

## CHAPTER 7

# Finite volume methods for hyperbolic equations

### 1. Basic aspects

We have seen that the appearance of discontinuities even when starting from smooth initial data is a generic situation for non-linear hyperbolic PDE's. To define what is meant by a solution in such cases the concept of weak solutions was introduced which involved integrating the discontinuous solution over some domain. This suggests that it might be advantageous to construct a numerical method which involves an integration step. It is also the case that a large class of PDE's of practical interest are derived from conservation laws in which a direct expression of the quantity being conserved might prove useful in a numerical algorithm. To evaluate a conserved quantity an integration step is again required.

Let us construct a numerical method based upon the above observations. Consider a scalar hyperbolic equation in conservation form

$$(1.1) \quad q_t + f(q)_x = 0$$

for which initial data

$$(1.2) \quad q(x, t = 0) = q_0(x)$$

is given over some domain, say  $0 \leq x \leq 1$ . We divide the domain into a number of cells

$$(1.3) \quad C_i = [x_{i-1}, x_i], \quad i = 1, \dots, n$$

with  $0 = x_0 < x_1 < \dots < x_n = 1$ . From our knowledge of the properties of equations of form (1.1), specifically the existence of characteristics, we expect an explicit time marching scheme to be an efficient numerical procedure. The basic problem faced in constructing such a procedure is to advance from time level  $t^n$  to time level  $t^{n+1}$ . Since we wish to involve conserved quantities in our formulation we are led to integrating (1.1) over  $[x_{i-1}, x_i] \times [t^n, t^{n+1}]$

$$(1.4) \quad \int_{x_{i-1}}^{x_i} \int_{t^n}^{t^{n+1}} [q_t + f(q)_x] dt dx = 0 .$$

This leads to

$$(1.5) \quad \int_{x_{i-1}}^{x_i} [q(x, t^{n+1}) - q(x, t^n)] dx + \int_{t^n}^{t^{n+1}} [f(q(x_i, t)) - f(q(x_{i-1}, t))] dt = 0 .$$

We introduce the quantities

$$(1.6) \quad Q_i^n = \frac{1}{h_i} \int_{x_{i-1}}^{x_i} q(x, t^n) dx, \quad i = 1, \dots, n$$

with  $h_i = x_i - x_{i-1}$ . Each  $Q_i^n$  expresses the total amount of the physical quantity  $q$  within cell  $C_i$  at time  $t^n$ . We also introduce

$$(1.7) \quad F_i^n = \frac{1}{k} \int_{t^n}^{t^{n+1}} f(q(x_i, t)) dt$$

with  $k = t^{n+1} - t^n$  which denotes the flux of quantity  $q$  through the interface between two cells at  $x = x_i$  over the time interval  $[t^n, t^{n+1}]$ . We thus obtain the scheme

$$(1.8) \quad Q_i^{n+1} = Q_i^n - \frac{h_i}{k} (F_i^n - F_{i-1}^n), \quad i = 1, \dots, n.$$

Such a scheme is known as a *finite volume scheme* and it satisfies our requirements that conserved physical quantities are used and that the derivation of the method allows for discontinuous functions  $q(x, t)$ . The quantities  $Q_i^n$  are known as *cell averages*, and they are taken to represent the value of  $q(x, t^n)$  somewhere within cell  $C_i$ , typically at the midpoint  $(x_{i-1} + x_i)/2$ . Our representation of the solution can be interpreted simply as a piecewise constant approximation

$$(1.9) \quad q(x, t^n) \cong Q_i^n, \quad x_{i-1} < x < x_i.$$

One ingredient is still required to obtain a complete method: a procedure to compute the fluxes  $F_i^n$  must be given. We know that hyperbolic equations have finite domains of dependence. If we choose a small enough time step the fluxes should only depend on the cell averages to the left and the right of the interface

$$(1.10) \quad F_i^n \cong \mathcal{F}(Q_i, Q_{i+1}), \quad i = 1, \dots, n-1.$$

The function  $\mathcal{F}(Q_i, Q_{i+1})$  defines an approximation of the true flux and is known as the *numerical flux* at interface  $x_i$ . The basic task faced in the construction of a finite volume method is to specify how the numerical fluxes are computed starting from knowledge of the cell average data.

## 2. Godunov methods

One important aspect of finite volume methods is that their construction follows quite closely the physical behavior of the problem solution. This can be exploited further in constructing numerical fluxes. Consider that we have the cell average data  $\{Q_i^n\}$ . If we interpret this as specifying a piecewise-constant approximation of  $q(x, t)$  we note that at each interface  $x = x_i$  a Riemann problem has been specified. If the solution  $q^\gamma(x_i, t)$  to the Riemann problem can be determined then an exact evaluation of the fluxes is possible by evaluating the physical flux function  $f$  at  $q^\gamma(x_i, t)$ .

$$(2.1) \quad F_i^n = f(q^\gamma(x_i, t^n)).$$

Such methods that use the true physical solution in construction of the numerical fluxes are known as Godunov methods and were introduced to study gas dynamics problems.

A very simple example is given by the scalar, constant-velocity, advection equation

$$(2.2) \quad q_t + uq_x = 0$$

solved on a uniform grid  $x_i - x_{i-1} = h$ . The solution to the Riemann problem is immediate

$$(2.3) \quad q^\gamma(x_i, t^n) = \begin{cases} Q_i^n & u > 0 \\ Q_{i+1}^n & u < 0 \end{cases}$$

leading to the numerical flux

$$(2.4) \quad f(q^\gamma(x_i, t^n)) = \begin{cases} uQ_i^n & u > 0 \\ uQ_{i+1}^n & u < 0 \end{cases}$$

and the scheme

$$(2.5) \quad Q_i^{n+1} = \begin{cases} Q_i^n - \frac{uh}{k} (Q_i^n - Q_{i-1}^n) & u > 0 \\ Q_i^n - \frac{uh}{k} (Q_{i+1}^n - Q_i^n) & u < 0 \end{cases}.$$

We recognize this as the upwind scheme derived in our study of finite difference methods. The interpretation is however a bit different. Whereas  $Q_i^n$  in the upwind finite difference scheme denoted the value of  $q$  at cell nodes  $x_i$ , here  $Q_i^n$  is the value at cell centers. The difference is however non-essential and we find that a large number of finite volume schemes have a close finite difference equivalent. This is especially useful in determining stability restrictions since we can apply the theory derived for finite difference methods.

In a Godunov method we use the exact physical flux  $f(q)$  to evaluate the numerical flux  $\mathcal{F}(Q_i, Q_{i+1})$ . This still involves an approximation, namely the piecewise constant approximation of  $q(x, t^n)$ , thus limiting the basic Godunov method to first order accuracy,  $\tau_i^n = O(h, k)$ . To obtain better accuracy we can introduce more accurate representations of  $q(x, t^n)$ . An obvious idea is to use a piecewise linear approximation

$$(2.6) \quad q(x, t^n) = Q_i^n + \sigma_i^n (x - x_{i-1/2}), \quad x_{i-1} < x < x_i$$

with  $x_{i-1/2} = (x_{i-1} + x_i)/2$ . Here we have assumed that  $Q_i^n$  represents the value of  $q(x, t^n)$  at the midpoint  $x_{i-1/2}$ . The slope  $\sigma_i^n$  may be constructed by interpolation between adjoining cell average values. A number of possibilities exist:

(1) *Downwind* or *Lax-Wendroff* slope

$$(2.7) \quad \sigma_i^n = \frac{Q_{i+1}^n - Q_i^n}{h}$$

(2) *Upwind* or *Beam-Warming* slope

$$(2.8) \quad \sigma_i^n = \frac{Q_i^n - Q_{i-1}^n}{h}$$

(3) *Centered* or *Fromm* slope

$$(2.9) \quad \sigma_i^n = \frac{Q_{i+1}^n - Q_{i-1}^n}{2h}$$

The names used for the slopes refer to the fact that when applied to the constant velocity advection equation each choice of slope leads to the corresponding finite difference scheme.

One problem associated with the desire for higher accuracy is that discontinuities can lead to non-physical oscillations in the numerical approximation. The

FIGURE 1. Non-physical oscillations introduced by piecewise-linear reconstruction. At  $t = t^n$  a shock profile at  $x = x_i$  is reconstructed using centered slopes. The profile is then advected to the new time level  $t^{n+1}$  and used to construct new cell averages, some of which are in error (open circles).

difficulty is easily understood if we consider the Riemann problem for constant-velocity advection

$$(2.10) \quad q_t + uq_x = 0$$

$$(2.11) \quad q(x, 0) = \begin{cases} q_l & x > x_i \\ q_r & x < x_i \end{cases}$$

Reconstruction of the sharp discontinuity leads to overshoots not present in the initial condition. These are then advected downstream and contaminate the numerical solution as shown in Fig. 1.

Non-physical oscillations can lead to a breakdown of the entire computation in some applications. A common example is encountered when  $q$  is some positive quantity physically but the numerical scheme oscillations lead to negative values. Typically the physical hypotheses used in constructing the algorithm are no longer valid and a runtime error results. This has led to the search for *high-resolution* algorithms that exhibit higher order accuracy (typically  $O(h^2, k^2)$ ) away from discontinuities and capture discontinuities without oscillations. Typically the accuracy of the algorithm is only  $O(h, k)$  near a discontinuity but this is not a problem since first-order accuracy is all we could expect. The procedure used in constructing such algorithms rests upon the identification of discontinuities in the initial data. If a discontinuity is identified a low-order, piecewise constant reconstruction of  $q(x, t)$  is used. Otherwise a higher-order, say piecewise-linear, reconstruction is used. The technique is known as a *slope-limiter method*, since it attempts to maintain zero slope near discontinuities. Various slope-limiters have been proposed and analyzed:

(1) *minmod limiter*

$$(2.12) \quad \sigma_i^n = \text{minmod} \left( \frac{Q_i^n - Q_{i-1}^n}{h}, \frac{Q_{i+1}^n - Q_i^n}{h} \right)$$

where the minmod function is defined by

$$(2.13) \quad \text{minmod}(a, b) = \begin{cases} a & \text{if } |a| < |b|, ab > 0 \\ b & \text{if } |b| < |a|, ab > 0 \\ 0 & ab \leq 0 \end{cases}$$

(2) *monotonized central-difference*

$$(2.14) \quad \sigma_i^n = \text{minmod} \left( \frac{Q_{i+1}^n - Q_{i-1}^n}{2h}, 2 \frac{Q_i^n - Q_{i-1}^n}{h}, 2 \frac{Q_{i+1}^n - Q_i^n}{h} \right)$$