of the time step?

[6] This smoothing effect is nicely illustrated for the heat equation in Chap. 1, Exercise 1.10.

[7] The heat equation with negative diffusivity has physically unstable solutions!

The results obtained with an artificial dissipation term are displayed in Fig. 10.6. Numerical oscillations are reduced near the shock and expansion waves, but large dissipation is also introduced in other regions of the solution. In particular, the contact discontinuity (see the graph for $\rho(x)$) is considerably smeared. Increasing the value of D allows one to completely remove the oscillations, but the overall accuracy is not satisfactory. More sophisticated methods have been proposed (see the references at the end of the chapter) to render the dissipation more selective with respect to the nature of discontinuities, but the general tradeoff between damping the oscillations and overall accuracy suggests that the artificial dissipation does not bring a real solution to the problem. A different approach, including more physics in the numerical approximation, is presented in the next section.

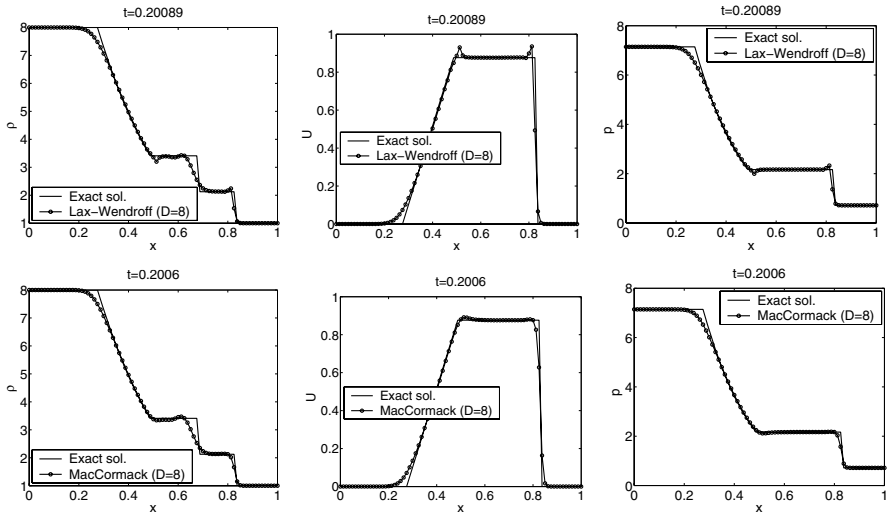

Fig. 10.6. Numerical results for the shock tube problem (Sod's parameters) using centered schemes with artificial dissipation. Lax–Wendroff scheme (up) and MacCormack scheme (down).

10.3.2 Upwind Schemes (Roe's Approximate Solver)

The origin of the numerical oscillations generated by the centered schemes discussed in the previous section comes from complete ignorance of the hyperbolic character of the Euler system of PDEs, in particular the propagation of the information along characteristics. These important (physical) features will be considered in deriving upwind schemes.

Physical information can be introduced at different levels of the numerical approximation. We distinguish between:

1. *flux splitting upwind schemes*, which use different directional discretization of the flux $F(W)$, depending on the sign of the eigenvalues λ of the Jacobian matrix (10.8); since λ corresponds to the propagation speed of the associated characteristic, these schemes include only the information on the direction of propagation of waves (up- or downstream);
2. *Godunov-type schemes*, which introduce a higher level of physical approximation by considering a discretization based on the exact solution of the Riemann problem at each interface between computational points; when the local Riemann problem is solved approximatively, we talk about Riemann solvers.

The following sections present the basic principle of Godunov schemes and the Riemann approximate solver of Roe.

Godunov-Type Schemes

The basic principle of a Godunov-type scheme is the following: the solution W^n is considered to be piecewise constant over each grid cell defined as the interval $]x_{j-1/2}, x_{j+1/2}[$; this allows us to define locally a Riemann problem at each interface between the cells; each local Riemann problem is solved **exactly** to calculate the solution W^{n+1} at the next time level.

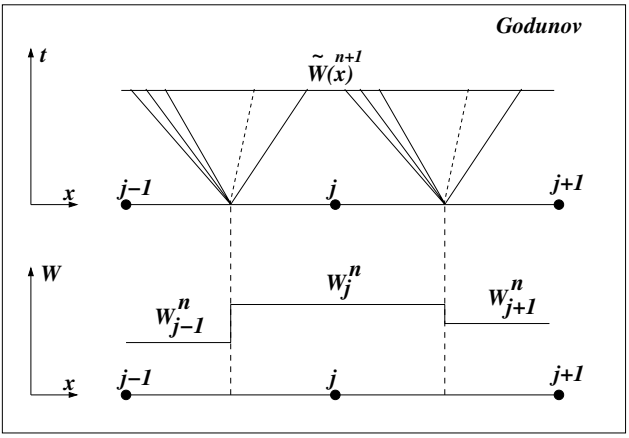

Fig. 10.7. Principle of a Godunov-type scheme.

More precisely, the numerical solution is advanced from time level $t_n = n\delta t$ to $t_{n+1} = t_n + \delta t$ in three steps (see Fig. 10.7):

Step 1. Using the known values W_j^n, define the piecewise constant function

$$W^n(x) = W_j^n, \quad x \in](j - 1/2)\delta x, (j + 1/2)\delta x[\,. \tag{10.44}$$

Step 2. Calculate the solution function $\tilde{W}^{n+1}(x)$, $x \in](j-1/2)\delta x, (j+1/2)\delta x[$ by gathering the exact solutions of the two Riemann problems defined at interfaces $\left(j-\frac{1}{2}\right)$ and $\left(j+\frac{1}{2}\right)$. This step requires that the waves issued from the two neighboring Riemann problems not intersect. This implies that the time step should be limited such that

$$\max_j \left(|U|+a\right)_{j+1/2}^n \frac{\delta t}{\delta x} \leq \frac{1}{2}. \tag{10.45}$$

Step 3. Obtain the solution $W^{n+1}(x)$, which is also a piecewise constant function, by averaging $\tilde{W}^{n+1}(x)$ over each cell:

$$W_j^{n+1} = \frac{1}{\delta x} \int_{(j-1/2)\delta x}^{(j+1/2)\delta x} \tilde{W}^{n+1}(x) dx. \tag{10.46}$$

We can show that the Godunov scheme can be written in the following *conservative* form:

$$\frac{W_j^{n+1} - W_j^n}{\delta t} + \frac{\Phi(W_j^n, W_{j+1}^n) - \Phi(W_j^n, W_{j-1}^n)}{\delta x} = 0, \tag{10.47}$$

where the flux vector is generally defined as

$$\Phi(W_j^n, W_{j+1}^n) = F(\tilde{W}_{j+1/2}^{n+1}). \tag{10.48}$$

The advantage of the conservative form is that it is valid over the entire domain of definition of the problem, even though the solution is discontinuous. This form is also used to derive approximate Riemann solvers. The exact form of the flux vector will be presented in the next section for the Roe solver.

Roe's Approximate Solver

The approximate solver of Roe is based on a simple and ingenious idea: the Riemann problem (10.7) at interface $\left(j+\frac{1}{2}\right)$ is replaced by the *linear* Riemann problem

$$\frac{\partial \tilde{W}}{\partial t} + A_{j+1/2} \frac{\partial \tilde{W}}{\partial x} = 0, \quad \tilde{W}(x, n\delta t) = \begin{cases} W_j^n, & x \leq \left(j+\frac{1}{2}\right)\delta x \\ W_{j+1}^n, & x > \left(j+\frac{1}{2}\right)\delta x \end{cases} \tag{10.49}$$

The first question raised by this approach is how to properly define the matrix $A_{j+1/2}$, which depends on W_j^n and W_{j+1}^n. This matrix is a priori chosen such that:

1. The hyperbolic character of the initial equation is conserved by the linear problem; hence $A_{j+1/2}$ admits a decomposition similar to (10.12):

$$A_{j+1/2} = P_{j+1/2}\, \Lambda_{j+1/2}\, P_{j+1/2}^{-1}. \tag{10.50}$$

In order to take into account the sign of the propagation speed of characteristic waves, it is useful to define the matrices following:

- $\text{sign}(A_{j+1/2}) = P_{j+1/2}\,(\text{sign}(\Lambda))\,P^{-1}_{j+1/2}$, where $\text{sign}(\Lambda)$ is the diagonal matrix defined by the signs of the eigenvalues λ_l: $\text{sign}(\Lambda) = \text{diag}(\text{sign}\lambda_l)$.
- $|A_{j+1/2}| = P_{j+1/2}|\Lambda|P^{-1}_{j+1/2}$, where $|\Lambda| = \text{diag}(|\lambda_l|)$.

2. The linear Riemann problem is consistent with the initial problem, i.e., for all variables u,
$$A_{j+1/2}(u,u) = A(u,u). \tag{10.51}$$
3. The numerical scheme is conservative, i.e., for all variables u and v,
$$F(u) - F(v) = A_{j+1/2}(u,v)(u-v). \tag{10.52}$$

For the practical calculation of the matrix $A_{j+1/2}$, the original idea of Roe was to express the conservative variables W and conservative fluxes $F(W)$ in (10.18) as quadratic forms of the components of the column vector $Z = \sqrt{\rho}(1,U,H)^t = (z_1,z_2,z_3)^t$:

$$W = \begin{pmatrix} z_1^2 \\ z_1 z_2 \\ \frac{1}{\gamma} z_1 z_3 + \frac{\gamma-1}{2\gamma} z_2^2 \end{pmatrix}, \quad F(W) = \begin{pmatrix} z_1 z_2 \\ \frac{\gamma-1}{\gamma} z_1 z_3 + \frac{\gamma-1}{2\gamma} z_2^2 \\ z_2 z_3 \end{pmatrix}. \tag{10.53}$$

Using the following identity, valid for arbitrary quadratic functions f, g,

$$(fg)_{j+1} - (fg)_j = \bar{f}(g_{j+1} - g_j) + \bar{g}(f_{j+1} - f_j), \quad \text{where} \quad \bar{f} = \frac{f_{j+1} + f_j}{2},$$

we can find two matrices $\bar{B}$ and $\bar{C}$ such that

$$\begin{cases} W_{j+1} - W_j = \bar{B}(Z_{j+1} - Z_j), \\ F(W_{j+1}) - F(W_j) = \bar{C}(Z_{j+1} - Z_j). \end{cases} \tag{10.54}$$

This implies that

$$F(W_{j+1}) - F(W_j) = (\bar{C}\ \bar{B}^{-1})(W_{j+1} - W_j), \tag{10.55}$$

which corresponds exactly to (10.52). Consequently, a natural choice for the matrix $A_{j+1/2}$ will be

$$A_{j+1/2} = \bar{C}\ \bar{B}^{-1}. \tag{10.56}$$

A remarkable property of this matrix (the reader is invited to derive it as an exercise!) is that it can be calculated from (10.8) by replacing the variables (ρ, U, H) with the corresponding *Roe's averages*

$$\bar{\rho}_{j+1/2} = R_{j+1/2}\rho_j,\ \bar{U}_{j+1/2} = \frac{R_{j+1/2}U_{j+1} + U_j}{1 + R_{j+1/2}},\ \bar{H}_{j+1/2} = \frac{R_{j+1/2}H_{j+1} + H_j}{1 + R_{j+1/2}},$$

$$\bar{a}^2_{j+1/2} = (\gamma - 1)\left(\bar{H}_{j+1/2} - \frac{\bar{U}^2_{j+1/2}}{2}\right), \text{ where } R_{j+1/2} = \sqrt{\frac{\rho_{j+1}}{\rho_j}}. \tag{10.57}$$

It is also remarkable that eigenvalue and eigenvector formulas (10.10) and (10.11) still apply to $A_{j+1/2}$ if one uses the corresponding Roe's averaged variables. This considerably simplifies the calculation of matrices $\text{sign}(A_{j+1/2})$ and $|A_{j+1/2}|$, which accounts for the popularity of Roe's approximate solver.

Once the matrix $A_{j+1/2}$ is defined, the upwinding in Roe's scheme follows the general principle of first-order upwind schemes applied to *linear* systems (see, for instance, Hirsch (1988) for more details). The flux in the general conservative form (10.47) becomes for Roe's solver

$$\Phi(W_j^n, W_{j+1}^n) = \frac{1}{2}\left\{F(W_j^n) + F(W_{j+1}^n) - \text{sign}(A_{j+1/2})[F(W_{j+1}^n) - F(W_j^n)]\right\}, \tag{10.58}$$

or, if we use (10.52),

$$\Phi(W_j^n, W_{j+1}^n) = \frac{1}{2}\left\{F(W_j^n) + F(W_{j+1}^n) - |A|_{j+1/2}[W_{j+1}^n - W_j^n]\right\}. \tag{10.59}$$

To summarize, Roe's scheme will be used in the form

$$W_j^{n+1} = W_j^n - \frac{\delta t}{\delta x}\left[\Phi(W_j^n, W_{j+1}^n) - \Phi(W_j^n, W_{j-1}^n)\right], \tag{10.60}$$

with the flux Φ given by (10.59); the matrix $|A_{j+1/2}| = P_{j+1/2}|\Lambda|P_{j+1/2}^{-1}$ will be calculated using Roe's averages (10.57) in (10.12) and (10.13).

Remark 10.2. Roe's scheme is first-order accurate in time and space.

Exercise 10.4. Use Roe's scheme (10.60) to solve numerically the shock tube problem (Sod's parameters). Compare to the numerical results previously obtained using centered schemes.

The results obtained using Roe's scheme are displayed in Fig. 10.8. Compared to centered schemes, the numerical solution is smooth, without oscillations. The shock wave is accurately and sharply captured, but the scheme proves too dissipative around the contact discontinuity, which is strongly smeared.

More accurate Riemann solvers can be derived in the framework of Godunov-type schemes by increasing the space accuracy. For example, we can use piecewise linear functions in steps 1 and 3 of the Godunov scheme to obtain solvers of second order in space. Several other approaches have been proposed in the literature to include more physics in the numerical discretization, leading to other classes of numerical methods, including TVD (total variation diminishing) and ENO (essentially nonoscillatory) schemes, which are now currently used to solve hyperbolic systems of PDEs. The reader who wishes to pursue the study of upwind schemes beyond this introductory presentation is referred to more specialized texts such as Fletcher (1991); Hirsch (1988); LeVeque (1992); Saad (1998).

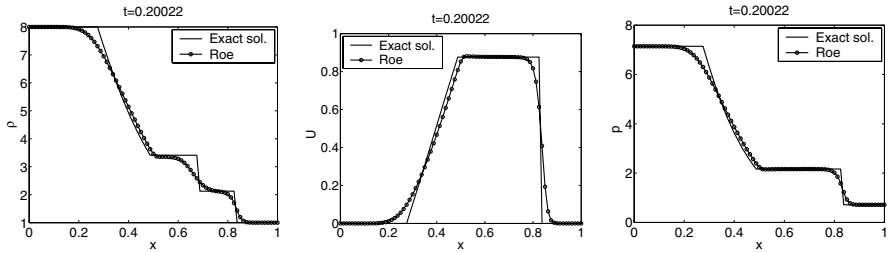

Fig. 10.8. Numerical computation of the shock tube problem (Sod's parameters) using Roe's approximate solver.

10.4 Solutions and Programs

The exact solution of the shock tube problem for a given time value t is computed in the script *HYP_shock_tube_exact.m*. The compatibility relation (10.29) is implemented as an implicit function (i.e., $f(x) = 0$) in the script *HYP_mach_compat.m*; this function is used as the first argument of the MATLAB built-in function `fzero` to compute the root corresponding to the value of M_s. The final solution, containing the discrete values for (ρ, U, p), is computed according to relations in Sect. 10.2.2. Note the use of the MATLAB built-in function `find` to compute the abscissas x separating the different regions of the solution.

The main program resulting from successively solving all the exercises of this project is *HYP_shock_tube.m*. After defining the input data (which are the parameters of regions (L) and (R)) as global variables, the space discretization is built and the solution is initialized using Sod's parameters. Three main arrays are used for the computation:

`usol(1:3,1:M)` to store the nonconservative variables $(\rho, U, p)^t$,
`w(1:3,1:M)` for the conservative vector $W = (\rho, \rho U, E)^t$,
and `F(1:3,1:M)` for the conservative fluxes $F(W)$.

The program allows one to choose among three numerical schemes: Lax–Wendroff, MacCormack, and Roe. When a centered scheme is selected, the value of the artificial dissipation is requested. The numerical solution is superimposed on the exact solution using the function `HYP_plot_graph` implemented in the script *HYP_plot_graph.m*. The most important functions called from the main program are:

- `HYP_trans_usol_w`: computes $W = (\rho, \rho U, E)^t$ from $usol = (\rho, U, p)^t$;
- `HYP_trans_w_usol`: computes $usol = (\rho, U, p)^t$ from $W = (\rho, \rho U, E)^t$;
- `HYP_trans_w_f`: computes $F = (\rho U, \rho U^2 + p, (E + p)U)^t$ from $W = (\rho, \rho U, E)^t$;
- `HYP_calc_dt`: computes $\delta t = \text{cfl} \cdot \delta x / (|U| + a)$ from $W = (\rho, \rho U, E)^t$.

All these functions are written with a concern for transparency with respect to the mathematical formulas. For this purpose, the vectorial programming

capabilities of MATLAB were used. Let us explain in detail this technique for the predictor step of the Lax–Wendroff scheme (see Fig. 10.4):
• the flux $F(W)$ is computed from W values for all $j = 1, \ldots, M$ components

```
F = HYP_trans_w_f(w);
```

• the artificial dissipation vector is added following (10.42); we use the MATLAB built-in function `diff` to compute differences $W_j - W_{j-1}$; these differences are computed along the rows of the array `w` and only for $j \geq 2$; according to the left-boundary conditions, the artificial dissipation vector will be completed by zeros for $j = 1$:

```
F = F-Ddx*[zeros(3,1) diff(w,1,2)];
```

• the intermediate solution $\tilde{W}$ is computed only for the components $j = 1, \ldots, M-1$:

```
wtilde=0.5*(w(:,1:M-1)+w(:,2:M))-0.5*dt/dx*(F(:,2:M)-F(:,1:M-1));
```

A similar MATLAB code will be written for the corrector step, having in mind that for this step, right-boundary conditions apply, and consequently, only the components $j = 2, \ldots M-1$ of W^{n+1} are computed:

```
Ftilde = HYP_trans_w_f(wtilde);
Ftilde=Ftilde-Ddx*[diff(wtilde,1,2) zeros(3,1)];
w(:,2:M-1)=w(:,2:M-1)-dt/dx*(Ftilde(:,2:M-1)-Ftilde(:,1:M-2));
```

Particular attention was devoted to the implementation of Roe's scheme, which requires a separate function `HYP_flux_roe` to compute the conservative flux Φ. In order to reduce memory storage, the flux at the interface $\left(j + \frac{1}{2}\right)$ is computed using this once (and once is not habit!) a `for` loop and several local variables that can be easily identified from mathematical relations. Note also that the analytical form (10.13) for $P^{-1}_{j+1/2}$ was used instead of the (time-consuming) MATLAB built-in function `inv`, which calculates the inverse of a matrix.

Chapter References

C. A. J. Fletcher, *Computational Techniques for Fluid Dynamics*, Springer-Verlag, 1991.
E. Godlewski and P.-A. Raviart, *Numerical Approximation of Hyperbolic Systems of Conservation Laws*, Springer-Verlag, 1996.
C. Hirsch, *Numerical Computation of Internal and External Flows*, John Wiley & Sons, 1988.
R. LeVeque, *Numerical Methods for Conservation Laws*, Birkhäuser, 1992.
M. Saad, *Compressible Fluid Flow*, Pearson Education, 1998.